

# using science to create a better place

## Ambient background metal concentrations for soils in England and Wales

Science Report: SC050054/SR

The Environment Agency is the leading public body protecting and improving the environment in England and Wales.

It's our job to make sure that air, land and water are looked after by everyone in today's society, so that tomorrow's generations inherit a cleaner, healthier world.

Our work includes tackling flooding and pollution incidents, reducing industry's impacts on the environment, cleaning up rivers, coastal waters and contaminated land, and improving wildlife habitats.

This report is the result of research commissioned and funded by the Environment Agency's Science Programme.

Published by:  
Environment Agency, Rio House, Waterside Drive, Aztec  
West, Almondsbury, Bristol, BS32 4UD  
Tel: 01454 624400 Fax: 01454 624409  
[www.environment-agency.gov.uk](http://www.environment-agency.gov.uk)

ISBN: 1844326101

© Environment Agency

November 2006

All rights reserved. This document may be reproduced with prior permission of the Environment Agency.

The views and statements expressed in this report are those of the author alone. The views or statements expressed in this publication do not necessarily represent the views of the Environment Agency and the Environment Agency cannot accept any responsibility for such views or statements.

This report is printed on Cyclus Print, a 100% recycled stock, which is 100% post consumer waste and is totally chlorine free. Water used is treated and in most cases returned to source in better condition than removed.

Further copies of this report are available from:  
The Environment Agency's National Customer Contact Centre  
by emailing:  
[enquiries@environment-agency.gov.uk](mailto:enquiries@environment-agency.gov.uk)  
or by telephoning 08708 506506.

**Author(s):**

S. P. McGrath and F. J. Zhao

**Dissemination Status:**

Publicly available

**Keywords:**

Ambient background concentration, risk assessment, soil, trace metals

**Research Contractor:**

Rothamsted Research, Harpenden, Hertfordshire AL5 2JQ  
01582 763133

**Environment Agency's Project Manager:**

G. Merrington, Walingford

**Science Project Number:**

SC050054

**Product Code:**

SCHO1106BLPV-E-P

# Science at the Environment Agency

Science underpins the work of the Environment Agency. It provides an up-to-date understanding of the world about us and helps us to develop monitoring tools and techniques to manage our environment as efficiently and effectively as possible.

The work of the Environment Agency's Science Group is a key ingredient in the partnership between research, policy and operations that enables the Environment Agency to protect and restore our environment.

The science programme focuses on five main areas of activity:

- **Setting the agenda**, by identifying where strategic science can inform our evidence-based policies, advisory and regulatory roles;
- **Funding science**, by supporting programmes, projects and people in response to long-term strategic needs, medium-term policy priorities and shorter-term operational requirements;
- **Managing science**, by ensuring that our programmes and projects are fit for purpose and executed according to international scientific standards;
- **Carrying out science**, by undertaking research – either by contracting it out to research organisations and consultancies or by doing it ourselves;
- **Delivering information, advice, tools and techniques**, by making appropriate products available to our policy and operations staff.

Steve Killeen

**Head of Science**

# Executive summary

Ambient background concentrations of trace metals in soils vary widely and therefore for any risk assessment of metals, consideration of a local or site specific background is important. The objective of this project was to investigate whether ambient background concentrations of Cd, Co, Cr, Cu, Ni, Pb and Zn in soils can be predicted from other semi-conservative soil properties such as Al and Fe concentrations or soil texture information.

Almost 6,000 soil samples were collected between 1978 and 1983 for the National Soil Inventory (NSI) for England and Wales at the intersects of a 5-km orthogonal grid. A suite of soil properties was determined, including clay content organic carbon (OC) content, pH and the concentrations of 17 elements. Statistical analysis was performed for this report on the NSI dataset.

A strong geochemical association was found between the trace metals Co, Cr and Ni and the major elements Al and Fe in the NSI dataset. Multiple regressions linking Co, Cr or Ni with Al and Fe concentrations were obtained which explained 62–85 per cent of the variation in the trace metal concentrations. Site-specific ambient background concentrations of Co, Cr or Ni in topsoils could be predicted reasonably well from Al and Fe concentrations.

Consistent with geochemical principles, there was no strong association between the trace metals Cd, Cu, Pb or Zn with Al and Fe. Therefore, ambient background concentrations of these trace metals could not be predicted using the regression approach with Al and/or Fe.

An alternative approach using a probability graph method was used to identify an 'uncontaminated' population of soils from the whole dataset. This approach yielded estimates of the geometric mean and the 10th to 90th percentile range of trace metals that might be associated with the 'uncontaminated' population. However, this method may be appropriate only for Pb because, for other trace metals, the 'uncontaminated' population was over-represented by soils with light textures.

A further approach was to explore the variation of metal concentrations with other soil properties. Soil texture has a major influence on trace metal concentrations. Concentrations of Cd, Co, Cr, Cu, Ni and Zn show an increasing trend from light to heavy textured soils, whereas peaty soils have higher Pb concentrations than other soil types. Median concentrations of trace metals were calculated for different soil texture groups and these can be used as ambient background concentrations that are specific to a particular soil texture.

# Acknowledgements

We thank Rodger White for statistical advice.

# Contents

<b>Executive summary</b>	<b>4</b>
<b>Acknowledgements</b>	<b>5</b>
<b>Contents</b>	<b>6</b>
<b>1 Introduction</b>	<b>8</b>
<b>2 Methods</b>	<b>10</b>
2.1 Data	10
2.2 Statistical analysis	11
2.2.1 Summary statistics and regression analysis	11
2.2.2 Probability plots	12
2.2.3 Frequency distribution according to texture groups	13
<b>3 Results and discussion</b>	<b>14</b>
3.1 Summary statistics	14
3.2 Correlation and principal component analysis	14
3.3 Regression analysis	18
3.4 Probability graph approach to deriving ambient background concentrations of trace metals in soils	22
3.5 Ambient background concentrations of trace metals in soils according to texture group	24
<b>4 Conclusions</b>	<b>28</b>
<b>References &amp; Bibliography</b>	<b>29</b>
<b>List of abbreviations</b>	<b>31</b>

## List of tables

Table 2.1	Groups and classes of soil texture from field assessment
Table 2.2	Summary statistics for seven trace metals in NSI soils (mg/kg) (n = 5,691)
Table 3.1	Correlation coefficients between elemental concentrations (mg/kg), OC content (%) and pH (n = 5,651)
Table 3.2	Correlation coefficients between seven trace metals concentrations (mg/kg) and clay content (%) and eCEC (cmol/kg) (n = 4,842)
Table 3.3	Latent vector loadings for the first four principal components
Table 3.4	Coefficients for regression equations relating Co, Cr or Ni to Al and/or Fe
Table 3.5	Predicted ambient background concentrations of Co, Cr and Ni as a function of Al and Fe concentrations in soil1

Table 3.6 Summary statistics of trace metal concentrations according to soil texture group

**List of figures**

Figure 3.1 Boxplots of  $\log_{10}$ -transformed concentrations of Cd, Co, Cr, Cu, Ni, Pb and Zn

Figure 3.2 Latent vector loadings for 19 variables used in principal component analysis

Figure 3.3 Relationships between Co and Al (a) or Fe (b)

Figure 3.4 Relationships between Cr and Al (a) or Fe (b)

Figure 3.5 Relationships between Ni and Al (a) or Fe (b)

Figure 3.6 Probability graph for soil Pb concentration in the NSI dataset

# 1 Introduction

Trace metals in surface soils are derived from both parent materials and anthropogenic activities. The latter means it is often difficult to quantify the natural background concentrations of metals in soils. The usual, or ambient, concentration of a metal in soil therefore consists of a natural pedo-geochemical fraction and an anthropogenic fraction (ISO 2004). In this definition, the anthropogenic fraction refers to moderate diffuse inputs into the soil and not the inputs from local point sources that generally result in a much elevated concentration. According to the ISO definition, the 'usual background value' of a substance in soils can be set at a chosen parameter from the frequency distribution of usual concentrations. This report uses the term 'ambient background concentration' (ABC) with the same meaning as that of 'usual background concentration'.

ABCs of metals in soils may vary depending on soil types and properties. For any risk assessment of metals, consideration of a local or site-specific ABC is essential. For example, Hamon *et al.* (2004) sampled sites in Namibia, Australia and Asia and used relationships with Fe or Mn to make estimates of the background concentrations of As, Co, Cr, Cu, Ni, Pb and Zn, assuming that the soils in a sub-set from 'remote areas' were uncontaminated. But in long-industrialised regions like the UK, this assumption cannot be made and the concentrations of metals in relatively uncontaminated soils can best be referred to as ABCs.

Given an appropriate dataset, various approaches to deriving ABCs for different soils in England and Wales can be tested. Rothamsted Research holds such a database of the concentrations of elements and other soil properties obtained from the National Soil Inventory (NSI).

The soil samples for the NSI were collected between 1978 and 1983 at the intersects of a 5-km orthogonal grid based on the Ordnance Survey National Grid across England and Wales. The sampling was therefore unbiased and not targeted at any particular land use; it included all types of land including agriculture, forests, woodlands, parks and moors. No attempt was made to devise a sampling strategy to cover urban areas adequately, so the coverage is predominantly non-urban. This 5-km grid yielded almost 6,000 samples, which were analysed by Rothamsted Research. Data have been collected in similar inventories for Scotland and Northern Ireland.

Summary statistics and maps for many of the variables were published in the *Soil Geochemical Atlas of England and Wales* (McGrath and Loveland 1992). A further project looked at the detailed spatial relationships within the NSI data (the geostatistics) and their application to the design of a soil monitoring network for England and Wales (Oliver *et al.* 2002). However, none of the publications or work on the NSI data have examined the aspects related to using the information and the relationships between variables to determine soil-specific ABCs.

In terms of environmental risk assessment, it is also important that all the data available for metal toxicity assessment for invertebrates, plants and soil microbes use ambient background soil concentrations as a 'control' to which metal doses are added in the testing procedure (Smolders *et al.* 2003; see also, for example, Oorts *et al.* 2006, Rooney *et al.* 2006). The Predicted No Effect Concentrations (PNECs) which are derived through statistical processes from these data can, for example, be added to the existing ambient background concentrations in soil to give a total acceptable concentration.

It is thus important to establish ambient concentrations of metals in soils. Because pristine soils are unlikely to be found in the UK, the work should:

- concentrate on rural soils;



- avoid areas with significant point source pollution such as mining and smelting.

The initial aim of this project was to investigate whether ABCs of Cd, Co, Cr, Cu, Ni, Pb and Zn in soils can be predicted from other semi-conservative soil properties such as Al and Fe concentrations. During the course of this project, however, it became apparent that not all trace metals showed a sufficiently close association with Fe and/or Al to allow a reasonable prediction. Two alternative approaches to estimating ABCs of these trace metals in soils were therefore explored. One was the probability graph method and the other used the percentile distribution of concentrations for different soil texture classes. The analysis was based on the NSI dataset.

The results obtained are valid for the soils in England and Wales, and extrapolation to other regions should be made with caution.

# 2 Methods

## 2.1 Data

The data used in the present study are from the National Soil Inventory of England and Wales. A total of 5,691 soils (0–15 cm) were sampled between 1978 and 1983 from the intersects of 5-km orthogonal grids. Soil samples were analysed for a wide range of physical and chemical properties. Details of sample collection and preparation, analytical methods and quality control, summary statistics and maps for many of the measured soil properties are given in the *Soil Geochemical Atlas of England and Wales* (McGrath and Loveland 1992).

There are 5,691 data points in the NSI dataset for each of the following elements: Al, Ba, Cd, Ca, Cr, Co, Cu, Fe, Pb, Mg, Mn, Ni, P, K, Na, Sr and Zn. The focus of this study is on seven trace metals – Cd, Co, Cr, Cu, Ni, Pb and Zn.

The elemental concentrations were determined following aqua regia digestion (McGrath and Cunliffe 1985); these are often referred to as pseudo-total concentrations. Certified reference soils from the Community Bureau of Reference (BCR) were used for quality control as these had certified results for the concentrations of metals in aqua regia digests.

There are 5,651 data points for pH and organic carbon (OC) and 4,842 data points for clay content. Clay content was not determined in peaty soils (generally with OC >12 per cent). Organic carbon was determined either by a dichromate digestion method for soils containing <20 per cent carbon or by loss-on-ignition for soils containing >20 per cent (McGrath and Loveland 1992).

Soil texture information was available from field assessments carried out at the time of soil sampling. There were six texture groups plus a group of peaty soils, with each texture group consisting of various texture classes (Table 2.1). In the present study, analysis was performed at the texture group level.

**Table 2.1 Groups and classes of soil texture from field assessment**

Texture group	Class
Sandy	Coarse sand, loamy coarse sand, medium sand, loamy medium sand, fine sand, loamy fine sand
Coarse loamy	Sandy loam, sandy silt loam
Fine loamy	Sandy clay loam, clay loam
Coarse silty	Silt loam
Fine silty	Silty clay loam
Clayey	Sandy clay, silty clay, clay
Peaty	Peaty

Soil cation exchange capacity (CEC) was calculated from the following equation (Helling *et al.* 1964):

$$\text{CEC (cmol/kg)} = [\text{OC}\% \times (51 \times \text{pH} - 59) \div 10 + \text{Clay}\% \times (4.4 \times \text{pH} + 30) \div 10] \div 10$$

This equation is based on the contribution to CEC from soil organic matter and clay as a function of soil pH. Unlike CEC measured at a buffered pH value (usually pH 7), the CEC derived from the above equation is an estimate at the measured soil pH and can be referred to as effective CEC (eCEC). Recent studies have shown that toxicity threshold data for several metals correlate strongly with soil eCEC (Smolders *et al.* 2004, Oorts *et al.* 2006, Rooney *et al.* 2006).

## 2.2 Statistical analysis

### 2.2.1 Summary statistics and regression analysis

The range, mean, median, skewness and kurtosis for each variable were computed for both transformed and log<sub>10</sub>-transformed data (except for pH, which is already log<sub>10</sub>-based). In the case of Cd, 74 data points with zero values were excluded prior to log transformation.

Boxplot analysis was performed for the Cd, Co, Cr, Cu, Ni, Pb and Zn data, based on both untransformed and log<sub>10</sub>-transformed values. This analysis identifies (Tukey 1977):

- the median;
- values for the 25th and 75th percentile (i.e. inter-quartile range);
- the upper and lower whisker. The lower and upper whiskers are 1.5 times the inter-quartile range.

Correlation analysis was performed on the log<sub>10</sub>-transformed data, except for pH and clay content as the latter showed a normal distribution without transformation. Next, principal component analysis (PCA) was performed on the correlation matrix to provide an overall view of the relations among variables. For PCA, the full dataset including all elemental concentrations, OC and pH (representing H<sup>2+</sup> activity) was used (n = 5,651).

Simple or multiple linear regression analysis was performed on the log<sub>10</sub>-transformed data. It was decided to exclude those data that were below the detection limits or outliers above an upper limit ('upper outliers') (Table 2.2). According to the EU Technical Guidance Document (TGD) for risk assessment (European Commission 2003), data used for assessing environmental concentrations in a regional scenario should not be influenced directly by point source emissions. The TGD recommends that, if the dataset is sufficiently large, the upper whisker for the log<sub>10</sub>-transformed data is taken as the cut-off level for the high-concentration outliers. Thus, the high-concentration outliers are defined as:

$$\log_{10}(x_i) > \log_{10}(p75) + 1.5 [\log_{10}(p75) - \log_{10}(p25)]$$

where  $x_i$  is the outlier cut-off value, and  $p25$  and  $p75$  are the 25th and 75th percentiles of the distribution respectively.

For Co, Cr and Ni for which regression analysis was applied, only very small numbers of samples were excluded because they were either below their respective detection limits (0.2–1.1 per cent of the dataset) or above the outlier cut-off values (0.2–0.5 per cent of the dataset) (Table 2.2). Exclusion of these data points benefited regression analysis because of a considerable reduction of skewness and kurtosis (see below).

All statistical analyses were carried out using Genstat version 8.2 (VSN International Ltd, Hemel Hempstead, UK).

**Table 2.2 Summary statistics for seven trace metals in NSI soils (mg/kg) (n = 5,691)**

	Cd	Co	Cr	Cu	Ni	Pb	Zn
<i>Untransformed data</i>							
Minimum <sup>1</sup>	<0.2	0.2	0.2	1.2	0.8	3.0	5.0
Maximum	41	322	838	1508	440	16338	3648
Mean	0.8	10.6	41.2	23.1	24.5	74.0	97.1
Median	0.7	9.8	39.3	18.1	22.6	40.0	82.0
Skewness	17.5	14.4	9.5	21.2	7.0	42.7	13.6
Kurtosis	573	436	206	655	133	2452	300
<i>log<sub>10</sub>-transformed data</i>							
Mean	-0.19 <sup>2</sup>	0.91	1.53	1.25	1.29	1.66	1.90
Median	-0.15	0.99	1.59	1.26	1.35	1.60	1.91
Skewness	-0.64	-1.30	-1.61	0.41	-1.09	1.09	0.09
Kurtosis	1.39	2.30	5.63	3.11	2.08	3.06	3.09
Detection limit (DL)	0.2	0.5	1	0.75	2	4.7	0.7
No. of samples <DL	450	39	9	0	60	9	0
Upper outlier cut-off <sup>3</sup>	2.8	51.2	147.1	69.2	114.1	238.8	267.5
No. of outliers	82	12	28	131	14	213	139

Notes:

<sup>1</sup> For some elements, minimum concentrations in the dataset were below the limits of quantification ('detection limit').<sup>2</sup> 74 data points with zero Cd were excluded prior to log<sub>10</sub> transformation.<sup>3</sup> Corresponds to the upper whisker in the boxplot shown in Figure 2.1 after antilog back transformation.

## 2.2.2 Probability plots

Geochemists have used probability graph paper to determine the threshold or 'inflection point' between anomalously high and background geochemical data (Sinclair 1974). Davies (1983) applied this method to estimate the mean and the range of Pb concentrations likely to be associated with uncontaminated soils using survey data from four areas in England and Wales.

In this method, log<sub>10</sub>-transformed soil metal concentrations (*y*-axis) are plotted against their cumulative frequency distributions on the probability scale (*x*-axis). A straight line means that the data are log-normally distributed, which is usually considered to be the case if all data points are from a single population of samples. But because a country-wide or regional survey may cover soil samples belonging to both uncontaminated and contaminated populations, the relationship between log<sub>10</sub> metal concentration and cumulative frequency presented on a probability graph is more likely to be a curve. Mixed populations may also be present, giving rise to curves with inflexion points related to the proportion of each population present.

The linear portion associated with low metal concentrations is taken as representing the distribution of the uncontaminated soils (Davies 1983). Here 'uncontaminated' soils should be taken as those not impacted by point source contamination.

More details of the method are given by Sinclair (1974) and Davies (1983). In this study, the method was applied to derive mean and the 10th to 90th percentile range for seven trace metals in the 'uncontaminated' population within the NSI dataset, although a detailed examination showed that this approach was not suitable for metals other than Pb (see discussion below).

### 2.2.3 Frequency distribution according to texture groups

Data were sorted according to the seven texture groups (Table 2.1). The number of samples varied from 182 for the coarse silty group to 2,002 for the fine loamy group.

Frequency distributions of the concentrations of the seven trace metals within each texture group could not be fitted satisfactorily with parametric distribution functions (e.g. normal, log-normal, gamma or Weibull). Therefore, non-parametric distributions were used.

Data were ranked in ascending order and the cumulative probability of a data point  $x_i$  was calculated according to the following formula (Cullen and Frey 1999):

$$F(x_i) = (i - 0.5)/n$$

where  $i$  is for the rank order of the sorted data and  $n$  for the total number of data points within each texture group.

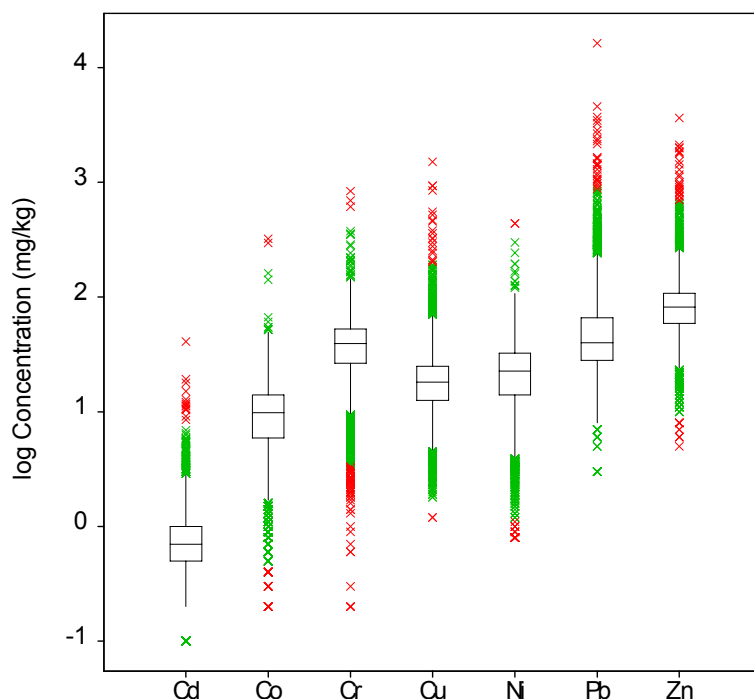
Median and percentile values were obtained from this empirical distribution function.

# 3 Results and discussion

## 3.1 Summary statistics

Table 2.2 shows the range, mean, median, skewness and kurtosis of the untransformed and  $\log_{10}$ -transformed data of Cd, Co, Cr, Cu, Ni, Pb and Zn. Boxplots for these metals are shown in Figure 3.1.

Concentrations of all seven metals in soils varied widely. The untransformed data are far from normally distributed. To stabilise the variances, logarithmic ( $\log_{10}$ ) transformation of these data was performed. The transformation substantially reduces the coefficients of skewness and kurtosis (Table 2.2), although the distributions of the  $\log_{10}$ -transformed data are still not quite normal. Excluding those data below the detection limit and above the upper outlier cut-off value (Table 2.2) further reduced the skewness and kurtosis of the data distribution (results not shown).



**Figure 3.1** Boxplots of  $\log_{10}$ -transformed concentrations of Cd, Co, Cr, Cu, Ni, Pb and Zn

The rectangular box represents the 25th to 75th percentile range, the horizontal line inside the box the median, vertical lines outside the box are lower and upper whisker (each being 1.5 times the inter-quartile range). Outliers are shown as crosses.

## 3.2 Correlation and principal component analysis

Table 3.1 shows the correlation coefficients between elemental concentrations, pH and soil OC content. Correlations between metal concentrations and clay content or eECE were performed separately using a smaller number of soils ( $n = 4,842$ , Table 3.2) because clay content and eCEC were not available for many peaty soils.

There were strong correlations between Co, Cr or Ni and Fe or Al ( $r > 0.7$ ). Zinc and Cd correlated strongly with each other ( $r = 0.73$ ), but not so strongly with Al or Fe ( $r = 0.5-0.56$ ). Lead showed a positive correlation with OC ( $r = 0.56$ ) and a negative correlation with pH ( $r = -0.39$ ), which can be explained by the observation of elevated Pb concentrations in acidic upland peaty soils (McGrath and Loveland 1992). Copper showed no strong correlations with any variable except with Zn ( $r = 0.67$ ). There were no strong correlations between the seven trace metals and clay content or eCEC (Table 3.2).

PCA identified four principal axes with latent roots (or eigenvalues) of 9.1, 2.7, 2.1 and 1.0, accounting for 47.9, 14.4, 10.8 and 5.3 per cent of the total variance respectively. The latent vectors (or eigenvectors) describe the contribution to the variation of each variable on each of the principal axes (Table 3.3).

The larger the absolute values of the latent vectors, the greater the proportion of the variation a particular variable accounts for on a given axis. For the first principal axis, Al, Fe, Co, Cr, Ni, K and Mg had latent vector values  $> 0.25$ , suggesting that this component represents variation in the structural elements of soil minerals (Al, Fe and K) and the metals associated them. The principal component 2 was dominated by OC, Pb, Cu and pH (latent vector values  $> 0.25$ ), suggesting the influence of acidic peaty soils and the association of Pb and Cu with these soils. The variables with latent vector values  $> 0.25$  for principal component 3 included Ca, Sr, pH and P, suggesting that this component represents variation in acidity and possibly also the influence of agricultural practices. The concentration of Na had a particularly large value of latent vector (0.65) in principal component 4, suggesting that this might represent a marine influence.

The latent vector loadings for the first three principal components are shown in Figure 3.2, which shows the similarity among Al, Fe, K, Mn, Cr, Co and Ni. This association can be explained by classical geochemistry. For example, Co and Ni are classified as siderophilic (iron-loving) and trivalent forms of Cr, Fe and Al also tend to associate with each other (Goldschmidt 1954). Similarly, Sterckeman *et al.* (2006) reported a close association between Co, Cr and Ni with Al and Fe in a survey of 52 surface soils from northern France. Other associations include Zn and Cd, which often occur together in minerals. The association among Ca, Sr and pH is easily understood because of their relationship with acidity.

Figure 3.2 confirms the tendency for Pb and OC to be associated with each other due to the generally large Pb concentrations in upland peaty soils in northern and western England, where there is a long history of lead mining and smelting, and fallout of lead from the atmosphere to acid ombrotrophic peat may be enhanced by high rainfall (McGrath and Loveland 1992).

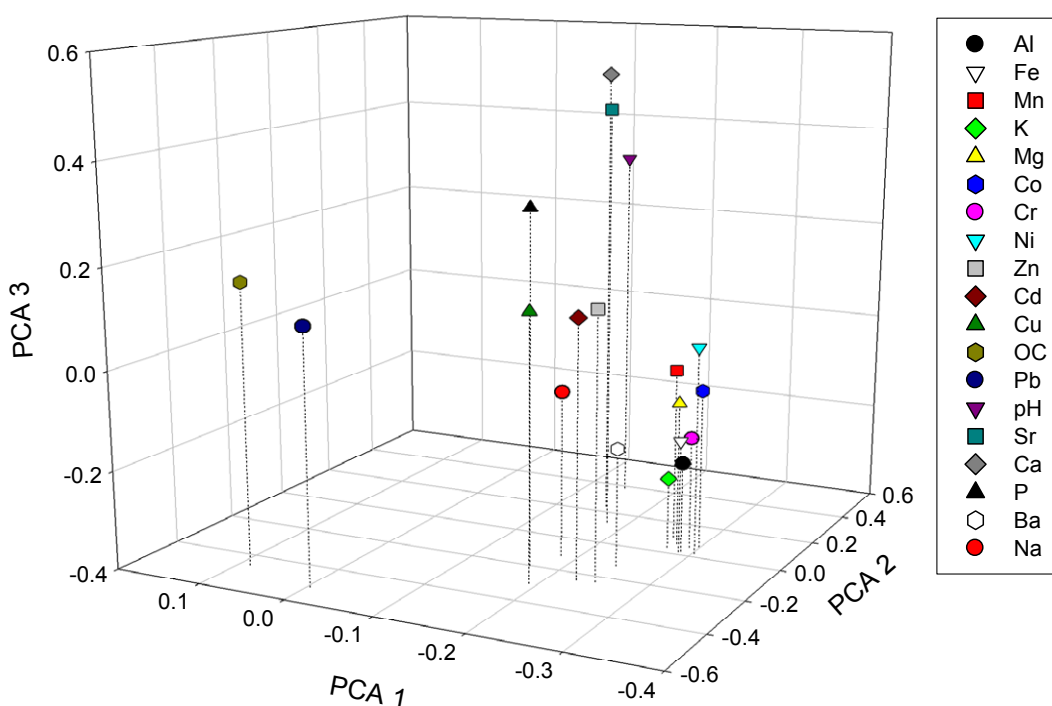




**Table 3.3 Latent vector loadings for the first four principal components<sup>1</sup>**

PCA	1	2	3	4
Al	<b>-0.287</b>	0.022	-0.215	0.152
Cd	-0.229	-0.242	0.119	-0.081
Co	<b>-0.298</b>	0.066	-0.074	-0.204
Cr	<b>-0.290</b>	0.053	-0.169	0.009
Cu	-0.192	<b>-0.304</b>	0.135	-0.153
Fe	<b>-0.285</b>	0.017	-0.169	-0.045
Mn	<b>-0.262</b>	0.104	-0.048	-0.205
Ni	<b>-0.302</b>	0.020	0.024	-0.167
Pb	-0.004	<b>-0.518</b>	0.112	-0.203
Zn	-0.249	-0.237	0.138	<b>-0.291</b>
Ba	-0.246	-0.125	-0.159	-0.057
Ca	-0.153	0.236	<b>0.522</b>	0.047
K	<b>-0.270</b>	0.032	<b>-0.254</b>	<b>0.266</b>
Mg	<b>-0.274</b>	0.062	-0.103	0.139
Na	-0.182	-0.116	-0.062	<b>0.648</b>
OC%	0.088	<b>-0.457</b>	0.168	<b>0.301</b>
P	-0.167	-0.201	<b>0.310</b>	0.141
Sr	-0.177	0.135	<b>0.462</b>	<b>0.270</b>
pH	-0.143	<b>0.384</b>	<b>0.332</b>	-0.094

Notes: <sup>1</sup> Absolute values > 0.25 are in bold. .



**Figure 3.2 Latent vector loadings for 19 variables used in principal component analysis**

### 3.3 Regression analysis

Informed by the results from correlation analysis and PCA, regression analysis was performed between Co, Cr, Ni and Fe or Al. Aqua regia-soluble Fe and Al can be considered as semi-conservative properties of soils; their concentrations are related more to the chemical composition of the soil-forming parent materials and the degree of weathering than to the anthropogenic influence.

Unlike Hamon *et al.* (2004), only weak associations were found between Zn, Cd, Cu or Pb and Fe (or Al, Mn). This and the strong relationship between Co, Cr, Ni and Fe or Al described above are in accord with classic geochemistry (Goldschmidt 1954). Therefore, regression was only performed between the latter group of elements.

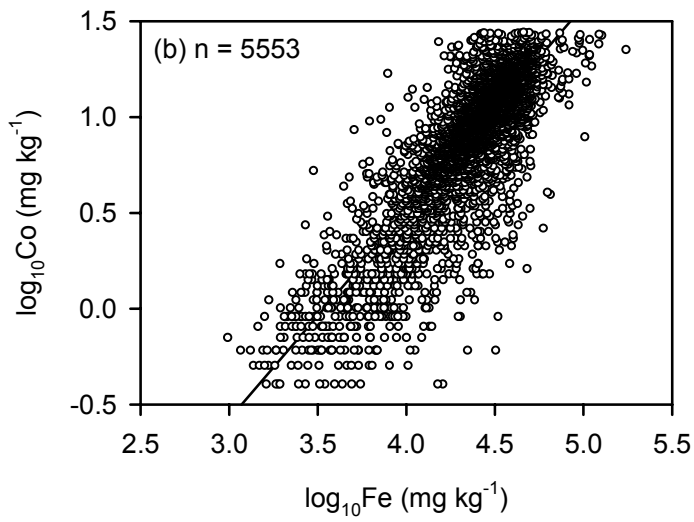
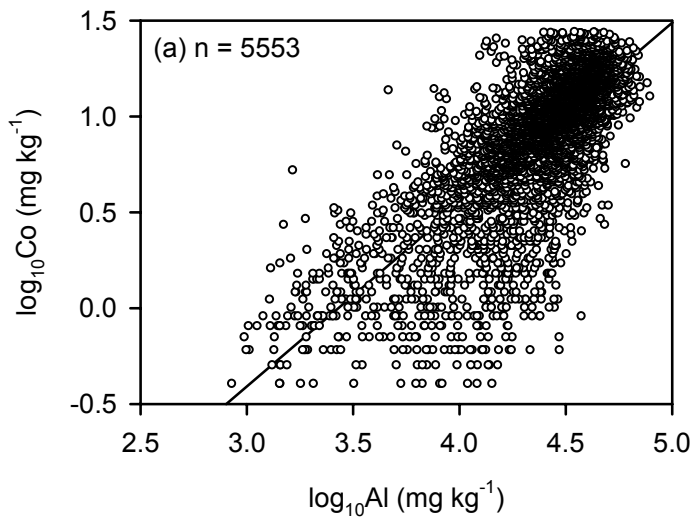
Figures 3.3–3.5 show the relationships between Co, Cr or Ni and Al or Fe respectively (excluding those data below the detection limit or above the upper outlier cut-off value). Most of the data points lie in the dark areas close to the regression lines. Coefficients for simple regressions are presented in Table 3.4. These regressions explained 53–82 per cent of the variation in the trace metal concentrations, which is considered satisfactory for such a large dataset. Linear regression based on  $\log_{10}$ -transformed data is appropriate because no non-linear trends were observed in the relationships (Figures 3.3–3.5) and the distribution of residuals (not shown) is considered acceptable for such a large dataset. Regression with Al was better than with Fe for Cr; the opposite was true for Co and Ni (Table 3.4).

For all three trace metals, a multiple regression including **both** Al and Fe as independent variables improved the model fit, resulting in a decreased standard error of observation and increased the percentage of variance accounted for ( $R^2_{adj}$ ) to 71, 85 and 62 per cent for Co, Cr and Ni, respectively (Table 3.4). It is therefore recommended that the multiple regression equations should be used in the prediction of ABCs of Co, Cr and Ni in topsoils.

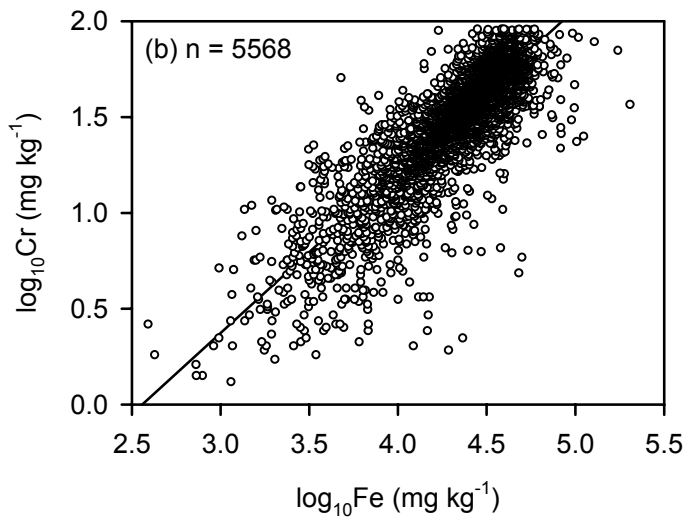
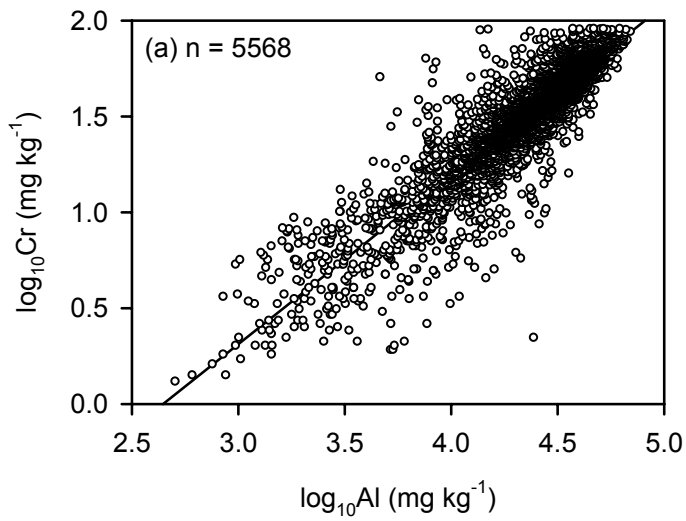
The multiple regression equations were then used to predict the ABCs of Co, Cr and Ni at different combinations of Al and Fe concentrations (Table 3.5). The concentrations of Al and Fe used represent realistic values likely to be encountered for soils in England and Wales, although different soils will have different combinations of Al and Fe concentrations.

The 95 per cent confidence range (CR) for each predicted mean is also given in Table 3.5. This was calculated approximately from the standard error (SE) of observation (see footnote to Table 3.5) because the SE values associated with regression coefficients (Table 3.4) were almost negligible compared with the SE for observation. The calculated 95 per cent CR already includes the variance for future observations and is therefore suitable for forecasting new observations. As expected, the predicted ABCs of Co, Cr and Ni increase with Al and Fe concentrations.

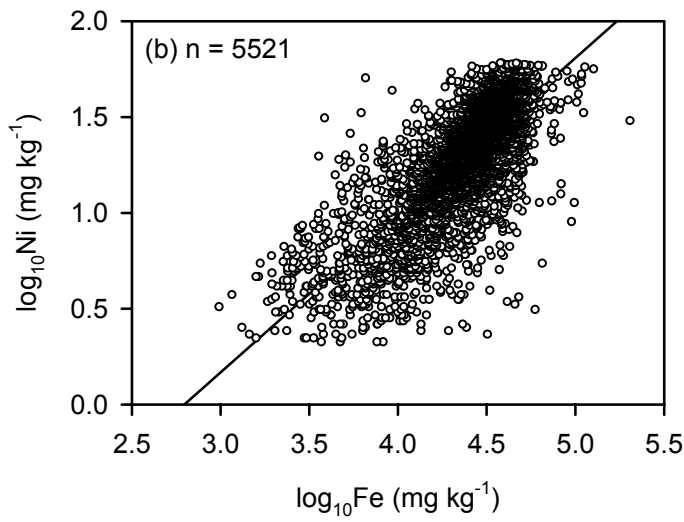
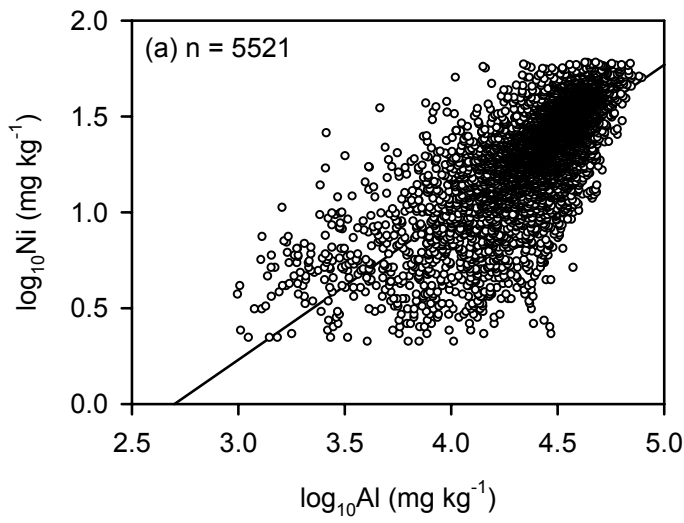
The approach of Hamon *et al.* (2004) was slightly different. They removed the upper 5 per cent data points, which gave rise to the largest positive residuals from the regression line, and derived the line encompassing the upper 95th percentile of the data. This was then used to predict the background metal concentrations. This approach has an inherent weakness in that the size of the 95 per cent CR is determined by the size of the SE of observation; thus, the poorer the fit between a trace metal and Al and Fe concentrations, the larger the 95 per cent CR relative to the predicted mean and the larger the soil background value for the trace metal. Deriving a larger soil background because of a poor fit is not a conservative approach.



**Figure 3.3 Relationships between Co and Al (a) or Fe (b)**



**Figure 3.4 Relationships between Cr and Al (a) or Fe (b)**



**Figure 3.5 Relationships between Ni and Al (a) or Fe (b)**

**Table 3.4 Coefficients for regression equations relating Co, Cr or Ni to Al and/or Fe**

Dependent variable (y)	Independent variable (x)	Regression coefficients			n	R <sup>2</sup> <sub>adj</sub>	SE of observation
		Constant	Slope (x <sub>1</sub> )	Slope (x <sub>2</sub> )			
log <sub>10</sub> Co	log <sub>10</sub> Al	-3.15 ± 0.048	0.93 ± 0.011		5,623	0.56	0.226
	log <sub>10</sub> Fe	-3.72 ± 0.041	1.06 ± 0.0094		5,623	0.69	0.188
	log <sub>10</sub> Al (x <sub>1</sub> ), log <sub>10</sub> Fe(x <sub>2</sub> )	-3.92 ± 0.042	0.26 ± 0.015	0.84 ± 0.016	5,623	0.71	0.183
log <sub>10</sub> Cr	log <sub>10</sub> Al	-2.37 ± 0.025	0.89 ± 0.0056		5,653	0.82	0.121
	log <sub>10</sub> Fe	-2.20 ± 0.031	0.85 ± 0.0071		5,653	0.72	0.149
	log <sub>10</sub> Al (x <sub>1</sub> ), log <sub>10</sub> Fe(x <sub>2</sub> )	-2.64 ± 0.024	0.63 ± 0.0090	0.33 ± 0.0092	5,653	0.85	0.109
log <sub>10</sub> Ni	log <sub>10</sub> Al	-2.13 ± 0.043	0.78 ± 0.0099		5,612	0.53	0.199
	log <sub>10</sub> Fe	-2.35 ± 0.041	0.83 ± 0.0094		5,612	0.58	0.187
	log <sub>10</sub> Al (x <sub>1</sub> ), log <sub>10</sub> Fe(x <sub>2</sub> )	-2.63 ± 0.041	0.35 ± 0.015	0.55 ± 0.015	5,612	0.62	0.179

**Table 3.5 Predicted ambient background concentrations of Co, Cr and Ni as a function of Al and Fe concentrations in soil<sup>1</sup>**

Al (mg/kg)	Fe (mg/kg)	Co (mg/kg)		Cr (mg/kg)		Ni (mg/kg)	
		Mean	95% CR <sup>2</sup>	Mean	±95% CR <sup>2</sup>	Mean	±95% CR <sup>2</sup>
5,000	5,000	1.4	0.6–3	9	5–14	5	2–11
10,000	20,000	5	2–12	21	13–34	14	6–31
20,000	40,000	12	5–26	41	25–66	25	11–57
30,000	80,000	23	10–53	66	40–108	43	19–97
40,000	120,000	35	15–80	90	55–148	59	27–133
50,000	200,000	57	25–130	123	75–201	85	38–191

Notes:

<sup>1</sup> Mean values predicted according to the multiple regression equations presented in Table 3.4.<sup>2</sup> 95% CR is calculated according to:  $10^{\wedge}(\text{predicted mean in log}_{10} \pm 1.96 \times \text{SE of observation})$ .

### 3.4 Probability graph approach to deriving ambient background concentrations of trace metals in soils

The regression approach described above was not found to be appropriate for trace metals such as Cd, Cu, Pb and Zn due to the lack of a close relationship with Al or Fe. The probability graph approach, which has been applied to derive background metal concentrations in soil (Sinclair 1974, Davies 1983), was therefore tested.

Figure 3.6 shows the cumulative frequency distributions for log<sub>10</sub>-transformed concentration of Pb. At low metal concentrations, the frequency distribution curve generally appeared to be linear, except for a very small proportion of data with the lowest concentrations that tended to lie off the straight line.

The linearity of the lower part of curve is consistent with a hypothesis that the metal concentrations represent uncontaminated soils in which the frequency distribution is lognormal (Davies 1983). The departure from linearity at the lowest level is likely to be due to the problem of metal concentrations being around or below the detection limit; this small portion of data can be ignored.

The next step was to identify the percentage cumulative frequency at which the straight line changes into a complex curve, i.e. the inflexion point (indicated by an arrow in Figure 3.6).

The inflexion point for Pb was at 30 per cent, implying that 30 per cent of all soil samples are from 'contaminated' soils and 70 per cent from 'uncontaminated' soils.

A new cumulative frequency distribution for the 'uncontaminated' soils ( $F'$ ) was then calculated using the equation:

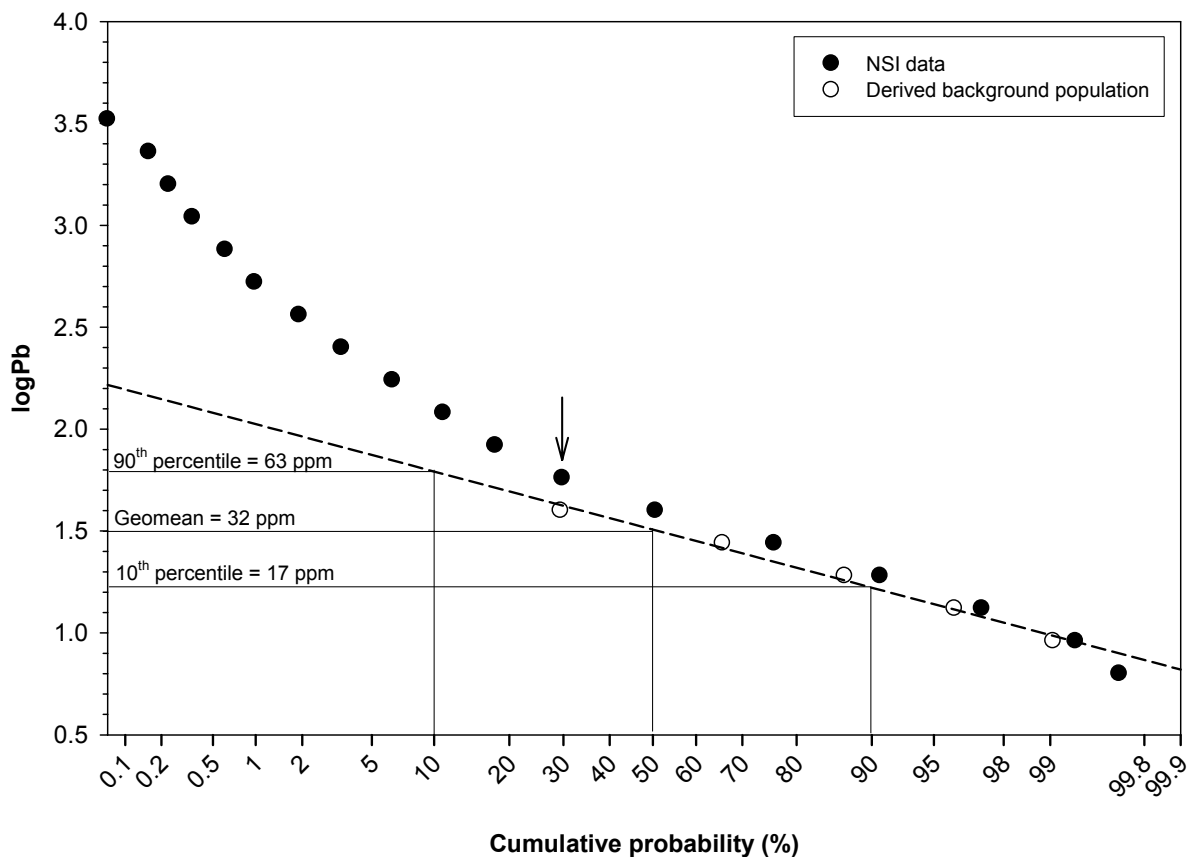
$$F' = 100 - (100 - F) \times (100/70)$$

where  $F$  is the cumulative frequency for the original data (Sinclair 1974).

A new line was then plotted of the relationship between  $\log_{10}\text{Pb}$  and  $F'$  (Figure 3.6), which represents the log-normal distribution of the 'uncontaminated' population. The mean  $\log_{10}\text{Pb}$  value corresponds to 50 per cent cumulative frequency on the new straight line and its antilog is the geometric mean (32 mg/kg Pb) of the arithmetic data. The 10th to 90th percentile values (17–63 mg/kg Pb) of this population of 'uncontaminated' soils can also be readily determined.

However, the probability graph approach has several shortcomings. First, it generates a single mean value and a single percentile range for the whole database instead of site-specific values. Secondly, the assumption that only the population representing the lowest metal concentrations is 'uncontaminated' may be too simplistic. There may be several populations in the whole dataset, with more than one (i.e. the lowest concentration) population being associated with the relatively 'uncontaminated' category.

When the probability graph method was applied to the data for Cd, Co, Cr, Cu, Ni and Zn, the 'uncontaminated' population was overly represented by soils of light textures, suggesting that their relatively low concentrations are partly due to the effect of soil type. Only in the case of Pb did the 'uncontaminated' population consist of different texture groups in similar proportions to those in the whole dataset; except that peaty soils are under-represented due to the elevated Pb concentrations in many upland peats (referred to above). Therefore, use of the probability graph approach may be suitable for Pb but not for other trace metals.



**Figure 3.6 Probability graph for soil Pb concentration in the NSI dataset**

Arrow indicates the inflexion point.

### 3.5 Ambient background concentrations of trace metals in soils according to texture group

Although it was shown that ambient background concentrations of Co, Cr and Ni can be derived using relationships based on geochemical associations and that the probability graph approach can be used for Pb, a different approach is needed for Cd, Cu and Zn.

Analysis of variance based on log-transformed data shows that soil texture has a highly significant influence on the concentrations of all seven trace elements. The percentages of variance accounted for by the texture groups are 27.8, 43.2, 46.3, 10.2, 40.3, 17.4 and 20.0 for Cd, Co, Cr, Cu, Ni, Pb and Zn respectively. There is a clear tendency for metal concentrations to increase from light to heavy textured soils – particularly for Cd, Co, Cr, Cu, Ni and Zn.

The International Organization for Standardization (ISO) recommends that the ambient background of a substance in soils can be set at a chosen parameter from the frequency distribution of usual concentrations (ISO 2004). Table 3.6 shows the range, median, 10th percentile and 90th percentile values for different soil texture groups.

The median concentrations of Cd, Co, Cr, Cu, Ni and Zn differ 3–5 fold between the clayey and sandy texture groups. In the case of Pb, sandy and peaty soils have the lowest and highest median values respectively, whereas the other texture groups show similar median values (Table 3.6). The median values for each soil texture group can be considered to represent the median ambient background concentrations (MABCs) of the group.



This approach, although not perfect, represents a significant refinement in that it takes into account a key soil property which is easily measurable in a site investigation. Alternatively, the ambient background concentration could be set at the 90 percentile value for each texture group (Table 3.6) (ISO 2004).

**Table 3.6 Summary statistics of trace metal concentrations according to soil texture group**

Soil texture group	No. of samples	Minimum	10th percen	Median	90th perce	Maximum
<b>Cd</b>						
Clayey	483	<0.2	0.6	1.0	1.6	19.2
Fine loamy	2,002	<0.2	0.4	0.8	1.3	40.9
Fine silty	1,061	<0.2	0.5	0.9	1.7	17.8
Coarse silty	182	<0.2	0.3	0.7	1.4	3.8
Coarse loamy	1,141	<0.2	<0.2	0.5	1.0	10.5
Sandy	228	<0.2	<0.2	<0.2	0.6	3.7
Peaty	557	<0.2	<0.2	0.6	1.2	5.8
<b>Co</b>						
Clayey	483	3.9	9.6	14.2	19.7	54.5
Fine loamy	2,002	<0.5	5.8	10.9	18.7	141.7
Fine silty	1,061	0.6	6.3	12.6	20.4	54.6
Coarse silty	182	1.3	5.2	10.2	18.2	42.4
Coarse loamy	1,141	<0.5	2.8	7.1	13.8	296
Sandy	228	<0.5	0.6	3.1	7.6	23.1
Peaty	557	0.8	0.7	1.7	9.3	321.8
<b>Cr</b>						
Clayey	483	18.7	42.8	59.1	82.8	837.8
Fine loamy	2,002	5	28.6	43.5	63.8	692.9
Fine silty	1,061	6	28.7	48.1	66.0	285.4
Coarse silty	182	6.3	23	39.1	66.2	143.5
Coarse loamy	1,141	<1	16.3	27.4	45.3	356.3
Sandy	228	<1	5.7	13.2	27.0	91.5
Peaty	557	<1	3.5	12.2	38.8	153.7
<b>Cu</b>						
Clayey	483	6.9	16.2	23.2	39.4	157.4
Fine loamy	2,002	3.7	11.2	18.8	36.3	933.5
Fine silty	1,061	4	10.7	18.6	32.4	322.6
Coarse silty	182	5.8	10.5	18.6	34.5	74.1
Coarse loamy	1,141	1.8	7	14.7	37.9	1507.7
Sandy	228	1.2	3.3	7.4	20.6	253.3
Peaty	557	1.9	6.4	15	43.7	930
<b>Ni</b>						
Clayey	483	10.5	25	38.2	51.1	194.6
Fine loamy	2,002	1.6	14	25.3	41.2	439.5
Fine silty	1,061	2.3	14.5	28.2	44.2	298.8
Coarse silty	182	4.5	13	22.4	39.2	89.7
Coarse loamy	1,141	<2	6.8	15.9	29.3	436.4
Sandy	228	<2	2	7.5	18.3	74.3
Peaty	557	<2	2.9	6.6	24.4	123.9
<b>Pb</b>						
Clayey	483	6	21	35	71	451
Fine loamy	2,002	<4.7	21	39	99	16338
Fine silty	1,061	<4.7	24	39	92	3461
Coarse silty	182	13	20	42	103	479
Coarse loamy	1,141	5	18	36	114	3697
Sandy	228	<4.7	10	22	54.8	275
Peaty	557	8	35	116	314	1647

**Table 8 (continued)**

<b>Soil texture group</b>	<b>No. of samples</b>	<b>Minimum</b>	<b>10th percen</b>	<b>Median</b>	<b>90th perce</b>	<b>Maximum</b>
<b>Zn</b>						
Clayey	483	36	75	106	157	962
Fine loamy	2,002	16	54	89	151	3648
Fine silty	1,061	19	57	90	156	2125
Coarse silty	182	18	50	78	145	299
Coarse loamy	1,141	10	33	65	132	1985
Sandy	228	5	12	35	86	2017
Peaty	557	12	23	52	104	1818

# 4 Conclusions

- Strong geochemical associations were found between the trace metals Co, Cr and Ni and the major elements Al and Fe in the NSI dataset.
- Multiple regressions linking the log-transformed concentrations of Co, Cr or Ni with Al and Fe were obtained which explained 62–85 per cent of the variation in the trace metal concentrations.
- Site-specific ambient background concentrations of Co, Cr or Ni in topsoils could be predicted reasonably well from Al and Fe concentrations.
- There was no strong association between the trace metals Cd, Cu, Pb or Zn with Al and Fe. This is consistent with geochemical principles. Therefore, ambient background concentrations of these trace metals could not be predicted using regressions with Al and/or Fe.
- A probability graph approach was used to identify an ‘uncontaminated’ population of soils from the whole dataset. This approach yielded estimates of the geometric mean and the 10th to 90th percentile range of trace metals that might be associated with the ‘uncontaminated’ population. However, this method may be appropriate only for Pb because, for other trace metals, the ‘uncontaminated’ population was over-represented by soils with light textures.
- Soil texture has a major influence on trace metal concentrations, with heavy textured soils having higher concentrations of Cd, Co, Cr, Cu, Ni and Zn than light soils, and peaty soils having higher Pb concentrations than other soils. Median concentrations of trace metals were calculated for seven soil texture groups and can be used as ambient background concentrations for specific soil textures.

# References & Bibliography

- Cullen A C and Frey H C, 1999 *Probabilistic Techniques in Exposure Assessment. A Handbook for Dealing with Variability and Uncertainty in Models and Inputs*. New York: Plenum.
- Davies B E, 1983 A graphical estimation of the normal lead content of some British soils. *Geoderma*, 29, 67-75.
- European Commission Joint Research Centre., 2003 *Technical Guidance Document on Risk Assessment*. 2nd edn. Luxembourg: Office for Official Publications of the European Communities. Available from: <http://ecb.jrc.it/Technical-Guidance-Document/> [Accessed 25 September 2006].
- Goldschmidt V M, 1954 *Geochemistry*. Oxford: Clarendon Press.
- Hamon R E, McLaughlin M J, Gilkes R J, Rate A W, Zarcinas B, Robertson A, Cozens G, Radford N and Bettenay L, 2004 Geochemical indices allow estimation of heavy metal background concentrations in soils. *Global Biogeochemical Cycles*, 18, GB1014.
- Helling C S, Chesters H G and Corey R B, 1964 Contribution of organic matter and clay to soil cation-exchange capacity as affected by the pH of the saturating solution. *Soil Science Society of America Proceedings*, 28, 517-520.
- International Organization for Standardization (ISO), 2004 *Soil quality: guidance on the determination of background values*. ISO CD 19258 (draft document). Geneva: ISO.
- McGrath S P and Cunliffe C H, 1985 A simplified method for the extraction of the metals Fe, Zn, Cu, Ni, Cd, Pb, Cr, Co and Mn from soils and sewage sludges. *Journal of the Science of Food and Agriculture*, 36, 794-798.
- McGrath S P and Loveland P J, 1992 *The Soil Geochemical Atlas of England and Wales*. London: Blackie Academic & Professional.
- Oliver M A, Loveland P J, Frogbrook Z L, Webster R and McGrath S P, 2002 *Statistical and geostatistical analysis of the National Soil Inventory of England and Wales*. Final Report to Department of Environment, Food and Rural Affairs. Silsoe, Bedford: National Soil Resources Institute, Cranfield University.
- Oorts K, Ghesquière U, Swinnen K and Smolders E, 2006 Soil properties affecting the toxicity of CuCl<sub>2</sub> and NiCl<sub>2</sub> for soil microbial processes in freshly spiked soils. *Environmental Toxicology and Chemistry*, 25, 845-853.
- Rooney C P, Zhao F J, and McGrath S P, 2006 Soil factors controlling the expression of copper toxicity to plants in a wide range of European soils. *Environmental Toxicology and Chemistry*, 25, 726-732.
- Sinclair A J, 1974 Selection of threshold values for geochemical data using probability graphs. *Journal of Geochemical Exploration*, 3, 129-149.
- Smolders E, McGrath S P, Lombi E, Karman C C, Bernhard R, Cools D, Van Den Brande K, Van Os B and Walrave N, 2003 Comparison of toxicity of zinc for soil microbial processes between laboratory-contaminated and polluted field soils. *Environmental Toxicology and Chemistry*, 22, 2592-2598.
- Smolders E, Buekers J, Oliver I and McLaughlin M J, 2004 Soil properties affecting toxicity of zinc to soil microbial properties in laboratory-spiked and field-contaminated soils. *Environmental Toxicology and Chemistry*, 23, 2633-2640.

Sterckeman T, Douay F, Baize D, Fourrier H, Proix N, Schwartz C and Carignan J, 2006 Trace element distributions in soils developed in loess deposits from northern France. *European Journal of Soil Science*, 57, 392-410.

Tukey J W, 1977 *Exploratory Data Analysis*. Reading, MA: Addison-Wesley.

# List of abbreviations

ABC	ambient background concentration
CEC	cation exchange capacity
CR	confidence range
eCEC	effective cation exchange capacity
ISO	International Organization for Standardization
MABC	median ambient background concentration
NSI	national soil inventory
OC	organic carbon
PCA	principal component analysis
PNEC	predicted no effect concentration
SE	standard error
TGD	EU Technical Guidance Document

We are The Environment Agency. It's our job to look after your environment and make it **a better place** – for you, and for future generations.

Your environment is the air you breathe, the water you drink and the ground you walk on. Working with business, Government and society as a whole, we are making your environment cleaner and healthier.

The Environment Agency. Out there, making your environment a better place.

Published by:

Environment Agency  
Rio House  
Waterside Drive, Aztec West  
Almondsbury, Bristol BS32 4UD  
Tel: 0870 8506506  
Email: [enquiries@environment-agency.gov.uk](mailto:enquiries@environment-agency.gov.uk)  
[www.environment-agency.gov.uk](http://www.environment-agency.gov.uk)

© Environment Agency

All rights reserved. This document may be reproduced with prior permission of the Environment Agency.