**Authors response to Topical Editor:**

**The Editor, SOIL Discuss., Dear Dr. Olivier Evrard,**

**Thank you for your comments to MS:**

**Ref. No.: soil-2015-30,**

**Interactive comment on “Local versus field scale soil heterogeneity characterization – a challenge for representative sampling in pollution studies” by Z. Kardanpour et al.**

**Topical editors comments**

**Received and published: 08 Oct 2015**

**We have addressed each issue raised by the reviewer. We report the changes made in the MS, and use “track-changes” on the master manuscript.**

**Reviewer comments:**

Topical Editor Initial Decision: Reconsider after minor revisions

Thanks to the authors for their reply to the reviewers' comments. The discussion could be prolonged, but I think that the clarifications added by the authors in the text are sufficient at this stage.

However, I think that English should be improved throughout the manuscript before considering the final acceptance of the paper.

**A thorough grammer/spell check has been carried out and some language modifications have been performed all included in the MS with track changes (attached).**

I also recommend the authors to include the text provided as footnotes in the main text when revising their manuscript.

**The footnotes have been added in the main text.**

Comments of the co-authors should also be removed from the revised version of the text... I'm looking forward to receiving the revised version of your manuscript.

**Comments of co-authors have been removed from MS.**

**We hope the revision has improved the manuscript sufficiently to allow publication.**

**Sincerely yours,**

**Zahra Kardanpour, Ole Stig Jacobsen, Kim H. Esbensen**

# Local versus field scale soil heterogeneity characterization - a challenge for representative sampling in pollution studies

Zahra Kardanpour[1,2], Ole Stig Jacobsen[1,] Kim H. Esbensen[1,2]

1. Geological Survey of Denmark and Greenland (GEUS)
2. ACABS research group, University of Aalborg, campus Esbjerg (AAUE)

Abstract

This study is a contribution to development of a heterogeneity characterisation facility for 'next generation' soil sampling for example aimed at more realistic and controllable pesticide variability in laboratory pots in experimental environmental contaminant assessment. The role of soil heterogeneity on quantification of a set of exemplar parameters is described, including a brief background on how heterogeneity affects sampling/monitoring procedures in environmental pollutant studies. The Theory of Sampling (TOS) and variographic analysis has been applied to develop a more general fit-for–purpose soil heterogeneity characterization approach. All parameters were assessed in large-scale transect (1-100 m) vs. small-scale (0.1 –0.5 m) replication sampling point variability. Variographic profiles of experimental analytical results from a specific well mixed soil type concludes that it is essential to sample at locations with less than a 2.5 meter distance interval to benefit from spatial auto-correlation and thereby avoid unnecessary, inflated compositional variation in experimental pots; this range is an inherent characteristic of the soil heterogeneity and will differ among other soils types. This study has a significant carrying-over potential for related research areas e.g. soil science, contamination studies, and environmental monitoring and environmental chemistry.

Keywords: Heterogeneity characterisation, soil, variogram, large-scale, representative sampling, Theory of Sampling (TOS), MCPA, biomass, CFU

## 1. Introduction

All parameters for realistic, effective integration of variability over different scales are directly related to soil heterogeneity. There is a growing need for an integrated understanding of contaminant behaviour  in soil pollution studies (Arias-Estévez et al. 2008; Crespin et al. 2001; Johnsen et al. 2013; Li et al. 2006; Rodriguez-Cruzet al. 2006; Sørensen et al. 2006; Torstensson and Stark 1975; Rasmussen et al. 2005). In this context there is a missing link in the form of soil heterogeneity and its effective characterization, a feature often overlooked. Heterogeneity characterisation is the first, and in some cases the most important step, in soil contaminant studies, with relationships to various other aspects of environmental research and monitoring. A result of introducing more valid soil heterogeneity characterisation will be improved soil sampling procedures (Kardanpour et al. 2014; Kardanpour et al. 2015a,b), which in turn will contribute towards improved environmental fate study reliability (Boudreault et al. 2012; Chappell and Viscarra Rossel 2013; de Zorzi et al. 2008; Lin et al. 2013; Mulder et al 2013; Totaro et al. 2013).

Even in simple systems, the variability and risk for misinterpretation may have strong effect on parameterisation of processes relating to compound fate studies. These latter issues are being increasingly more recognised, as is the lack of appropriate methods to ensure documented representativity of the experimental batch volumes/masses with respect to the surrounding geology and biotic/abiotic soil characteristics. There is an urgent need for scientifically based experimental approaches, scale-up procedures and attendant principles for parameterisation of variability in these types of natural systems(Kardanpour et al. 2014; Adamchuk et al. 2011; Chappell and Viscarra Rossel 2013; de Zorzi et al. 2008).

Of particular interest will be a newly developed facility for empirical variability characterisation, which allows heterogeneity to be mapped at problem-dependent scale hierarchies. Based on this, it is possible to devise optimised sampling strategies that will allow fit-for-purpose representativity with respect to laboratory experiments depending of similar (or at least comparable) soil samples (pots). For this purpose the Theory of Sampling (TOS) delivers benchmarks measures expressing acceptable maximum heterogeneity limits and in the case of violations/transgressions furthers a complete understanding of how to identify and eliminate the detrimental sampling errors and provides tools for unambiguous mixing effectiveness. Combining these tools with specific knowledge on the relevant contaminant processes and compound properties, it will be possible to address the critical scale-dependent variability with increased confidence based on more realistic environmental parameter delineation.

We here introduce the variographic approach mainly for the cases of 1-D as a means of characterising the heterogeneity in one transect direction. Compared to the typical major variability in the Z-direction of soil depth profiles (soil horizons, layers and~~,~~ geological formations), the linear (1-D) or 2-D heterogeneity *within* soil horizons is significantly smaller, although this is exactly the kind of heterogeneity the present study aims at controlling.

Contrary to depth profile zonation a.o. the within-horizon 1-D and 2-D heterogeneity complies with the requirements of both TOS and geostatistics, i.e. spatial heterogeneity can be modelled variographically w.r.t. a physically meaningful *average level* (the inherent *stationarity assumption* in geostatistics). It is not meaningful to apply variographic characterisation on measurement series which contain discontinuous shifts, upsets or other disrupt, level changes, as is the prime characteristicon of soil depth zonations. The geostatistical tradition of modelling 2-D patterns based on projection onto a 1-D transect is also not free from debatable issues.~~[1]~~ The present authors do not wish to reject the 2-D geostatistical tradition with this statement, but in relation to the present matters this issue is better deferred to another occasion in which the 2-D modelling issue can be presented and discussed in full - this issue is a legitimate and interesting area for a fruitful debate. Entering into a 3-D geostatistical modelling realm, there are also here issues that in need of further discussion, e.g. the required minimum number of samples (measurements) needed for meaningful, and stable variogram calculation, The present foray only aims at presenting the power of a simple 1-D variogram characterisation operator based on TOS, upon which several versions of potential follow-up generalisations to 2-D and 3-D cases may be entertained.~~.~~ In the present context all *isotropic* 2-D heterogeneity patterns can be characterised comprehensively by a *randomly selected* 1-D direction (transect). In all sampling operations there should preferentially always be some sort of random selection involved, unless compelling geo-science reasons exists for choosing a direction related to the genesis of the specific heterogeneity met with, e.g. choosing a 1-D transect either along a dominant plow direction.

This study focuses on development of the necessary heterogeneity characterisation for sampling/monitoring and multi-parameter modelling practices, allowing implementation of realistic pesticide variability in experimental environmental contaminant assessment studies. The study has a significant carrying-over potential for related research areas e.g. soil science, contamination studies, and environmental monitoring.

We here focus on characterization of soil heterogeneity in terms of soil moisture, organic matter (LOI), biomass, microbiology, MCPA sorption and mineralization. The measured parameters are here used to illustrate effective management of heterogeneity; this particular location has been studied before in its own right. Following two earlier complementary studies, the focus below is on the necessary representativity demands when facing compound fate and mineralization studies (Kardanpour et al. 2014; Kardanpour et al. 2015). Field

~~[1] The present authors do not wish to reject the 2-D geostatistical tradition with this statement, but in relation to the present matters this issue is better deferred to another occasion in which the 2-D modelling issue can be presented and discussed in full - this issue is a legitimate and interesting area for a fruitful debate. Entering into a 3-D geostatistical modelling realm, there are also here issues that in need of further discussion, e.g. the required minimum number of samples (measurements) needed for meaningful, and stable variogram calculation, The present foray only aims at presenting the power of a simple 1-D variogram characterisation operator based on TOS, upon which several versions of potential follow-up generalisations to 2-D and 3-D cases may be entertained.~~

observation indicates a very well mixed sandy soil with almost no visual heterogeneity features. But the main issue is: does this apparent uniformity extend to all fate compounds? How is it possible to document that small sample masses, as typically used in pot experiments, are representative of their entire parent field, or to which sub-field scale? In other words, how can results and conclusions from laboratory experiments be reliably scaled-up and generalized to larger scales?

## 2. Materials and Methods

### 2.1. Location and sampling pattern

Fladerne Bæk is situated on the Karup peri-glacial outwash plain, Jutland, Denmark (56°N, 9°E) South West of Karup airport. The substratum is an arable sandy soil which has been tilled and cropped for more than 100 years, mainly supporting barley and potatoes during last 30 years. Thus this is a typical “very well mixed” soil type compared to the much more heterogenous~~eity~~ glacial clayey soil types treated in (Kardanpour et al. 2014). Soil samples were collected from the topsoil (A-horizon) in cylindrical cores; the present samples cover depth interval from 0-15 cm. The 60 m long sampling transect was roughly N-S. Each field sample included 200-300 grams of fresh soil. At the center of this transect at point 29, seven additionally samples form a roman grid (3 x 3) replication experiment with 0.3 meter equidistance.

The sampling rationale aimed at variographic fate characterization commensurate with a long profile at a scale length between 1m and 60 m; the roman square was intended as a basis for conventional statistical treatment (average and, standard deviation). This central sample layout serves as a small scale local ‘replication experiment’ compared with the transect dimensions (Kardanpour et al. 2014). In total 64 samples were collected, 57 samples from the long profile and nine samples of the small grid (two samples identical to two from the transect), one in between and three more in each side of transect with the same distance as the first three in the center of transect. The original fresh soil was kept frozen until use.

The primary sampling was specifically intended to correspond to current sampling traditions in the soil and microbiology communities. In other studies efforts have been made to optimize each individual field sample, for example with respect to the famous “Gy’s formula”, from which control over the so-called Fundamental Sampling Error is often sought. However, in the present study it is a major point to outline how the variographic approach a.o. lead to a procedure with which to characterize the magnitude of the total sampling-plus-analytical error and thus to be warned of the need to control (better) all the inherent sampling errors, see e.g. (DS 3077 ~~(~~2013) for a comprehensive introduction.

### 2.2. Theory of sampling and variographic analysis

The Total Analytical Error (TAE) is most often under acceptable control in the analytical laboratory as regards to both accuracy and precision. A sampling procedure must be both *correct* (ensures accuracy) and *reproducible* (ensures precision); TOS defines *representativity* in a rigid conceptual and mathematical approach. The critical issue is always, even for TOS-compliant sampling, that analytical results are but an *estimate* of the true (average) analytical grade of the lot sampled, because the aliquot is based on only a miniscule mass (0.5 – 2.0 g) compared to the entire field topsoil layer it is supposed to represent (typical mass/mass sampling ratios range $1:10^3$ to $1:10^9$). The full sampling-analysis process and its characteristics is therefore the only guarantee for the relevance and reliability of the aliquot brought forth for analysis. The fundamental TOS principles need to be applied to all appropriate scales along the entire 'field-to-aliquot' pathway, not only to the primary sampling, but in particular also to the successive stages of mass reduction in the laboratory before the ultimate analytical aliquot extraction. The only change in this multi-stage sampling chain is the operative scale (TOS principles and unit operations are scale-invariant). A comprehensive overview of all subsampling issues (laboratory mass reduction) has been published in (Petersen et al. 2004), which does not include the 'coning-and-quartering' approach, despite the fact that this approach has enjoyed some popularity e.g. for certain field applications to soils (Gerlach et al. 2002). However the coning-and-quartering approach has been severely criticized in the professional TOS literature, e.g. most recently in (Esbensen and Wagner 2014); from a representativity point of view ~~coning~~ this mass reduction approach must be strongly discouraged.

On the basis of a correct sampling and mass reduction regimen, it is possible to characterize the inherent auto-correlation between units of a process/lot or along 1-D transect (or transect). The *semi-variogram* (in this work referred to simply as the 'variogram') is employed to describe the variation observed between sample pairs as a function of their internal distance.

To calculate a variogram ~~a~~ sufficient number of units (increments/samples) are extracted equidistantly, spanning the process interval of interest, or the full transect length, as needed. The variogram is a function of a dimensionless, relative lag parameter, j, which is this distance between two units, the analytical results of which are compared. Full details of the variographic approach are described in (DS3077 2013; Esbensen et al. 2007; Esbensen et al. 2012a; Esbensen et al. 2012b; Gy 1998; Minkkinen et al. 2012; Petersen and Esbensen 2006; Petersen et al. 2005). Variograms may have apparent different specific appearances, but three fundamental characterizing features carry all the important information related to sampling errors and the heterogeneity along the transect in any-and-all variogram: the *sill*, the *range*, and the y-axis intercept, termed the *nugget effect*. Definitions of these features are given below.

The Sill is the y-axis value at which the variogram levels off and becomes horisontal. The Sill represents the total variance calculated from all experimental heterogeneity values. The sill corresponds to the overall maximum variance for the data series if/when calculated *without* taking their ordering into account.

The Range is the lag distance beyond which the variogram v(j) levels off and reaches a stable, constant Sill. Samples taken at lags below the Range are auto-correlated to a larger and larger degree as the lags gets smaller and smaller. The range carries critical information as to the local heterogeneity with respect to the objective of the present method development.

The Nugget Effect indicates the amount by which the variance differs from zero when a variogram is extrapolated backwards so as to correspond to what would have been a lag = 0. A lag equal to zero has no physical meaning, but it represents the hypothetical case of two samples extracted at the same time and location (indeed from exactly the same physical volume of the lot). Thus although ‘true replicates’ from the exact same soil location (volume) are not physically possible, the nugget effect never-the-less allows to estimate the corresponding discontinuous variance difference. This can be viewed as a collapse of the 1-D sampling situation (profile, transect) to a stationary sampling situation (small lots, 2-D and 3-D lots), see (DS3077 2013; Esbensen et al. 2007, 2012a, 2012b) for further descriptions.

The nugget effect has a special interest, it contains all sampling, - sample handling/processing and analytical errors combined, which makes up the total measurement uncertainty. A variogram with a high nugget effect w.r.t. the sill signifies a measurement system not in sufficient control (DS3077 2013; Esbensen and Wagner 2014).

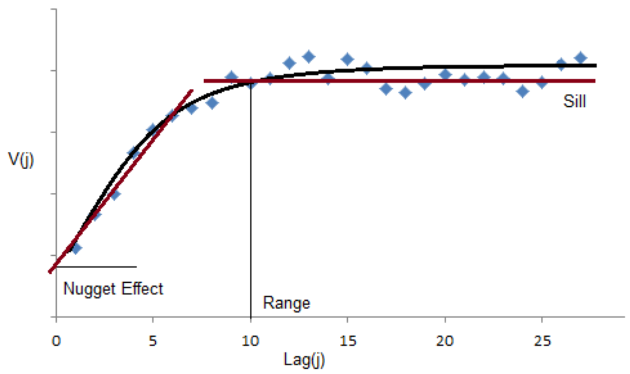


Figure 1. A generic variogram, schematically defining nugget effect, sill, and range. The illustration depicts an *increasing variogram*, which is the most often occurring type of variogram in the case of significant auto-correlation (for lags below the range)(Kardanpour et al. 2014). The nugget effect magnitude relative to the sill in this illustration is significant of an acceptable total measurement system, < 20%.

Variogram calculations are strongly influenced by *outliers* and/or *trends*. A valid variographic analysis often necessitates outlier deletion after proper recognition and description and occasionally also de-trending of the raw transects data if/when trends are dominant or severe. In this study the raw data transect was de-trended using a simple regression slope subtraction from the data set where needed.

### 2.3. Mass reduction/subsampling procedure

After the stored samples were thawed and accommodated for 20 °C for a week, before being processed further,~~. T~~the primary field sample size (200-300 gram) must be reduced to the analytical sample size (1-2 gram), not at all a trivial mass-handling issue. In order to provide representative sub-samples, TOS principles were applied scrupulously to all mass reduction steps. Thus samples were dried and macerated, or ground, where appropriate, and subsequently deployed in a longitudinal tray, forming a 1-D lot, using the soil-adapted bed-blending/cross-cut reclaiming technique described in detail in (Petersen et al. 2004) and (Kardanpour et al. ~~(~~2015~~b~~). These pre-blended micro-beds were cut by 10 randomly selected transverse increments along the elongated dimension which were aggregated, resulting in subsamples of 20-30 gram each. The exact same procedure was repeated in a secondary mass reduction step further down ending up with the final analytical mass (2 gram) for the wet samples analyses. This procedure has been ~~honed~~ applied to provide full representativity in ~~the course of this project specifically so as to do away with~~ samples and to exclude all of the post-primary-sampling errors in order better to be able to focus in the latter and the variogram deployment, *ibid*.

The remainders of the secondary sub-samples were air-dried for four days in lab temperature (20 °C), to be used in parallel sorption experiments. As a further scale-down iteration, a similar bed–blending/cross-cut reclaiming were used to provide analytical samples of 2 gram, also based on 10 increments each.

(Kardanpour et al. ~~(~~2015~~b~~) describe the “from-field-sampling-to-aliquot” pathway in full details, complete with an exhaustive pictorial exposé.

## 2.4. Analytical experiment methods

### MCPA Sorption

The sorption experiment started in glass vials with Teflon caps containing 1 g of the respective soils, and 9 ml of Milli-Q water. The vials were kept for 24 hours and then shaken in a horizontal, angled shaker prior to addition of 1 mL$^{14}$C-MCPA stock solution, with 10,000 dpm in each individual vials. Sorption experiments were performed with two initial concentrations: 1 and 100 mg MCPA/L. Sorption was determined for MCPA in all off the 64 soil samples, using $^{14}$C-labeled MCPA.

After adding the stock solution, the vials were incubated in the shaker for 48 hours and then placed vertically for another 48 hours, all at 20 °C. Subsequently 2 mL of the solution were transferred to the 2 mL Eppendorf micro-centrifuge tubes and centrifuged at 14,500x g for 7 min. Radioactivity in 1.5 mL supernatant was determined using a Wallac 1409 Liquid Scintillation Counter after mixing it with 10 mL OptiPhase Hisafe3 scintillation cocktail.

### MCPA Mineralization

Mineralization experiments were carried out in100mL glass jar with air tight lid. Two gram soil (wet weight) was placed in small plastic vials before adding 0.5 mL of $^{14}$C -labeled MCPA (5 mg MCPA kg$^{-1}$ soil) with a radioactivity of 2,000 dpm. In the glass jar a LSC vial was also placed containing 2 mL 0.2 M of NaOH as a $CO_2$ trap. The jars were incubated at 20°C for 14 days. Mineralization encountered as %-evolved $^{14}CO_2$ was measured at day 3, 7 and 14. The $CO_2$-traps were changed and replaced with a fresh trap at each sampling date.$^{14}C$ in the NaOH was measured as described in the sorption experiment by Liquid Scintillation Counting.

Biomass; substrate induced respiration (SIR)

The same set up as used for MCPA was used for the glucose mineralization with adding 0.5 mL$^{14}C$ -labeled glucose with 5000dpm to the 2 gram of soil. All other set up details, equipment and experimental design wereidentical. Alkaline traps were replaced with fresh alkaline traps and measured after 4 and 24 hours considering the rapid respiration of the glucose and $^{14}C$ measured as described in the sorption experiment by Liquid Scintillation Counting. Conversion into biomass were according to ( Dictor et al 1998; Tate et al. 1988).

Microbiology, Bacteria Colony Formation Units (CFU)

A suspension was made with 2 gram of soil into 200 mL sterile water and after shaking for 15 minutes, diluted with sterilized water ended in two different dilutions for each sample; with three and four order of magnitude To measure the soil microbiology, 1 mL of each sample were placed on a Petrifilm® (3M, Saint Paul, Minnesota, USA) sheet and CFU was counted after 3 and 7 days of incubation at 20°C.

# 3. Result

## 3.1. Geochemical profiling

In order to show the natural soil heterogeneity in a comparable format, Figures 2-5 illustrates the individual large-scale parameter transects; concentration vs location of the samples taken from the transect in Fladerne field. Also shown is the variation of the central *small-scale* replication samples is shown as mean concentration ± 2 SD with dashed horizontal lines in the figures. The large-scale variation of the soil moisture, loss on ignition (LOI) and the biomass content are to be compared to the small scale replication result for the same parameter in each graph, Figure 2.

The same comparison graph illustrated for the MCPA sorption in Figure 3 for two different initial MCPA concentrations, as it is clear, the soil sorption behavior shows different variation with different concentrations. The results of the MCPA mineralization of the soil in Figure 4 also show different variability ~~with~~ in different mineralization steps. The transect of the MCPA mineralization is illustrated for different mineralization steps: first three days, four to seven days and eight to fourteen days. The two latter periods shows rather a similar variation because these two periods are in the final part of the mineralization development, Figure 6.

The soil microbiology (Log (CFU/g soil)) transect after seven days of incubation is also illustrated in Figure 5.

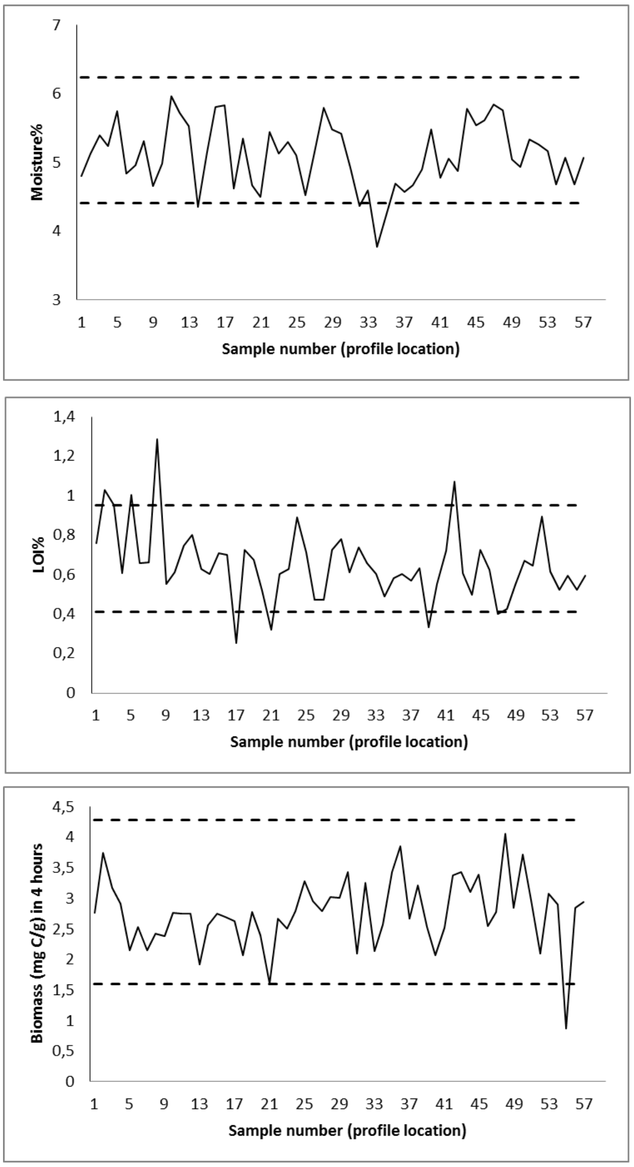


Figure 2. Fladerne Bæk, transects of soil moisture (%), LOI, and biomass (mg C/g); soil biomass vs. sample number (transect location). Dashed lines represent mean ± 2 SD of the small-scale replication experiment.

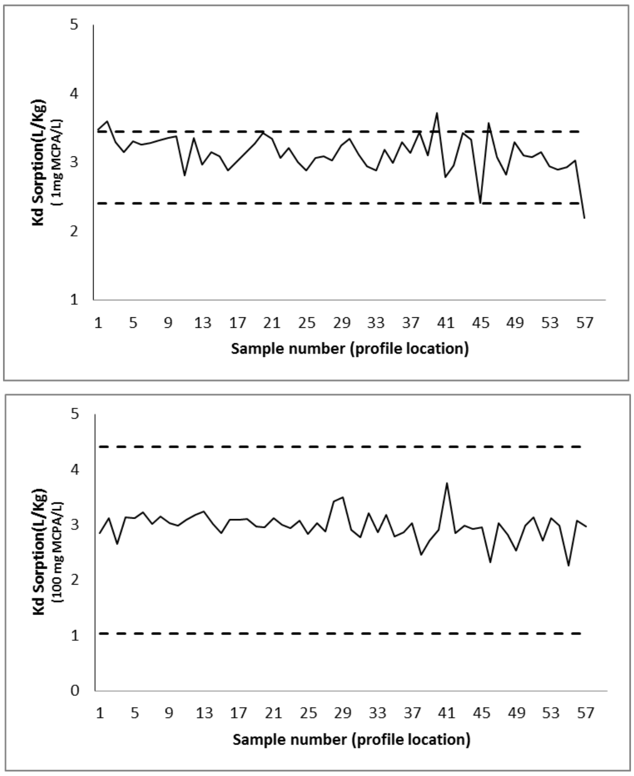


Figure 3. Fladerne Bæk, transects of $K_d$ MCPA sorption vs sample number (transect location), $K_{d,1}$: MCPA (1 mg/ L), $K_{d,100}$: MCPA (100 mg/ L). Dashed lines represent mean ± 2 SD of the small-scale replication experiment.

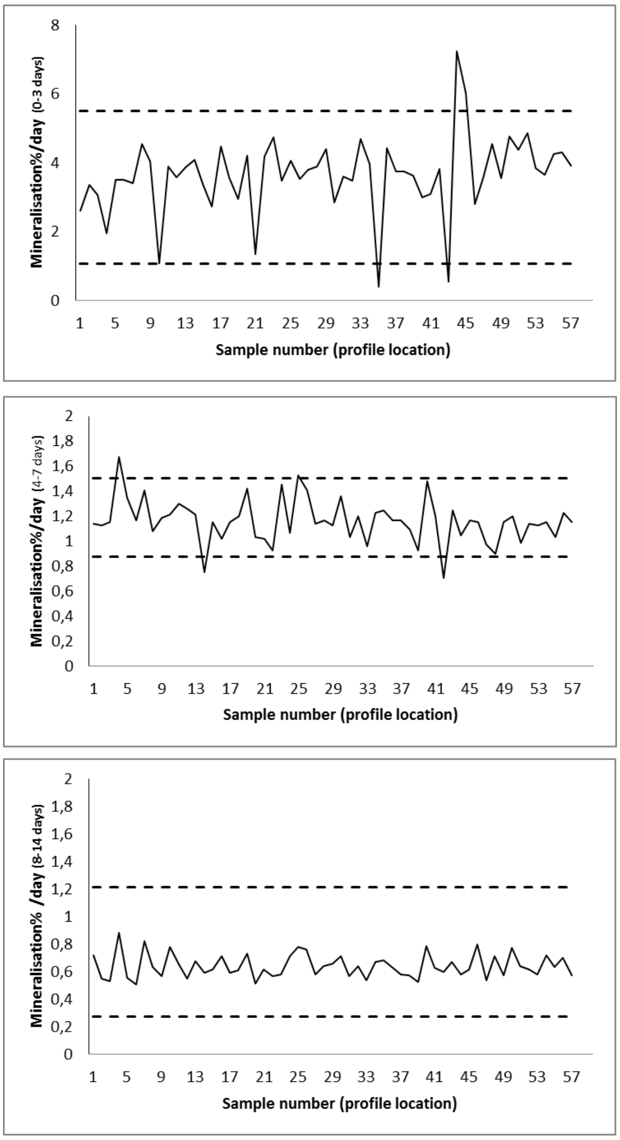


Figure 4. Fladerne Bæk, transects of MCPA mineralization in three different periods: 0-3days, 4-7days, 8-14 days vs. sample number (transect location). Dashed lines represent mean ± 2 SD of the small-scale replication experiment.

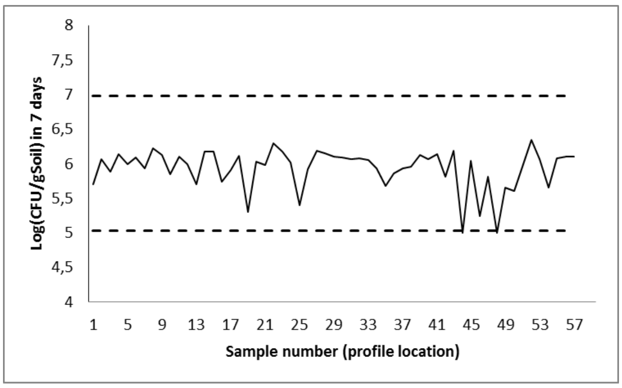


Figure 5. Fladerne Bæk, transects of log (CFU/g soil) vs sample number (transect location)

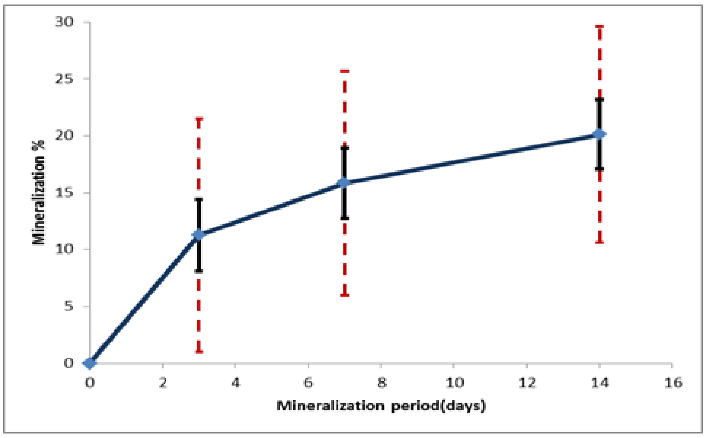


Figure 6. Average mineralisation rate for all 57 samples: Error bars are based on the standard deviation (solid bars) and the range of the whole sample set (stippled bars)

The Fladerne case represents an inherently very well mixed soil type, which has been under the plow for up to 100 years[2]. The consequence of taking care of this, low-heterogeneity end of the spectrum, is that there is a limit to the degree of transect heterogeneity to be expected, as indeed witnessed in Fig.s 2-5, where concentrations only comparatively rarely deviate outside the +/- 2 STD of the central Roman square design employed. This specific soil- and tilling history feature must not lead to untoward confusion and illegitimate generalizations however. It is the general applicability of the variographic approach which is illustrated here, as it happens, on a very well-mixed substratum. Our parallel study showcases the approach on a significantly more heterogeneous case, in which the central Roman square does not bracket most of the transect concentration manifestations.This case was selected to represent the one (almost extreme) end of a spectrum (only little inherent heterogeneity) from which to compare a whole spectrum of increasingly more heterogeneous soil types, horizon and geological formations. Our own studies went a fair distance in this direction as possible with the (Kardanpour et al. 2015), but obviously many, even more heterogeneous cases exist and are on record in the literature.

## 3.2. Experimental variograms

Prior to variogram calculation, all parameters have been checked for outliers and trends, Figures 2-5. Variograms have been calculated with using large scale experimental transects without model fitting of the variogram parameters. This is common in geostatistics, but not used here as TOS' variogram approach is not used for kriging but solely for heterogeneity

[2] ~~The consequence of taking care of this, low-heterogeneity end of the spectrum, is that there is a limit to the degree of transect heterogeneity to be expected, as indeed witnessed in Fig.s 2-5, where concentrations only comparatively rarely deviate outside the +/- 2 STD of the central Roman square design employed. This *specific* soil- and tilling history feature must not lead to untoward confusion and illegitimate generalizations however. It is the general applicability of the variographic approach which is illustrated here, as it happens, on a very well-mixed substratum. Our parallel study showcases the approach on a significantly more heterogeneous case, in which the central Roman square does not bracket most of the transect concentration manifestations.~~

characterization and interpretation.

Two different behaviors can be observed as displayed by two parameters groupings, the increasing Min1, LOI and Biomass variograms at the top, versus the reminder of parameters, which show a strongly similar form and behavior, Figure 7. As the sill levels represent the maximum parameter variation along the transect, parameters Min1, LOI and Biomass clearly display the highest transect variability. All variograms are of the increasing type with a distinct nugget effect. Following (DS3077 2013), the %-age nugget effect in relation to the sill, termed $RSV_{1\text{-dim}}$, is an expression of the total measurement uncertainty MU including TSE (Esbensen and Wagner 2014). In the present study this $MU_{total}$ quality index ranges from 15% ($K_d$, 100) to 75% (Min1). There is thus an appreciable difference concerning the possibility to measure and characterize soil heterogeneity along the transect, ranging from very good to very poor. This facility for total measurement uncertainty validation is a powerful TOS benefit, with a wide carrying-over potential to many other sciences and application fields. This feature ~~was~~ is described in full in (Esbensen and Romanach, 2015~~);~~ ~~((~~Kardanpour et al. 2015) in which, by the way, the 1-D transect of the present study appears in the form of a 1-D industrial process measurement series, illustrating the surprising generality of the variogram approach - modeling and interpretation ~~of the variogram from such disparate data types are identical,~~ and showing the way for application ~~also~~ to natural process in the geo-science and environmental science realms.

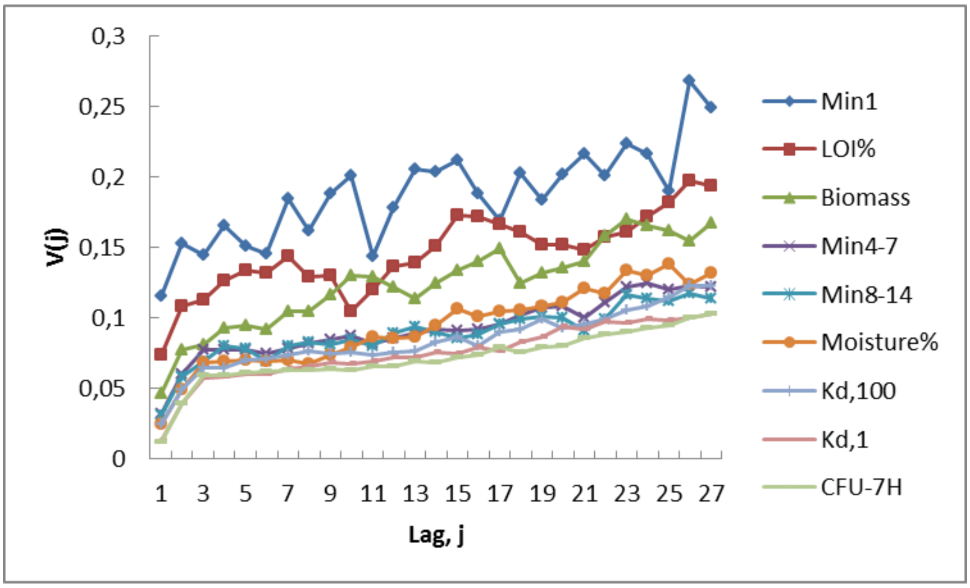


Figure 7. Synoptic variogram of all parameters in the present study comparing nugget effect, sill and range levels

Applying the multivariate data analysis approach developed in the former studies (Kardanpour et al. 2014; Kardanpour et al. 2015), i.e. using the variograms as the input (X-matrix) to a Principal Component Analysis (PCA) with no centering and no scaling (see further below), the first component is found to represent 99% of the total variogram variance over all parameters, making it easy to find the average range characterizing the heterogeneity of the Fladerne transect, ca. 5 meter. Figure 8 shows the loadings for PC components 1 and 2, displayed in a fashion that mimics a *spectrum*. As expected the PC-1 loadings delineates a general variogram shape, in fact presenting the *average* of all variograms in Figure. 7. The PC-2 loadings accounts for deviations herefrom, as caused by the individual variograms (mainly expressing a higher or lower average slope), a general feature, markedly overprinted by random deviations. This component models the set of different slopes of the individual variograms, and it accounts for less than 1% total variance, but never-the-less lends itself easily to be interpreted as the well-known spectroscopic 'tilting' signature, (Martens & Næs 1991).

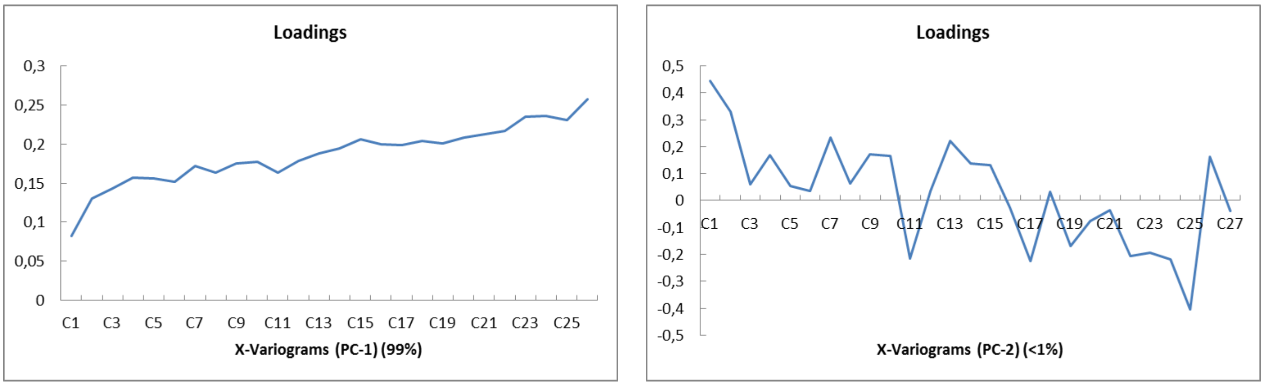


Figure 8. PCA ($X_{variogram}$) loading plot for PC-1 (left) and PC-2 (right). The $X_{variogram}$ matrix has not been subjected to pre-treatment before PCA (no centering, no scaling). The range of the average variogram shape as represented by the PC1 loadings is ca. 5meters.

In our earlier studies, (Kardanpour et al. 2014), can be found a discussion *pro et contra* pre-treatment of an X-matrix made up of variograms. When basing variograms on heterogeneity contributions (a one-to-one transformation of the original analytical concentrations), this issue becomes moot, as this transformation is already performing what amounts to scaling. In the present paper we therefore did not apply centering, opting for the easily interpreted and useful appearance of the average variogram shape, Figure 8 (left).

## 4. Discussion

Aiming for a general approach to soil heterogeneity characterisation, a set of naturally occurring organic, anthropogenic and biota parameters were studied at scales from 1 to 60 ~~(100)~~ m to be compared with other, for example minerogenic parameters (see further below). The first step is always inspection of the raw data set with respect to potential outliers and/or trends. In the present study the geochemical parameter transects show no outliers and no strong trends, Figures 2-5.

The experimental design allows comparison of the small-scale replication variability (classic statistics) and large-scale variability. All transects can for example be directly compared with the level and variation at the small-scale experiment (less than 1 meter), by the pertinent mean ±2 SD. In figures 2-5 the variation of the parameters in any selected small scale window cannot be overestimated to the large scale, indeed it cannot be also obtained from a small scale replication study deviation estimate. This is just for visual orientation however and not to be confused with the nugget effect, a much more general characterisation of the small(est) scale variability pertaining to below lag = 1, summing up and averaging this information for all the sample pairs in the transect.

Any short interval on a transect Figures 2-5 can be considered as a small scale study in its own right. In this context there is a clear difference between the empirical variability in different segments *along* each transect: the local variability does not necessarily extend to larger scales. This has an important practical conclusion: any local small-scale sample collection cannot be generalised to larger scales. Unwitting or un-reflected scaling-up of small scale experimental organic, anthropogenic and biota fate and mineralization results will bring an inflated uncertainty outside experimental control. The mineralisation parameters which show different variation behaviour in the different mineralisation steps send an important message regarding studies concerning time-dependent characterisations. A similar difference is observed for MCPA sorption with different concentrations, i.e. when studies are concerned with concentration-dependent phenomena.

The *general* local variability behaviour is however well captured as the below-range part of the general variogram loading spectrum for PC1. The variogram is able to generalise the common local scale behaviour. With TOS, there is synoptic information residing in the range, sill and nugget effect for each individual parameter. Whenever heterogeneity variograms display a range, this relates to the ease and risk associated with attempting to secure field samples with minimum variability: Sampling with smaller inter-increment lag distances than the range makes it possible to use the inherent auto-correlation between samples in a beneficial fashion.

From the earlier studies (Kardanpour et al. 2014; Kardanpour et al. 2015) the overall conclusion was only to employ *composite sampling*. In the present context this means that, wherever practically possible, increments should only be collected with a maximum of half the observed range as a means to avoid unnecessary compositional variability effects due to the inherent soil scale heterogeneity. It follows that in order to minimize the total sampling error, increments must be sampled with a maximum lag of 0.5*range, *preferentially* smaller. In the present soil variograms a general range of 5 meters is observed for multivariate variographic approach of the parameters, Figure 8. It is evident that a thorough mixing of the selected set of increments is mandatory to sample locations with less than 2.5 meters distance in between; for other soil types/analytes other numerical magnitudes apply.

The variograms show different behaviour with respect to mineralisation stages. This is expected from the slower rate of the mineralisation in the latter stages, Figure 6. The later stages display a flat variogram that only represent little auto-correlation between sample locations, Figure 7, and the low sill level representing low variation along the transect. As it is common in environmental studies, results of the mineralisation are mostly reported in terms of the accumulated mineralisation rate (see Figure 6 as an example), i.e. results that are mostly affected by the first stages of the mineralisation.

Most of the variograms level off quickly after only a few lags (range ca. 5 meters) followed by a flat (or slightly increasing) trend, while first step of MCPA mineralisation, biomass and LOI show more markedly increasing variograms, Figure 7.

The CFU sill level is lower than natural organic and anthropogenic compounds indicating lower variability of soil microbiology at the large scale(s). This can be compared with results from a series of other large-scale studies on different microbial communities for different anthropogenic and natural compound mineralization, which also showed that microbial biomass seem to be stable intrinsic parameter of longer periods. (Sørensen et al. 2003; Bending et al. 2001; Bending et al. 2003; Walker et al. 2001).

It is always a matter for discussion when theoretically anticipated correlations between the physiochemical/microbial activities fail to appear in specific real-world case studies. The more complex compounds have shown a more irregular, patchy fashion of decaying due to more specific microbial communities (but still generally isotropic in nature). Analysis of soil parameters rarely gives a clear pattern; this seems to be associated to a number of not-included or unknown parameters, resulting, in some cases in a high degradation potential, but low elsewhere (Sørensen et al. 2003; Rasmussen et al. 2005; Bending et al. 2001; Walker et al. 2001). Upon reflection this is no mystery however, but simply a result of local soil heterogeneity, which cannot be formulated or predicted based on the physiochemical biological or microbial correlation of the properties of soil in large scale studies. A variographic heterogeneity characterization at all scales is thus a beneficial pilot experiment able to focus on the relevant heterogeneities characterizing individual, or group of parameters in their proper scale-dependent relationships.

Summing up the results of all measured parameters studied here, for environmental purposes and objectives related to soil parameters at field scale, it is advantageous to employ a variographic heterogeneity characterisation as a pilot study. Results here ~~from~~ will lead to a comprehensive understanding of the spatial variability and auto-correlation of the parameters in the field.

The results from the present study show that for well-mixed sandy soil it is recommended to sample locations with less than 2.5 meters inter-distance in between, preferentially smaller. It is necessary to conduct a similar variographic pilot experiment in order to outline the relevant scale-heterogeneity characteristics for other soil types, which unavoidably will tend to show more irregular spatial heterogeneity patterns – each principal soil type will in principle be characterised by a specific range, but there is a further caveat. Each analyte may in fact display its own, more or less specific range, as witnessed above, as well as by a plethora of studies in the literature. When controlling the spatial heterogeneity is of the essence, the logical solution is to design the sampling according the analyte with the *smallest* range, i.e. the most heterogeneously distributed analyte – this will by necessity also take care of all other analytes with higher ranges. If emphasis is on sampling costs (a not totally unlikely alternative scenario that may, or may not clash with other requirements of which only one really matters though: representativity) it is a comforting thought that all analytes are measured on the same final aliquot. By carefully optimising the primary field sampling according to the principles presented here, all analytes will be measured with the same, optimal relevance, indeed w.r.t. the same representativity. If sampling is done right from the start, there are no extra costs – while the opposite is a very different case, as should be abundant clear.

Results from a parallel study on the *minerogenic* compounds for the same Fladerne field (Kardanpour et al. 2014) show a similar soil heterogeneity compared to the present *anthropogenic* compounds. The nugget effect for most of the minerogenic compounds are of the same order of magnitude as those for the anthropogenic compounds, i.e. the total measurement system and procedures (sampling/handling/processing/analysis) pass all the quality criteria for representative sampling established in the recent sampling standard (DS3077 2013).

In cases where the next step in studies might be assessment of the main factors driving the spatial heterogeneity of soil contamination analytes for example, the 1-D (or 2-D X-Y) approach advocated here, will only serve as a basis for proper selection of experimental material to be taken to the laboratory - upon which further considerations will focus on, say, the potential factors involved in contaminant input and transport a.o. Note that these latter processes manifest themselves primarily in the Z-direction, where it is by no means a given that application of the same variographic approach (or geostatistical modelling) will necessary give meaningful results. ~~(see earlier footnote).~~

## 5. Conclusions

A pilot experiment aimed at an intrinsic 1-D soil heterogeneity characterization is a critical success factor for laboratory studies relying on field samples to provide the experimental pots, which for replicate and comparative study objectives need to be as similar as at all possible. As a case study the variographic results for sandy soils show that the distance between two sample spot must be less than 2.5 meters for the present set of organic compounds and soil type. Specific soil types and/or other analytes will in principle display different ranges and nugget effects, and hence our call for systematic deployment of the *variographic pilot experiment*, from which can be derived all necessary information for designing an optimal sampling plan e.g. identifying the analyte with the smallest range (for significantly correlated analytes). For the case of well-mixed soil components, a general PCA-approach for modelling a whole set of variograms may be useful in addition to individual analyte consideration.

Without this types of information, experimental fate study work is essentially devoid a valid basis as regards interpretation, scale-up and scientific generalisation of the experimental results back to the field scale. A large-scale 1-D transect sampling can reveal the inherent heterogeneity at all scales from the smallest local sampling equidistance up to the maximum experimental length scale studied. Variographic analysis was here employed successfully to soil heterogeneity at scales between 1 and 100 meters, other scenarios may require other numerical parameters, while the general approach remains identical.

The TOS-guided variogram pilot study approach illustrated here has a substantial carrying-over potential to geochemistry and environmental science, as well as other application areas. It is even applicable to *dynamic systems*, i.e. to natural or technological processes in these realms.

Acknowledgements

The authors gratefully acknowledge the Danish Research Council for Ph.D. stipend funding (Stipend No. 562/06-18-10028(6)) to ZK and valuable laboratory services and assistance provided by GEUS, Dept. of Geochemistry. The authors are grateful to GEUS' management for a positive attitude with respect to development and application of TOS' principles of proper representative sampling, which are not always recognized in geosciences communities to the extent necessary in general, alas neither at GEUS specifically.

References:

Adamchuk, Viacheslav I., Raphael a. Viscarra Rossel, David B. Marx, and Ashok K. Samal. 2011. Using Targeted Sampling to Process Multivariate Soil Sensing Data. *Geoderma* 163 (1-2): 63-73. ¨

Arias-Estévez, Manuel, Eugenio López-Periago, Elena Martínez-Carballo, Jesús Simal-Gándara, Juan-Carlos Mejuto, and Luis García-Río. 2008. The Mobility and Degradation of Pesticides in Soils and the Pollution of Groundwater Resources. *Agriculture, Ecosystems & Environment* 123 (4): 247–60.

Badawi, Nora, Anders R Johnsen, Jan Sørensen, and Jens Aamand. 1999. Centimeter-Scale Spatial Variability in 2-Methyl-4-Chlorophenoxyacetic Acid Mineralization Increases with Depth in Agricultural Soil. *Journal of Environmental Quality* 42 (3): 683–89.

Bending, Gary D, Suzanne D Lincoln, R Sebastian, J Alun W Morgan, Jens Aamand, Sebastian R Sørensen, and Allan Walker. 2003. In-Field Spatial Variability in the Degradation of the Phenyl-Urea Herbicide Isoproturon Is the Result of Interactions between Degradative Sphingomonas Spp . and Soil pH. *App.Environ. Microbiol.* 69(2): 827-834.

Bending, Gary, Eve Shaw, and Allan Walker. 2001. Spatial Heterogeneity in the Metabolism and Dynamics of Isoproturon Degrading Microbial Communities in Soil. *Biology and Fertility of Soils* 33 (6): 484–89.

Boudreault, Jean-Philippe, Jean-Sébastien Dubé, Mirela Sona, and Eric Hardy. 2012. Analysis of Procedures for Sampling Contaminated Soil Using Gy's Sampling Theory and Practice. *The Science of the Total Environment* 425: 199–207.

Chappell, Adrian, and Raphael a. Viscarra Rossel. 2013. The Importance of Sampling Support for Explaining Change in Soil Organic Carbon. *Geoderma* 193-194: 323-325.

Crespin, M a, M Gallego, M Valcárcel, and J L González. 2001. Study of the Degradation of the Herbicides 2,4-D and MCPA at Different Depths in Contaminated Agricultural Soil. *Environmental Science & Technology* 35 (21): 4265–4270.

De Zorzi, Paolo, Sabrina Barbizzi, Maria Belli, Ales Fajgelj, Radojko Jacimovic, Zvonka Jeran, Umberto Sansone, and Marcel van der Perk. 2008. A Soil Sampling Reference Site: The Challenge in Defining Reference Material for Sampling. *Applied Radiation and Isotopes :* 66 (11): 1588–1591.

Dictor Marie- Christine, Laurent Tessier and Guy Soulas. 1998. Reassessment of theK EC Coefficient of the Fumigation ± Extraction Method in a Soil Profile. *Soil Biology and Biochemistry* 30 (2): 119-127.

DS3077. 2013. Representative Sampling/ Horizontal Standard. *Danish Standard Authority* 44:1–38.

Esbensen, Kim H., Hans Henrik Friis-Petersen, Lars Petersen, Jens Bo Holm-Nielsen, and Peter P. Mortensen. 2007. Representative Process Sampling — in Practice: Variographic Analysis and Estimation of Total Sampling Errors (TSE). *Chemometrics and Intelligent Laboratory Systems* 88 (1): 41–59.

Esbensen, Kim H., Claudia Paoletti, and Pentti Minkkinen. 2012a. "Representative Sampling of Large Kernel Lots I. Theory of Sampling and Variographic Analysis." *TrAC Trends in Analytical Chemistry* 32: 154–164.

Esbensen, Kim H., Claudia Paoletti, and Pentti Minkkinen. 2012b. Representative Sampling of Large Kernel Lots III. General Considerations on Sampling Heterogeneous Foods. *TrAC Trends in Analytical Chemistry* 32: 178–184.

Esbensen, Kim H., and Claas Wagner. 2014. Theory of Sampling (TOS) versus Measurement Uncertainty (MU) – A Call for Integration. *TrAC Trends in Analytical Chemistry* 57: 93–106.

Esbensen, Kim H., and Rodolfo J. Romanach. (2015) Counteracting soil heterogeneity sampling for environmental studies (pesticide residues, contaminants transformation) - TOS is critical. Proceedings 7.th World Conference on Sampling and Blending (WCSB7), p.205-209.

Gerlach, Robert W., David E. Dobb, Gregory a. Raab, and John M. Nocerino. 2002. Gy Sampling Theory in Environmental Studies. 1. Assessing Soil Splitting Protocols. *Journal of Chemometrics* 16 (7): 321–328.

Ghafoor, Abdul, Nicholas J. Jarvis, Tomas Thierfelder, and J Stenström. 2011. Measurements and Modeling of Pesticide Persistence in Soil at the Catchment Scale. *The Science of the Total Environment* 409 (10) 1900-1908.

Gy, P. M. 1998. *Sampling for Analytical Purposes*. 1st ed. Chichester, West Sussex, UK: John Wily & Sons, 172pp, ISBN: 978-0-471-97956-2.

Johnsen, Anders R, Philip J Binning, Jens Aamand, Nora Badawi, and Annette E Rosenbom. 2013. The Gompertz Function Can Coherently Describe Microbial Mineralization of Growth-Sustaining Pesticides. *Environmental Science & Technology* 47 (15): 8508–8514.

Johnsen, Anders R, Bjarne Styrishave, and Jens Aamand. 2014. Quantification of Small-Scale Variation in the Size and Composition of Phenanthrene-Degrader Populations and PAH Contaminants in Traffic-Impacted Topsoil. *FEMS Microbiology Ecology,* 88: 84-93.

Kardanpour, Zahra, Ole S. Jacobsen, and Kim H. Esbensen. 2014. Soil Heterogeneity Characterization Using PCA (Xvariogram) – Multivariate Analysis of Spatial Signatures for Optimal Sampling Purposes. *Chemometrics and Intelligent Laboratory Systems* 136:24–35.

Kardanpour, Zahra, Ole S. Jacobsen, and~~Jakobsen, O.S. &~~ Esbensen, Kim.H. 2015 Counteracting soil heterogeneity sampling for environmental studies (pesticide residues, contaminants transformation) - TOS is critical. Proceedings 7.th World Conference on Sampling and Blending (WCSB7), p.205-209.

Li, B G, J Cao, W X Liu, W R Shen, X J Wang, and S Tao. 2006. Geostatistical Analysis and Kriging of Hexachlorocyclohexane Residues in Topsoil from Tianjin, China. *Environmental Pollution* 142 (3): 567–575.

Lin, Qinghuo, Hong Li, Wei Luo, Zhaomu Lin, and Baoguo Li. 2013. Optimal Soil-Sampling Design for Rubber Tree Management Based on Fuzzy Clustering. *Forest Ecology and Management* 308: 214–222.

Martens, Harald, and Tormod Næs. 1991. *Multivariate Calibration*. Edited by John Wiley & Sons. Chichester, West Sussex, UK, 438p, ISBN: 978-0-471-93047-1.

Minkkinen, Pentti, Kim H. Esbensen, and Claudia Paoletti. 2012. Representative Sampling of Large Kernel Lots II. Application to Soybean Sampling for GMO Control. *TrAC Trends in Analytical Chemistry* 32: 165–177.

Mulder, V.L., S. de Bruin, and M.E. Schaepman. 2013. Representing Major Soil Variability at Regional Scale by Constrained Latin Hypercube Sampling of Remote Sensing Data. *International Journal of Applied Earth Observation and Geoinformation* 21: 301–310.

Petersen, Lars, Casper K. Dahl, and Kim H. Esbensen. 2004. Representative Mass Reduction in Sampling—a Critical Survey of Techniques and Hardware. *Chemometrics and Intelligent Laboratory Systems* 74 (1): 95–114.

Petersen, Lars, and Kim H Esbensen. 2006. Representative Process Sampling for Reliable Data Analysis — a Tutorial. *Journal of Chemometrics 19:* 625–647.

Petersen, Lars, Pentti Minkkinen, and Kim H. Esbensen. 2005. Representative Sampling for Reliable Data Analysis: Theory of Sampling. *Chemometrics and Intelligent Laboratory Systems* 77 (1-2): 261–277.

Rasmussen, Jim, Jens Aamand, Per Rosenberg, Ole S Jacobsen, and Sebastian R Sørensen. 2005. Spatial Variability in the Mineralisation of the Phenylurea Herbicide Linuron within a Danish Agricultural Field: Multivariate Correlation to Simple Soil Parameters. *Pest Management Science* 61 (9): 829–837.

Rodriguez-Cruz, M. Sonia, Julie E. Jones, and Gary D. Bending. 2006. Field-Scale Study of the Variability in Pesticide Biodegradation with Soil Depth and Its Relationship with Soil Characteristics. *Soil Biology and Biochemistry* 38 (9): 2910–2918.

Rosenbom, Annette E, Philip J Binning, Jens Aamand, Arnaud Dechesne, Barth F Smets, and Anders R Johnsen. 2014. Does Microbial Centimeter-Scale Heterogeneity Impact MCPA Degradation in and Leaching from a Loamy Agricultural Soil?. *The Science of the Total Environment* 472:90–98.

Sørensen, Sebastian R, Gary D Bending, Carsten S Jacobsen, Allan Walker, and Jens Aamand. 2003. Microbial Degradation of Isoproturon and Related Phenylurea Herbicides in and below Agricultural Fields. *FEMS Microbiology Ecology* 45 (1): 1–11.

Sørensen, Sebastian R, Anne Schultz, Ole S Jacobsen, and Jens Aamand. 2006. Sorption, Desorption and Mineralisation of the Herbicides Glyphosate and MCPA in Samples from Two Danish Soil and Subsurface Profiles. *Environmental Pollution* 141 (1): 184–194.

Tate K.R., Ross D. J. and Feltham C.W. 1988. A Direct Extraction Method to Estimate Soil Microbiology C: Effects of Experimental Variables and Some Different Calibration Procedures. *Soil Biology and Biochemistry* 20: 329–335.

Thurman, Elisabeth A Scribner. 2008. A Decade of Measuring , Monitoring , and Studying the Fate and Transport of Triazine Herbicides and Their Degradation Products in Groundwater , Surface Water , Reservoirs , and Precipitation by the US Geological Survey. *The Triazine Herbicides*, 451-475.

Torstensson, N T L, and J Stark. 1975. The Effect of Repeated Applications of 2 , 4-D and MCPA on Their Breakdown in Soil. *Weed Research* 15: 159–164.

Totaro, Sara, Paola Coratza, Caterina Durante, Giorgia Foca, Mario Li Vigni, Andrea Marchetti, Mauro Marchetti, and Marina Cocchi. 2013. Soil Sampling Planning in Traceability Studies by Means of Experimental Design Approaches. *Chemometrics and Intelligent Laboratory Systems* 124:14–20.

Walker, A, M Jurado-Exposito, G D Bending, and V J Smith. 2001. Spatial Variability in the Degradation Rate of Isoproturon in Soil. *Environmental Pollution* 111 (3): 407–415.