Evaluation of the reproducibility of sampling and analysis in a large scale study on dioxin-like chemicals in soils in Australia

Jochen Mueller[1], Renee Muller[1], Katrina Goudkamp[1], Melanie Shaw[1], Munro Mortimer[2], David Haynes[3], Debbie Burniston[6], Robert Symons[6], Ray Correll[4], Michael R Moore [5]

[1]EnTox, The University of Queensland, Brisbane
[2]EPA Queensland, Brisbane
[3]Great Barrier Reef Marine Park Authority, Townsville
[4]CSIRO, Adelaide
[5]Queensland Health Scientific Services, Brisbane
[6]AGAL, Sydney

Introduction

Soil and sediments are important sinks for dioxin-like chemicals in the environment. Hence, analysis of soil samples has been used to evaluate potential emission sources of these chemicals to the environment[1, 2]. Interpretation of analytical results however, requires a good understanding of the accuracy and precision of the analytical procedures as well as a sampling program that is designed for the purpose of the specific study[3].

In 2002 the Department of Environment and Heritage, Australia, commissioned studies as part of the National Dioxin Program (NDP) that aimed to evaluate the background levels of dioxin-like chemicals in soils nationally. The key stratification criteria for the sampling strategy included regions (north, south-east and south-west) as well as different land-use types (i.e. urban, industrial, agricultural and remote)[4]. Considering the financial and time constraints of the study, in essence the study required collection and analysis of about 80 representative samples to cover the Australian which is about 20 times the size of the unified Germany and only around 20 % smaller than the USA.

While the primary aim of the soil study was to evaluate concentrations of dioxin-like chemicals in Australian soil, a key element was an assessment of the accuracy and reproducibility of both the sampling scheme and analytical results, focussing on data that were unexpected in one way or the other. The aim of this sub-project was therefore to evaluate the analytical accuracy and reproducibility as well the overall reproducibility of the sampling strategy (i.e. how well samples represent a specific sampling site category).

Methods and Materials

Reproducibility of sampling site selection
Soil samples were collected from more than 80 locations representing urban, industrial, agricultural and remote sites across Australia, covering a range of climates and soil types. A hierarchical set of terms were defined that describe a sample:

i) A *location* is represents both a given geographical region (e.g a city or agricultural district) and a specific land use within that region (e.g. Brisbane Urban). Within each *location*, two *samples* are collected.
ii) A *sample* is one of two replicates representing a *location* and is comprised of a soil pooled from 3 *sub-sampling sites*.
iii) A *sub-sampling site* represents an area within the location such as for example a park in the city or an agricultural feild (shaded ovals in Fig.1). 6 soil *cores* are collected from each.
iv) A soil *core* is an individual soil sub-sample.

18 *cores* from 3 *sub-sampling sites* are pooled to make a *sample*. Two samples are collected at each *location*. Analysis of both samples from given location provides information regarding the reproducibility of the sampling site selection.

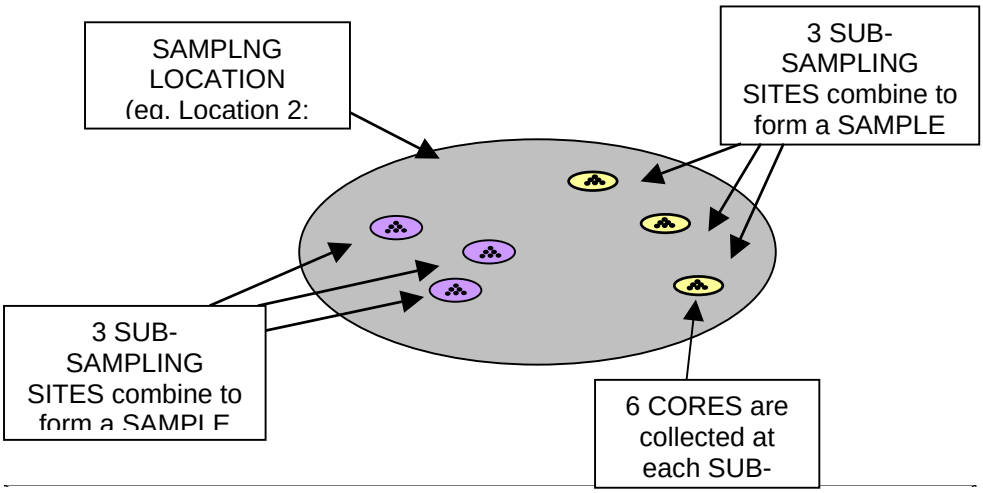

Figure 1: Schematic representation of the sampling strategy to evaluate sampling reproducibility.

While duplicate samples were obtained from all locations, the reproducibility of the sampling strategy was assessed by analysing both replicate samples from only 18 of the 86 locations. As only a selection of second replicate samples were analysed, these were chosen once analytical results from the initial samples were obtained. Hence, while the sampling was usually performed concurrently, replicate samples were necessarily analysed in separate batches. The decision to analyse a second sample from a location was based on unexpected or elevated results in the

analysis of the first replicate and/or detectable levels of a range of different dioxin-like chemicals in the soils. Sampling replication can therefore be considered as biased towards identifying samples with relatively large sampling errors.

Analytical reproducibility and interlaboratory comparison
To establish the analytical reproducibility of the study, six samples from different analytical batches were selected for re-analysis. These samples were again selected partly as a result of elevated levels in the first analysis. The evaluation of the analytical reproducibility included replicate analysis of both samples from one location. This allowed a direct comparison of the reproducibility of the selection of the sampling sites with the analytical reproducibility.

Interlaboratory calibration
An interlaboratory calibration was also conducted to assess analytical accuracy, as requested by the Department of Environment and Heritage, Australia. For this purpose, 8 of the samples that showed detectable levels of dioxin-like chemicals were sent to a second laboratory (Ontario Ministry of the Environment, Laboratory Services Branch, Toronto, Ontario, Canada. M9P 3V6) for re-analysis.

Analytical methods
The analytical methodology for the determination of PCDD/Fs and PCBs was based on quantification of the analytes through isotope dilution techniques and is modified from those described by the USEPA methods 1613B and 1668A, respectively.

In brief, samples were freeze dried and mixed to produce a homogenous sample. A sub sample was removed and spiked with a range of isotopically labelled surrogate standards, and then extracted with toluene using an accelerated solvent extractor. Clean up was effected by partitioning with sulfuric acid then distilled water. Sulfur was removed using copper or silver nitrate dispersed on silica gel. Further purification was performed using column chromatography on acid and base modified silica gels, neutral alumina and carbon dispersed on celite. After cleanup, the extract was concentrated to near dryness. Immediately prior to injection, recovery standards were added to each extract, and an aliquot of the extract was injected into the gas chromatograph. The analytes were separated by the GC and detected by a high-resolution (≥10,000) mass spectrometer. The quality of the analysis was assured through reproducible calibration and testing of the extraction, cleanup, and GC/MS systems.

Evaluation of differences between two samples/analytical results
In this report, comparisons between replicate samples or replicated analysis have been made by calculating the normalised difference. The normalised difference between two samples was calculated from

Results and Discussion

The analytical reproducibility was evaluated by repeated analysis of 6 selected samples at AGAL: Cairns (U1A), Sydney (U3A), Sydney (U3B), Melbourne (U2A), Sydney (I2A) and Adelaide (I1A). The results indicate good agreement in the replicate analysis of the samples with respect to

the concentration expressed as TEQ, ∑PCDD/Fs or ∑PCBs (Table 1). Overall, the mean normalised difference between congeners detected in both replicates varied between 14 % in sample Sydney U3B and 39 % in Sydney U3A. Furthermore, considering all data, on average the difference between replicate analyses was about 25 % and for more than 90 % the mean normalised difference was less than 67 %, indicating that for 90 % of detectable congeners the difference between replicate analysis was less than a factor of two.

Table 1 Summary of analytical reproducibility for replicate analysis of six soil samples.

Sample	Land-use	Concentrat. $TEQ_{(DF\&P)}$	∑ PCDD/Fs	∑PCBs	No. Detect.	Mean Norm. Diff.
Cairns	Urban	10	68400	ND	23	25 %
(U1A)		7.4	53600	ND	15	
Sydney	Urban	4.7	12000	31	39	39 %
(U3A)		3.7	12100	92	29	
Sydney	Urban	9.3	42400	74	36	14 %
(U3B)		9.7	42200	98	32	
Melbourne	Urban	6.1	13000	44	30	18 %
(U2A)		6.2	11300	37	32	
Sydney	Industrial	11	11000	86	32	38 %
(I2A)		9.2	10800	24	30	
Adelaide	Industrial	3.7	550	820	37	17 %
(I2A)		2.6	610	690	27	

Interlaboratory calibration results

Based on all results that were detectable by both laboratories (i.e. including all congeners and homologues for all eight samples), the mean normalised difference was less than 30 %, which should be considered a good result for an interlaboratory calibration. Most importantly, we could not find a systematic difference between the two laboratories (i.e. neither of the laboratories were consistently higher or lower for any compounds) in a given sample or for all samples. Hence, it seems that the variability is more related to the usual uncertainties in the normal laboratory procedure.

Reproducibility of sampling

Eighteen of these 'B' samples were analysed for comparison with results of previously analysed 'A' samples. For one of the samples the difference was so large that the 'A' sample was recollected and reanalysed and the latter result varied greatly from the first result. The reasons remain unclear, but it was decided to exclude this result from further analysis. For the remaining 17 samples, results showed that the sampling reproducibility was highly variable. For example, for 8 of the 17 samples the difference was greater than a factor 2 and for a further 3 samples the difference was very close to a factor 2. Using all congeners detectable in both sampling replicates for 9 of the 17 duplicated samples the mean normalised difference of all congeners detected in both samples was greater than a factor of 2 (>67%). The difference between location replicates representing Wollongong U1 (urban), Sydney I2 (industrial), Brisbane U1 (urban), Cairns U1 (urban) and Latrobe I1 (industrial) was greater than 3 (>100%). This suggests that the samples are grossly different in contamination and may indicate that an historical or current point source exists near one of the sub-sampling sites where samples were collected. Notably, for the purpose of the study we devised contamination categories (<0.2 pg TEQ/g dmt; 0.2 –1 pg TEQ/g dmt, 1-5 pg TEQ/g dmt 5-40 pg TEQ/g dmt and >40 pg TEQ/g dmt) and for most samples the A and B samples still fell into the same contamination category.

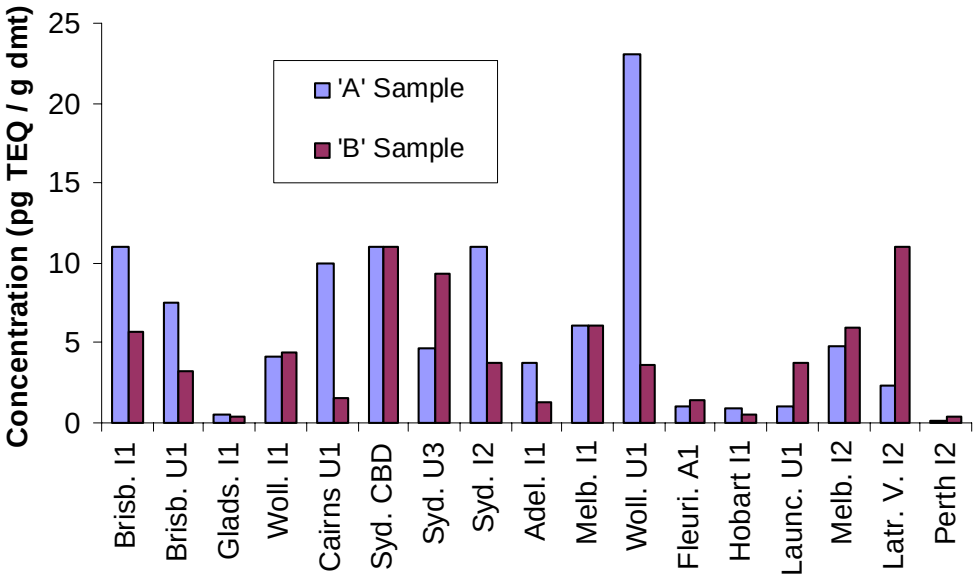

Figure 2 Comparison of the concentration (expressed as pg TEQ/g dmt) between the A and B samples from 17 different sampling LOCATIONS around Australia

Direct comparison of analytical and sampling reproducibility

A direct comparison between the analytical reproducibility and the sampling reproducibility is possible for the location Sydney U3 (urban) where both 'A' and 'B' samples were analysed twice (Figure 3.1). The results show that the concentration at the 'B' site, whether expressed as TEQs or pg g^{-1} dwt, was more than double that at the 'A' site. However, as shown by the standard deviation bars, the analytical reproducibility is relatively high.

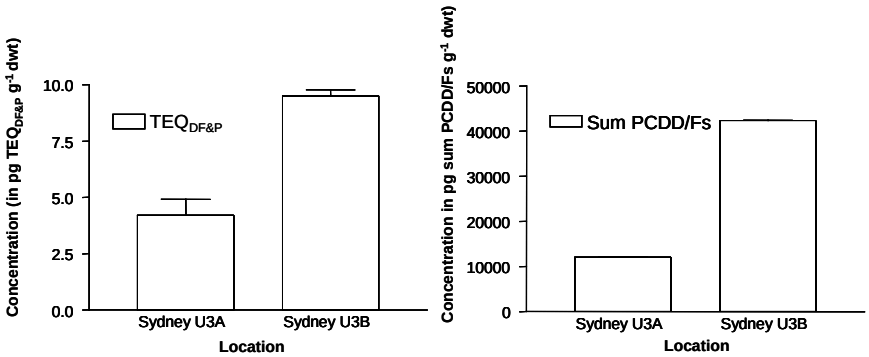

Figure 3. Results for replicate analysis of soil samples Sydney U3A and Sydney U3B.

The last 20 years have seen substantial progress in the analysis of trace organic chemicals such as dioxins and recent laboratory comparisons suggest that the quality of analytical results are generally very high. In contrast to the progress that has been made in the sample analysis, the sampling strategy seems to often been neglected or at least receives less attention with respect to QC/QA procedures. Our study suggests that for inhomogeneous matrices such as soils, the sampling strategy is the weakest point of the study.

Acknowledgements

The authors would like to thank all the colleagues and field staff as well as some volunteers for assisting with the sample collection and organizing the collection in the different regions. Without their input this study would have been impossible or unaffordable. The studies on which this paper is based were funded under the National Dioxins Program administered by the Australian Government Department of the Environment and Heritage. The views expressed herein are not necessarily those of the Commonwealth of Australia. That National Research Centre for Environmental Toxicology is co-funded by Queensland Health.

References

[1] B. D. Eitzer, *Environmental Science and Technology*, 1993, **27**, 1632.
[2] W. Rotard, W. Christmann, and W. Knoth, *Chemosphere*, 1994, **29**, 2193.
[3] L. H. Keith, G. L. Patton, D. L. Lewis, and P. G. Edwards, in 'Determination what kinds of samples and how many samples to analyse', ed. L. H. Keith, Washington DC, 1998.

[4] Müller J.F., Muller R, Goudkamp K, Shaw M, Mortimer M, Haynes D, Burniston D, Symons R, and Moore M, 'Dioxins in Soils in Australia; National Dioxins Program Technical Report No. 5', Australian Government Department of the Environment and Heritage, Canberra, 2004.