# Quality Assessment of Soil Pollution Monitoring: Focus on Representativeness<sup>1</sup>

André Desaules<sup>1</sup>

<sup>1</sup> Soil Monitoring Consultant, Attinghausenstr. 15, CH-3014 Bern, Switzerland

Correspondence: André Desaules, Attinghausenstr. 15, CH-3014 Bern, Switzerland. Tel: 41-31-331-5719. E-mail: andre.desaules@bluewin.ch

Received: January 2, 2016Accepted: January 22, 2016Online Published: February 26, 2016doi:10.5539/ep.v5n1p51URL: http://dx.doi.org/10.5539/ep.v5n1p51

# Abstract

Soil monitoring data can be no better than the quality of the monitoring system they stem from. Quality assessment (QA) of soil monitoring requires reliable and comprehensive quality assessment and quality control (QA/QC) schemes including (1) the selection of parameters and measurement quality related to (2) space and (3) time. It can be presented by a synoptic diagram with three axes based on a table with quality criteria. The two major quality parameters are the degrees of resolution (precision) and representativeness (bias), whereas the latter does not yet include parameter selection and soil sampling.

As a result the quality of soil monitoring is usually greatly overestimated. This finding is supported by examples and practical recommendations are given. Since full representativeness for the three aspects of soil monitoring is a fiction in practice, their biases have to be quantified completely, continuously and reliably. The most important challenges are to quantitatively assess and control the representativeness of primary soil sampling and to improve it.

Keywords: soil monitoring, soil sampling, quality assessment, representativeness, bias, soil contaminants

## 1. Introduction

Environmental modelling and risk assessment require sufficient reliable data. Soil chemical monitoring produces such data – but are its quantity and quality sufficient? Since monitoring programs are only as good as the information (data) they produce (Houston & Hiederer, 2009), comprehensive and reliable quality assessment of soil monitoring is crucial.

Quality is a relative notion defined by the degree of suitability for a specific purpose. The present purpose is the spatio-temporal monitoring of soil contaminants. The more specifically the purpose is defined, the more precisely the quality requirements of a particular soil monitoring system can be determined and its quality performances assessed. For environmental monitoring they include the three aspects (1) object of monitoring – here soil contaminants, (2) space and (3) time.

The objective is to outline a scientifically based, comprehensive, coherent and consistent framework for the assessment of soil monitoring for contaminants. Major problems of soil monitoring quality are presented and discussed and some recommendations are made.

## 2. Soil Monitoring Quality: Criteria and Representation

To comprehensively assess soil monitoring, quality assessment criteria for all three monitoring aspects (object, space and time) are required. These are quantified by QA parameters where available. QA criteria initially developed for soil sampling by Nothbaum et al. (1994) are outlined below:

- Resolution: What is the density of parameters and measurements? QA parameters for this are: (1) Numbers of parameters, (2), precision (= random errors) of the measured parameters, (3) spatial measurement density (= support) and (4) temporal measurement density (= periodicity).

<sup>&</sup>lt;sup>1</sup> Based on a presentation given at the International Workshop on Monitoring Soil Properties at Different Scales; Thünen Institute, Braunschweig, 19<sup>th</sup>/20<sup>th</sup> November 2014.

- Representativeness: To what extend do the parameters and measurements represent the defined monitoring purpose, area and time span? The QA parameter for the degree of representativeness is known as bias and can be related to (1) parameters, (2) space and (3) time. Bias is a measure of proportionality between the monitoring entity and the parameters or measurements.

- Indicator value: What is the degree of indication, sensitivity or robustness of parameters and measurements?

- Sufficiency: Are the numbers of parameters and measurements adequate through space and time?

- Reliability: To what degree are the selected parameters and measurements trustworthy? Quantifiable QA parameters are measurement uncertainty (= precision and bias) and comparability.

- Validity: What is the scope of application of the parameters and measurements? Here methodological and spatio-temporal boundary conditions have to be considered.

- Efficiency/Economy: What is the relation between investment and results, can it be optimized?

- Overall QA: Weighted estimate of all QA parameters.

The degrees of resolution (precision) and representativeness (bias) are the two major QA parameters. The other QA criteria are further QA specifications.

Figure 1 shows a representation of an environmental QA as a scenario. The QA consists of a quality diagram with three axes to visualize the monitoring quality and eventually to compare the quality of different monitoring systems. The basis is the table with the rating of the QA criteria. The optimum monitoring quality (= 100 %) is earmarked and has to be specified previously explicitly for all three monitoring aspects (parameters, space, time). The general mandate e.g. to monitor soil contamination country-wide is not specific enough. The assessment and weighting of the individual QA criteria and parameters are fundamental problems which require further development and objectification. However, the final choices and decisions are generally not scientific but of a pragmatic or normative nature.

## 3. Choice of Parameters and Measurement Quality

On the basis of the scientific criteria reproducibility and traceability alone, objects such as soil contamination cannot be monitored in a holistic way, but need to be fragmented into measureable parameters or indicators. There is the scientific necessity of simplification with the inherent danger of oversimplification. For example, in the current practice of soil contamination monitoring, some priority contaminants are pragmatically selected from a myriad of (partly still unknown) soil contaminants on the basis of their toxicity and frequent high concentration levels. However, total concentrations of trace metals in soil are poor predictors of toxicity (Smolders et al., 2009) because trace metal concentrations are expressed as chemical elements and not as effectively occurring chemical species which are toxicologically relevant. Furthermore, buffer and aging effects in soils and the exposure of target organisms and organs to contaminants often remain unconsidered. The same holds for cumulative and antagonistic effects of soil contaminants.

The assessment of the chemical measurement quality in soils raises the problem of adequate QA/QC. Initially QA/QC schemes were developed for industrial production, to ensure and trace a required minimum product quality. Only later were QA/QC schemes adopted for environmental purposes, but they are still largely inadequate, because they usually cover only the measurements in the laboratory as a routine procedure and exclude the primary soil sampling in the environment, which is an essential part of the whole measurement chain. An exception are the studies performed by Gustavsson et al. (2006) and Boudreault et al. (2012). It is a serious problem for QA/QC that, in contrast to industrial production, the natural environment cannot be standardized. This makes the development of reliable environmental QA/QC schemes much more complex, labour-intensive and costly. Reliable QA/QA schemes are required to be complete chains of evidence. Weak chain links diminish the validity of conclusions and missing chain links are a source of misinterpretations and faulty conclusions.

The measurement quality is quantified by the two QA parameters precision (random error) and bias (systematic error) for each measurement step and summed up as variances and mean bias to the overall uncertainty. Note that the bias – at least for soil sampling – is not a constant (systematic) value as the misleading term "systematic error" suggests (Gy, 1982, 2004; Pitard, 1993). The precision is determined by sufficient replicate samples and replicate measurements. The bias as a measure of the degree of representativeness or the deviation from the "true" mean or consensus value, respectively, has to be determined by reference values, since all environmental measurements are relative measurements which require reference values for calibration.



0 – 33% poor quality; 34 – 66% moderate quality; 67 – 100% good quality or "good soil monitoring practice"

Figure 1. Representation of a comprehensive environmental quality assessment (scenario)

For chemical soil analysis in the laboratory more or less suitable certified reference materials exist. But for soil sampling there are no certified reference sampling plots. This is the reason why the soil sampling bias cannot be determined directly by measurements, but can only be estimated by means of Pierre Gy's theory of correct soil sampling (Gy, 1982, 1989, 2004; Pitard, 1993; François-Bongarçon; Gy, 2002). Table 1 gives an overview of the quantification procedures of the measurement quality in chemical soil monitoring.

TT 1 1 1	010			1	· ·	•	1 . 1	• 1	•,	•
Table I	()A/(	)( measur	ements	and	estimates	1n	chemical	SOIL	monite	oring
10010 1.	V11/V	20 mousui	ententes	unu	cotiniates		chienneur	5011	monne	Jung

Measurement steps	Precision	Bias		
	(Replicate measurements)	(Reference values)		
In the field:				
Primary soil sampling	Measurements	Estimates		
In the laboratory:				
Subsampling	Measurements	Estimates		
In the laboratory:				
Chemical analysis	Measurements	Measurements		

Due to missing reference values to determine the sampling bias, in practice the primary sampling bias is usually not taken into account. As a result the overall measurement uncertainty is expressed one to two orders of magnitude smaller than it really is, because by far the biggest measurement errors stem from non representative primary soil sampling (Gy, 2004; Gustavsson et al., 2006; Boudreault et al., 2012). It is a serious shortcoming that in the QA/QC handbook of "Measurement uncertainty arising from sampling" (Ramsey & Ellis, 2007) not one example considers the primary soil sampling bias. QA/QC schemes which do not include the primary soil sampling bias massively overestimate chemical soil data quality.

Correct representative primary soil sampling is very demanding and means that the chemical proportionality of the sampling area is reflected in the mean value of the analysed aliquot. This can imply a proportional mass reduction of more than ten orders of magnitude (Desaules, 2012). Although a soil chemical value can never be better than the quality of the previous soil sampling (Thompson & Ramsey, 1995), in current practice the quality of primary soil sampling is still massively deficient compared to chemical measurements in laboratories.

#### 4. Spatial Measurement Quality

The spatial distribution of soil chemical concentrations is neither random nor systematic, but corresponds to a complex natural- and human impact-induced distribution pattern. To be representative a correct soil sampling procedure has to respect this distribution pattern. An approximation to the real distribution pattern is only possible via preliminary variographic studies. These are random soil sampling transects covering the monitoring area to detect correlations of chemical soil concentrations with environmental factors such as, for example, geology or land use. On this basis adequate stratified soil sampling strategies can be derived. The best possible strategy consists of stratified systematic soil sampling with adequate spatial resolution based on variographic studies. A resulting mean value is considered experimentally as free of bias or in other words, representative if it is robust compared to different adequate soil sampling strategies.

A soil sampling site or sample is representative, if the mean concentration of a particular analyte (e.g. Cd) is free of bias for a specific area. The bias is a quantitative measure of the degree of representativeness. Figure 2 exhibits plots of maximum degree of representativeness or minimum bias, respectively for Zn, Cd and Pb from a test area of a comparative soil sampling study (Desaules et al., 2001).



Figure 2. Sampling plots with minimum deviations of the mean topsoil concentrations for Zn, Cd and Pb from a test area of 5°200 m<sup>2</sup> with two land-use types as a quantitative measure for minimum bias or maximum degree of representativeness, respectively: Compound samples of 25 increments from 0 to 20 cm soil depth for each plot of 10 by 10 m (compiled from Desaules et al., 2001)

Only a few sampling plots are fully representative (bias = 0) either for the whole test area or for one of the two land-use types and the plots are usually representative for only one single trace metal. A representative soil monitoring site for multiple chemical substances is a fiction. A monitoring site or sampling plot can *a priori* be characteristic (typical) for a certain area, if it is specified accordingly (e.g. deciduous forest site on loess), but whether the site is representative, can only be validated *a posteriori*. Unfortunately even in environmental science the term "representativeness" is often ominously confused with "characteristic" or "typical".

#### 5. Temporal Measurement Quality

The primary prerequisite for reliable measurement times series is long-term stable measurement systems. To monitor temporal chemical changes in soils, a cost- and labour-intensive method was developed with temporal replicate analysis of site-specific reference samples which ensures measurement stability of the laboratory analysis (Desaules, 1998, 2012). The reliability and validity of the method was confirmed by a subsequent report (Meuli et al., 2014).

However, so far unresolved remains the problem of the temporal primary soil sampling bias, (Desaules, 2012, 2012a), which is by far the main contribution to the overall measurement uncertainty and it usually remains undetected because of incomplete QA/QC schemes and practice, as stated above. Especially fatal is the fact that, unlike laboratory analysis, primary soil sampling is a destructive process and cannot be corrected retrospectively. That means that primary soil sampling errors in measurement time series are irreversible.

A comparative study of two time series with different measurement periodicities at the same monitoring sites revealed that the effect of measurement periodicity exceeded that of measurement period (Desaules et al., 2004), as is illustrated by the example in Figure 3. This finding is a strong indicator for artefacts produced by soil sampling. This means that measured trends in time series studies may be a result of soil sampling artefacts, as long as this cannot be refuted by a comprehensive and reliable QA/QC.



Figure 3. Time series comparisons for Pb and Hg of different lengths and measurement periodicities (10 years with 3 measurements and 3 years with 6 measurements) on the same soil monitoring site under permanent grassland (from Desaules et al., 2004)

Since the soil conditions are more or less unstable through time at monitoring sites (e.g. soil moisture conditions, soil swelling and shrinking, oxidation of soil organic matter), this leads - despite stabilized laboratory measurement systems - to the measurement of temporal chemical soil concentration changes which have nothing to do with soil contamination.

The concentrations of chemical substances as contaminants have different depth gradients in soils. For this reason temporal differences in sampling depth and/or sample material integrity (more or less complete samples) result in measurable soil chemical changes (Desaules, 2012), as shown in Figure 4.

Furthermore, dynamic soil processes can lead to measurable changes in chemical soil concentrations within a relatively short time. This is especially the case for soil erosion or accumulation, but also for many soil mixing processes such as mechanical soil tillage (e.g. ploughing), bioturbation (e.g. by earth worms, voles, plant roots), kryoturbation (alternating freezing and thawing) (Desaules, 2012). Figure 5 shows the complexity of temporal chemical soil concentration changes due to different depth gradients and soil mixing depth.



Figure 4. Temporal soil sampling bias by selective loss of sample material (from Desaules, 2012)



Figure 5. Temporal chemical concentration changes in the reference topsoil (0-20 cm) by pedoturbation at different depths and concentration gradients (from Desaules, 2012a)

These are all reasons, except for measurement errors, why the validation of measured temporal chemical soil concentration changes by flux balances (inputs by atmospheric deposition, agrochemicals and cropping outputs) is not reliable (Desaules, 2012 a). The comparison of soil measurements and flux balance for Cu in Figure 6 agrees only at two out of six monitoring sites.

Finally, Table 2 shows effects of changing boundary conditions on the temporal change of measured Cu concentrations.



guide value = 40 mg/kg

Figure 6. Comparison of soil monitoring measurements (soil depth 20 cm) with flux balances for Cu on six grassland sites after 10 years (from Desaules, 2012a)

Table 2. Measured Cu concentrations in topsoils (0-20 cm) at five vineyard monitoring sites and concentration changes after 5 years (from Desaules, 2012a)

Sampling campaign	Mean	Mean Vineyard sites						
	value	#5	#55	<b>#96</b>	#101	#106		
_		(mg/kg)						
t1	433	314	244	654	520			
t2	301	313	223	456		213		
Conc. changes	-132	-1	-21	-198	-307			

The relatively important decreases of Cu concentrations measured in the topsoil (0-20 cm) had the following causes: #55 selective loss of sample material (Figure 4), #96 dilution of topsoil Cu concentration by deep ploughing (Figure 5), destruction of site #101 and substitution by site #106. Due to all these causes the mean value of the five sites decreased by 30 % in only five years. The effect on the mean value of temporal changes of boundary conditions at single monitoring sites is known as the "Will-Rogers phenomenon". The given example reveals how erroneous interpretations of mean value changes can be without adequate metadata (Desaules, 2012a).

#### 6. Conclusions and Recommendations

Biased soil monitoring is neither representative nor reliable. As long as soil monitoring is not under complete and traceable QA/QC, it cannot be claimed to be unbiased, representative and reliable. Current QA/QC schemes of regional soil chemical monitoring systems are usually not complete, because they do not explicitly consider either the representativeness (bias) of primary soil sampling, or the selection of monitoring parameters. This leads to a massive and uncontrolled overestimation of the monitoring quality. The current soil monitoring practice is (still) inadequate for assessing soil contamination, not only at a particular site (Desaules, 2012b) but also through space and time. Insufficient or altogether unknown representativeness (bias) is the greatest obstacle to "good soil monitoring practice" and at the same time the greatest challenge.

Some recommendations are given below on how to cope with the three fundamental biases in soil monitoring which have been discussed above:

1) Parameter bias: The use of bio-indicators is a method to generally screen eco-toxic symptoms, but they do not allow an explicit identification of the specific cause. The paradigm of Paracelsus from the  $16^{th}$  century – "All substances are poisonous and nothing is without poison; the dose alone makes the poison." – needs to be refined for the current practice of risk assessment. Toxicity is not only a matter of substance concentration (e.g. chemical elements) but also of the occurring chemical species and their particular availability to specific organs and organisms. Thus the safest way to avoid parameter bias is to limit risk assessment to specific chemical species,

soil conditions and organisms and to avoid uncontrolled extrapolation and generalization. The chronic effects of soil contamination in the ubiquitously contaminated environment are yet to be identified.

2) Spatial bias: As demonstrated above (Figure 2) representative soil monitoring sites for multiple chemical substances are a fiction which should be abandoned in practice. A solution would be to use carefully selected, well characterized and documented intensive soil monitoring sites with reliable cause/effect analysis (process studies). On this basis, assisted by soil contamination functions and appropriate metadata (Desaules, 2012a), spatial pedotransfer functions (e.g. McBratney et al., 2002) can be developed for specific chemical substances and randomly validated by reliable soil chemical measurements.

3) Temporal bias: One possible measure to prevent bias by different soil sampling depths and/or sample materialisation (Figures 4 and 5) is the conditioning of soil moisture by irrigation before soil sampling. More accurate but even much more invasive, labour-intensive and costly would be the extraction of drilling cores from previously *in-situ* frozen soil (Hofmann, 2000). Relevant temporal environment or management changes and the successive improvement of measurement time series lead to inconsistent time series which are other sources of temporal bias (Table 2). To detect sources of temporal bias it is important to produce short range test time series (Figure 3), to validate the soil monitoring data with an alternative, reliable and independent method (Figure 6) and to collect sufficient appropriate and reliable metadata. So far 31 sets of metadata have been identified for soil contamination monitoring (Desaules, 2012a).

Since perfect representativeness (bias = 0) is a fiction in practice for all three soil monitoring aspects discussed (object, space, time), it is essential to quantify the correspondent biases completely, continuously and reliably. The priority challenge is to improve and quantitatively assess and control the degree of soil sampling representativeness (bias). This would represent a change of paradigm in current soil monitoring practice.

## Acknowledgments

I am indebted to the Swiss Federal Offices for the Environment and Agriculture, where I coordinated the Swiss Soil Monitoring Network (NABO) for 27 years. I agree with *Karl Popper* that all I actually know about soil monitoring is based on a succession of empirical falsifications. Therefore I want to thank all the numerous colleagues and authors who induced me to make errors and still more the ones who helped me to detect and to correct them. Most stimulating among the latter are my former colleagues *Ruedi Dahinden, Konrad Studer, Thomas Keller* and *Pierre Gy* with his numerous writings on sampling theory and sampling practice.

# References

- Boudreault, J. P., Dubé, J. S., Sona, M., & Hardy, E. (2012). Analysis of procedures for sampling contaminated soil using Gy's Sampling Theory and Practice. *Sci. Tot. Environ.*, 425, 199-207. http://dx.doi.org/10.1016/j.scitotenv.2012.03.013
- Desaules, A. (1998). Zeitliche Veränderungen von Schadstoffgehalten im nationalen Bodenbeobachtungsnetz (NABO) der Schweiz. *Mitt. Dtsch. Bodenkundl. Ges.*, 87, 319-322.
- Desaules, A. (2012). Measurement instability and temporal bias in chemical soil monitoring: sources and control measures. *Environ. Monitoring Assessment, 18*, 487-502. http://dx.doi.org/10.1007/s10661-011-1982-1
- Desaules, A. (2012a). The role of metadata and strategies to detect and control temporal data bias in environmental monitoring of soil contamination. *Environ. Monitoring Assessment, 184*, 7023-7039. http://dx.doi.org/10.1007/s10661-011-2477-9
- Desaules, A. (2012b). Critical evaluation of soil contamination assessment methods for trace metals. *Sci. Tot. Environ.*, 426, 120-131. http://dx.doi.org/10.1016/j.scitotenv.2012.03.035
- Desaules, A., Keller, A., Schwab, P., & Dahinden, R. (2004). *Analysen von Zeitreihen und Ursachen gemessener Konzentrationsveränderungen von Schwermetallen und Phosphor in Böden auf Dauerbeobachtungsflächen.* Agroscope FAL Reckenholz, Eidg. Forschungsanstalt für Agrarökologie und Landbau, 8046 Zürich. Retrieved from http://www.bafu.admin.ch/bodenschutz/10161/10187/index.html? lang=de
- Desaules, A., Sprengart, J., Wagner, G., Muntau, H., & Theocharopoulos, S. (2001). Description of the test area and reference sampling at Dornach. *Sci. Tot. Environ.*, 264, 17-26. http://dx.doi.org/10.1016/S0048-9697(00)00609-4
- François-Bongarçon, D., & Gy, P. (2002). The most common error in applying 'Gy's Formula' in the theory of mineral sampling, and the history of the liberation factor. *J. South African Institute Mining Metallurgy*, 475-479.

- Gustavsson, B., Luthbom, K., & Lagerkvist, A. (2006). Comparison of analytical error and sampling error for contaminated soil. J. Hazardous Materials, B138, 252-260. http://dx.doi.org/10.1016/j.jhazmat.2006.01.082
- Gy, P. M. (1982). Sampling of particulate materials theory and practice (2nd ed.). Amsterdam: Elsevier.
- Gy, P. M. (1998). Sampling methods for analytical purposes. Chichester: Wiley.
- Gy, P. M. (2004). Sampling of discrete materials a new introduction to the theory of sampling: I Qualitative approach. *Chemometrics and Intelligent Laboratory Systems*, 74, 7-24. http://dx.doi.org/10.1016/S0169-7439(04)00167-4
- Hofmann, B. A. (2000). In situ ground freezing to obtain undisturbed samples of loose sand for liquefaction assessment. J. Geotech. and Geoenv. Engineering, 126, 979-989. http://dx.doi.org/10.1061/(ASCE)1090-0241(2000)126:11(979)
- Houston, T. D., & Hiederer, R. (2009). Applying quality assurance procedures to environmental monitoring data: a case study. *J. Environ. Monitoring*, *11*, 774-781. http://dx.doi.org/10.1039/b818274b
- McBratney, A. B., Minasny B., Cattle, S. R., & Vervoort, R. W. (2002). From pedotransfer functions to soil inference systems. *Geoderma*, 109, 41-73. http://dx.doi.org/10.1016/S0016-7061(02)00139-8
- Meuli, R.G., Schwab, P., Wächter, D., & Ammann S. (2014). Nationale Bodenbeobachtung (NABO) 1985–2004. Zustand und Veränderungen der anorganischen Schadstoffe und Bodenbegleitparameter. Bundesamt für Umwelt, Bern. Umwelt-Wissen Nr. 1409: 94 S. Retrieved from www.bafu.admin.ch/publikationen
- Nothbaum, N., Scholz, R., & May, T. W. (1994). *Probenplanung und Datenanalyse bei kontaminierten Böden*. Berlin: Erich Schmidt Verlag.
- Pitard, F. F. (1993). Pierre Gy's sampling theory and sampling practice. Boca Raton: CRC Press.
- Ramsey, M. H., & Ellison, S. L. R. (2007). Measurement uncertainty arising from sampling: a guide to methods and approaches. Eurachem/CITAC Guide, Eurachem. Retrieved from https://www.eurachem.org/index.php/publications/guides/musamp
- Smolders, E., Oorts, K., Van Sprang, P., Schoeters, I., Janssen, C. R., McGrath, S. P., & McLaughlin, M. J. (2009). Toxicity of trace metals in soil as affected by soil type and aging after contamination: Using calibrated bioavailability models to set ecological soil standards. *Environ. Toxicology and Chemistry*, 28, 1633-1642. http://dx.doi.org/10.1897/08-592.1
- Thompson, M., & Ramsey, M. H. (1995). Quality concepts and practices applied to sampling an exploratory study. *Analyst, 120,* 261-270. http://dx.doi.org/10.1039/an9952000261

#### Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (http://creativecommons.org/licenses/by/3.0/).