Normal background concentrations of contaminants in the soils of England. Available data and data exploration

Land Use Planning and Development Programme

Commissioned Report CR/11/145 N

BRITISH GEOLOGICAL SURVEY

LAND USE PLANNING AND DEVELOPMENT PROGRAMME
COMMISSIONED REPORT CR/11/145 N

Normal background concentrations of contaminants in the soils of England. Available data and data exploration

E L Ander, M R Cave, C C Johnson and B Palumbo-Roe

Keywords

contaminant, normal concentration, soil, background.

Bibliographical reference

ANDER, EL, CAVE, MR, JOHNSON, CC AND PALUMBO-ROE, B. 2011. Normal background concentrations of contaminants in the soils of England. Available data and data exploration. *British Geological Survey Commissioned Report*, CR/11/145. 124pp.

Keyworth, Nottingham British Geological Survey 2011

BRITISH GEOLOGICAL SURVEY

The full range of our publications is available from BGS shops at Nottingham, Edinburgh, London and Cardiff (Welsh publications only) see contact details below or shop online at www.geologyshop.com

The London Information Office also maintains a reference collection of BGS publications, including maps, for consultation.

We publish an annual catalogue of our maps and other publications; this catalogue is available online or from any of the BGS shops.

The British Geological Survey carries out the geological survey of Great Britain and Northern Ireland (the latter as an agency service for the government of Northern Ireland), and of the surrounding continental shelf, as well as basic research projects. It also undertakes programmes of technical aid in geology in developing countries.

The British Geological Survey is a component body of the Natural Environment Research Council.

British Geological Survey offices

BGS Central Enquiries Desk
Tel 0115 936 3143 Fax 0115 936 3276
email enquiries@bgs.ac.uk

Kingsley Dunham Centre, Keyworth, Nottingham NG12 5GG
Tel 0115 936 3241 Fax 0115 936 3488
email sales@bgs.ac.uk

Murchison House, West Mains Road, Edinburgh EH9 3LA
Tel 0131 667 1000 Fax 0131 668 2683
email scotsales@bgs.ac.uk

Natural History Museum, Cromwell Road, London SW7 5BD
Tel 020 7589 4090 Fax 020 7584 8270
Tel 020 7942 5344/45 email bgslondon@bgs.ac.uk

Columbus House, Greenmeadow Springs, Tongwynlais, Cardiff CF15 7NE
Tel 029 2052 1962 Fax 029 2052 1963

Maclean Building, Crowmarsh Gifford, Wallingford OX10 8BB
Tel 01491 838800 Fax 01491 692345

Geological Survey of Northern Ireland, Colby House, Stranmillis Court, Belfast BT9 5BF
Tel 028 9038 8462 Fax 028 9038 8461

www.bgs.ac.uk/gsni/

Parent Body

Natural Environment Research Council, Polaris House, North Star Avenue, Swindon SN2 1EU
Tel 01793 411500 Fax 01793 411501
www.nerc.ac.uk

Website www.bgs.ac.uk
Shop online at www.geologyshop.com

Foreword

This report presents the results from the first two work packages of a Defra-funded Science and Research project to establish normal background contaminant concentrations in the soils of England. The project (Project reference SP1008: *Establishing data on normal/background levels of soil contamination in England*) commenced 5th October 2011 and is scheduled to end 31st March 2012. Work Package 1 (WP1) is concerned with a review of existing data and Work Package 2 (WP2) an exploration of the data. A methodology to determine these concentrations will be reported as part of Work Package 3 (WP3, due in January 2012). Technical guidance in the use of normal background concentrations will be written for contaminants for which NBCs can be determined by the end March 2012 (Work Package 4, WP4).

Acknowledgements

The following licences have been used with respect to some of the data sets investigated:

- Environmental Change Network (ECN): CEH – ECN:CJ10/11 – issued 6th December 2011 for one year
- Countryside Survey 10 km grid references for CS plot: CEH – issued 1st December 2011 for one year
- NSI Topsoil data – issued 12th December 2011 for one year (Cranfield ref: WU13058V)
- For the mineralisation and mining data layer based on the Department of Environment Ove Arup “Mining Instability in Britain” contract, copyright ownership of the “derived product” - the ArcView mining instability dataset - is vested in NERC/BGS (BGS IPR reference IPR35-10)

Contents

FIGURES

TABLES

Summary

This draft report describes the completed Work Package 1 and 2 tasks of the Defra commissioned Science and Research Project SP1008 - *Establishing data on normal/background levels of soil contamination in England*. Work Package 1 (WP1) has investigated the available soil data sets that can be used to determine the Normal Background Concentrations (NBCs) for contaminants in soils from England. A database of available data set information has been created investigating the sample coverage, the sampling strategy and methods, the laboratory procedures, and the robustness, format and availability of the data. Key data sets are the British Geological Survey's G-BASE topsoils for urban and rural areas and the National Soil Inventory (XRFS reanalysed) data which, although at a reduced sampling density compared to G-BASE, has complete coverage for both England and Wales. Supplementary minor data sets are also included in the Project's extensive bibliographic database. There are also large data sets available that specifically target a particular land use or activity, *i.e.* not systematically collected, *e.g.* the BGS Mineral Reconnaissance Programme (MRP) soils. Such data will not be used to estimate NBCs as they will cause a substantial bias in the determinations.

This data and information gathering phase will underpin subsequent work packages of this project. Work Package 1 has also addressed some key questions such as, "*what are the priority contaminants*?" and "*what is meant by a normal background concentration*?"

By exploring the data sets available, gaps in information can be identified. Whilst the key data sets give good coverage for the inorganic elements (with the exception of mercury (Hg)), there is less data available on the distribution of organic contaminants in the environment. To some extent the key data sets can be supplemented by other smaller data sets such as the UK Soil and Herbage Pollutant Survey, the Countryside Survey (both providing data on Hg and organics) and some European-wide geochemical mapping projects such as the FOREGS and GEMAS data. There are also a number of peer-reviewed publications that provide valuable additional information for the less widely surveyed contaminants (such as organic substances and Hg). Such peer-reviewed publications are included in the Project's EndNote bibliographic database.

Work Package 2 (WP2) has focused on exploring the data identified in WP1 and, in particular, investigated in detail four contaminants – arsenic (As), lead (Pb), benzo[*a*]pyrene (BaP) and asbestos. However, given the very limited amount of information on naturally occurring asbestos minerals in soils, very little data exploration of this contaminant is possible.

WP2 has defined the domains to use in the methodology proposed for Work Package 3 (WP3). These will be used to establish NBCs where it is possible to take into account more local influences on contaminant concentrations in the soil. Three key spatial data sets have been identified for use in the WP3 methodology – the BGS Soil-Parent Material Model (SPMM) for classifying the most important geogenic and pedological impacts on contaminants in soil; the 'Ove Arup' (originally Department of Environment sponsored) mineralisation and historical mining database; and the Office of National Statistics (ONS) Generalised Land Use Database (GLUD) which can be used to determine an urbanisation index (UI) for England at a sufficiently resolved scale. The first two databases can be used to relate 'natural' contaminant concentrations to the normal background, whereas the third data set gives a good indication of areas likely to be impacted by diffuse anthropogenic pollution.

Data exploration of As, Pb and BaP has produced much new quantifiable evidence concerning these contaminants that will give useful information to include in the technical guidance for Work Package 4. Summary data tables, plots and maps generated by WP2 will be extremely useful in future publications concerning the NBCs of As, Pb and BaP. The exploration has

demonstrated how the distribution of these contaminants across England can be quantified. Arsenic has significant areas where underlying ironstones supply high levels of natural As, as do some, but not all of the metalliferous mineralised/mining areas. The anthropogenic input of As in urban areas is much less significant than for Pb, which also has association with some notable natural areas of mineralisation but not such a strong association with underlying parent material. Data for BaP is more sparse, but the anthropogenic diffuse input in urban areas, declining through semi-urban to rural regions is well demonstrated.

WP2 also investigates other non-spatial aspects of the data sets. ‘Total’ and ‘partial’ analytical method results are compared for As and Pb giving a linear regression equation that demonstrates how such values can be compared, though also demonstrating some of the flaws with the partial analysis results. The importance of organic matter for fixing contaminants in the soil is also demonstrated.

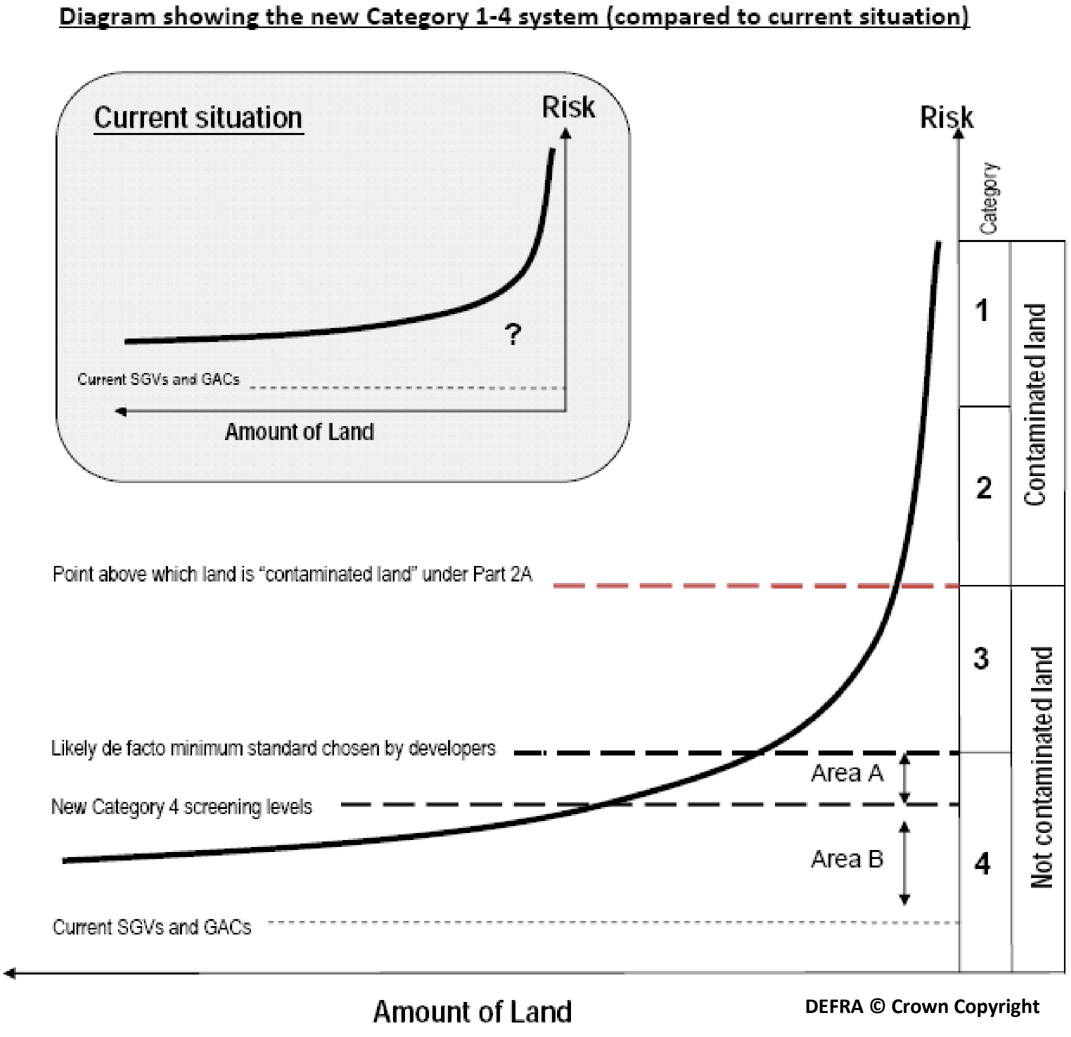
The work of WP1 has demonstrated how difficult some of the data from the key data sets is to obtain. This is data that can usefully be used by those tasked with assessing potentially contaminated land and ways to make the data far more accessible should be explored.

The work of WP1 and WP2 has been conducted to a large extent in parallel with the methodology of WP3. The Project is now ready to proceed with WP3 to determine the NBCs using robust statistical procedures for key contaminants in the important domains defined by WP2.

1 Introduction

1.1 BACKGROUND

The work described here is part of an initiative to simplify the contaminated land regime for England and Wales where there is a legacy of land contamination from industrial activity and urbanisation and areas where there are high natural levels of some contaminants. Statutory Guidance is issued by the Secretary of State for the Department of Environment, Food and Rural Affairs (Defra) in accordance with section 78Y of the Environmental Protection Act 1990. Section 57 of the Environment Act 1995 created Part 2A of the Environment Protection Act 1990 establishing a legal framework for dealing with contaminated land (DETR, 2000). The Statutory Guidance is intended to explain how the contaminated land regime should be implemented. However, the Guidance, which is supposed to explain when land does (and does not) need to be remediated, has created significant uncertainties. Therefore, revision of the Statutory Guidance intends to make it more usable for those working with contaminated land and remediation (DEFRA, 2011a). A new four category test is proposed to help decide when land is, and is not contaminated land (Figure 1). Category 1 describes land which is clearly problematic, for example, because similar sites are known to have caused a significant problem in the past. Category 4 describes land that is clearly not contaminated. Categories 2 and 3 is land for which categorisation is less straightforward.



DEFRA © Crown Copyright

Figure 1: The new four category system to help test when land is, and is not contaminated (DEFRA, 2011b)

This project is one of many steps to realise the revision of the Statutory Guidance and will produce technical guidance to describe “normal” levels of contamination. This technical guidance, along with new category 4 screening levels (C4SLs), will help to define clearly Category 4 land.

1.2 PROJECT OBJECTIVES – WP1 AND WP2

The project objectives covered by this report are detailed in the project proposals (BGS, 2011) which form part of the contract of work.

1.2.1 Review of existing data - Work Package 1 (WP1)

This is the initial phase of the project that underpins subsequent objectives:

- Tabulation of existing data sets;
- Maps of England showing sample sites from main data sets (ArcGIS layers);
- Tabulation of contaminants assessing coverage/gaps in the evidence base;
- Glossary of terms (to ensure consistent use of terminology throughout project); and
- EndNote bibliography of key reports/publications (to underpin project’s knowledge base).

1.2.2 Exploration of the data - Work Package 2 (WP2)

This is the second phase of the project following on from the initial data gathering phase (WP1). Investigation of the data will explore the statistical methodology to be used in WP3, by focussing on a suite of key contaminants that have contrasting sources (natural/diffuse anthropogenic) and spatial controls on their distribution. The contaminants will be explored spatially by areas defined as domains. Component parts of WP2 (delivered in draft report):

- Definition of classifiers, classes and domains and assignment of contaminants to domains;
- Descriptive statistics for contaminant domains;
- Worked examples on use of proxy data for areas where there are gaps in data; and
- Draft report for WP1 & 2.

1.3 SCOPE OF THIS WORK - LIMITS

The preceding discussions cover the scope of this project with the objectives for the initial phases clearly set out in Section 1.2. It is equally important to understand some of the limits of the work. The normal background concentrations:

- are not a definition of the C4SLs - but could feed into their derivation - as they provide information expressed as quantitative statistically derived values;
- are not trigger levels to be used in any risk assessments and are derived independently of any measure or assessment as to whether there is a potential risk or harm to receptors; and
- are best estimates derived from robust statistical analysis based on our current knowledge. As more data becomes available they will need to be refined.

2 Review of Existing Data

2.1 INTRODUCTION

The chemical and physical properties of a soil are the result of complex interactions of natural and anthropogenic processes over a period of time. In particular, since the industrial revolution at the end of the 18th century, significant areas of soil have been affected by direct or less direct (diffuse) pollution, and many urban areas have been left with legacy of contamination. However, for the overwhelming majority of English soils, the single-most important factor in determining their chemical composition is the parent material from which it has been derived. This is demonstrated, for example, in the BGS environmental geochemical atlas of Central and Eastern England (BGS, 2010) and forms the basis for identifying the important domains with which to attribute Normal Background Concentrations (NBCs) of contaminants in soil. England has a varied geology - the geological map of England and Wales (Figure 2) reflects the presence of many different formations, and with this comes a similar variability in the chemical composition of the soils.

Both industrial and agricultural activities have added an increasingly diverse range of contaminants to the environment. Even the apparently most pristine of rural environments will have experienced elevated levels of contaminant concentrations above their natural values. Contaminants have been dispersed at a national and global scale as diffuse pollution through atmospheric deposition, for example, coal burning. However, it must be remembered that a soil acts not only as a sink for contaminants but also as a natural source.

A soil is likely to contain most of the elements of the Periodic Table with concentrations ranging from ultra-trace amounts to a presence measured in weight percent. Elements combine with other elements to form minerals or other chemical substances, so any particular element may be present in the soil in many different forms, including as something classified as a contaminant. Furthermore, certain soil chemical parameters, such as pH (acidity/alkalinity) and organic matter, can be important in controlling the mobility of contaminants, which in turn may influence the hazardous properties of soil.

The distribution of chemical elements within soil is of interest to a wide variety of scientific disciplines. Soil science is concerned with the soil as a natural resource and in their management. Their chemical properties, along with the physical and biological properties, need to be mapped and understood. Geochemists are also interested in mapping the behaviour and distribution of chemical elements at the earth's surface for a variety of sample types including soil. Indeed, as with many scientific challenges in the management and exploitation of resources, and concerns regarding the environmental and health impacts of changes to the chemical surface environment, many of the traditional sciences now work together under the umbrella of environmental science. As a result, in the review of existing soil chemical data, information from many scientific disciplines has been investigated.

In a review of the soil chemical data suitable for estimating NBCs, certain criteria need to be established before the value of the data is known. This includes a definition as to what is meant by "normal background", what are the priority contaminants, and how the contaminant concentrations have been determined and mapped.

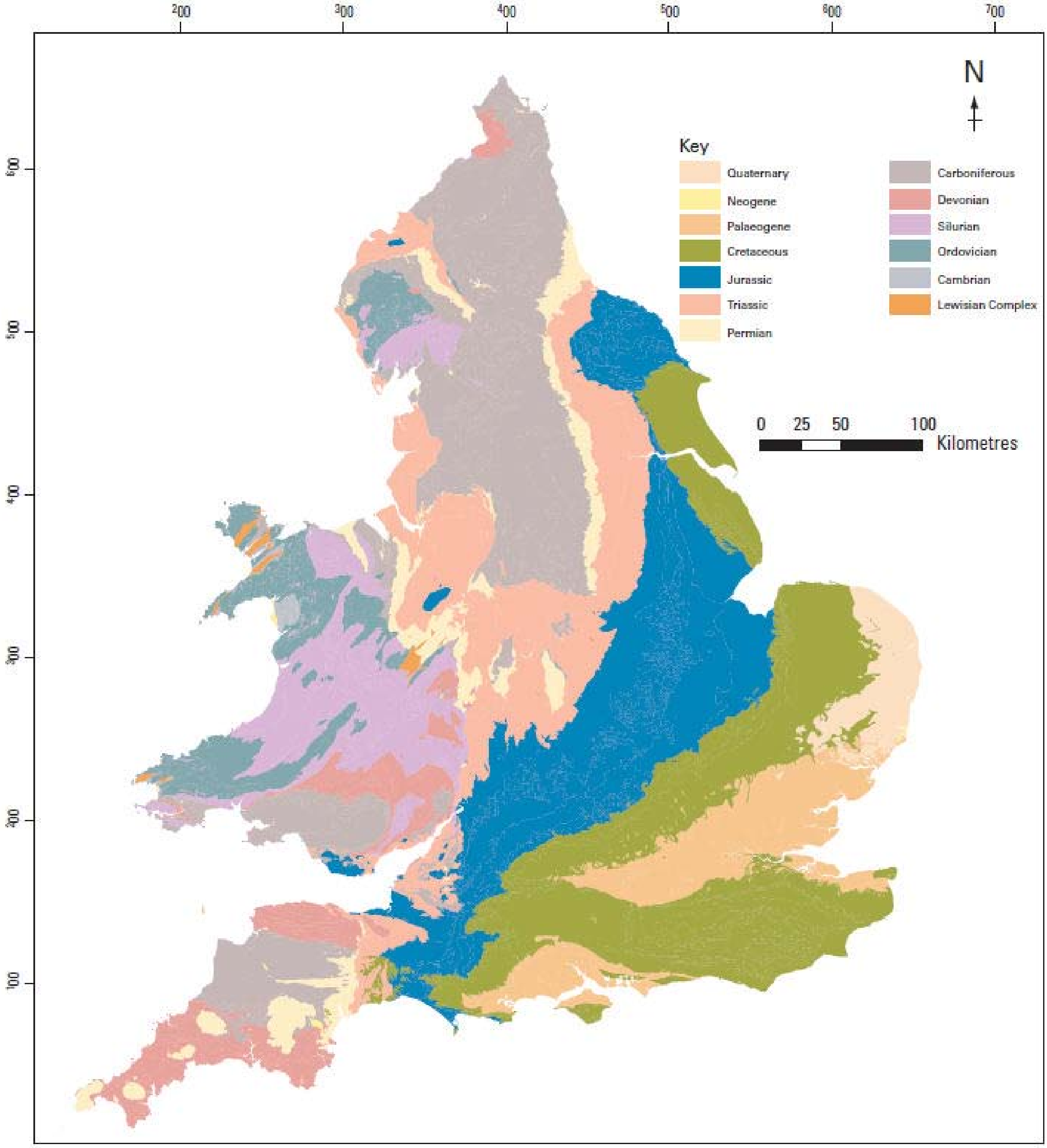


Figure 2: A simplified geological map of England and Wales (after Rawlins *et al.,* 2012)

2.2 NORMAL BACKGROUND CONCENTRATION (NBC)

A review of how NBCs have previously been defined and determined will be given in the WP3 report on *"Methodology for the determination of normal background contaminant concentrations in English soils"*. However, in this current report it is important to clearly define what is meant by the terms "Normal" and "Background" and why it is appropriate to use these two words together.

2.2.1 Definitions

There are a number of terms that are used to convey the expected concentrations of a contaminant in soil. These include: normal, typical, baseline, ambient, characteristic, natural, background and widespread. There are some subtle differences between these terms, they can mean different things in different disciplines, and can be confused with other alternate uses. For example, a statistician would associate the word “normal” when used in context of defining the spread of a population of results, *i.e.* a normal (or Gaussian) distribution. The terms normal, typical, characteristic and widespread are more or less synonymous. The important thing is that normal is the term used in the draft Statutory Guidance (Defra, 2011b) and so will be the term used throughout this project. In Section 3 of the Guidance it is noted that “normal” presence/levels of contaminants:

- *should not be considered to cause land to qualify as contaminated land, unless there is a particular reason to consider otherwise (Guidance, Section 3.22);*
- *may result from the natural presence of contaminants at levels that might reasonably be considered **typical** in a given area and have not been shown to pose an unacceptable risk to health or the environment (Guidance, Section 3.23(a)); and*
- *are caused by low level diffuse pollution, and common human activity other than specific industrial processes (Guidance, Section 3.23(b)).*

The last bullet point illustrates the need to refer to contaminant levels as “normal background”.

The term background has a more complex and varied usage than the term normal and is used differently in different areas of science, for example:

- in exploration geochemistry the term background has been long-established and defines an area of normal element concentrations distinguished from anomalously high concentrations (that may indicate the presence of metalliferous mineralisation) by a threshold value;
- in environmental geochemistry background is a relative measure to distinguish between natural element or compound concentrations and anthropogenically-influenced concentrations in real sample collectives (Matschullat *et al*., 2000); and
- in the ISO 19258:2011 (ISO, 2011) guidance on soil background the content of a substance in a soil resulting from both natural geological and pedological processes and including diffuse source inputs.

Background as used in the draft Statutory Guidance is therefore similar to the ISO 19258:2011 guidance. It is important that normal concentrations are qualified by the term background so as to capture that it includes both natural and anthropogenic diffuse pollution. In this way the term “normal background” embraces Section 3.23(b) of the Guidance:

For the purpose of this Guidance, “normal” levels of contaminants in soil may result from:

> *(b) The presence of contaminants caused by low level diffuse pollution, and common human activity other than specific industrial processes. For example, this would include diffuse pollution caused by historic use of leaded petrol and the presence of benzo[a]pyrene from vehicle exhausts, and the spreading of domestic ash in gardens at levels that might reasonably be considered typical.*

2.3 CONTAMINANTS

2.3.1 Definition

A contaminant can be defined in many ways and in its broadest sense it is "something that is an unwanted constituent". In the draft Statutory Guidance (DEFRA, 2011b) the terms "contaminant", "pollutant" and "substance" are used with the same meaning, that is, " **a substance relevant to the Part 2A regime (DETR, 2000) which is in, on or under the land and which has the potential to cause significant harm to a relevant receptor, or to cause significant pollution of controlled waters**". This is the definition used for the purpose of this report. Detailed clarification of what is meant as significant harm and relevant receptor is also given in the draft Statutory Guidance, receptors falling into three categories: human, ecological systems and property (*e.g*. buildings, crops and livestock).

Literature covering aspects of contamination give less robust definitions of a contaminant. For example, ISO 19258:2011 (ISO 2011) defines a contaminant as substance or agent present in the soil as a result of human activity, and notes that there is no assumption in this definition that harm results from the presence of the contaminant. Therefore, this definition does not include natural contaminants. Cole and Jeffries (2009) in their report on using soil guideline values (SGVs) say a contaminant is a substance that is in, on or under the land and has the potential to cause harm.

2.3.2 Classification and Prioritisation

There are thousands of potential contaminants which might be present on various sites around England (although a smaller sub-set probably drives the risk on most sites) (DEFRA, 2008). The Defra-EA report on "Potential Contaminants for the Assessment of Land" (DEFRA-EA, 2002) identified the priority chemicals for the development of SGVs. This was based on the chemicals likely presence in sufficient concentrations on affected UK sites that were considered to pose a risk to humans, buildings, water resources or ecosystems. This report classified the chemicals into five categories: metals, semi-metals/non-metals, inorganic chemicals, organic chemicals and others (asbestos and pH). An updated priority chemicals list is presented by Martin and Cowie (2008). This list is given in Table 1 and contains fifty six chemicals, fourteen of which are chemical elements plus cyanides (an inorganic substance) and asbestos (mineralogically defined). The remaining contaminants can be classified as organic substances and these in the UK soil environment will be overwhelmingly associated with (though not exclusively) anthropogenic activity. Radioactive elements are considered to be outside the scope of this work. Globally there is generally good agreement as to what are priority contaminants, though there are national differences. In Finland, for example, the "Government Decree on the Assessment of Soil Contamination and Remediation" (Finnish Government, 2007) lists eleven inorganic elements: antimony (Sb), arsenic (As), mercury (Hg), cadmium (Cd), cobalt (Co), chromium (Cr), copper (Cu), lead (Pb), nickel (Ni), vanadium (V) and zinc (Zn). Two elements, Sb and Co, are not on the UK list.

A broader classification of contaminants is discussed by Smith *et al*. (1999) who note that land contamination manifests itself in a wide range of physical, chemical or biological forms:

1. Chemical contaminants;
    a. organics
    b. inorganics
    c. munitions
    d. salinity
2. Biological pathogens;

3. Radiochemicals; and
4. Physical contaminants, including made ground and mine workings.

Smith *et al*. (1999) also reiterate that the presence of a contaminant has no assumption of any resulting harm, simply that it may affect the quality of the land.

These work packages are concerned with the chemical contaminants and will focus primarily on those listed in Table 1, particularly As, Pb, benzo[*a*]pyrene (BaP) and asbestos as these were named in the Project’s proposals (BGS, 2011). Natural and diffuse pollution occurrences and sources of these and further contaminants shown in Table 1 are summarised in Appendix 1.

Inorganic	***Organic***	
Arsenic (As)	acetone	fenitrothion
Beryllium (Be)	aldrin	hexachlorobuta-1,3-diene
Cadmium (Cd)	atrazine	hexachlorocyclohexanes
Chromium (Cr)	azinphos-methyl	malathion
Copper (Cu)	benzene	naphthalene
Lead (Pb)	Benzo[*a*]pyrene (BaP)	organolead compounds
Mercury (Hg)	carbon disulphide	organotin compounds
Molybdenum (Mo)	carbon tetrachloride	pentachlorophenol
Nickel (Ni)	chloroform	phenol
Selenium (Se)	chlorobenzenes	polychlorinated biphenyls (PCB)
Sulphur (S)	chlorophenols	polycyclic aromatic hydrocarbons (PAH)
Thallium (Tl)	chlorotoluenes	tetrachloroethane
Vanadium (V)	1,2-dichloroethane	tetrachloroethene
Zinc (Zn)	dichlorvos	toluene
cyanide	DDT	total petroleum hydrocarbons
asbestos	dieldrin	2, 1,1,1-trichloroethane
	dioxins and furans	trichloroethene
	endosulfan	trifluralin
	ethylbenzene	vinyl chloride
	explosives	xylenes

Table 1: List of priority contaminants (from Martin and Cowie, 2008)

2.4 CONCENTRATIONS

2.4.1 Terminology

There are a variety of terms used to describe the presence of contaminants in soils and these include: contents, levels, values, results and concentrations. In this report the term concentration is preferred, particularly when in the context of a defined quantity. However, when referring to contaminants in a qualitative or relative way, *e.g.* high or low, “levels” is an equally acceptable term. The presence of a contaminant in a soil is generally expressed in terms of the weight[1] of the chemical per a unit weight of the dried soil, *e.g.* mg/kg or mg/g. In geochemistry the terms parts per million (ppm is equivalent to mg/kg) and parts per billion

---

[1] *Asbestos is an exception. This is generally measured by the number of mineral fibres per volume of material*

(ppb is equivalent to mg/g) are still encountered though this usage is now discouraged in favour of the SI units.

Geochemists, by convention, will often express major element concentrations *e.g.* MnO, Fe_2O_3, in terms of oxide weight per cent, but as the inorganic elements under consideration here are trace elements (Table 1) this should not apply.

2.4.2 Measurement

Concentrations are measured by analytical methods which can be described as either destructive or non-destructive. In destructive determinations the soil is destroyed by extracting minerals into a solution (*e.g.* acid extraction or fusion/leaching) followed by measurements made on that solution (*e.g.* inductively coupled plasma mass spectrometry (ICP-MS)). In non-destructive determinations measurements are made directly on the soil (*e.g*. X-ray fluorescence spectrometry (XRFS) and neutron activation analysis (NAA)) which remains intact after the analysis, though XRFS may involve the addition of a binder to help create pellets. Portable XRFS devices are available for use in the field – all references to XRFS in this document refer to laboratory based analyses.

There are many issues regarding how contaminant concentrations are measured. These relate to how the soil sample is collected, how it is prepared for analysis, what and if any extraction procedure is used and what analytical instrument is used for the determination. No results for the concentration of a contaminant in soil can be interpreted without the provision of this information. Protocols for standardising sample collection and analysis are available from many sources (*e.g.* international geochemical sampling – Darnley *et al*. (1995); G-BASE sampling – Johnson (2005); UK Soil and Harbage Pollutant Survey – Woods *et al*. (2007); soil collection and storage - ISO:10381-2 (2002)) though standard protocols tend to vary between different scientific disciplines and the term “standard” does not imply any measure of acceptance and may be driven by economic rather than scientific considerations.

Many analytical measurements of contaminants are quoted as ‘total’ concentrations. For inorganic contaminants, determining true total concentrations usually involves analysis by non-destructive XRFS or NAA. Destructive analytical methods involving sample digestion in an acid solution, for example, a combination of hydrofluoric-nitric-perchloric acid, gives “near total” results. Less vigorous reagents used in the extraction, such as *aqua regia*, will yield more readily available concentration of contaminants in soils. In more recent years extraction procedures have been developed that mimic the uptake of certain contaminants into specified receptors, for example, the unified BARGE method Wragg *et al*. (2011) predicts the human bioaccessible fraction of a contaminant in a soil.

2.4.3 Quantifying a normal background concentration

When modelling the distribution of contaminants in soils it is necessary to analyse many soils collected over a region. The level of uncertainty attached to the model will depend on the density and numbers of soils collected and analysed. The statistical approach to this will be addressed in Work Package 3 (WP3). Contaminant analysis of samples between sites, within sites (*i.e.* duplicates), and the same sample (*i.e.* replicates) will be associated with a range of contaminant concentrations (known as sample, sampling and analytical variability, respectively). These concentrations will form a spread of values collectively known as a population and described by density distribution plots such as histograms or cumulative frequency plots. Selected contaminant population densities, with the aid of such plots, are described in Section 3. Different parts of the population can be assigned to differing causes, for example, if you consider the Pb in topsoils for the entire country (see Section 3.2.2), the results can be partitioned into populations associated with urban diffuse pollution, mineralisation and mining, and “other” areas. These populations may numerically overlap, or even be enclosed within another population.

Each population can be described either graphically (*e.g.* boxplot) or by statistically derived numbers (quantiles, mean, median, standard deviation, skewness *etc*.). It will be the objective of WP3 of this Project, through robust statistical methodology that partitions data populations into domains and removes outlier data points, to quantify NBCs. As normal background contaminant concentrations are to be provided to assessors, guidance in their use needs to be associated with a range of statistical parameters rather than a single value such as the mean. The provision of a single parameter is likely to lead to assessors using that value as a trigger value for contaminated land assessment.

2.5 AVAILABILITY OF DATA FOR CONTAMINANT CONCENTRATIONS IN ENGLISH SOILS

WP1 of this project is concerned with a review of the main data sources that are pertinent to the assessment of NBCs in soils. This is specifically English soils, though not exclusively, as data from other regions can provide useful supplementary information. The main data sets are summarised in Appendix 2 as tables derived from the Project’s MS ACCESS 2007 database. It is expected this database will be continually updated during the Project as new information becomes available. Supplementing this database of useful soil data sets is the Project’s EndNote[2] bibliography.

The soil data is summarised by source in three data tables with the following information:

1. **Basic Information**
    a. Data Set – Key field, label (usually abbreviation) identifying the data set
    b. Origin Organisation – the main organisation responsible for the data set
    c. Coverage – a description of the data set area with, where given, an indication of sampling density
    d. Class – a code used to classify the data set (R rural; U urban; I industrial; * major usefulness; # moderate usefulness; ~ minor usefulness; 0 no significant use)
    e. Period of collection – year or range of years of sample collection
    f. Number of samples – number of soil samples in data set (for England unless otherwise cited). Sometimes qualified to being the number of sites
    g. Sampling Strategy – the rationale behind the soil sampling programme
    h. Sampling Method – a summary of the information provided for the soil collection
    i. Reference – a record number referring to a supplementary bibliographic list given in record number order. These are references copied from the Project’s EndNote bibliography
2. **Analysis Information**
    a. Data Set - Key field, label (usually abbreviation) identifying the data set
    b. Sample Preparation – a summary of the information provided on the sample preparation of the soil
    c. Pre-analysis procedures – summarises laboratory methods used prior to contaminant determination, specifically any extraction procedures
    d. Analytical Method – analytical method used to do the determination (usually an abbreviation – see list of abbreviations at the end of this report)
    e. Analytes determined – a list of elements/substances determined
3. **Data Information**
    a. Data Set - Key field, label (usually abbreviation) identifying the data set

---

[2] EndNote is a widely used software package for storing bibliographic information. This information can be exported into a format usable in other software applications

b. Format – summarises how the data set is made available
c. Availability – information about how available the data is and where from
d. IPR/Copyright information – summarises IPR/Copyright issues
e. Email contact – an email address for data enquiries
f. Web page – a hyperlink to a web page where more information can be found
g. Robustness of data – comments on the robustness of the data set
h. Additional Information – a memo field where further useful information about the data is added

The spatial extents of some of the principal data sets investigated are shown in Figure 3.

2.5.1 Primary soil data sets

Whilst Figure 3 and Appendix 2 show and document some of the principal soil data sets investigated, there are a limited number that have been utilised for the data exploration phase of this project. Different data sets are more appropriate for particular contaminants than for other contaminants, and, whilst not spatially extensive, may provide valuable data that the larger data sets do not contain. The most useful data sets are those that:

- Include results for priority contaminants;
- Are associated with a systematic rather than a targeted sampling strategy so as to represent a broad range of land use types;
- Are spatially extensive across England with a good sample density;
- Are soils that have been collected and analysed to internationally recognised standards and have associated quality assurance;
- Unambiguously define total concentrations of contaminants;
- Are compatible with other available data sets; and
- Provide good resolution of the sample site coordinates.

Three primary data sets stand out as satisfying the majority of criteria listed above and are the best data available for exploring most of the inorganic contaminants. These are the G-BASE rural, G-BASE urban and NSI(XRFS) topsoil data sets. It is also very important to be aware that the soil data sets used are topsoils (*c.*0 - 15 cm) so as to be representative of both the anthropogenic and parent material contribution of the contaminant. Surface soils (0 - 2 cm) will over-represent airborne contaminants and possibly contain too much organic litter. Deeper soils (>30 cm), whilst capturing some historical anthropogenic contamination, are likely to over-represent the parent material contribution (see also Section 3.2.1.1).

The G-BASE rural and urban data are sampled in a consistent manner and provide results for some fifty chemical elements including contaminants As, Cd, Cr, Cu, Cu, Pb, Mo, Ni, Se, S, Tl, V and Zn. The high density of sampling (one site every two kilometre squares for rural and four sites every kilometre square in urban areas) enables interpretations to be made down to a local area scale. The combined G-BASE rural and urban data (*c.*37,000 topsoil samples) gives a data set ten orders of magnitude bigger than the next largest data set (NSI). G-BASE is the only programme to have systematically mapped the chemical baseline of urban areas and the “London Earth” sub-project of G-BASE has provided chemical information for the capital city representing the largest urban geochemical mapping project in the world. Unfortunately, the G-BASE project has not yet completed a soil baseline for all of England, currently covering mainly central and eastern England (see Figure 3). However, the uncovered areas can be supplemented by the NSI(XRFS) data set for which topsoil samples, collected and prepared in a similar way to the G-BASE project, have been reanalysed at BGS to give total element concentrations. The earlier NSI data only contained a limited number of element results following determination by ICP-AES using an *aqua regia* extraction. The NSI data has a sampling density that approximates to 1 sample per 25 km^2. Notwithstanding the comments made in Section 4 regarding access to soil data, the G-BASE and NSI soil data is generally readily accessible with sample site coordinates to a good degree of resolution.

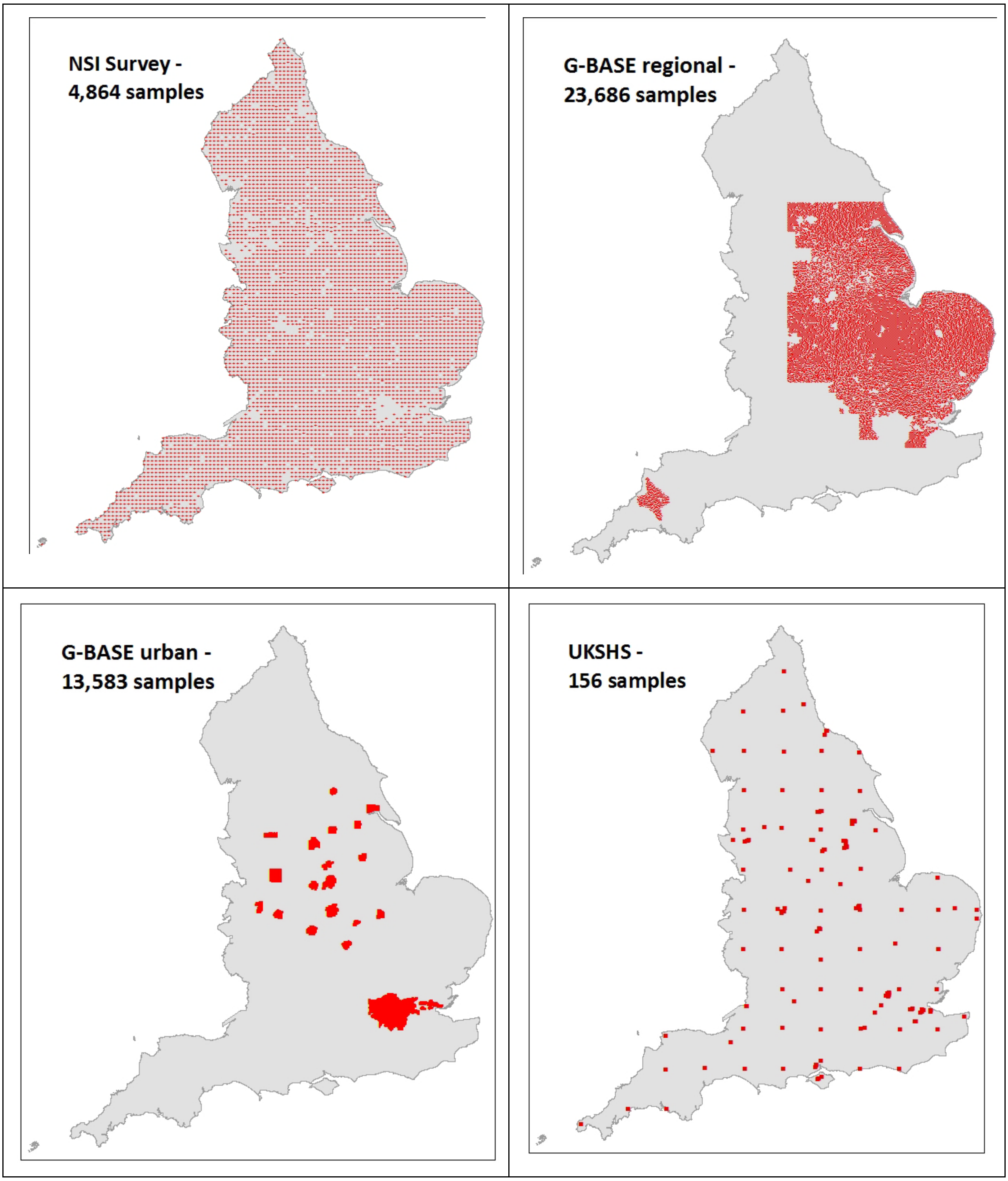


Figure 3: Maps showing the spatial extents of some of the principal soil data sets investigated

(……continued overleaf)

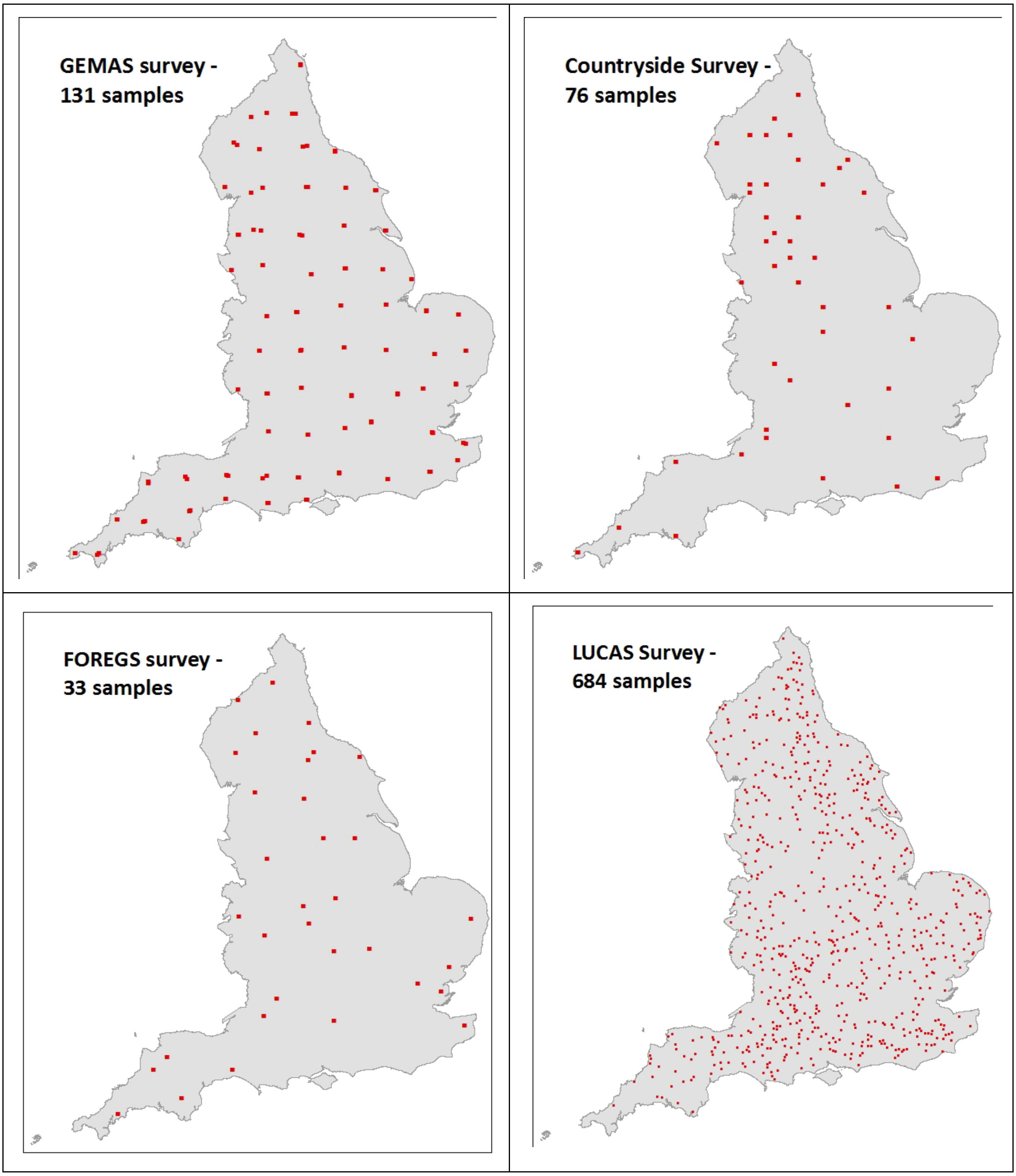


Figure 3: *continued*

2.5.2 Supplementary data sets

There are other small data sets that usefully supplement the G-BASE and NSI(XRFS) data with additional information that can be used to test methodology and modelling (GEMAS, FOREGS, UKSHS and Countryside Survey), providing additional contaminant data for Hg and organic substances. However, there are issues regarding the use of these data sets. For example, some only determine contaminants following an *aqua regia* extraction and so do not unambiguously represent total contaminant concentrations. Additionally, they also tend to have a much reduced sampling density and so will fail to capture local and some regional variability. The UKSHS and Countryside Survey 2000 data also suffer from the fact that site coordinates are degraded to the nearest 10 km in order to satisfy land access agreements and, in the case of Countryside Survey, the representativeness of the sampling sites for future surveys. At a local scale this prevents attribution of parent material polygons mapped in detail to a 1:50,000 scale.

Some other significant data sets are inappropriate to use as they target a specific land use or land group and would therefore bias any NBCs towards that particular land use. The largest of these is the BGS Mineral Reconnaissance Programme (MRP) soil analyses which have as many samples as the G-BASE programme. However, as these samples were collected in predominantly metalliferous mineralised areas, often associated with a long legacy of mining, sampling strategies were geared towards finding high results for metals. The MRP is also an example of a programme for which there was great variability in the sampling and analytical methodology used, so the data set cannot be analysed as a single entity. Site investigations targeting contaminated land will similarly produce data that cannot be used to establish normal backgrounds as results will predominantly be for contaminated soils, which is what is required when investigating a site, but not good for establishing a local or regional trend. Projects specific to a particularly land use and targeting the humus layer rather than mineral soil (*e.g.* the FOREST data set) are also of limited value to the project.

A big Europe-wide project – LUCAS (Land Use Coverage and Area frame Survey) (JRC laboratory of the EC) – is currently in progress with some 1,373 topsoil sample sites in the UK. Heavy metal analysis of top soils is proposed but not yet completed. Land access agreements may also prevent site coordinates being readily available if and when this project delivers some data.

Finally, an important source of information, particularly for Hg and organic contaminants is contained in peer-reviewed publications. Some of these by way of example, are Tipping *et al.* (2011), Cousins *et al.* (1997), and Jones *et al.* (1989) (contained in Appendix 2), others are listed in the Project's EndNote bibliography. Jones *et al.* is an example of a paper containing original data, with site coordinates, but for Wales rather than England, data which can be extrapolated to supplement sparse information for English soils (see Section 3.2.3). However, many publications contain just summary tables without site locations, and care has to be taken to note how much of the data is original or compiled from other publications. Additional information about potential natural contaminant distribution is also available in the form of maps. This is for asbestos (BGS, 2003a,b) and for sulphate (Forster *et al.*, 1995).

It is also worth noting that chemical data for inorganic elements in the surface environment are available for sample media other than soils. The collection of drainage sediments and waters from small streams is a well-established way of defining the geochemical baseline. Eighty five percent of England is covered by stream sediment sampling (Johnson *et al.*, 2005) at a very high density of approximately 1 site every 2 square kilometres. This sampling will be complete by 2013 and provides useful data to supplement soil information where it is spares in order to define area where normal levels of contaminants are high.

2.6 GAPS IN KNOWLEDGE

Available contaminant data to determine NBCs varies both spatially and from contaminant to contaminant. A gap in knowledge will apply where it is thought that there is insufficient information to define a NBC. A summary of available data, grouped by contaminants and with an assessment of knowledge gaps, is given in Table 2. Those domains that are attributed using a limited number of samples will have a greater uncertainty associated with the estimated NBC. This uncertainty will be defined as part of the methodology to be described in Work Package 3.

When this uncertainty becomes unacceptably high, then WP3 may identify further knowledge gaps.

Generally, most inorganic element contaminants in soils are well-mapped by a combination of the G-BASE rural/urban and the NSI(XRFS) data sets. It is only in areas where there is no G-BASE data and where the NSI data (sampling density of 1 site per 25 km^2) may fail to pick up local variability due to a complex underlying geology and/or mineralisation and mining (*e.g.* Cornwall). Certain domains defined by the parent material on which a soil has developed will have characteristically high levels of certain contaminants – this is discussed in Section 3.2 These only represent a relatively small area of England yet are an important natural contaminant source. Unsampled local areas over such parent material (*e.g.* ironstones, organic-rich shales and ultrabasic rocks) can be attributed with NBCs by extrapolating the knowledge from similar areas already mapped in detail. Organic-rich black shales (*e.g.* Kimmeridge Clay) are particularly noted for the potential contaminants they can contribute to the overlying soil including organic substances. Soils developed over such lithologies are worthy of further investigation to gain a better understanding of associated contaminants, particularly organic substances.

The systematic chemical mapping of urban areas by the G-BASE project has provided valuable background information in areas where population density and the potential for anthropogenic contamination is high. In such areas a good understanding of contaminant distribution is most relevant because of historical legacies and the risk of harm in densely populated areas. There are still many major population centres in England (*e.g.* Birmingham, Liverpool, Bristol and Newcastle) where such systematic mapping has not yet been done.

Mercury is one of the inorganic contaminants for which information in topsoils is more sparse having been only routinely determined by low density sampling (*e.g.* UKSHS, GEMAS and FOREGS). This data is supplemented by a number of peer-reviewed publications containing data (*e.g.* Tipping *et al*. (2011)) but any significant systematic mapping of Hg in urban areas is notably lacking for a contaminant that is potentially very hazardous in urban and industrialised areas. The G-BASE project only has a limited amount of Hg results for urban data for Stoke (747 samples – Fordyce and Ander, 2003) and London (473 samples - unpublished data).

Beryllium, cyanide and sulphur are inorganic substances for which, relative to the metallic elements, have poor coverage across England (see Table 2). However, since NBCs for these elements arising from natural soil forming processes and diffuse pollution are going to be significantly lower than concentrations arising from specific industrial processes associated with particular land uses, then this is not a significant knowledge gap.

Similarly, organic contaminants have only sparse coverage across England (mainly from UKSHS, CS:2000 and some peer-reviewed publications). The higher sampling and analytical costs of mapping organic chemicals in the environment makes a high density systematic survey like the G-BASE project prohibitively expensive. Nevertheless, evidence of the diffuse pollution of many very harmful organic substances in built-up areas means that it is important to have more comprehensive information on the levels organic contaminants in urban areas.

There is no data available about natural concentrations of asbestos minerals in soils. However, there are maps available that are based on the underlying geology and give an estimate on the potential for asbestos minerals to be present (BGS 2003a,b) (see Section 3.2.4). Unexploited areas of natural asbestos in England can only be considered to be a very minor source contributing to the normal background (Studds, undated). The lack of information on asbestos in soils, outside targeted site investigations, cannot be considered a significant gap in knowledge. For asbestos minerals the question is not what constitutes a NBC, rather whether such harmful minerals are present or absent.

CONTAMINANT	*KNOWLEDGE GAPS*
arsenic (As), cadmium (Cd), chromium (Cr), copper (Cu), lead (Pb), molybdenum (Mo), nickel (Ni), selenium (Se), thallium (Tl), vanadium (V) and zinc (Zn)	Generally excellent spatial coverage because of high density G-BASE urban and rural data supplemented by NSI(XRFS) data. **In the areas covered only by NSI(XRFS), where there is a significant geological variability and mining/mineralisation (*e.g*. Cornwall) greater density of soil sampling would improve NBCs estimation**. Also there are still **many major urban centres that have not systematically been mapped** (e.g. Birmingham, Bristol, Liverpool, Newcastle). Lower density sampling data sets provide supplementary information (UKSHS, GEMAS, FOREGS, CS;2000) but heavy metals are generally determined following *aqua regia* extraction and so are not truly total concentrations.
beryllium (Be)	This is not an element determined by XRFS so is absent from the major datasets. Beryllium is reported in GEMAS dataset (65 sites) and FOREGS (60 sites). **This is a contaminant for which the spatial coverage is not substantial.** However, given that even the natural elevated levels of this contaminant cannot be considered to pose a high risk, then **this cannot be seen as a significant knowledge gap**
mercury (Hg)	More recent G-BASE soils determined by ED-XRFS do report Hg, though the quality of results produced by this analytical method is deemed inadequate for this Project. The UKSHS (61 rural, 13 urban, 22 industrial sites) provides a low density national coverage, as does the FOREGS, GEMAS and CS:2000 data. G-BASE urban areas have limited Hg data (London & Stoke) and this is a contaminant potentially representing a significant hazard in built areas. **More data is required for Hg in topsoils from the systematic mapping of built areas.** There are a number of useful sources of Hg data in peer-reviewed publications.
sulphur (S)	Sulphur is an element determined by ED-XRFS and so is reported in the more recent G-BASE urban and rural topsoil results as well as by GEMAS and FOREGS. Natural and diffuse levels of S are highly unlikely to reach the levels found on industrial sites involving sulphur and its compounds. Of greater concern in the natural environment is the environmental risks caused by sulphur related acidification (acid mine waters) and the impact of sulphate on concrete structures. The BGS National Geotechnical Properties Database which primarily holds geotechnical information extracted from site investigation records provided by clients, consultants and contractors, and from field and, secondarily, from laboratory test results carried out by the British Geological Survey, contains a datasets on sulphur contents in soils and rocks in UK (Self *et al*. 2008). **There is no knowledge gap identified.**
cyanide (CN^-)	**No data has been found giving concentrations of cyanide in natural soils**. Cyanides are readily attenuated in anaerobic soils, mostly through biodegradation (Smith, 1994). Data available in literature related to made ground below gas works and other industrial sites are not considered relevant to the estimation of NBCs. Cyanide in soil is only likely to be a contaminant of concern in association with industrial sites.
asbestos (minerals)	**There is no systematic data on the concentration of asbestos fibres in natural soils.** There are available maps of possible occurrence of asbestos minerals in mafic and ultramafic rocks in UK which could find there way into overlying soils. The zones of asbestos potential shown on the maps are based on the simplified regional distribution of metamorphic zones based on the UK 1:250 000 geological map. As this is a contaminant very much associated with the built environment and specific historical land use, for which site investigations should provide much knowledge, **this cannot be considered a significant knowledge gap**. The impact of potential rock sources of asbestos minerals should only be of concern if an activity involves the extraction or tunnelling through potentially hazardous rocks.
Organic contaminants (PCBs, PAH, dioxins and furans)	There are data sets with low density sampling that can be used to describe the organic contaminant distribution on a national scale (UKSHS, CS:2000) and some peer-reviewed publications. Relative to inorganic contaminants, the special precautions required for sampling soils for organic contaminants and the analytical methods are very much more costly, and so have restricted the amount of available data. **Greater density of soil sampling over organic rich rock-types and in built areas would give better estimates for natural and diffuse pollution concentrations of these organic contaminants.**

Table 2: A list of contaminants with a summary of knowledge gaps

3 Exploratory Data Analysis

3.1 LANDSCAPE DATA

3.1.1 Soil-Parent Material Model

The Soil-Parent Material Model[3] (SPMM) (Lawley, 2011) has been developed by BGS, using as its basis the mapped boundaries of the national 1:50,000 superficial and bedrock geological data (DigMapGB-50[4]), and is used within a GIS environment. Soil 'Parent Material' is the first recognisably geological material found beneath a soil profile, and is the lithology on which that soil has developed (Avery, 1990). Soils thus inherit many properties, including chemical composition, from this material.

In the SPMM the geological data have been combined into one layer of information which indicates the rock/sediment formation mapped as directly underlying soil. Where this is a superficial deposit (such as alluvium, glacial deposits, peat), the data set also maintains the record of the solid geological formation first encountered beneath this surface sediment; such information is of benefit where the underlying solid geology imparts chemical (or other) characteristics into the overlying superficial deposits, and thus the soil. The information, which has historically routinely been attributed to the mapped digital polygons in DigMapGB, largely comprises lithological and chronological information. Augmenting this in the SPMM is additional information on texture, mineralogy and lithology, which is attributed in a hierarchical classification system. In the context of the present study this means that a higher level of aggregated characteristics can easily be applied to soil geochemical data than is possible solely using DigMapGB; for instance, retrieving all formations which are classed as 'ironstones' (irrespective of their formal name) and confers benefits from using the SPMM.

The scale of mapping (1:50,000) is also relevant. This is the scale at which much geochemical sampling is undertaken, and gives the user a reasonable feel for the degree of uncertainty on the data. Where data is provided at other common scales, such as 1:250,000 or 1:625,000 the boundaries and number of polygons are simplified and aggregated in order to provide generalised information at the national-scale. More detailed mapping, such as 1:10,000, is not available in a consistent format or as part of the SPMM data, and would imply greater certainty in sample locations and polygon boundaries than is appropriate from the data. Soil mapping is available at a national-scale (see *e.g.* NSRI NATMAP[5]) but this is not systematically mapped at 1:50,000 and would require attribution with the latest geological mapping data in order to retrieve information on key formations, and so has not been used in this study.

3.1.2 Land use

Normal background concentrations of contaminants result from natural geological and pedological processes and diffuse source inputs (ISO 2011). Whilst the domains of natural sources can be identified from geological and related data sources (see Sections 3.1.1 and 3.1.3), domains of diffuse anthropogenic pollution need to be defined by a different approach. Diffuse pollution arises where substances are widely used and dispersed over an area as a result of land use activities, often associated with urban development. These activities may be

---

[3] http://www.bgs.ac.uk/products/onshore/soilPMM.html

[4] http://www.bgs.ac.uk/products/digitalmaps/digmapgb_50.html

[5] http://www.landis.org.uk/data/natmap.cfm

recent or have been carried out in the past but cannot be tied down to a specific location or source. Examples of diffuse pollution include atmospheric deposition of contaminants arising from industry, domestic coal fires and traffic exhaust, and disposal of domestic coal ash. A simple approach to defining diffuse pollution domains is to link them to urbanisation.

The ONS (Office for National Statistics, 2011) has proposed a set of criteria for the definition of where the starting point for an urban area is. The identification of those areas is based on land use which is irreversibly urban in character. This comprises:

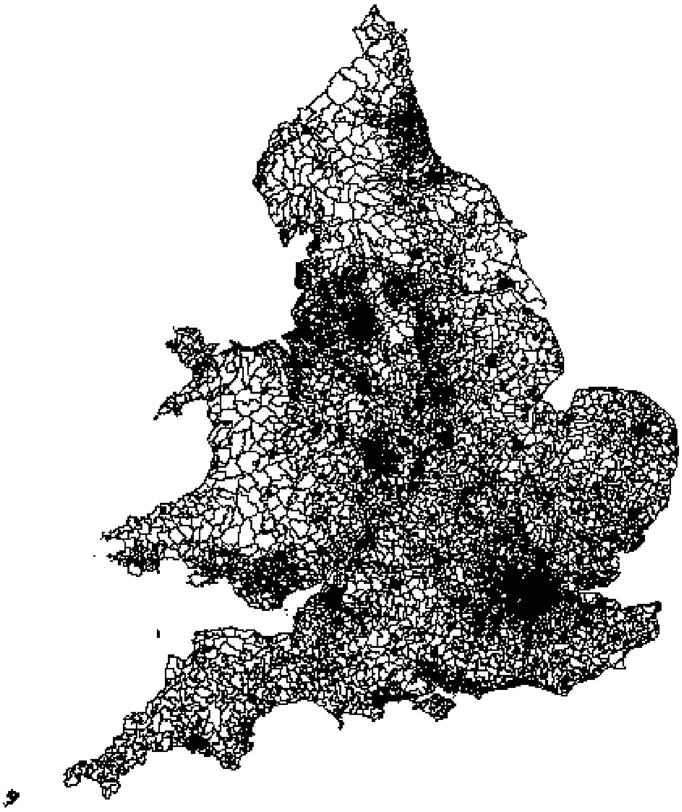
i) permanent structures and the land on which they are situated, including land enclosed by or closely associated with such structures;
ii) transportation corridors such as roads, railways and canals which have built-up land on one or both sides, or which link built-up sites which are less than 200 m apart;
iii) transportation features such as airports and operational airfields, railway yards, motorway service areas and car parks;
iv) mine buildings, excluding mineral workings and quarries; and
v) any area completely surrounded by built-up sites.

Areas such as playing fields and golf courses are excluded unless completely surrounded by built-up sites. The prerequisite for the recognition of an urban area is that the area of urban land should extend for 20 hectares or more. Separate areas of urban land are linked if less than 200 m apart. Land between built-up areas is not regarded as urban unless it satisfies one of the conditions listed above.

The definitive database for land use for England is the Ordnance survey MasterMap® (Ordnance Survey, 2011), however, this is a licensed product with a great amount of detail. The CEH Land Cover Map (LCM2000[6], and more recent version) are digital data sets that provide substantial land use information at a high resolution, again a product requiring a licence to use it. However, the ready availability and quantitative outputs of the Generalised Land Use Database (GLUD) Statistics for England 2005 (Communities and Local Government, 2007) make this particularly suitable for implementing a measure of urbanisation.

---

[6] http://www.ceh.ac.uk/LandCoverMap2000.html



Source: Office for National Statistics licensed under the Open Government Licence v.1.0

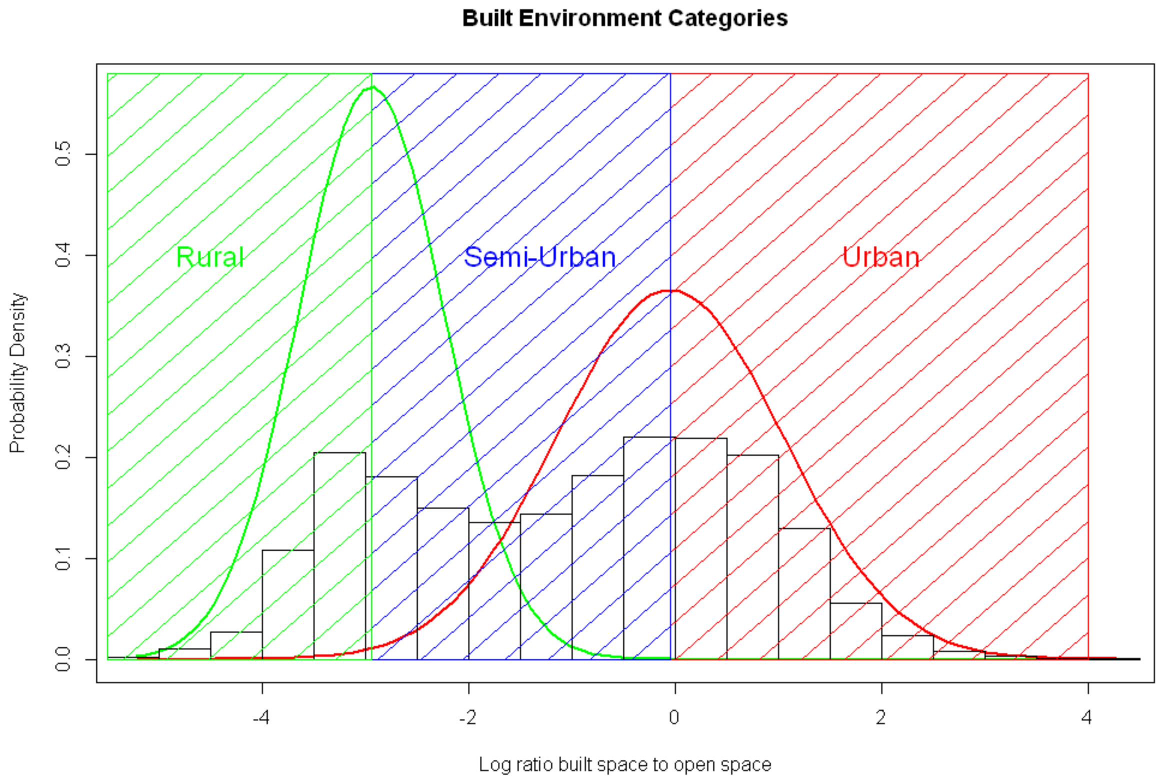
Figure 4: Census Area Statistical Wards 2003

3.1.2.1 DEFINING URBAN DOMAINS USING GLUD

The GLUD was produced using an automated methodology which has been developed to allocate all identifiable land features on Ordnance Survey's OS MasterMap® (Ordnance Survey, 2011) into nine simplified land categories and an additional 'unclassified' category. Full details of the background and methodology for production of the data set are available at the Communities and Local Government website (Communities and Local Government, 2007). The land use classifications are:

i. Domestic buildings;
ii. Domestic gardens;
iii. Non-domestic buildings;
iv. Roads;
v. Paths;
vi. Rail;
vii. Greenspace;
viii. Water;
ix. Other land uses (largely hard standing); and
x. Unclassified.

The area of each of these classifications subset to the 8850 Census Area Statistical Wards 2003 (CASW) which are supplied as a shape file by the ONS. These are shown in Figure 4. The GLUD only supplies land use statistics for wards in England.



Adapted from data from the Office for National Statistics licensed under the Open Government Licence v.1.0

Figure 5: Data distribution of the natural log of UI and the proposed domains

Using the ONS classification of the built environment as a measure of urbanisation, the ratio of built space to open space for each CASW was calculated to give an urbanisation index (UI) where:

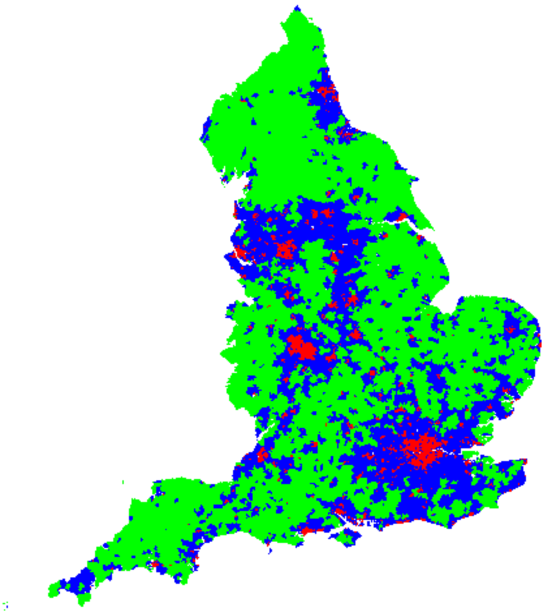
i) 'built space' was calculated as the sum of the area of Domestic buildings, Domestic gardens, Non-domestic building, Roads and Rail; and

ii) 'open space' was calculated as the sum of the area of Paths, Greenspace and Water.

Examination of the distribution of the UI for each CASW in England showed that the values are positively skewed. Taking the natural logarithm of the UI to make the distribution more symmetrical shows a bimodal distribution of data (Figure 5). Using the mixtools package (Young *et al*., 2009) from the R programming language (R Development Core Team, 2011) the natural log UI can be resolved into two Gaussian populations. The relative proportions of the mean, standard deviation of the two populations are given in Table 3.

Population	Rural	Urban	Rural	Urban
Proportion	0.35	0.65		
	Log values		Back transformed	
Mean	-2.94	-0.0424	0.053	0.958
SD	0.704	1.09		

Table 3: Summary statistics for rural and urban domains (mean and SD are log ratio)



Adapted from data from the Office for National Statistics licensed under the Open Government Licence v.1.0

Figure 6: Built Environment categories in England (red - Urban, blue - Semi-Urban and green - Rural)

Given that the two distributions overlap, the data were divided into three categories (Figure 5):

i) CASW with logUI greater than the mean of the highest value population to be designated as 'Urban' areas;

ii) CASW with logUI between the mean of the two populations to be designated as 'Semi-Urban' areas; and

iii) CASW with logUI less than the mean of the lowest value population to be designated as 'Rural' areas.

Figure 6 shows the spatial distribution of the three categories in England which are proposed as domains for defining the contribution of diffuse pollution to contaminant backgrounds.

3.1.2.2 EXPLORATORY ANALYSIS OF THE PROPOSED BUILT ENVIRONMENT DOMAINS FOR ARSENIC, LEAD AND BENZO[*a*]PYRENE

The suitability of the three categories of urbanisation was tested on three contaminants: As and Pb concentrations in topsoils from urban and rural samples from the BGS G-BASE project and the NSI (XRFS) data set; and BaP from the EA UKSHS data. For As and Pb, where the number of soil data sites were higher than the number of CASW (*c*.30,000 compared to *c*.8,000, respectively), the As and Pb values were interpolated to the CASW centres using inverse distance weighted (idw) interpolation using a power of 2 and a maximum distance of influence of 5 km. For the BaP data where the number of soil data sites (73) is very much less than the CASW the UI was interpolated to the locations of the BAP soil samples using a power of 2 and maximum distance of 7 km. The data for As, Pb and BaP are summarised as scatter plots of log (contaminant concentration) against log (UI) with a two dimensional probability density function superimposed as percentile contours (R statistical programming language package "sm" (Bowman and Azzalini, 1997)). In addition, each contaminant has been plotted as a boxplot classified on the three categories of urbanisation shown in Figure 5.

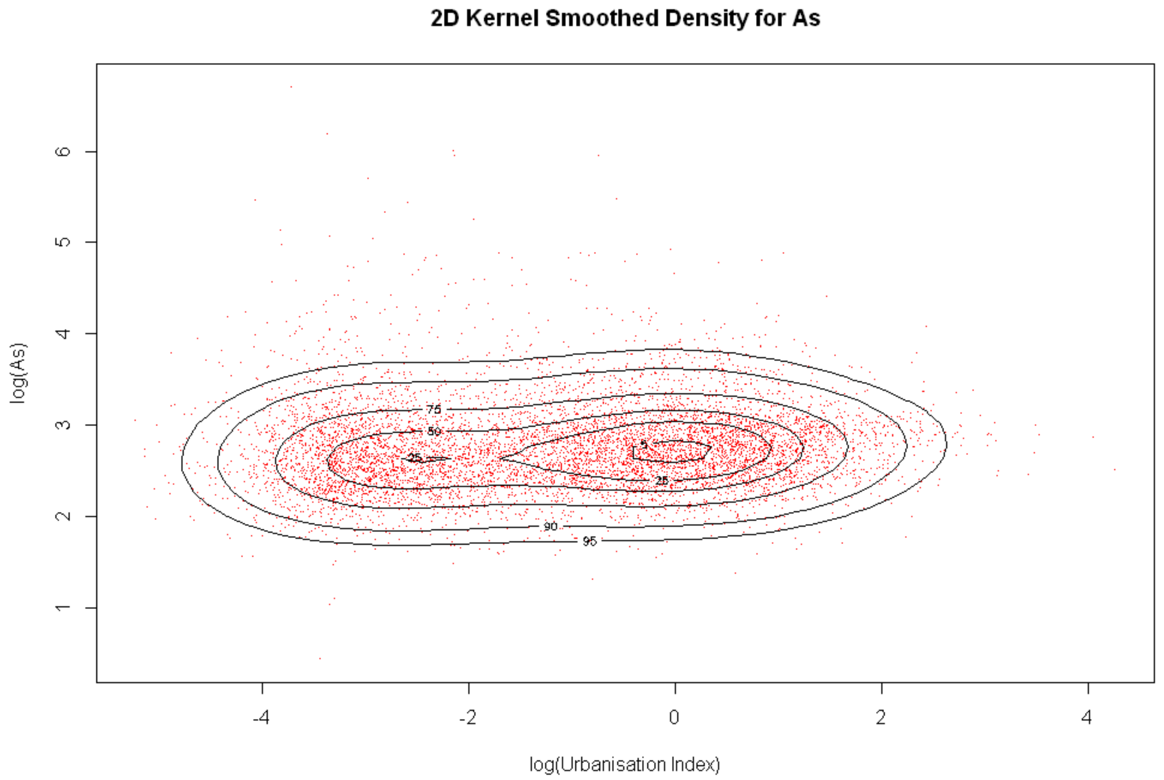


Figure 7: 2-D probability density plot for As and UI with percentile contours. Axis units as natural logs, As concentrations in mg/kg.

For As, the 2-D density plot clearly shows the bimodal distribution in the UI but there appears to be no observable trend in the As concentration between the rural and urban UI distribution (Figure 7). This is confirmed in the UI conditioned boxplot for As (Figure 8) where all three categories of urbanisation have very similar distributions. This suggests that, compared to natural sources, As in soils in England has a minimal contribution from diffuse pollution.

For Pb, the 2-D density plot (Figure 9) clearly shows increased Pb concentrations at higher UI values with a noticeable protuberance in the density contours at high UI values that is associated with data sites from the London area. These trends are also clearly shown in the Pb boxplot (Figure 10) with an increasing trend in Pb concentrations from the rural to urban categories. The boxplot also shows a number of high outliers in the rural and semi-urban categories which are likely to come from natural Pb sources in Derbyshire which occur in rural and urban locations. These exploratory plots confirm literature findings that the Pb content of soils has a significant diffuse pollution contribution (D'Arcy *et al.*, 2000; Rawlins *et al.*, 2005; Biasiolia *et al.*, 2006; Napiera *et al.*, 2008; Marchant *et al.*, 2011) and that the proposed categories for urbanisation provide a practical means for defining diffuse pollution domains.

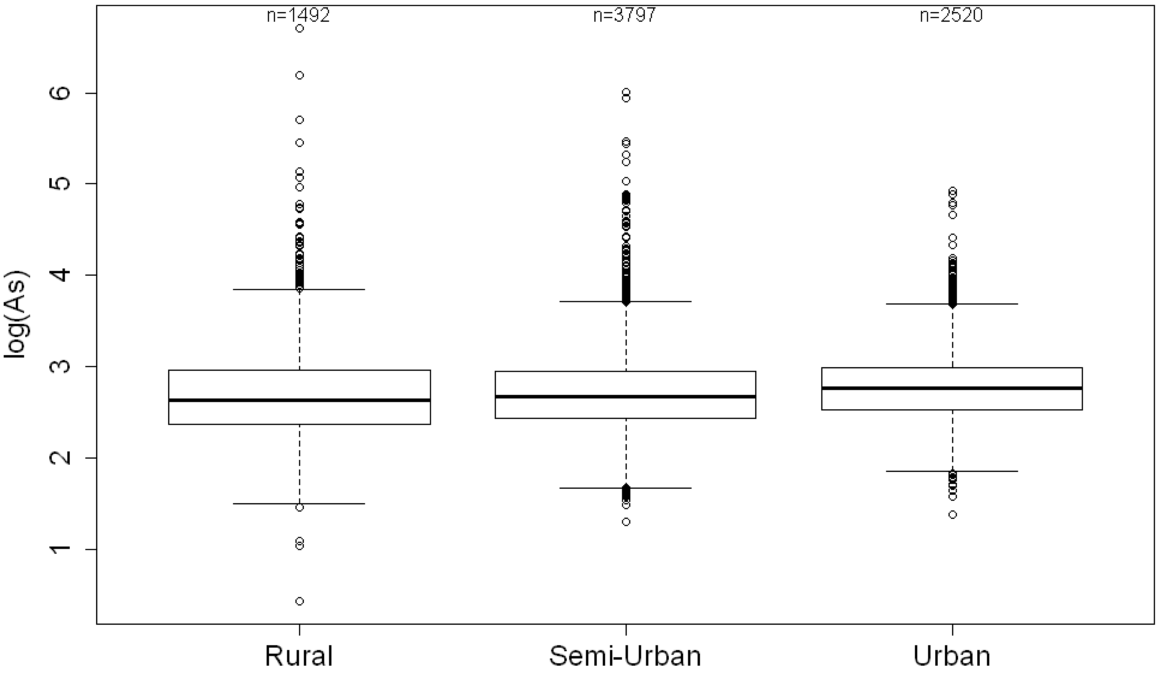


Figure 8: Boxplot of As in topsoil concentrations split by proposed diffuse pollution domains. n is the number of CASW. As concentrations in natural log(mg/kg).

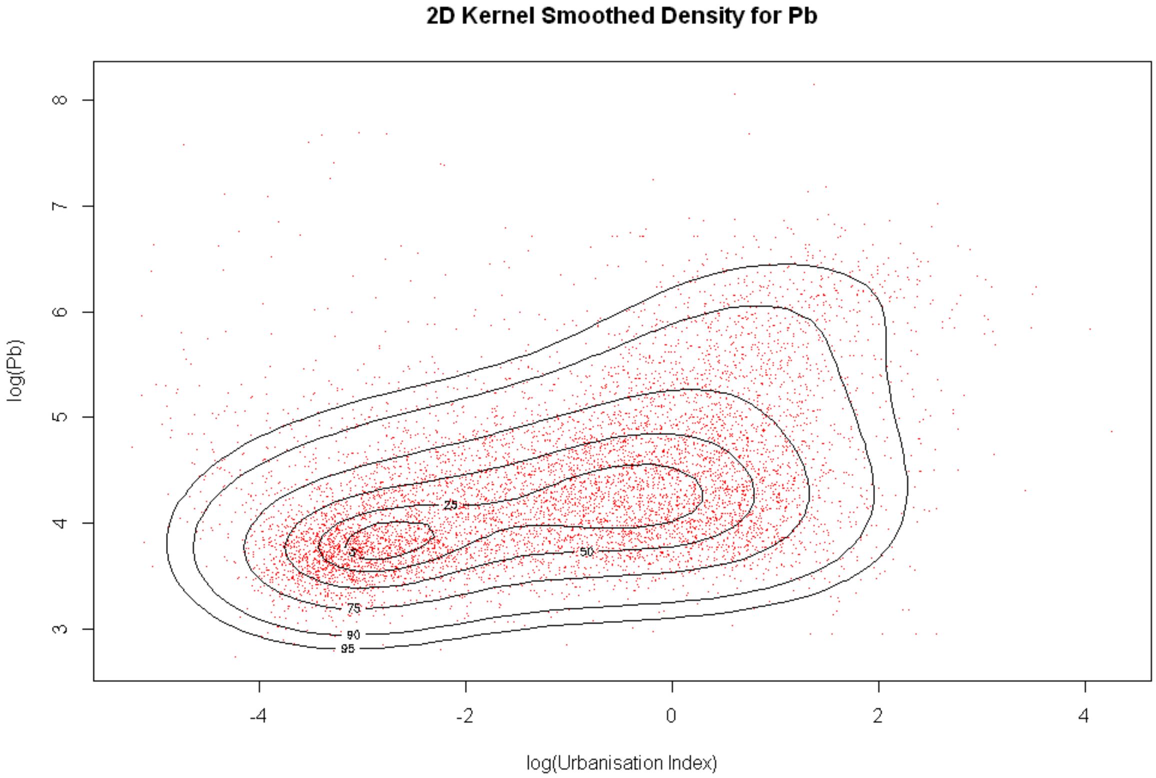


Figure 9: 2-D probability density plot for Pb and UI with percentile contours. Axis units as natural logs, Pb concentrations in mg/kg.

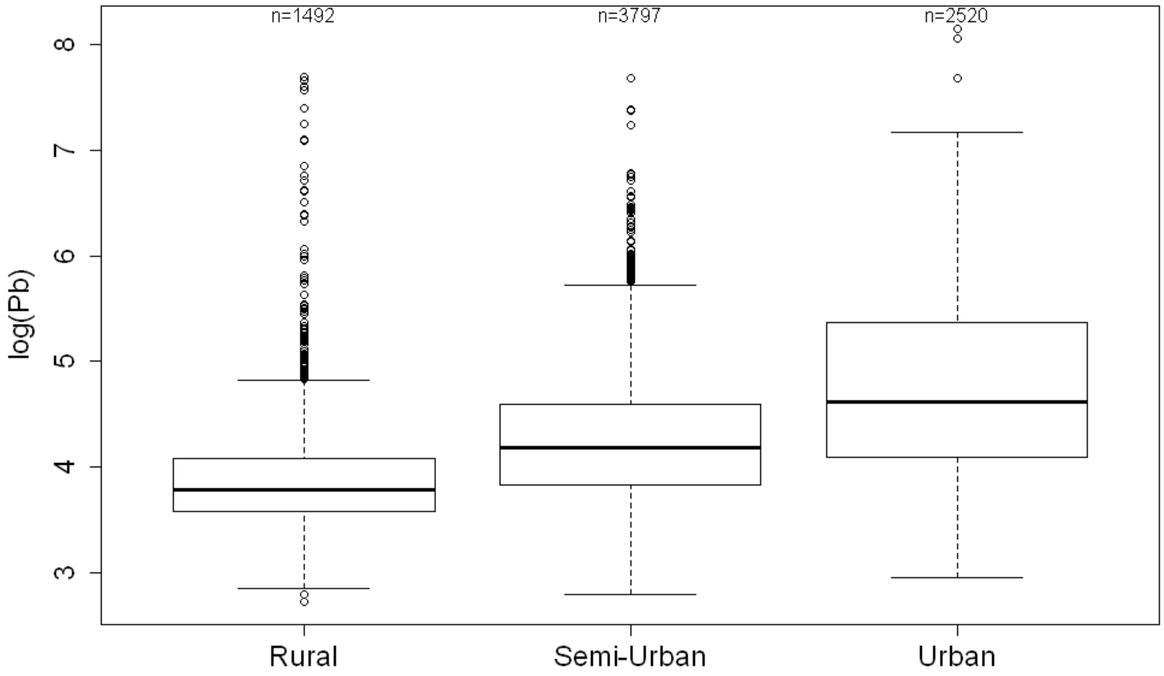


Figure 10: Boxplot of Pb in topsoil concentrations split by proposed diffuse pollution domains. n is the number of CASW. Pb concentrations in natural log(mg/kg).

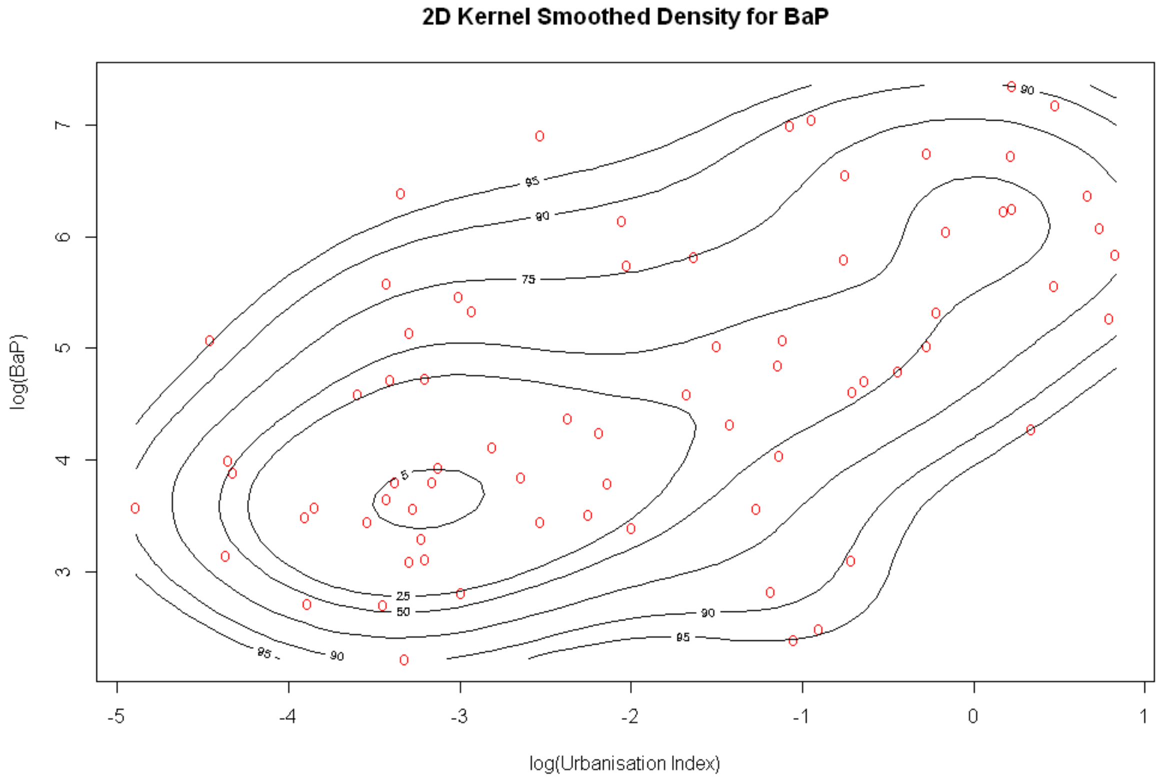


Figure 11: 2-D probability density plot for BaP and UI with percentile contours. Axis units as natural logs, BaP concentrations in µg/kg.

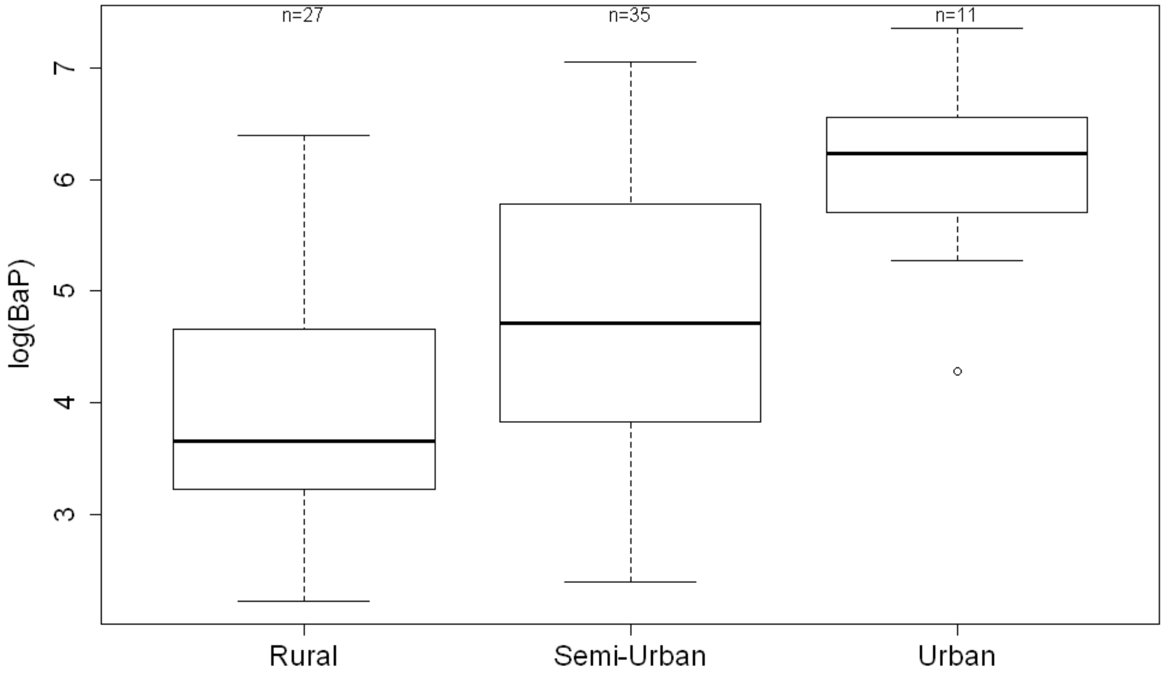


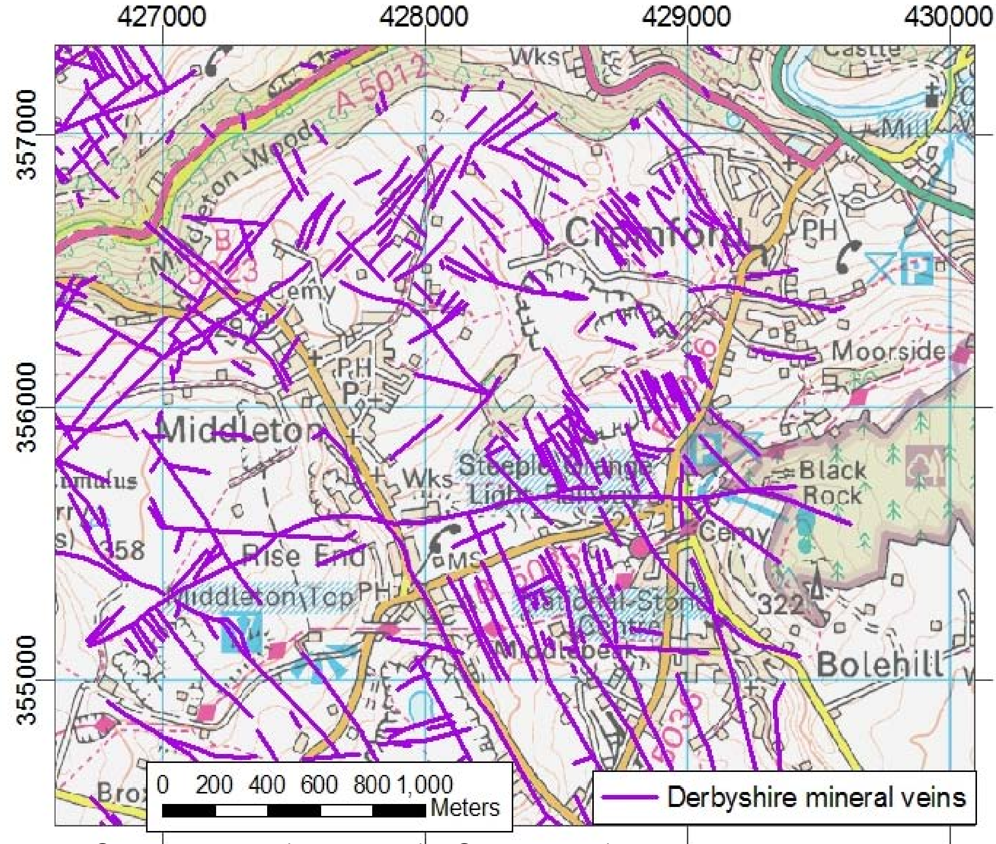
Figure 12 Boxplot of BaP concentrations in soil split by proposed diffuse pollution domains. n is the number of samples. BaP concentrations in natural log(µg/kg).

For BaP there is a very much reduced data set compared (73 data points compared to *c*.30,000 for As and Pb), however, the 2-D density plots and the boxplots (Figure 11 and Figure 12) show very similar trends to the Pb with increasing concentrations in BaP from rural to urban locations. This confirms literature findings *e.g.* Biasiolia and Ajmone-Marsan (2007) that BaP is mostly derived from anthropogenic inputs and its concentration in soils is controlled by diffuse pollution.

3.1.3 Mineralisation and mining geographical mapping

The historical mining industry in England was of huge economic importance in specific areas, often up until the early 20th Century. Non-ferrous metalliferous mining depends entirely on the natural occurrence of mineral deposits. Where these ores occur, they will act as a geochemically distinct parent material to the overlying soil and impart naturally high concentrations of the metals (or metalloids) contained in the particular area. The detail to which geological mapping of the surface occurrence of mineral veins has been undertaken will rely largely on the age of mapping and importance of the deposit; small, diffuse uneconomic veins are unlikely to be mapped in much detail. An example of an area of the Peak District with very detailed mineral vein mapping is shown in Figure 13, and gives an intuitive indication of the scales over which changes in this particular parent material may occur.

The history of the location and extent of some mining activities which have exploited this wealth in England is lost in antiquity, with many mining areas having evidence of extraction in Roman times, and the likelihood of working being much older than that. Even more recent mining activities, from the 1800s to the First World War have left an incomplete record of the spatial extent of land used. Likewise mapping of the location of mine shafts has been attempted in some areas, but again is constrained by the time available to compile such information and often can be shown to have missed some localities. Thus, whilst the general areas of mining are known, the accuracy of information at a detailed scale is not clear or consistent.



BGS © NERC. Contains Ordnance Survey data © Crown Copyright

Figure 13: Detailed mineral vein mapping from part of Derbyshire (BGS data)

The data set examined in this work is that of Metalliferous Mineralisation and Mining, originally produced in hard-copy by Ove Arup (1990) for Department of the Environment, but which has been 'cleaned' and turned into a polygon layer by BGS. The data for England has been further attributed in this study by giving a name to the major ore fields (Figure 14), to allow soil sample sites and geochemical data to be joined to the ore fields and separately analysed for typical soil concentrations. This mapping is generalised to 0.5 km grid squares, which is probably a suitable level of spatial resolution for this type of data for the reasons stated previously. Therefore, it should be expected that not every occurrence of mineralisation/mining has been located. Where soil geochemical data is encountered that falls outside a given mineralised domain, but is of a concentration expected for that contaminant within the local mineralised domain, and lies over the parent material which is known to be affected by mineralisation in that ore field, that high soil concentration can relate to natural processes, or historical mining.

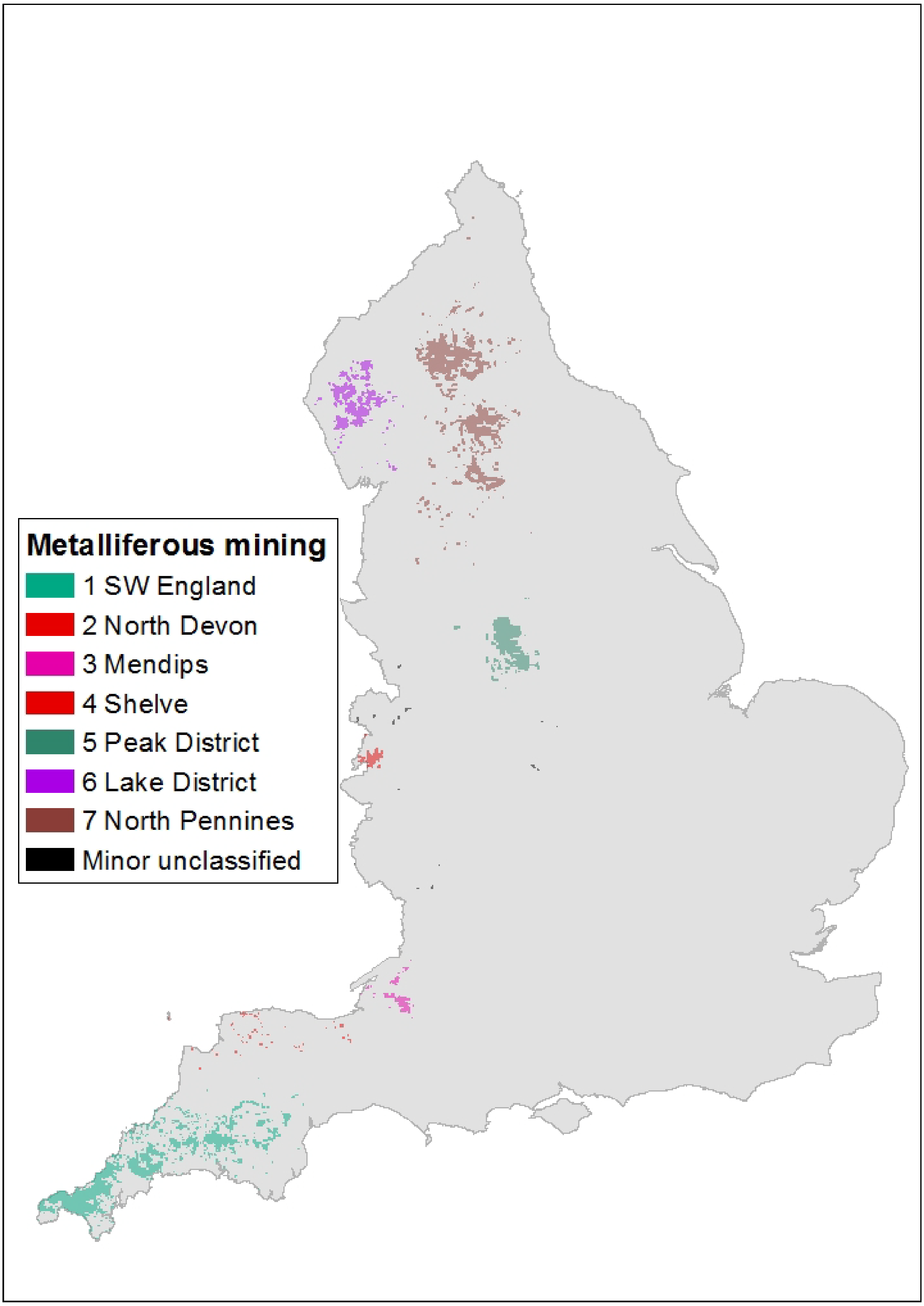


Figure 14: Metalliferous mineralised and mined areas of England, labelled with the categories used in this project

3.2 CONTAMINANT DATA

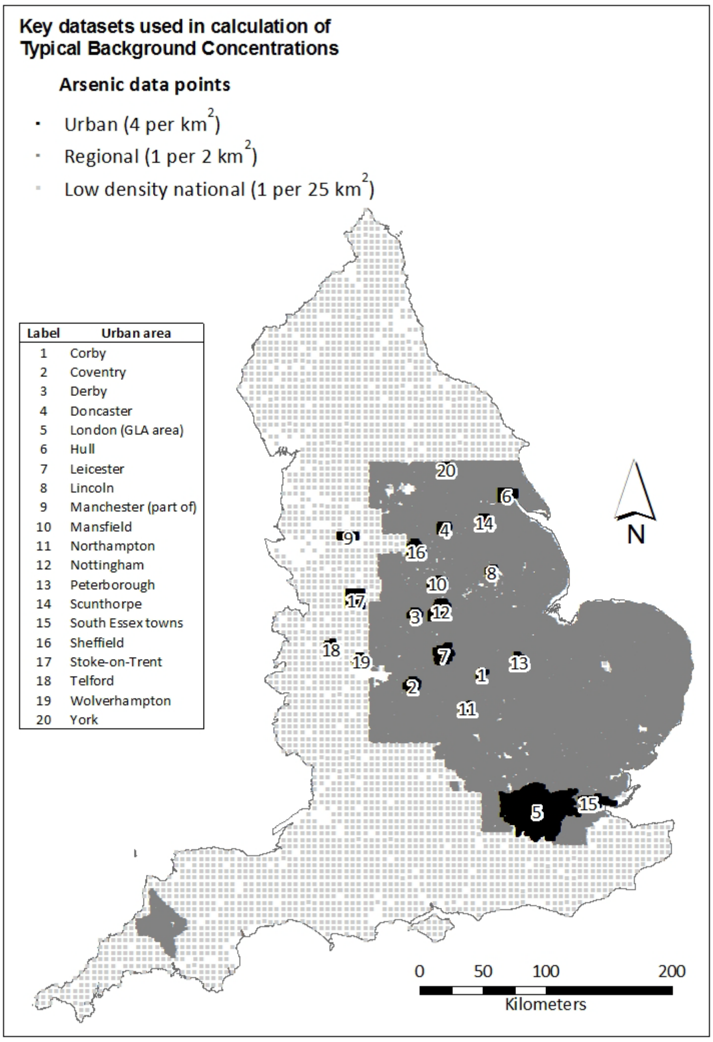


Figure 15: Location of As topsoil sample site data used in Figures 17 – 19 and Table 5 (the low density NSI(XRFS) sites exist within the G-BASE areas).

3.2.1 Arsenic

Soil As concentrations are widely recognised to reflect both natural and anthropogenic inputs. Natural sources are primarily from the weathering of underlying bedrock to form soil, whilst anthropogenic inputs, such as mine spoil dust, ore smelting fumes or coal-burning emissions, tend to be more local in occurrence than those which are geologically derived (Reimann *et al*., 2009). Arsenic occurs primarily bound to mineral phases, such as iron oxides and clays, or soil organic matter. Arsenic minerals, such as arsenopyrite, are found in some mineralised areas, where arsenic is sufficiently abundant, such as Cornwall (Aston *et al*., 1975).

3.2.1.1 AVAILABLE DATASETS

Arsenic is measured routinely by the large soil geochemical baseline projects (Table 4), so surface soil data from these have been used as the major ('core') data sets in this analysis. These data provide systematic coverage, through consistent sampling and analysis methods, at differing densities for the whole of England and many urban areas and are shown in Figure 15. All data is determined by XRFS, which has the benefit of being an absolute total concentration.

An area of central England covering *c.*30,000 km^2, with 14,229 samples sites have analyses for top- (5 – 20 cm, <2 mm size fraction) and deep soils (35 – 50 cm, <150 micron size fraction). Of these, 13,746 have both top- and deep soil As data reported. If these can be shown to be interoperable, the 1 per 2 km^2 data from G-BASE deep soils could be used in the area where there is no G-BASE topsoil data (n = 2,653), extending the area of higher density data coverage by 6,837 km^2. However, although most topsoil and deep soil samples are very similar, as shown by the regression (see Figure 16)

$$[As_{topsoil}] = 0.396 + (0.976 \times [As_{deep\ soil}]) \quad (n = 12{,}943;\ R^2 = 91\%,\ P<0.05),$$

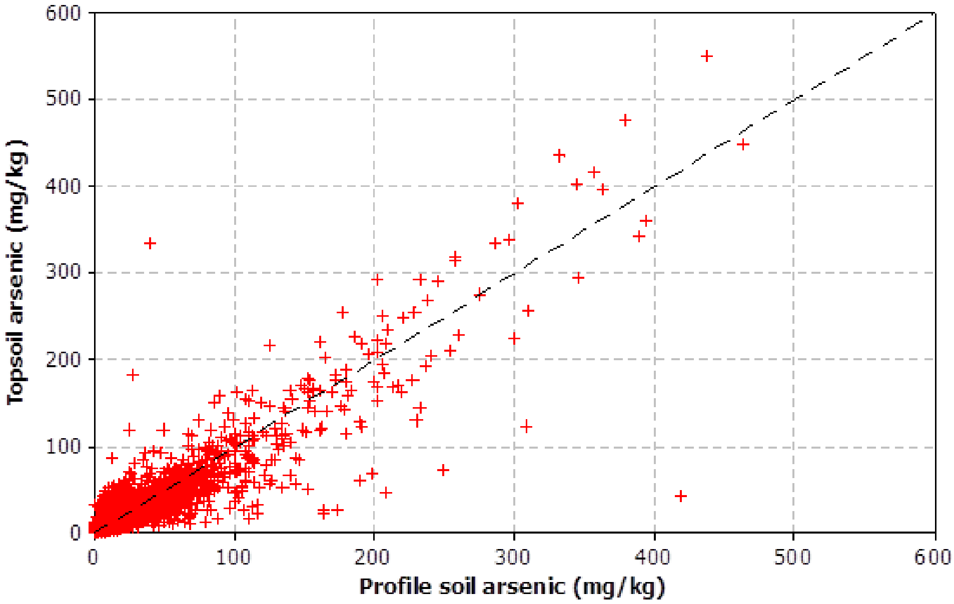


Figure 16: Comparison of deep (‘profile’) and topsoil As concentrations from the G-BASE project.

when data ratios (topsoil/deep soil) are compared, the data is negatively skewed (-3.7), implying that outliers caused by deep soil concentrations are greater than those from the topsoil. When the location of these was examined they were found to have a systematic distribution, closely related to the outcrop of specific formations. It is thought the large differences in the concentrations, where the deep soil As could be twice that of the topsoil As, could lead to unacceptably large errors in predicting topsoil concentrations, so the use of the G-BASE deep soil results was not pursued any further.

In order to prepare the three core datasets for use, a minor correction factor was applied to results from samples collected prior to 2006, taken from the BGS Corporate Geochemistry Database (1st October 2011), so as to ensure all data was levelled to the most recent certified reference material As concentrations.

The minor data sets from GEMAS and FOREGS are useful comparisons between XRFS and some common analytical digestion methods.

3.2.1.2 OVERALL DATASET PROPERTIES

The range of As in England’s soil is greater than 5 orders of magnitude (Table 5(a)), although the great majority of the data is in the much smaller range of 6.8 mg/kg (5th percentile) to 33 mg/kg (95th percentile) and a median of 14 mg/kg. When three different sources of systematic data are compared (Table 5(b); Figure 17) it can be seen that there is only a small statistical difference between the three, with a range of median concentrations 13.4 – 15.1 mg/kg. Figure 17 shows that the data sets only deviate in composition between the urban and other data at >95th percentile of the data and that the urban soils are of a typically lower concentration than the national/regional data. The location of these data is shown in Figure 15, with labelled urban areas for the data shown in Table 5(c).

There are few other UK datasets published with background As concentrations, but where they are, the English data is generally elevated in comparison to the other datasets. The pan-European data of FOREGS (Salminen *et al*., 2005) had a median of 7 mg/kg for an *aqua regia* digest of the samples, whilst the UKSHS found a median English concentration of 8.3 mg/kg, with that for the whole UK 7.1 mg/kg (Ross *et al*., 2007). The majority of references cited by Reimann *et al*. (2009) also have median concentrations <10 mg/kg, one significant exception being that of the G-BASE survey of Swansea with a median concentration of 53 mg/kg, due to the industrial legacy of the city (Fordyce *et al*., 2005).

Data set name	Samples in England	Data set use
G-BASE (rural)	23,686	Core
G-BASE (urban)	13,583	Core
NSI(XRFS)	4,864	Core
GEMAS	131	Minor
FOREGS	33	Minor
NSI	*No As data*	none
CS:2000	76	Minor
UKSHS	156	Minor

Table 4: Summary of data sets used in this study. See Appendix 2 for data set summaries and associated references

3.2.1.3 VARIATION BETWEEN AREAS

Table 5(c) shows that there are substantial differences in concentrations between the urban data sets, *cf*. Northampton and Coventry, which will be discussed further later. Whilst the urban data sets are not normally distributed (skewness <1 required), it can be seen that these are not as greatly skewed as the rural data sets. When the Tamar catchment data (Cornwall/Devon) is compared to that of the regional data from Eastern England, much higher As concentrations are found. The Tamar catchment has As-bearing mineralisation, particularly abundant in the lower catchment. This gives rise to naturally elevated baseline concentrations. In addition to the natural release of As from the underlying lithologies through the processes of soil formation, the extensive mining and roasting of the As ores in the latter half of the 19th century resulted in widespread physical and chemical dispersion of As around these areas. This combination of factors is the reason that this area has concentrations which are typically higher than the Eastern England regional G-BASE data.

Figure 18 shows the concentration distribution of As in soils in England, using an interpolated surface (inverse distance weighted - idw) thematically coloured according to the percentile values (5th, 10th, 15th, 25th, 50th, 75th, 90th, 95th, 99th) used routinely in BGS regional geochemical mapping (Johnson *et al*., 2005); these thresholds are used to enhance the visualisation of skewed datasets. The non-parametric measure of distribution (percentiles) has the benefit of compressing the relatively narrow interquartile range concentrations into two classes, whilst showing more classes at the high and low end of the concentration range, and is not affected by outliers, particularly with such a large sample set (here n >42,000).

Figure 18 demonstrates clear spatial dependence of the concentrations on location – greater spatial resolution is also apparent in this image in the area where there are samples collected at least 1 per 2 km^2, rather than the lower density NSI survey (see Figure 15). The spatial scale and controls on concentration can be examined *a priori* using methods such as the geostatistical technique of kriging, or in a more geochemical process-based approach by co-location with controlling factors, such as underlying geology, in determining normal concentrations. However, the huge data-processing that this would require in conjunction with a dataset such as the BGS SPMM (Section 3.1.1) means that an approach which prioritises the highest As concentration areas, that may form separate natural background concentration domains, is appropriate. This has the benefit of reducing the data to a limited number of domains, with the objective of producing a workable series of domains that can be applied by a user through Work Package 4 project outputs.

Such processes which examine the data set independent of sample location, have been widely used in establishing geochemical background/baseline and anomalous areas (Appleton, 1995; Davies, 1983; Sinclair, 1974). The approach chosen here is the k-means cluster analysis in the R statistical programme (Hartigan and Wong, 1979). The k-means method aims to partition the data into k groups such that the sum of squared Euclidean distances from points to the assigned cluster centres is minimized. It is not practical to require that the solution has minimal sum of squares for all data combinations, so algorithms have been produced which seek local "optima" solutions such that no movement of a point from one cluster to another will reduce the within-cluster sum of squares. There are a number of algorithms available for achieving this; the R statistical programming language uses the method of Hartigan and Wong (1979) which is generally considered to be the most efficient.

The thresholds derived from this analysis are shown in Figure 19, and have only been used in this exploratory data analysis to derive data populations for further analysis. The outcome of this is that it has greatly reduced the detail in the data, and clearly identified geographically coherent areas (Figure 19) of concentrations >27 mg/kg which merit further examination to resolve their spatial extent, data statistical characteristics and underlying causes.

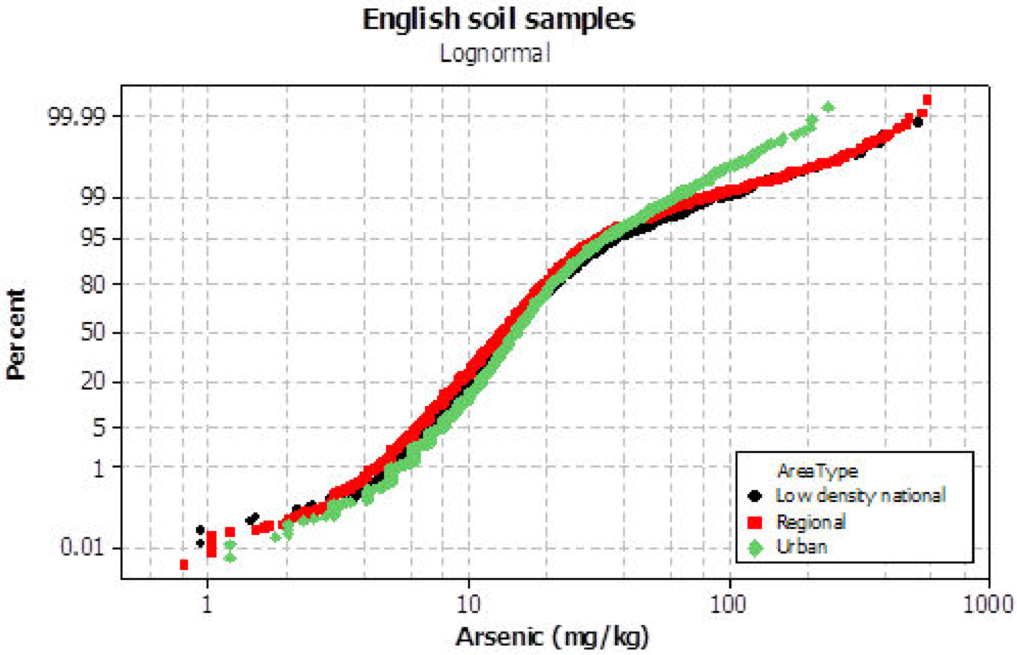


Figure 17: Probability plot of As in topsoils. Data categorised according to density of sampling programme and presented with a $\log_{10}$ x-axis to illustrate variation over the full range of concentrations

(a) As (mg/kg)	Number	Mean	Minimum	25th percentile	Median	75th percentile	Maximum	Skewness
All data	42133	17.6	<0.5	10.7	14.1	18.9	15110	186

(b) As (mg/kg) Area type	Number	Mean	Minimum	25th percentile	Median	75th percentile	Maximum	Skewness
Low density national (NSI XRFS)	4864	18.5	<0.5	10.8	14.4	19.9	536	10
Regional (G-BASE rural)	23686	17.5	<0.5	10.1	13.4	18.1	15110	143
Urban (G-BASE)	13583	17.5	1.20	12.1	15.1	19.5	1008	29

(c) As (mg/kg) Area name	Number	Mean	Minimum	25th percentile	Median	75th percentile	Maximum	Skewness
National (NSI XRFS)	4864	18.5	<0.5	10.8	14.4	19.9	536	10
Eastern England	23222	16.2	<0.5	10.0	13.2	17.9	555	12
Tamar catchment	464	80.4	6.90	16.5	22.3	42.3	15110	21
Corby	133	23.4	10.7	16.9	19.5	23.9	90	3
Coventry	390	9.9	2.03	7.1	9.1	11.1	105	9
Derby	275	15.8	5.66	11.3	13.5	16.7	63	3
Doncaster	279	15.3	2.03	10.1	13.1	17.1	75	3
Hull	407	24.1	3.04	15.1	20.2	26.2	207	6
Leicester	652	14.1	4.25	9.9	13.2	17.3	84	3
Lincoln	215	15.3	4.05	8.1	11.1	21.2	65	2
London (GLA area)	6494	17.1	1.20	12.9	15.5	18.9	161	5
Manchester (part of)	300	28.1	2.53	15.8	20.3	28.4	1008	16
Mansfield	257	13.9	3.04	7.1	11.1	16.1	72	3
Northampton	275	34.3	8.48	22.8	30.4	42.6	107	1
Nottingham	636	14.4	5.05	11.1	13.1	16.1	88	4
Peterborough	272	18.0	7.47	14.5	17.2	21.1	35	1
Scunthorpe	196	26.4	3.04	11.1	19.2	31.2	191	3
Sheffield	575	25.8	4.05	17.1	22.2	29.2	241	5
South Essex towns	715	14.5	4.62	11.4	13.4	16.2	82	4
Stoke-on-Trent	745	16.3	2.03	11.1	14.1	18.1	137	5
Telford	292	12.0	5.05	8.1	10.1	13.9	54	3
Wolverhampton	284	19.8	3.54	13.4	16.6	22.5	158	6
York	191	11.6	3.04	8.1	10.1	13.1	94	7

Table 5: Statistical summary of As in topsoil from the main datasets (a) whole dataset (b) by sampling density and (c) by local areas

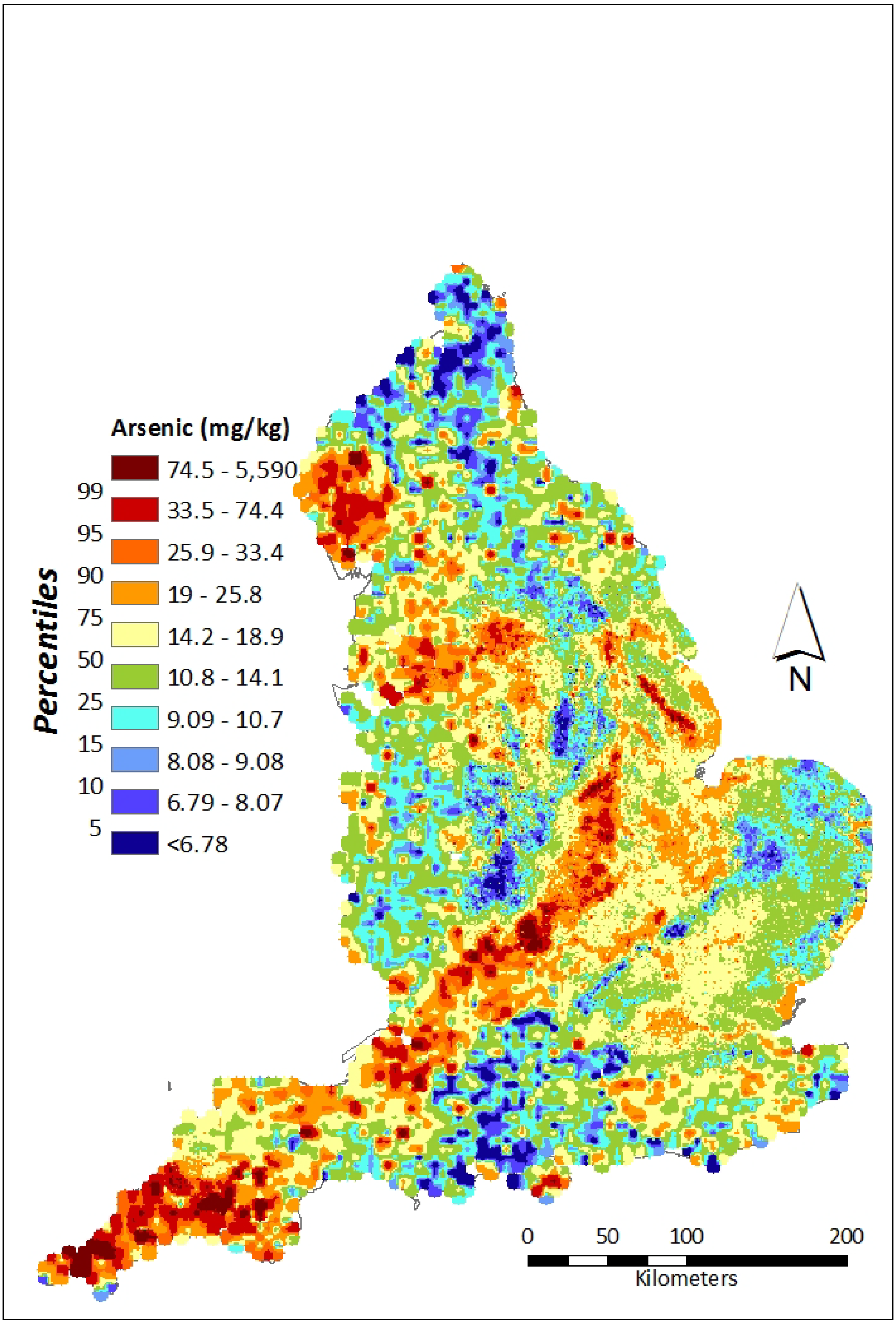


Figure 18: Interpolated map of topsoil As. Colour thresholds are designed for highly skewed geochemical data

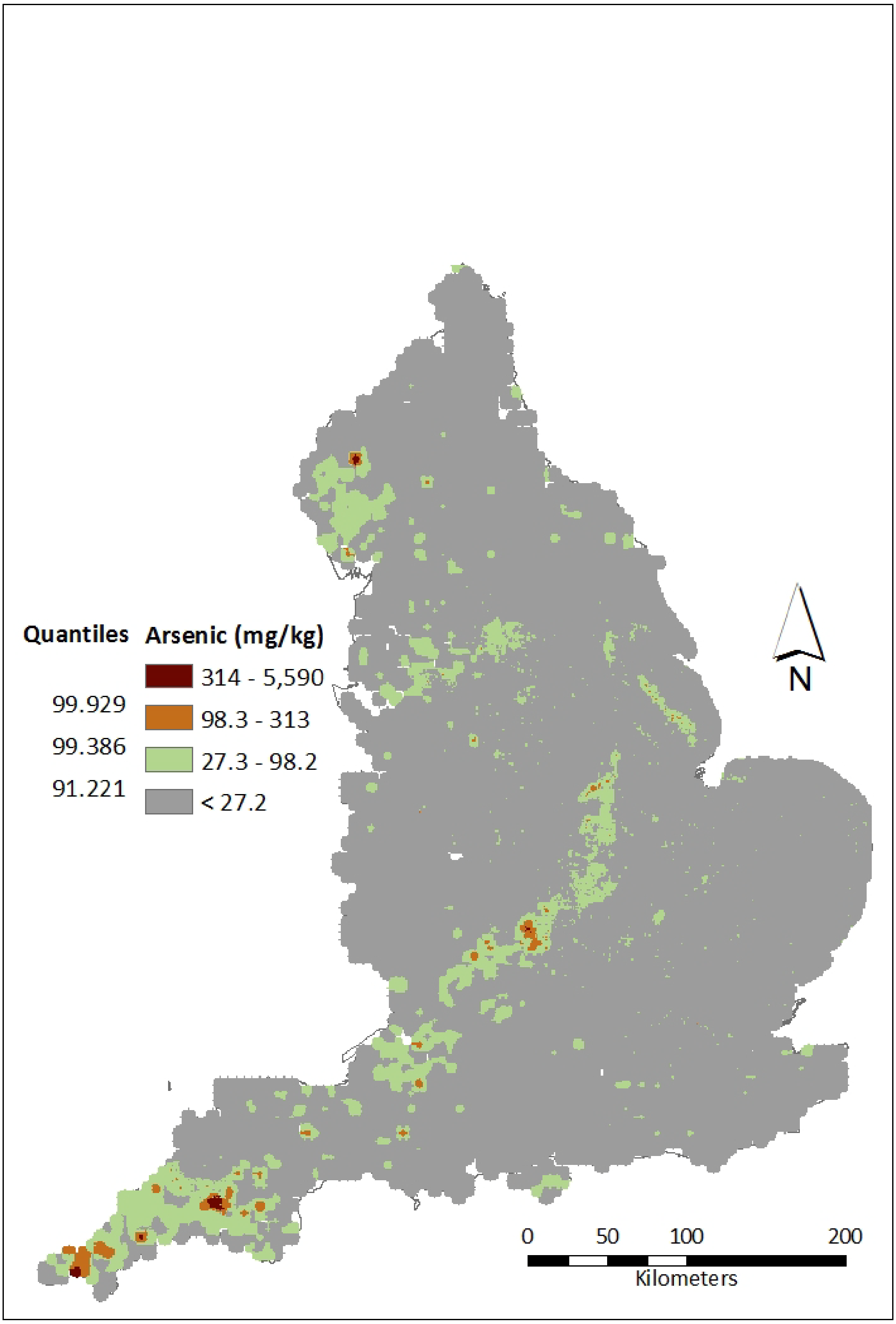


Figure 19: Interpolated map of topsoil As using k-means cluster analyisis

3.2.1.4 DEFINING ARSENIC DOMAINS

Geological controls on As in soils are seen in the geochemical data (Figure 18 and Figure 19). Highest (>19 mg/kg; 90th percentile) concentrations of soil As shown in Figure 18 are typically associated with (i) As-bearing ironstones, (ii) As-bearing mineralisation, and (iii) organic-rich soils (such as the East Anglian Fens); these are discussed further below for the areas where they fall above the lowest k-means cluster group (>27 mg/kg As).

The lowest concentrations (<10 mg/kg) are typically observed over the Chalk of southern England where there is no glacial till cover, the sands of Norfolk and Suffolk (including Breckland and coastal areas). In central England and into north Nottinghamshire the Mercia Mudstone Group is associated with particularly low soil As concentrations, whilst in northern England the Carboniferous Limestone and Millstone Grit successions give rise to soils naturally low in As. These are not considered any further in this work due to their low concentrations.

Ironstones

The location of the ironstones outcrop, as defined by the BGS Soil-Parent Material Model (SPPM v6), is shown in Figure 20; this occupies *c*.1% of the surface area of England (*c*.1,300 km^2). It can be seen that the major outcrop areas extend SW-NE in central England, with a further area in eastern England (Lincolnshire). Smaller outcrops can be seen in northern and southern England, including part of the Isle of Wight. The largest outcrop areas appear to approximately coincide with the area of higher data in Figure 19. The summary statistics for the data from soil samples which directly overlie the ironstones is shown in Table 6. It can be seen that in comparison to the overall data set (median 14 mg/kg) these concentrations are greatly elevated (median 45 mg/kg) and confirm that these data are generally not within the lowest k-means cluster (25th percentile of 28 mg/kg).

Since the G-BASE project reported high soil As concentrations over parts of central/eastern England (Breward, 2007; Rawlins *et al.*, 2003b) considerable effort has been made to understand the causes and implications of these data. The ironstones can be both silicate and carbonate cemented (SPMM classes), although this does not appear to have a strong influence on soil As concentrations (medians of 45 mg/kg in both cases). It is also worth noting that there is no apparent superimposition of any urbanisation effects increasing As concentrations in these data, with a rural median of 56 mg/kg (n = 253) and an urban median of 35 mg/kg (n = 162) in G-BASE data.

The ironstones by definition have >15% iron oxides, and these can exceed 30%; it is these to which the majority of the As is bound (Palumbo-Roe *et al.*, 2005). This strong association to iron mineral phases results in a very low proportion of the As being bioaccessible in *in vitro* human bioaccessiblity testing (Cave *et al.*, 2002), with 10% typical for these soils (compare with Table 6 results), and low stream water concentrations (Breward, 2007).

Area name	Number	Mean	Minimum	25th percentile	Median	75th percentile	Maximum	Skewness
All ironstone	437	72.9	4.05	27.8	45.0	83.4	555	2.8

Table 6: Summary As concentrations (mg/kg) in topsoils over ironstone parent material

The outcrop area of the ironstones can be compared with soil As concentrations. Figure 21 shows an example demonstrating that the outcrop area provides a very good boundary to these concentrations, where high density data is available. This is despite the fact that the boundary edges are highly complex, due to the tectonic history and surface erosion of these formations. It can be seen that the soil As concentrations vary as much as tenfold across the boundary, irrespective of how close they lie to that boundary. Sufficient data is also available on the Isle of Wight outcrop to give confidence in this boundary. A problem of sampling densities relatively low in relation to the outcrop areas of soil parent materials is shown in Figure 22; the low density (NSI) sample points around the ironstone outcrops are shown, but none of them intersect the outcrop. On the basis of the findings for the rest of the outcrops it assumed that this region's ironstone would form part of an ironstone domain for NBCs, but this requires further testing through sampling and analysis.

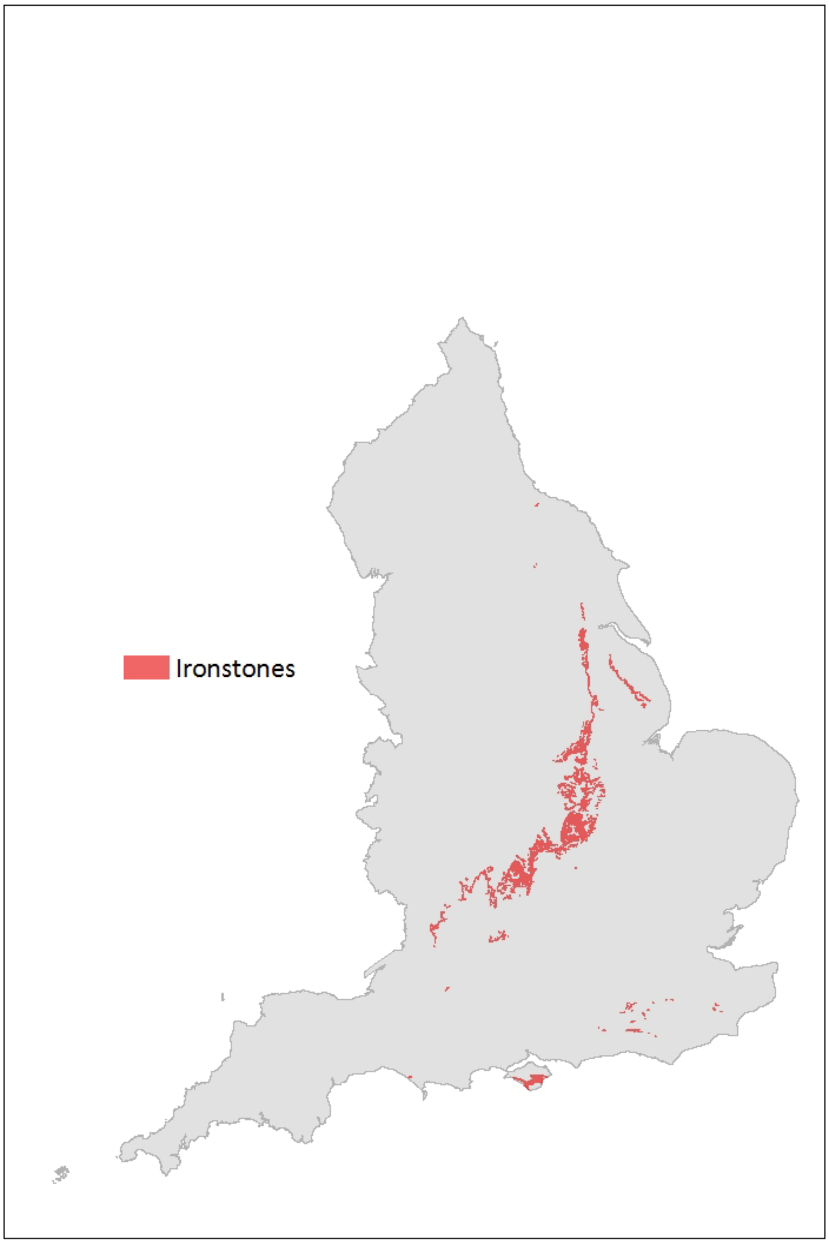
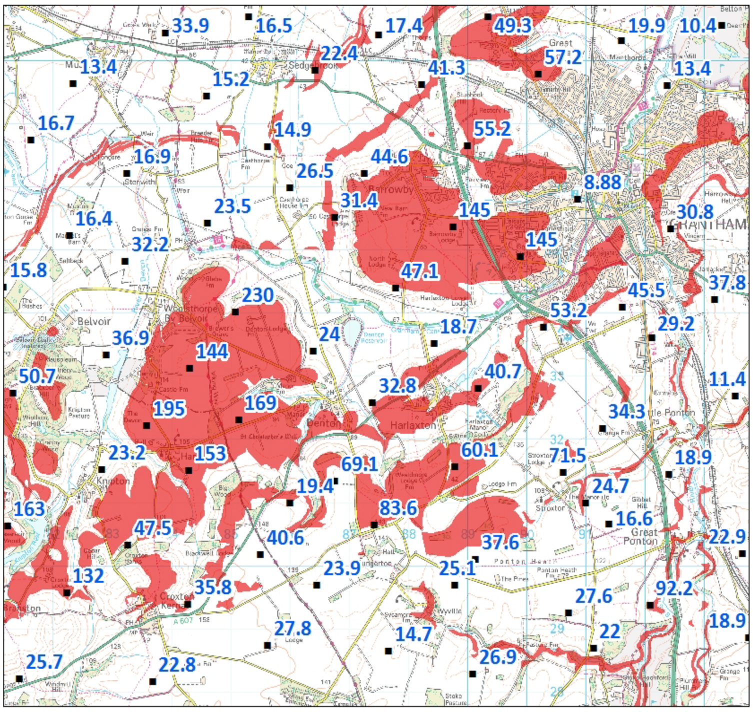
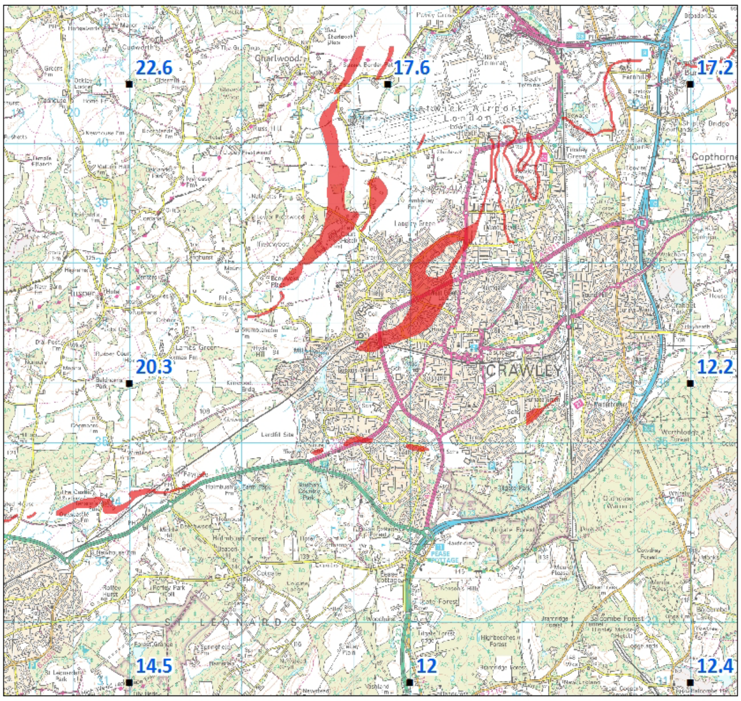


Figure 20: Map showing area of ironstone outcrop



BGS © NERC. Contains Ordnance Survey data © Crown Copyright

Figure 21: Example of topsoil As concentrations over and around ironstones (Grantham, Lincolnshire). Data in blue are As concentrations in mg/kg for sampling sites shown as black dots, ironstone outcrop area in red.



BGS © NERC. Contains Ordnance Survey data © Crown Copyright

Figure 22: Example of topsoil As concentrations close to, but not over ironstones (Crawley, West Sussex). Data in blue are arsenic concentrations in mg/kg for sampling sites shown as black dots, ironstone outcrop area in red.

Metalliferous mineralisation and mining (non-ferrous)

Mineralisation can lead to an entirely naturally enhanced concentration of soil As if the soil is formed over an As-containing mineral vein. The underlying precept of geochemical exploration, the foundation for all geochemical mapping techniques, is this property of naturally enhanced concentrations in environmental media near a mineral deposit. These concentrations can be very high, and of themselves potentially harmful (Runnells *et al.*, 1992). In addition to these naturally elevated concentrations, an ore field exploited for its minerals prior to pollution and prevention control measures, will tend to have resulted in point source and diffuse increases in concentration. This would be through: waste material deposits; wind/water dispersion of those wastes; and local processing of the economic ores to purify them to sell, further leading to atmospheric emissions and disposal of wastes to stream courses. An example of the relative variation which may be expected if traversing a mineral vein, or area of diffuse pollution, is shown in Figure 23.

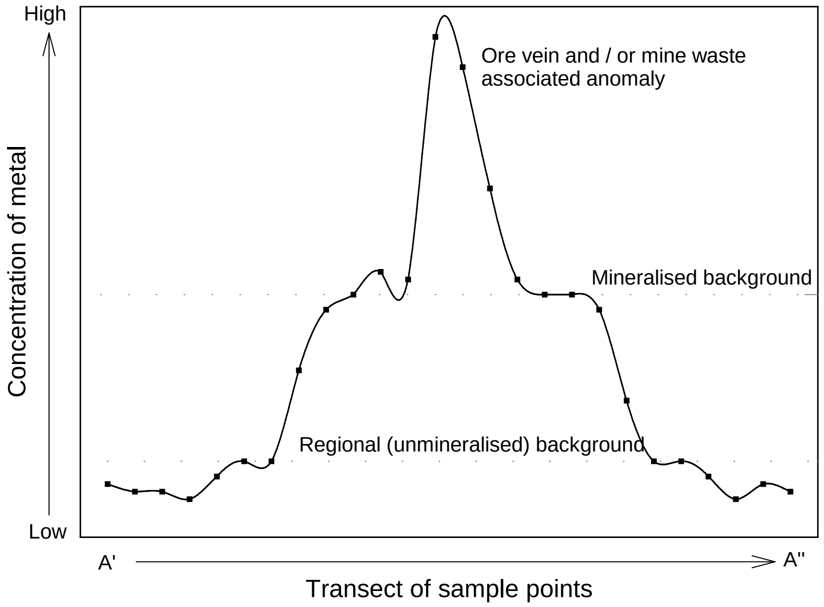


Figure 23: Idealised illustration of the naturally elevated baseline occurring over an area of mineralisation, and an anomaly associated with a mineral vein or mine wastes

The mapping of mineral veins will depend entirely on the degree to which an ore field has been surveyed in detail, and the size of the mineral vein which makes definitive attribution to this source difficult. The locations of many abandoned mine workings are also poorly mapped, which can also lead to problems trying to use such geographical data to locate mining areas. Much greater spatial uncertainty should be expected defining boundaries of these areas than with relationships with a more widespread and well-mapped parent material such as the ironstone.

When the topsoil As data from sites positioned within the generalised mapped areas of metalliferous mining in England are statistically summarised (Table 7) it can be seen that not all the ore fields are associated with enhanced topsoil As concentrations. This is a direct result of the presence or absence of As in the primary mineralisation, whether as the mined ore or as an accessory element in the deposit. In particular, it can be seen that the Cornwall/west Devon, Lake District and Mendip ore fields have high As concentrations in relation to the overall data set, whilst the others do not. Lead in minerals in the Mendips are reported to have quite high concentrations of As (*e.g.* pyromorphite[7]). Although the Peak District area has higher concentrations than some of the surrounding area (Figure 18), which is in keeping with

---

[7] http://www.bgs.ac.uk/mendips/minerals/Mins_Mines_1.htm

literature reports of enhanced As in the mineralised, mined and smelting area (Li and Thornton, 1993), these concentrations are still generally <25 mg/kg, which is why they do not appear as significantly different in Figure 24 and are not included in the As mineral area classification. The highest concentrations are found in Cornwall/Devon, which is consistent with the economic importance of that ore field for As and Cu mining in the 19th century, which resulted in the widespread extraction, processing and uncontrolled (and unmapped) waste disposal (Abrahams and Thornton, 1987; Aston *et al.*, 1975). Comparison with Figure 19 also shows that this is reflected in the outputs of the k-means cluster analysis.

Thus, the separate ore fields have been categorised accordingly, and only those which have high typical As concentrations are used in further analysis in preparation for a possible separate mineralised/mined spatial domain. The distribution of these ore fields is shown in Figure 25, whilst the data are presented in Table 8. With a median of 46 mg/kg, and other high concentration statistical properties, it can be seen that this data distribution is very different to that in the rest of the data set (median 14 mg/kg). The large positive skew in the data also reflects uncertainty in categorising the data. An example of this is shown for part of Devon/Cornwall in Figure 26 where it can be seen that high topsoil As concentrations, up to 300 mg/kg, occur out-with the mapped mineralised/mined zones. Some of these have been examined in more detail and it can be seen that where former mine shafts are not mapped at the 1:50,000 scale (but high soil As concentrations are found and shafts are mapped at 1:10,000) they have been missed.

This suggests that whilst the categorisation of these data using mining records is valuable in allowing potential data domains to be created, they are likely to be associated with false-positive (outside the mapped area) and false-negative (within an apparently impacted area) results more often than boundaries such as the ironstones, if used in this way. This is almost inevitable with the spatial data available, and within the general area of the ore fields described high soil As concentrations could be related to mineralisation/mining. A further uncertainty in this analysis lies in the fact that so few samples are located over the North Devon, Shelve and Mendips areas, as these are only covered by NSI sampling.

Arsenical and metalliferous ores tend to be associated with more soluble As than is found in the ironstone areas described above, with a positive relationship between soil/sediment concentrations and stream water concentrations in the Tamar area, for instance (Aston *et al.*, 1975; Rawlins *et al.*, 2003a) unlike that over ironstones (Breward, 2007). This is likely to be a result of an association of poorly crystalline iron oxyhydroxide phases and weathering products of the primary sulphide mineral ore phases, unlike the very insoluble iron oxides found in the ironstones. Bioaccessibility tests on soils from the Tamar area found that a median of 10% was bioaccessible (Palumbo-Roe and Klinck, 2007) though more recent bioaccessibility modelling indicates that about 25% is bioaccessible in the Tamar area.

Area	Samples	Mean	Minimum	25th percentile	Median	75th percentile	Maximum	Skewness
Non-mineralisation	41626	16.9	<0.5	10.7	14.1	18.8	1008	15.0
1 SW England	153	208.2	6.9	27.9	47.5	123.2	15110	12.1
2 North Devon	4	20.2	12.1	14.0	20.7	25.9	27	-0.5
3 Mendips	8	47.5	20.9	28.6	45.4	66.3	81	0.3
4 Shelve	3	15.5	12.8	12.8	13.5	20.2	20	1.7
5 Peak District	224	16.8	0.5	13.0	16.1	20.2	70	1.7
6 Lake District	26	58.1	12.0	26.7	40.9	51.5	536	4.9
7 North Pennines	89	14.4	5.0	8.8	13.2	16.3	74	3.6

Table 7: Summary statistics of topsoils As (mg/kg) in the main non-ferrous mining areas of England

Area	Samples	Mean	Minimum	25th percentile	Median	75th percentile	Maximum	Skewness
As - mineralised	187	181	6.9	27.8	45.6	106	15110	13
Not As-mineralised	41946	16.9	<0.5	10.7	14.1	18.8	1008	15

Table 8: Summary statistics of As mineralised area topsoils (mg/kg) compared with all other topsoil data

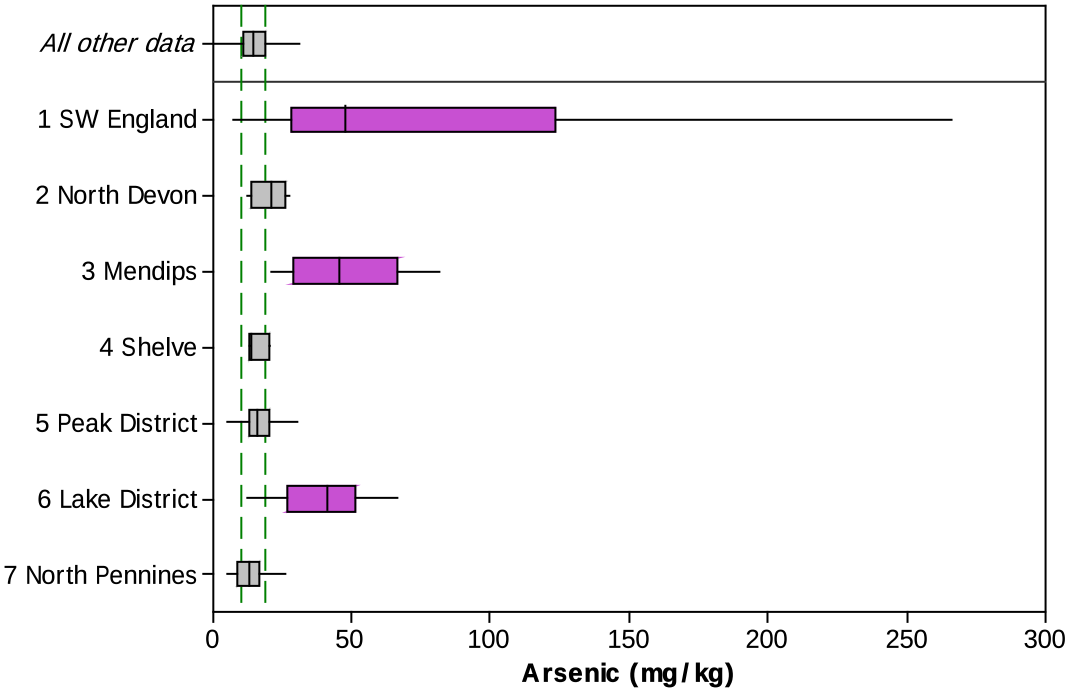


Figure 24: Boxplot of topsoil As concentrations over the main English ore fields, compared with the overall dataset. The boxes show the interquartile range (which is projected for 'all other data' using dashed lines, to facilitate comparison), with the median shown within the box

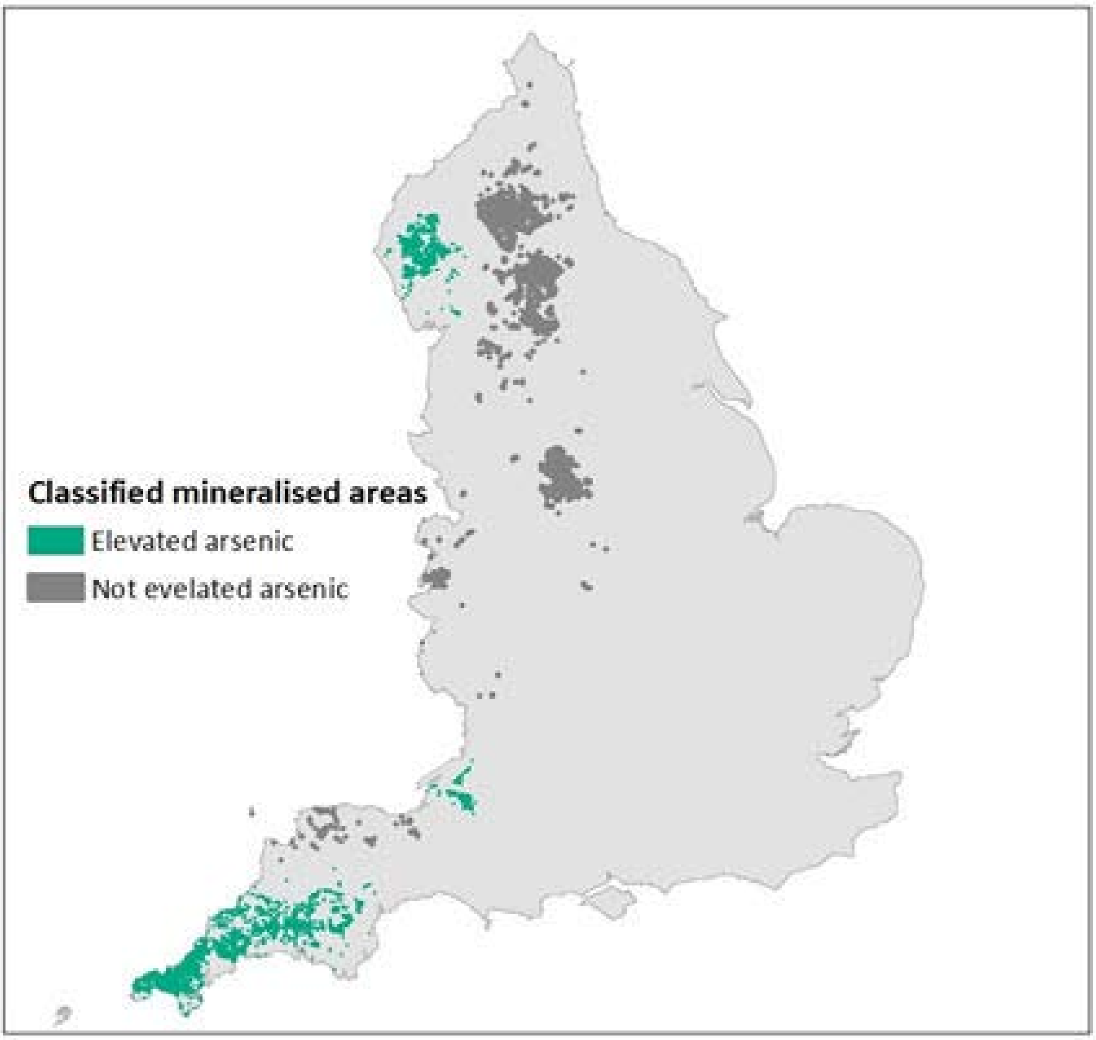
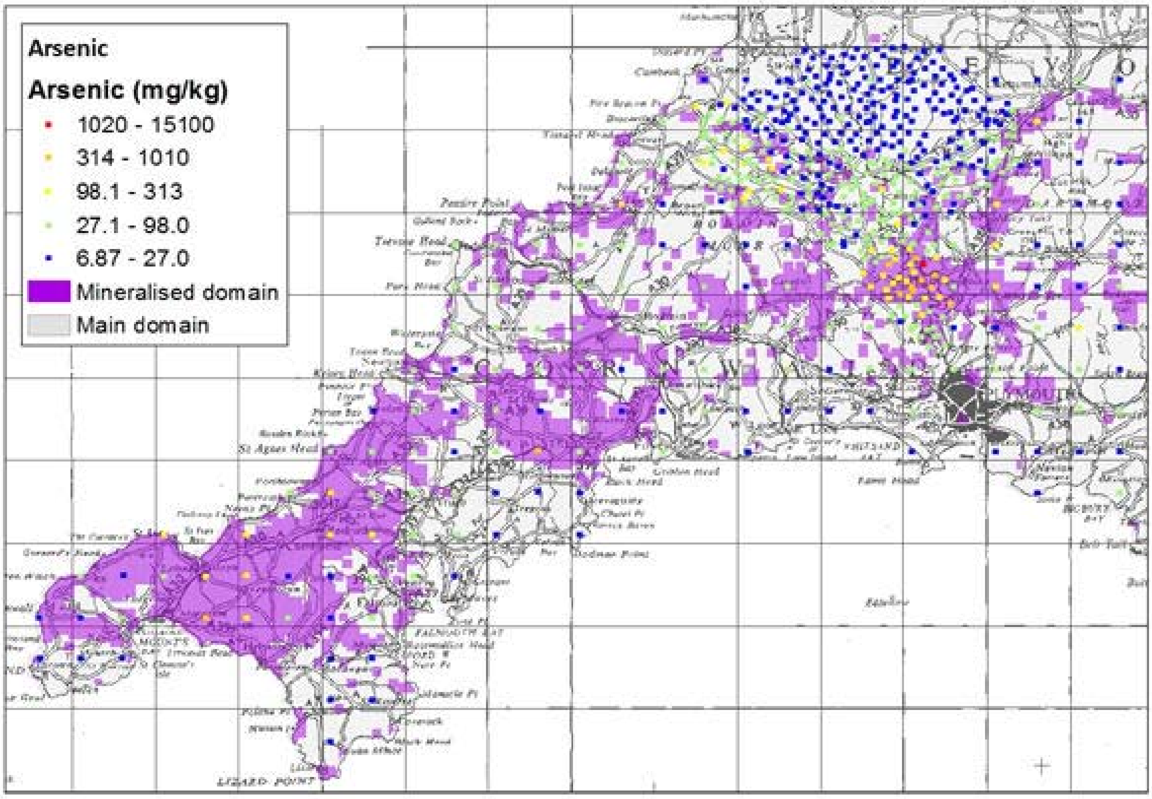


Figure 25: Map of metalliferous mining areas in England, thematically coloured depending on whether topsoil As concentrations are typically elevated



BGS © NERC. Contains Ordnance Survey data © Crown Copyright

Figure 26: Example of mineralised area mapping and topsoil As concentrations (thresholds from k-mean cluster analysis, Figure 16) in Cornwall

3.2.1.5 PROPERTIES OF THE DOMAINS

When the topsoil As data are split into the As mineralised and ironstone groups described above, with the residual data described as “principal” they have the statistical properties shown in Table 9, with the distribution of the data shown in Figure 27 and Figure 28. It can be seen that there is a large difference between the distributions of the data sets with enriched soil As when compared to that of the Principal Domain; there is a much smaller difference in the properties of the ironstone compared with the mineralised area. The relatively small skew in the Ironstone Domain is reflected in the distribution in Figure 28, and contrasts with that from the mineralised areas, where the mean is much greater than the 75th percentile (Figure 27). This difference is likely to result from the predominantly natural controls acting on the soils above ironstones as compared to the additional mining activities in the mineralised zones; there is also much greater uncertainty on the mapping boundaries for the mineralised zone as described above. For these reasons, and the expected differences in arsenic mineralogy, the Mineralised and Ironstone Domains should not be amalgamated, despite their statistical similarities. There are still high concentrations outliers in the Principal Domain (*c*.5% are >27 mg/kg; Figure 28), which Work Package 3 will investigate further to establish if they should be included in enriched domain classes.

.

Area	Samples	Mean	Minimum	25th percentile	Median	75th percentile	Maximum	Skewness
Principal	41509	16	<0.5	10.6	14.1	18.6	1008	18
Ironstone	437	73	4.1	27.8	45.0	83.4	555	3
Mineralised	187	181	6.9	27.8	45.6	105.5	15110	13

Table 9: Summary statistics for domain classified topsoil As data (mg/kg)

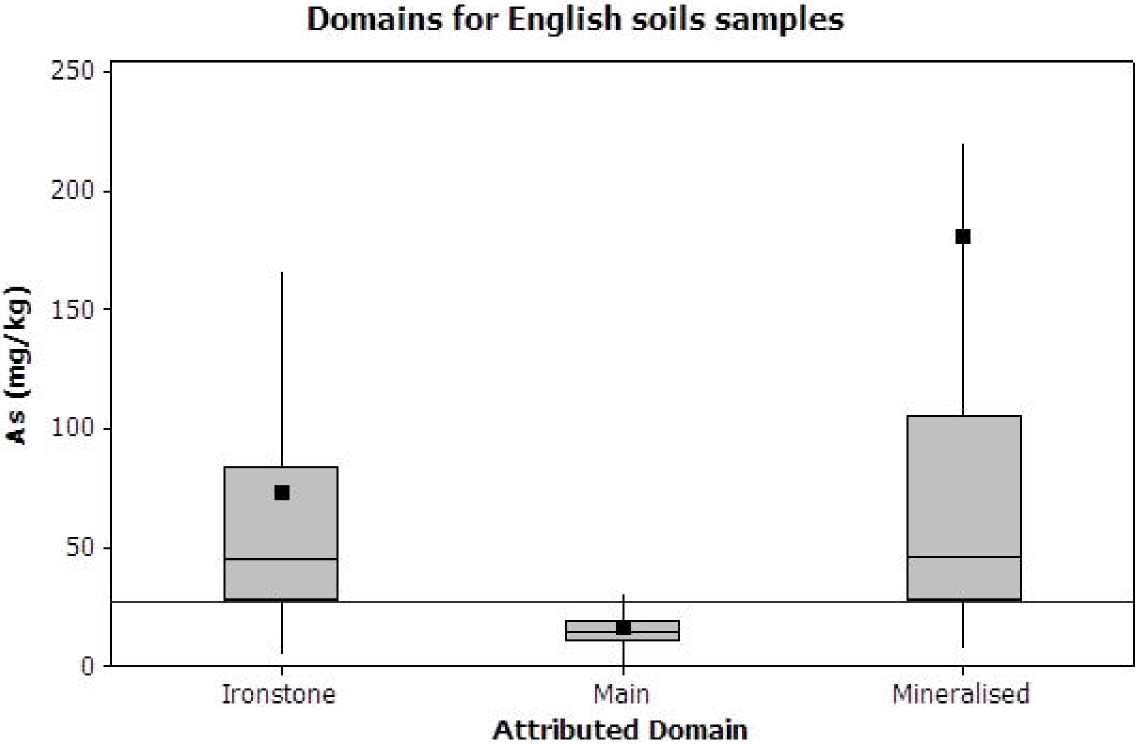


Figure 27: Boxplot of As topsoil results attributed to potential domains (Main = Principal Domain)

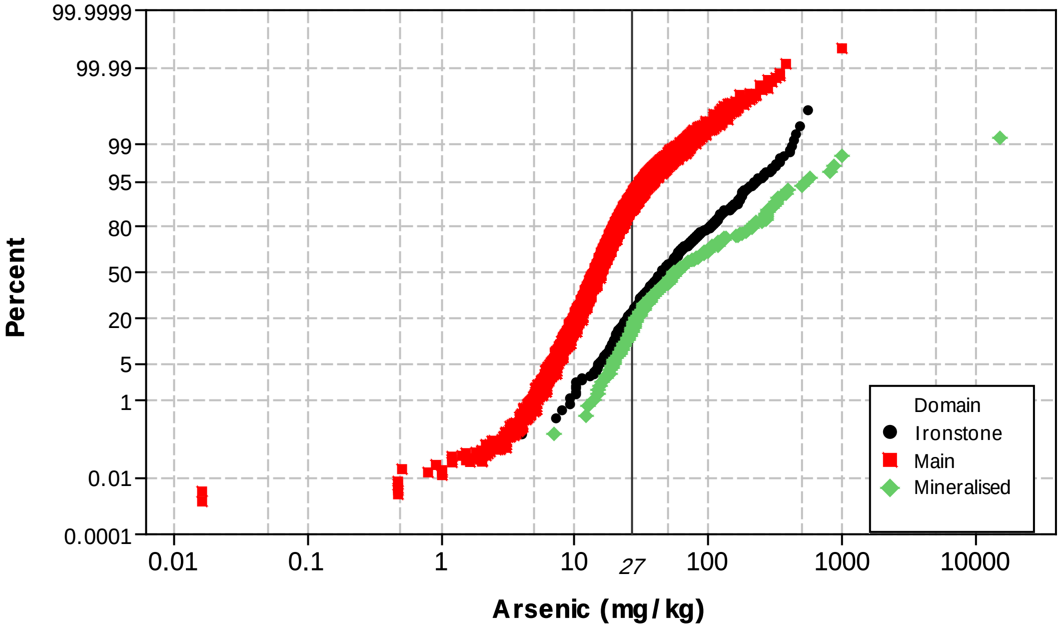


Figure 28: Probability plot of topsoil As results categorised by potential domains. The k-means threshold of 27 mg/kg is shown by the vertical line. (Main = Principal Domain)

3.2.1.6 COMPARISON WITH OTHER ANALYTICAL METHODS

The XRFS analysis, for all the topsoil results described so far in this section, present both an internally consistent data set, and an absolute total concentration for soil As. However, XRFS is not as widely used outside regional geochemical surveys as it is within them; more frequently acid digestion followed by ICP-MS determinations are used to measure soil As concentrations. Two data sets are available from systematic surveys which have both XRFS and ICP-MS analysis following *aqua regia* digest, and a further data set which has two acid digestion and analytical methods used to measure soil As.

GEMAS

The GEMAS survey (Reimann *et al.*, 2012) of grazing and arable soils (paired sites collected close to each other) has topsoil As data by XRFS and *aqua regia* digest followed by ICP-MS analysis. The data for England (n = 130) are shown in Figure 29, with additional information on whether they are arable or grassland pasture sites. It can be seen that there is no difference arising from land use, and that the data has a linear positive relationship which has a minor offset from the 1:1 line, with lower concentrations recovered from the acid digest than XRFS analysis. The relative importance of these factors has been tested using an analysis of variance (ANOVA) on log(10) transformed data, which produces a normal distribution (P<005). The ANOVA results show that 4% of the variance is accounted for by differences in analytical method, whilst <1% by land use; the remainder of the variance lies in geochemical variation between sites. The regression equation (determined using Minitab v16 software application) is:

$$[\mathrm{As}_{\mathrm{XRFS}}] = 2.58 + (1.13 \times [\mathrm{As}_{\mathit{aqua\ regia}}]) \text{ (n=130; } R^2 = 89\%\text{; and } P<0.05).$$

Two significant outliers can be seen where XRFS As is much greater than ICP-MS As.

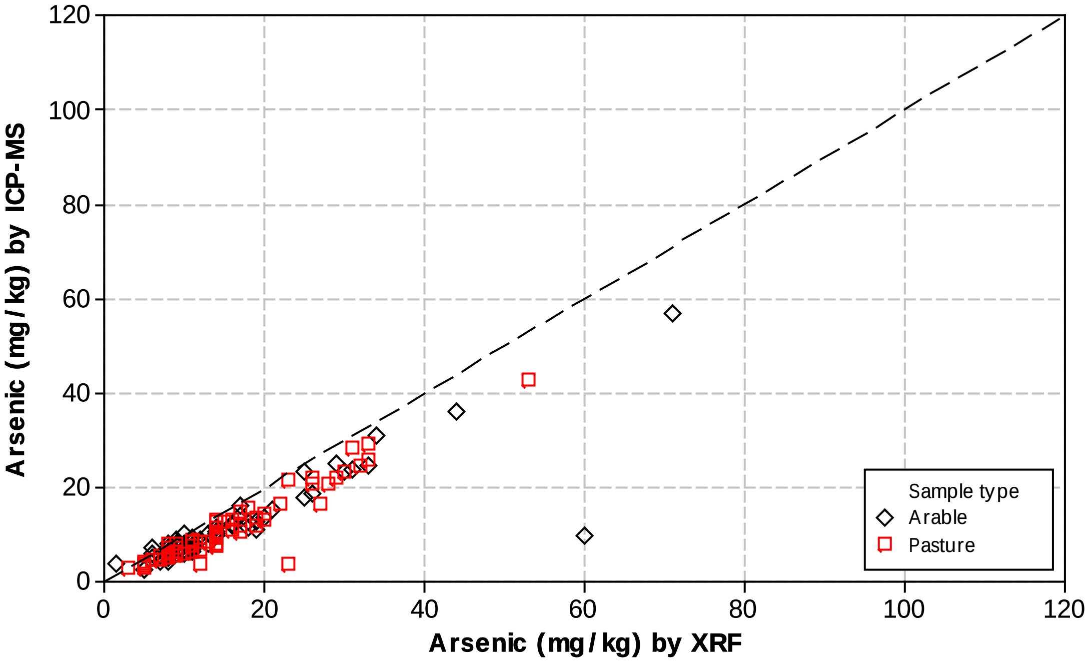


Figure 29: Comparison of GEMAS project topsoil As concentrations by XRFS and *aqua regia* digest ICP-MS in England

Tellus

The Tellus geochemical survey of Northern Ireland (Smyth, 2007) has analysed 6872 topsoil samples by XRFS and by ICP-MS after *aqua regia* digest (Figure 30). It can be seen that there is a close linear relationship between the data with a systematic bias to higher concentrations by XRFS; this would be expected from this total measurement, unlike the acid digest which will not attack a proportion of residual material. There would also be a systematic bias expected between two analytical measurement techniques.

The regression equation for these data (determined using Minitab v16 software application) is:

$[As_{XRFS}] = 3.75 + (1.06 \times [As_{aqua\ regia}])$ ($n=6872$; $R^2 = 94\%$; and $P<0.05$).

FOREGS

The FOREGS (Salminen *et al.*, 2005) topsoil As results were determined using two different digestion methods (and not by XRFS). These data are shown in Figure 31. The lower concentrations of the *aqua regia* digest are compromised by the relatively high detection limit of ICP-AES used to measure As, when compared with the 'total' digestion method which used ICP-MS analysis. When that effect is excluded, it can be seen that there is no significant difference between the methods.

These data suggest that the relationships found above between XRFS and *aqua regia* digest on two data sets, could be expected to hold true for any quasi-total acid attack followed by ICP-MS analysis.

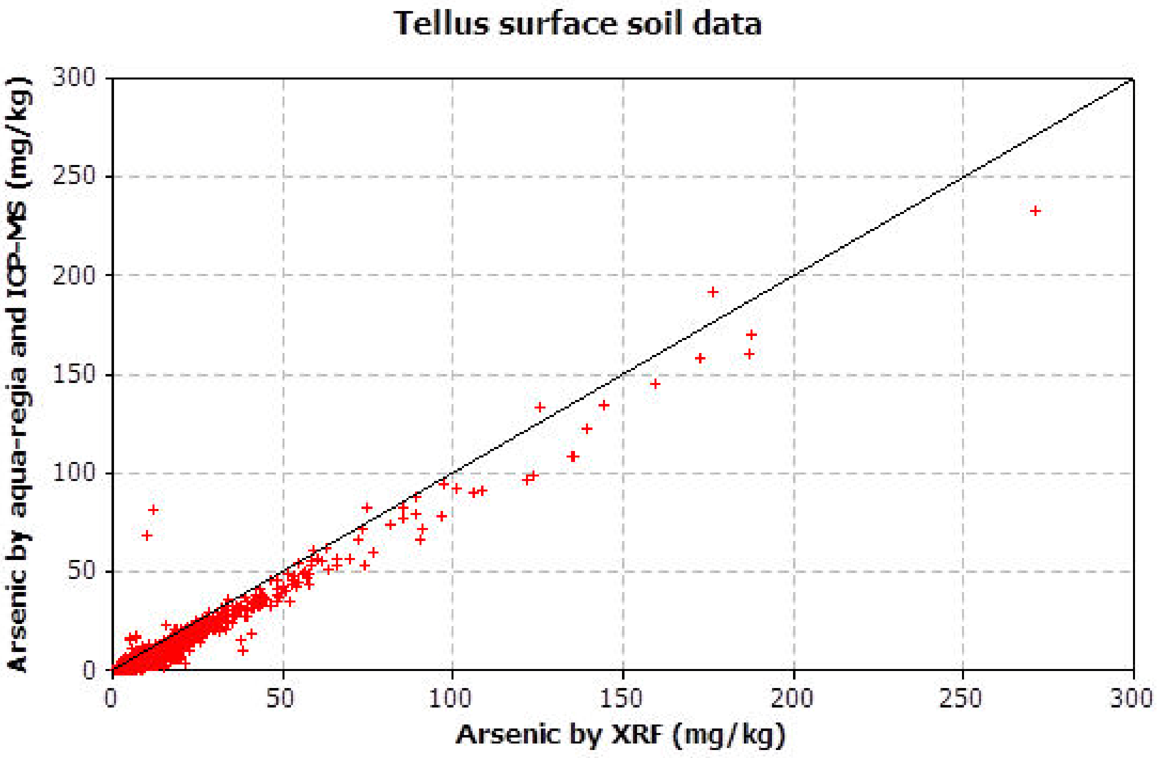


Figure 30: Comparison of topsoil As concentrations in Northern Ireland by XRFS and ICP-MS after *aqua regia* digest

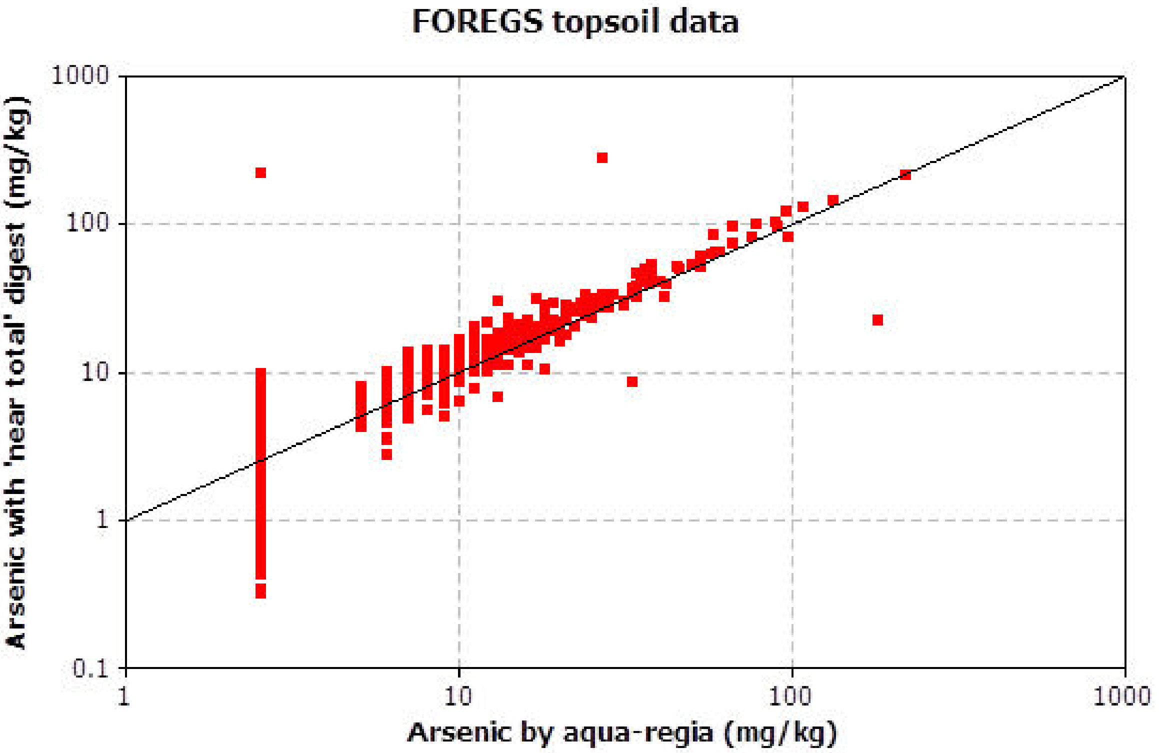


Figure 31: FOREGS topsoil As data from two different acid digest methods

3.2.2 Lead

Soil Pb concentrations are known to be affected by both underlying lithological Pb concentrations and diffuse anthropogenic release of Pb, primarily from combustion engine exhausts.

Lead mineralisation in soil parent material is one of the most significant sources of naturally high concentration soil Pb. Ore extraction and processing can further enhance concentrations and result in a much wider area being affected by diffuse contamination, as previously described for As. Lead is also a common trace element found in coal, leading to diffuse release from domestic and industrial combustion, (Rawlins *et al.*, 2002) but probably the most significant more recent source of diffuse Pb in the environment was that from the use of alkyl-Pb as a petrol additive. Concentrations in petrol were decreased from 0.4 g/L to 0.15 g/L in 1986 , with complete withdrawal of leaded-petrol in the UK in 2000. The isotopic signature of the alkyl-Pb used was sufficiently different to those from UK ores so that it has been possible to clearly demonstrate the significance of this source of Pb in soils (Bacon *et al.*, 1996; Bellis *et al.*, 2004), in addition to the evidence from the distance-decline from busy roads routinely observed in soil Pb concentrations (e.g. Kelly *et al.*, 1996).

3.2.2.1 AVAILABLE DATASETS

Lead is measured routinely by the largest soil geochemical baseline projects (Table 10), so topsoil data from these have been used as the major ('core') data sets in this analysis. These data provide systematic coverage, through consistent sampling and analysis methods, at differing densities for the whole of England and many urban areas and are shown in Figure 32. All data is determined by XRFS, which has the benefit of being an absolute total concentration. The data sets used are as described for As (Table 4).

In order to prepare the three core datasets for use, a minor correction factor was applied to data on samples collected prior to 2006, taken from the BGS Corporate Geochemistry Database (1st October 2011) to ensure the data was levelled to the most recent Certified Reference Material Pb concentrations.

The NSI survey original analysis (*aqua regia)*, and the minor data sets - GEMAS and FOREGS - are useful comparisons between XRFS and some common analytical digestion methods used in site investigations.

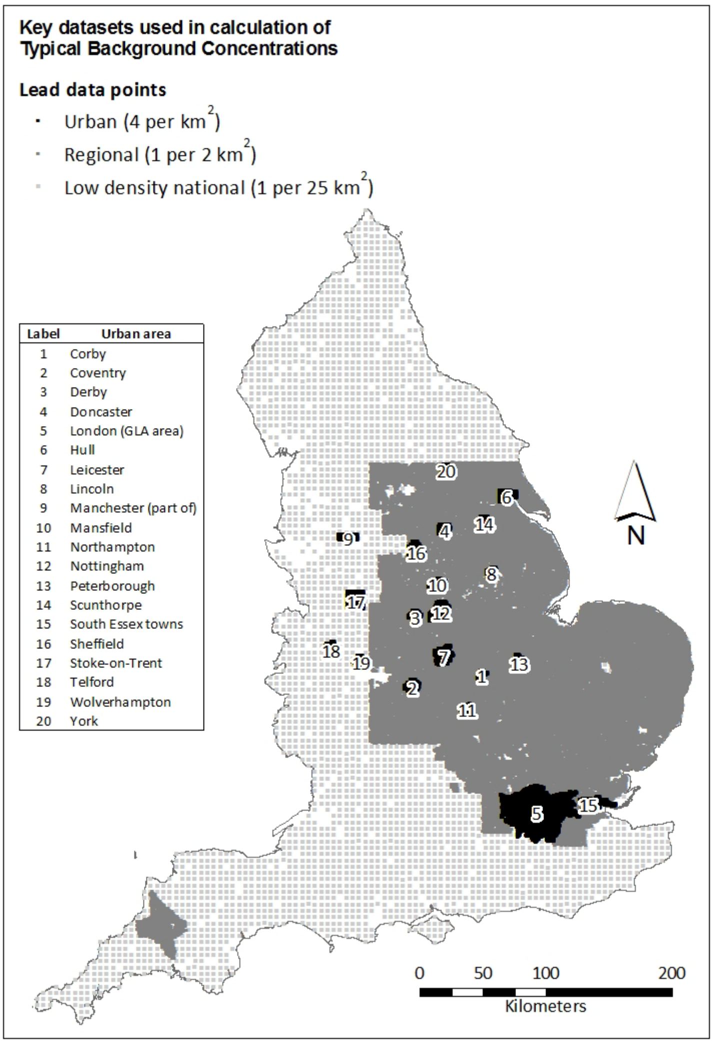


Figure 32: Location of Pb topsoil sample site data used in Figures 33 - 35 and Table 10 (the low density NSI(XRFS) sites exist within the G-BASE areas).

3.2.2.2 OVERALL DATASET PROPERTIES

The median concentration of Pb in all the topsoils is 47 mg/kg with an overall range exceeding a thousand-fold variation, and a positively skewed distribution (Table 10a). This median is similar to that from various previous robust estimates (*e.g.* median, geometric mean) of typical soil Pb concentrations: geometric mean of 42 mg/kg for four areas of England and Wales (Davies, 1983); cereal growing fields in Britain of 26 – 32 mg/kg (Zhao *et al*., 2004); rural soil median of 37 mg/kg in Britain (Spurgeon *et al*., 2008); and, 41 mg/kg for England from UKSHS (Ross *et al*., 2007). Where separate summaries of data from Scotland are provided the English data tend to be higher in comparison – *e.g*. median values of 20 mg/kg (Reaves and Berrow, 1984) and 28 mg/kg (Ross *et al*., 2007).

Whilst available data for England are reasonably consistent in central tendency, they are elevated when compared with published pan-European (FOREGS) data, where the median is 15 mg/kg (Salminen *et al*., 2005), and England can be seen to have concentrations in the upper end of the mapped data distribution.

3.2.2.3 VARIATION BETWEEN AREAS

A substantial difference in topsoil Pb concentrations between urban and rural areas is evident in Table 10(b) and Figure 33, where the median concentration for urban G-BASE data is

128 mg/kg, with a 75th percentile of 253 mg/kg. These areas are then subject to further large variations in typical concentrations, as has previously been observed in large urban studies in the UK (Culbard *et al*., 1988). Whilst the median of 155 mg/kg in Derby is related to the proximity of the Peak District ore field, many of the other variations in concentration will most likely relate to factors such as total population, population/transport network density, age of buildings and industrial history (Culbard *et al*., 1988; Kelly *et al*., 1996). There are also urban centres such as Corby and Peterborough, which have atypically low concentrations, with medians of 37 mg/kg and 39 mg/kg, respectively. Larger urban areas, and in particular London and its satellite towns, can be clearly identified in Figure 34. Otherwise high concentrations over large areas generally reflect ore fields where Pb-mineralisation is abundant, such as the Peak District, North Pennines and Lake District. Much of Norfolk is characterised by typically low concentrations (<30 mg/kg), but it is interesting to note that the impact of relatively small urban centres can still be picked out at the national scale in this area. Lead has a well recognised affinity for organic matter, which McGrath and Loveland (1992) noted through the strong relationship between Pb concentration in English and Welsh soils mapped as peaty (median 116 mg/kg; overall dataset median 40 mg/kg), as had Reaves and Berrow (1984) previously for Scottish soils. Within this data set, this relationship may be complicated by the proximity and abundance of ore fields in the larger extents of organic-rich soils in northern England, and it can clearly be seen that Pb is not present in enhanced concentrations in the extensive East Anglian fen peats which are found north of Cambridge.

In order to help the identification of significant domains needed to characterise topsoil Pb in England, the k-mean cluster approach has been used as has previously been described for As. The outcome is shown applied to the interpolated concentration map in Figure 35. This clearly shows the importance of mineralisation and mining areas, particularly in northern England, and the impact of urbanisation, most clearly seen in the London area. This information has been used to guide the domain characterisation in the following subsections.

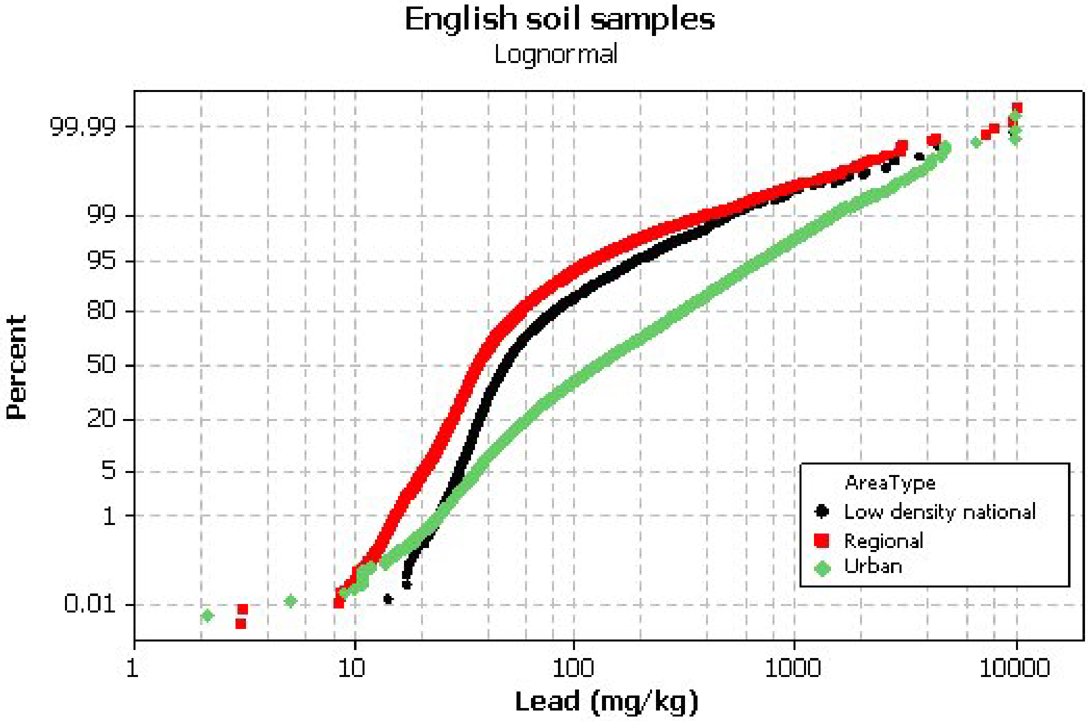


Figure 33: Probability plot of Pb in topsoils. Data categorised according to density of sampling programme and presented with a $\log_{10}$ x-axis to illustrate variation over the full range of concentrations

(a) Pb (mg/kg)	**Number**	**Mean**	**Minimum**	**25th percentile**	**Median**	**75th percentile**	**Maximum**	**Skewness**
All data	42133	113.6	2.12	33.2	47.1	99.5	10196	15

(b) Pb (mg/kg) **Area type**	**Number**	**Mean**	**Minimum**	**25th percentile**	**Median**	**75th percentile**	**Maximum**	**Skewness**
Low density national (NSI XRFS)	4864	79	14.1	38.5	47.9	70.4	9738	29
Regional (G-BASE rural)	23686	57	3.0	29.1	36.0	50.0	10196	34
Urban (G-BASE urban)	13583	225	2.1	68.2	127.7	253.3	10000	9

(c) Pb (mg/kg) **Area name**	**Number**	**Mean**	**Minimum**	**25th percentile**	**Median**	**75th percentile**	**Maximum**	**Skewness**
National (NSI XRFS)	4864	79	14.1	38.5	47.9	70.4	9738	29.0
Eastern England	23222	57	3.0	29.1	36.0	50.0	10196	34.1
Tamar catchment	464	49	14.6	29.2	35.6	49.6	519	5.7
Corby	133	49	17.9	31.5	36.7	44.4	438	5.4
Coventry	390	108	18.7	48.6	72.2	124.6	976	4.0
Derby	275	321	21.1	101.5	155.4	397.2	2052	2.1
Doncaster	279	123	18.7	49.8	77.1	130.7	1072	3.9
Hull	407	227	10.9	66.4	114.1	239.7	2825	4.0
Leicester	652	109	16.3	44.5	64.7	119.9	2053	6.5
Lincoln	215	109	15.8	34.3	53.7	103.4	1364	4.8
London (GLA area)	6494	296	10.8	97.2	180.1	340.3	10000	8.8
Manchester (part of)	300	270	20.1	145.1	213.7	313.4	2687	5.4
Mansfield	257	112	2.1	43.0	75.2	138.4	1285	4.4
Northampton	275	82	25.4	40.3	55.6	95.3	656	3.7
Nottingham	636	145	13.8	64.4	99.5	171.5	976	3.1
Peterborough	272	68	14.3	30.4	39.3	64.3	761	5.0
Scunthorpe	196	88	9.9	31.3	45.0	75.9	3214	11.3
Sheffield	575	239	19.6	113.1	160.8	260.1	4188	7.1
South Essex towns	715	140	18.9	59.6	92.9	163.9	1442	3.9
Stoke-on-Trent	745	173	10.9	65.4	91.7	156.0	4099	7.2
Telford	292	147	20.6	53.7	91.7	168.6	1205	3.6
Wolverhampton	284	243	27.4	104.4	154.5	279.6	2779	4.5
York	191	210	24.5	68.3	104.4	243.6	2338	3.9

Table 10: Statistical summary of Pb in topsoil from the main datasets (a) whole dataset (b) by sampling density and (c) by local areas

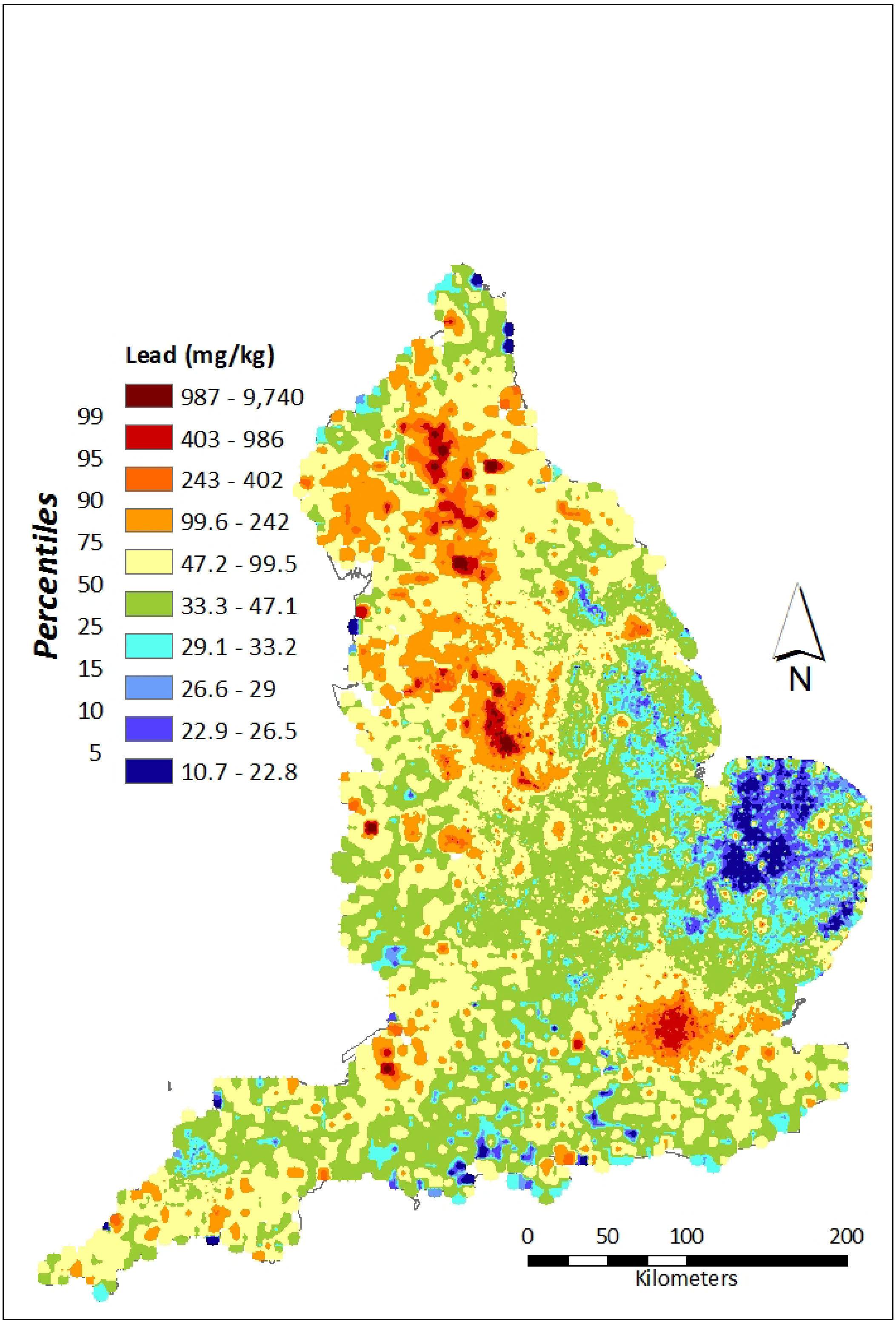


Figure 34: Interpolated map of topsoil Pb. Colour thresholds are designed for highly skewed geochemical data

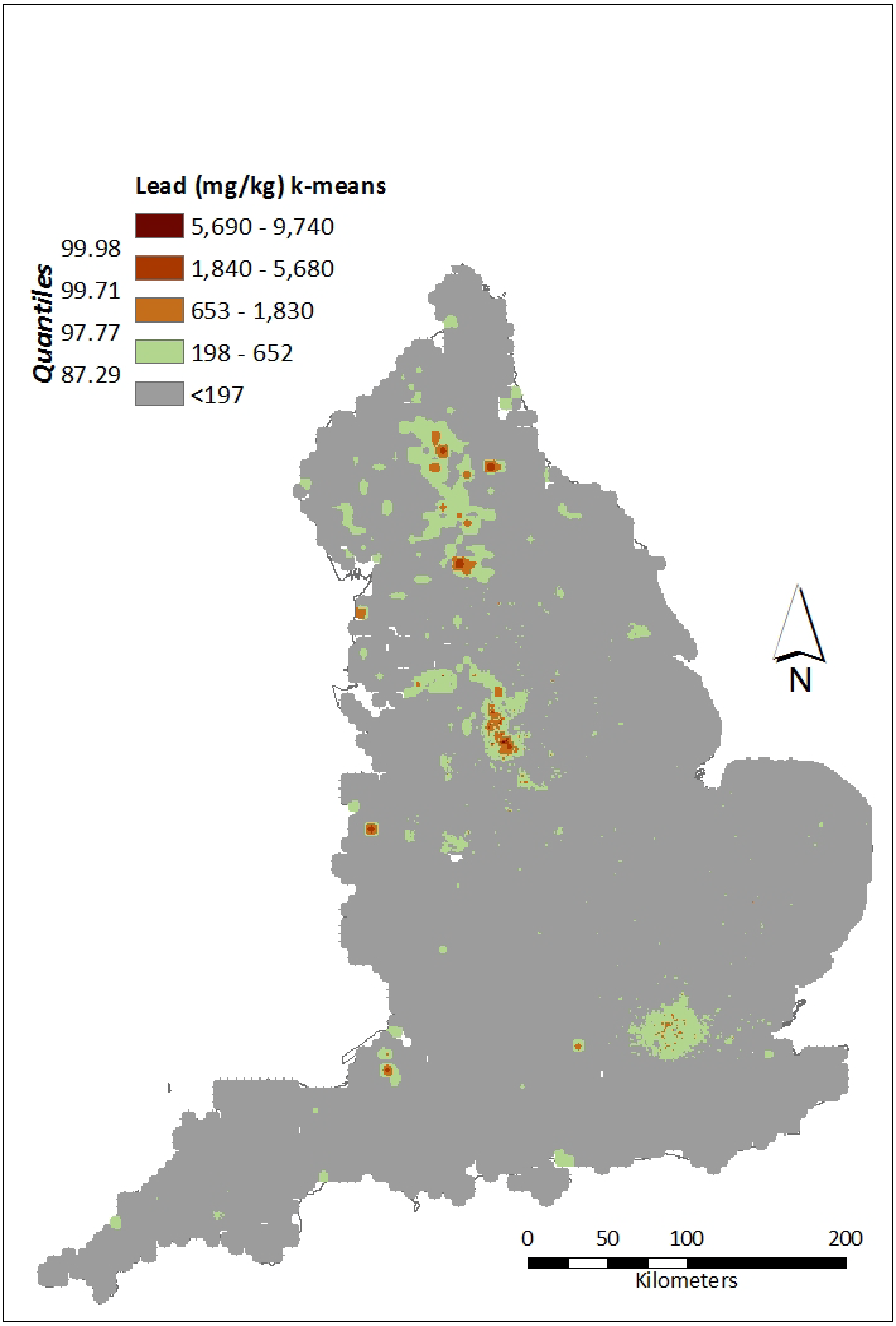


Figure 35: Interpolated map of topsoil Pb classified using k-means cluster analysis thresholds

3.2.2.4 DEFINING LEAD DOMAINS

Mineralisation/mining

There are ore fields formerly mined in England where Pb minerals were the principal economic ores. This is particularly the case in the Peak District, the North Pennines, Lake District, Mendips and Shelve areas, which are recognised to have given rise to areas impacted by diffuse pollution directly with mine waste and via ore-smelting emissions (*e.g*. Colbourn and Thornton, 1978; Fuge *et al*., 1991; Shepherd *et al*., 2009). The summary statistics of the topsoil sample points intersected with the largest non-ferrous metalliferous ore fields are shown in Table 11 and Figure 36. These clearly show how those ore fields with enhanced Pb, due originally to the occurrence of Pb in the deposit minerals, have higher Pb concentrations. It should be noted that there are very low sample numbers for North Devon, Mendips and Shelve; this may be the reason that Shelve is not found to have topsoil Pb concentrations as elevated as may be expected from the presence of Pb mineralisation. Alternatively, this may reflect a far lower environmental dispersion of Pb. This area has not been included in the proposed 'mineralised' domain for Pb. For the ore fields of south-west England and north Devon although Pb is found in the ore field it is in sufficiently low concentrations, or very localised, that it does not manifest in significantly elevated topsoil Pb concentrations.

When the four areas with enhanced topsoil Pb concentrations are combined into a Mineralisation Domain for Pb, and compared with all the other topsoil Pb data (Figure 37 and Table 12), this shows that this captures a data population with a higher concentration distribution. The areas which this relates to are shown in Figure 38.

Area	Samples	Mean	Minimum	25th percentile	Median	75th percentile	Maximum	Skewness
Non-mineralisation	41626	109	2	33	47	98	10000	13.1
1 SW England	153	88	23	46	64	99	519	3.2
2 North Devon	4	42	34	35	40	50	53	0.9
3 Mendips	8	484	103	116	223	591	2043	2.4
4 Shelve	3	94	60	60	110	113	113	-1.7
5 Peak District	224	858	36	205	401	896	10196	4.5
6 Lake District	26	156	64	104	155	194	281	0.3
7 North Pennines	89	344	35	104	201	361	4490	5.4

Table 11: Summary statistics of topsoil Pb (mg/kg) in the main non-ferrous mining areas of England

Area	Samples	Mean	Minimum	25th percentile	Median	75th percentile	Maximum	Skewness
Lead mineralised	347	665	35	151	290	638	10196	5
Not lead mineralised	41786	109	2	33	47	98	10000	13

Table 12: Summary statistics of Pb mineralised topsoil data (mg/kg) compared with all other topsoil data

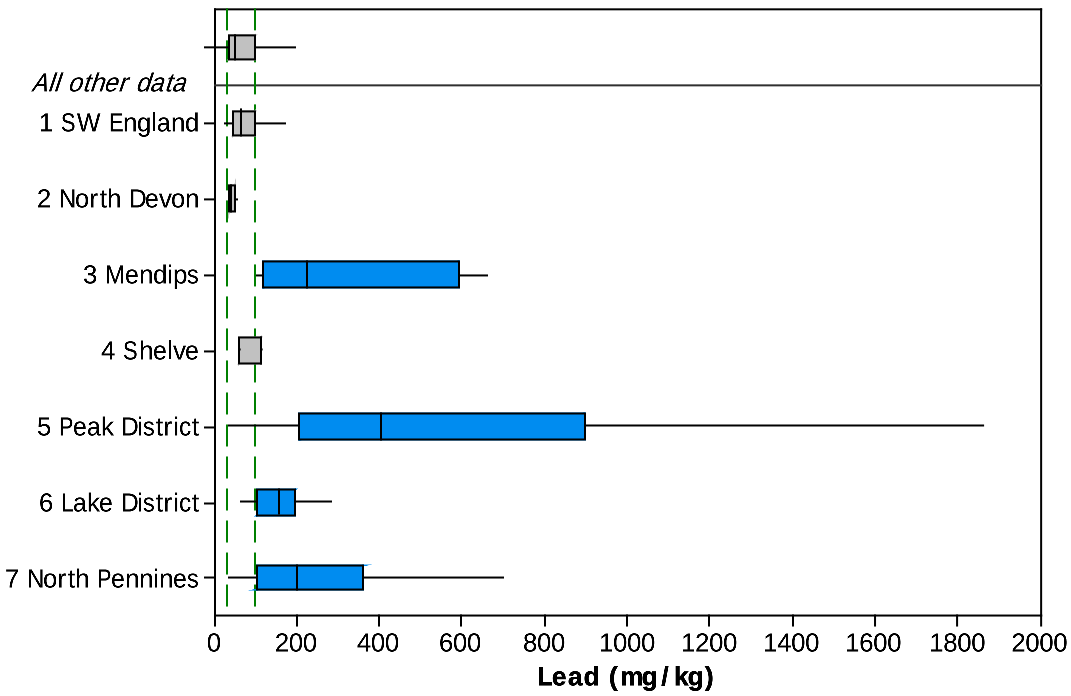


Figure 36: Boxplot of topsoil Pb concentrations over the main English ore fields, compared with the overall dataset. The boxes show the interquartile range (which is projected using dashed lines for 'all other data', to facilitate comparison), with the median shown within the box

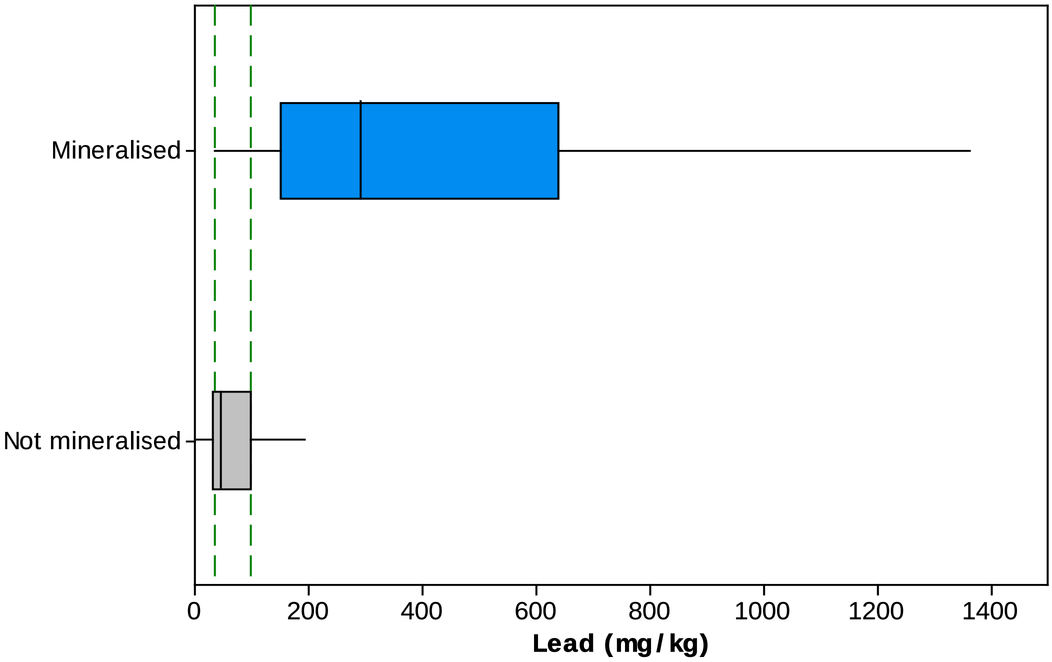


Figure 37: Boxplot of topsoil Pb concentrations for the proposed mineralised domain, in comparison with all other topsoil Pb data. The boxes show the interquartile range (which is projected using dashed lines for non-mineralised, to facilitate comparison), with the median shown within the box

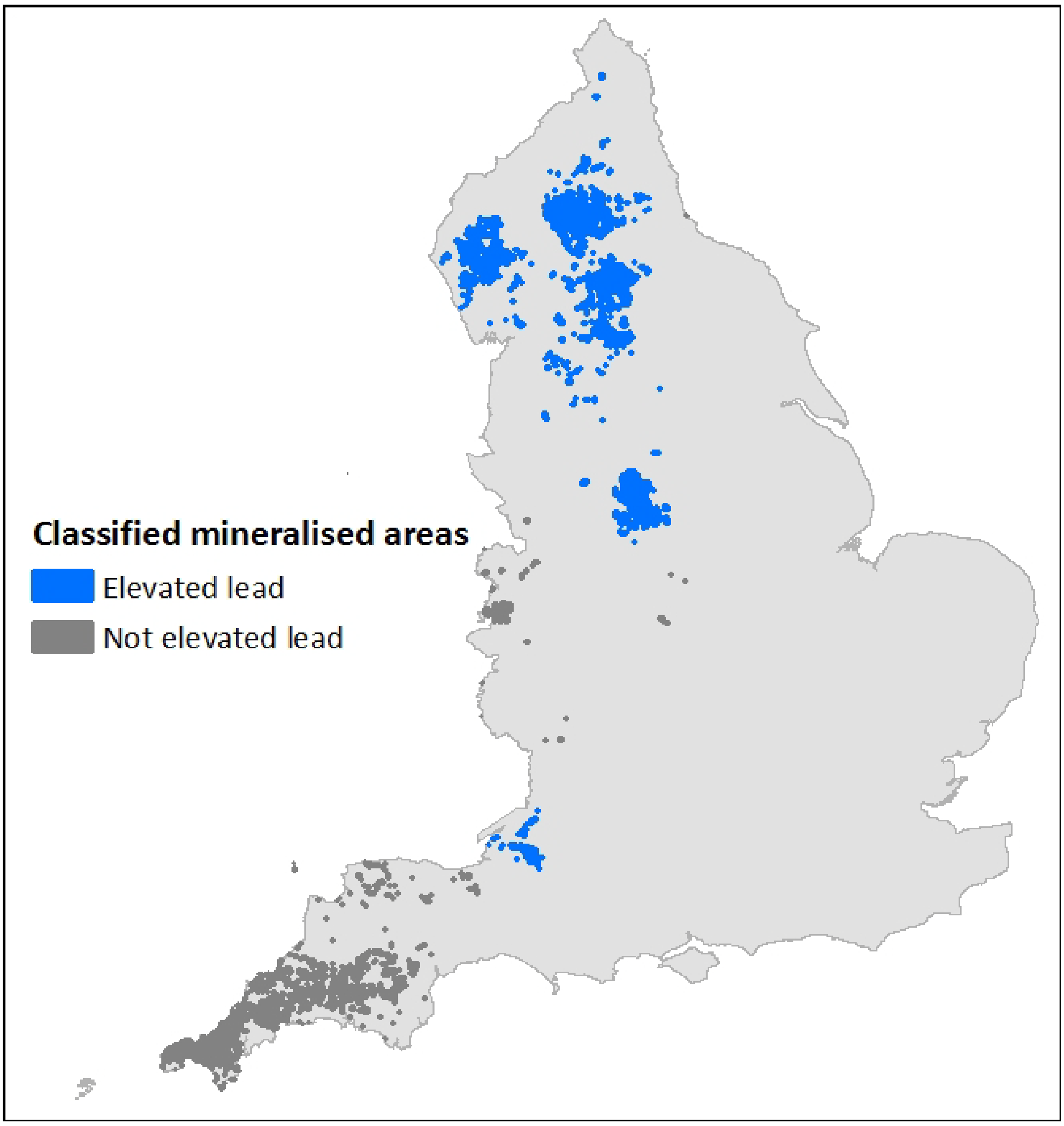


Figure 38: Map of metalliferous mining areas in England, thematically coloured depending on whether topsoil Pb concentrations are typically elevated

Urban

The topsoil samples which were collected as part of the BGS G-BASE urban surveys have substantial variations in concentration between urban areas (Figure 39; Table 10) as well as being substantially elevated in comparison to rural data. These differences between urban areas are likely to be impacted by the age and density of the built environment with additional impact from typical historical industrial activities, as has been found in other surveys of urban areas (Culbard *et al*., 1988; Kelly *et al*., 1996). Derby lies within the catchment of the River Derwent, which drains much of the Peak District ore field, and associated historical metallurgical works whilst the high median concentrations in larger cities such as London, Sheffield, Wolverhampton and Manchester are likely to relate to a combination of emissions from domestic, vehicle and industrial sources.

When the data from the three main datasets are attributed to the GLUD dataset (see Section 3.1.2), the variations in Pb data can be seen to reflect these categories of rural, semi-urban and urban with increasing median concentrations (Table 13; Figure 40), with an urban median of 166 mg/kg compared with a rural median of 35 mg/kg; the rural 75th percentile is lower than the urban 25th percentile. It should also be noted that the skewness of the data reduces in the semi-urban and urban categories. The higher skew in the rural and semi-urban data may reflect the inclusion of data from mineralised areas within these data. This lower skew is also reflected in the shape of the urban data distribution in Figure 40, which is closer to a straight line than the other datasets, indicating it is closer to a log-normal distribution. The areas defined by the urban area of GLUD are shown on the map in Figure 41.

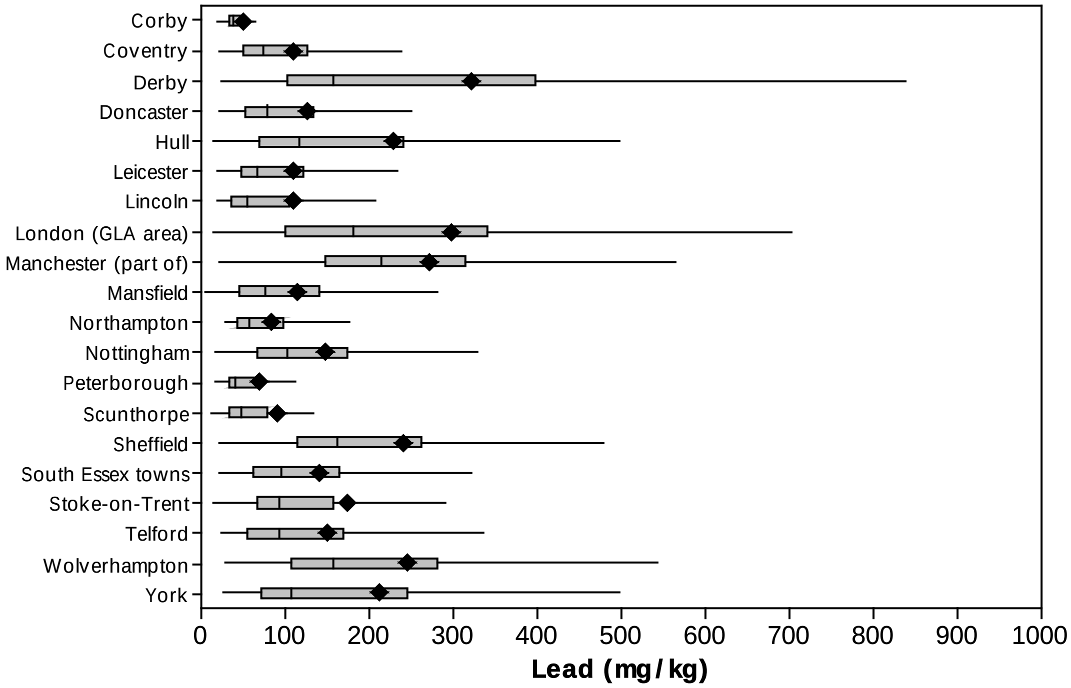


Figure 39: Boxplot of topsoil Pb concentration in G-BASE urban sampling areas

Area	Samples	Mean	Minimum	25th percentile	Median	75th percentile	Maximum	Skewness
Rural	18573	56	8.5	29	35	46	10196	34
Semi-urban	16031	104	3.1	38	57	100	10000	18
Urban	7475	277	2.1	89	166	322	10000	8

Table 13: Summary statistics of topsoil Pb concentrations attributed to GLUD categories

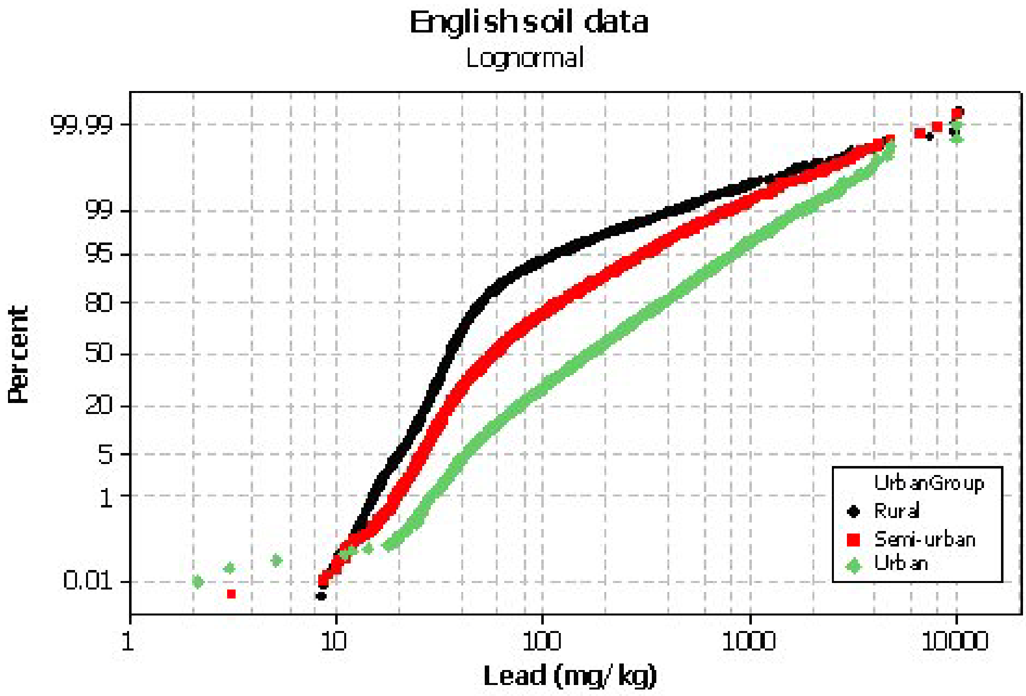


Figure 40: English topsoil Pb data categorised by urbanisation extent defined by GLUD

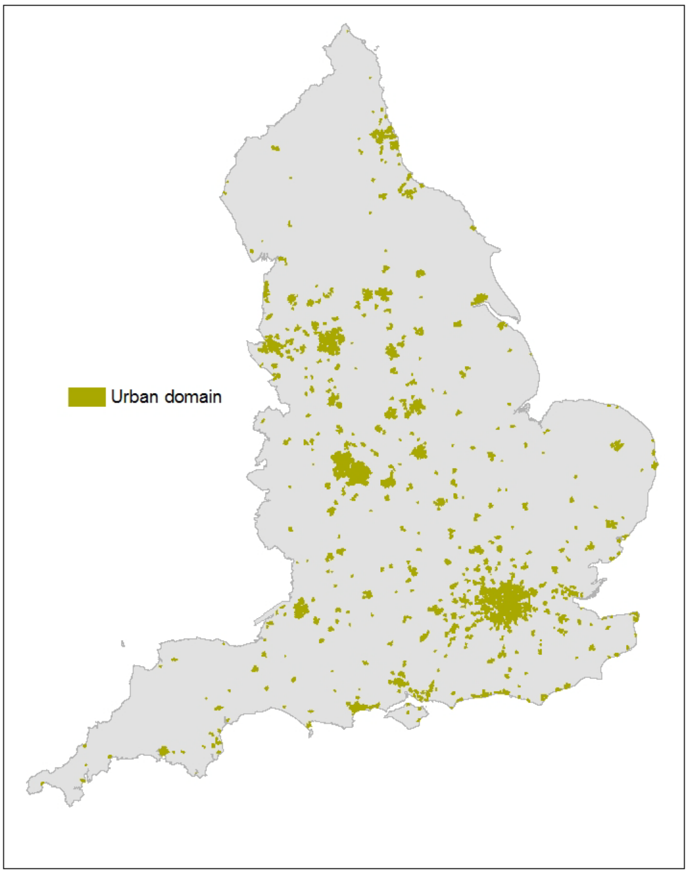


Figure 41: Urban domain derived from GLUD mapping

3.2.2.5 PROPERTIES OF THE DOMAINS

Topsoil data from the main data sets classified as to whether they lie in the urban, mineralised or residual ("principal") proposed domains have median concentrations which differ greatly (Table 14), and overall population properties which also differ greatly (Figure 42 and Figure 43). All the data are still positively skewed, but a relatively low skew is found for the defined domains of 'mineralised' and 'urban', which are both <1 when the data are log-transformed. The Principal Domain looks bimodal in distribution, with a break in slope at *c*.60 mg/kg, and 5% of the data exceeding the original minimum k-cluster means analysis threshold of 197 mg/kg. Many of the outliers lie close to mineralised areas, or within large open-spaces in the London area. Work Package 3 will examine these geographical and numerical outliers more closely to investigate whether they should be included in the other domains as appropriate.

Area	Samples	Mean	Minimum	25th percentile	Median	75th percentile	Maximum	Skewness
Principal	34257	72	3	31	41	66	10000	22
Mineralised	347	665	35	151	290	638	10196	5
Urban	7529	276	2	89	166	322	10000	8

Table 14: Summary statistics for domain classified topsoil Pb data (mg/kg)

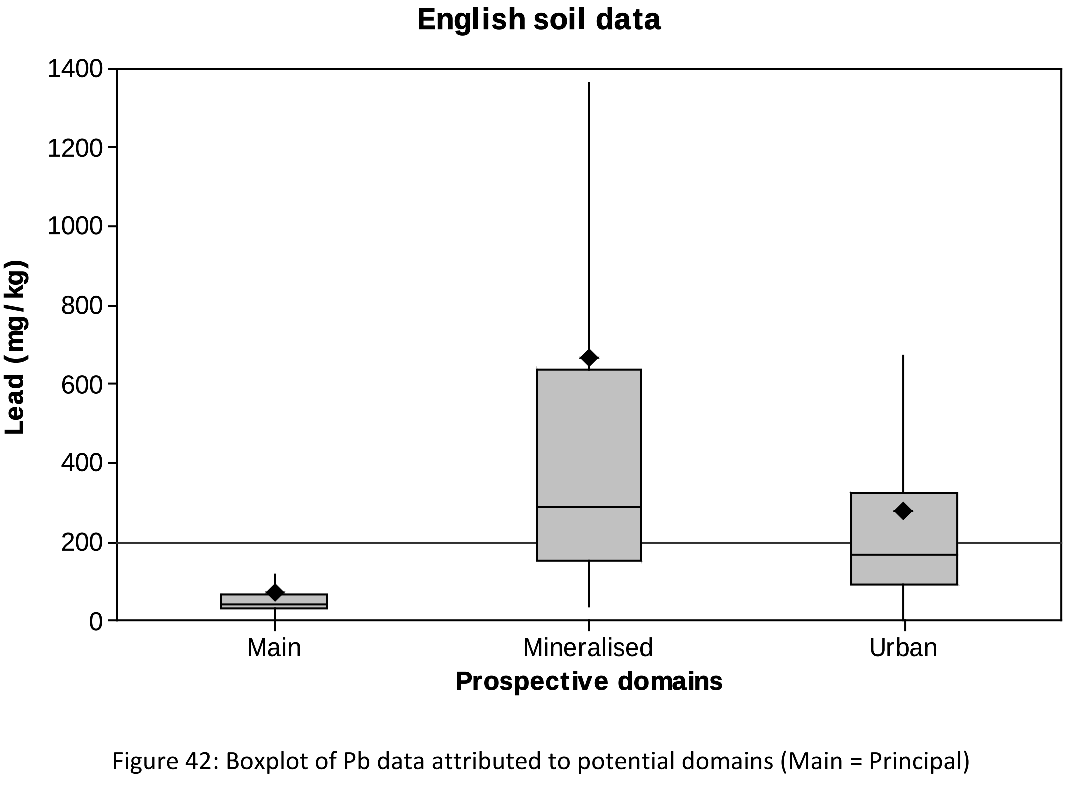


Figure 42: Boxplot of Pb data attributed to potential domains (Main = Principal)

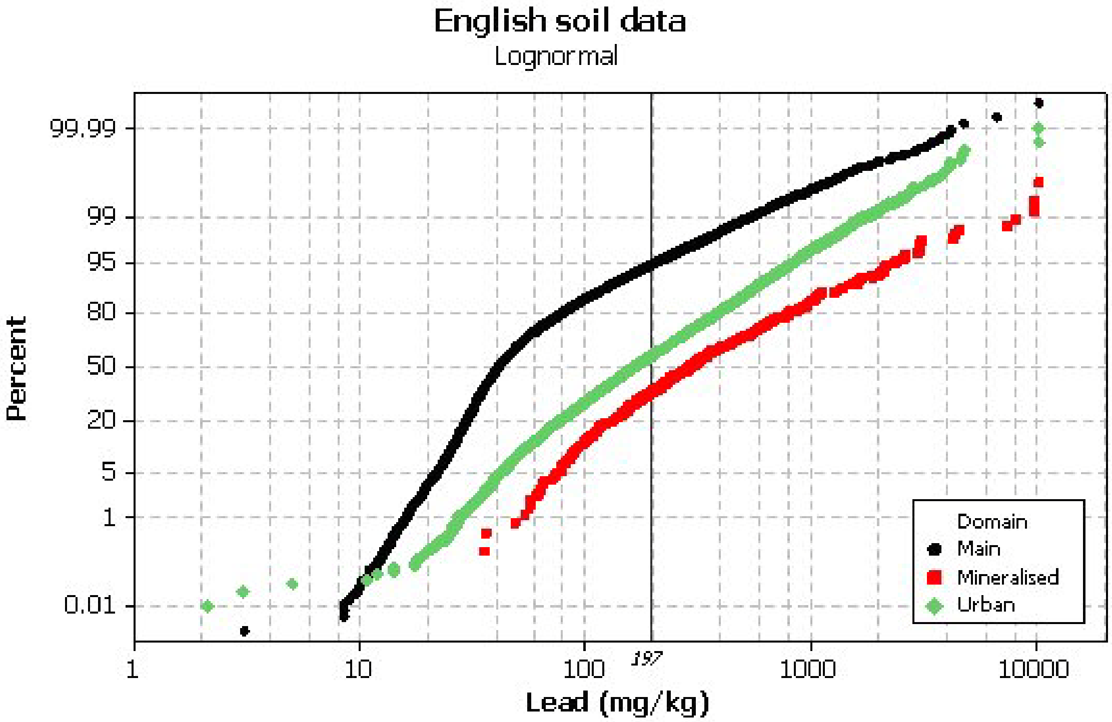


Figure 43: Probability plot of topsoil Pb data categorised by potential domains. The k-means threshold of 197 mg/kg is shown with a vertical line. (Main = Principal)

3.2.2.6 COMPARISON OF ANALYTICAL METHODS

The analysis of the NSI topsoil samples by ICP-AES following *aqua regia* digest allows the comparison of these methods on a one of the key data sets NSI(XRFS) used in this project. Comparisons are also made using the English GEMAS samples, as was also undertaken for As (Section 3.2.1.6). The FOREGS survey topsoil samples only had one analytical method for Pb, so no comparison can be made.

NSI acid-digest and XRFS analysis comparison

The analysis through the original programme of Pb concentrations measured by ICP-AES following *aqua regia* digest (McGrath and Loveland, 1992) allows comparison with the recent reanalysis by XRFS (Rawlins *et al.*, 2012). These data are shown in Figure 44 and it can be seen that there is a positive linear correlation between the datasets, which has the following regression:

$$[\mathrm{Pb}_{\mathrm{XRFS}}] = 30.7 + (0.668 \times [\mathrm{Pb}_{\textit{aqua-regia}}])\ (n=3956;\ R^2 = 95\%;\ \text{and}\ P<0.005).$$

The systematic offset between the data sets, 31 mg/kg, is large in relation to the concentrations (and in comparison with the GEMAS data, below). This may, however, in part reflect a relatively high detection limit in the ICP data (23 mg/kg) compared to the XRFS data (1.2 mg/kg) and the natural abundance of Pb (Table 10).

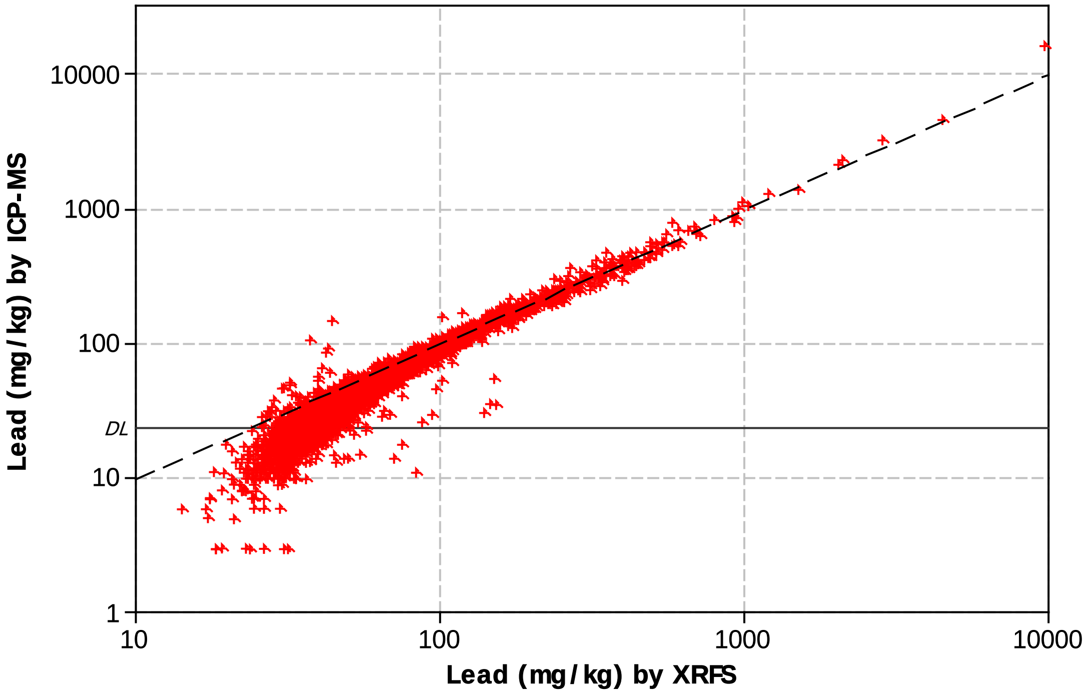


Figure 44: Comparison of topsoil Pb concentrations in NSI samples by XRFS and aqua regia acid digest (DL detection limit)

GEMAS acid digest and XRFS analysis comparison

The GEMAS data has both land use and analytical method comparisons as used for As, which are shown in Figure 45. These show that where the same sample is used for more than one analysis there is a very close linear correlation between Pb concentration data. For the arable topsoils the result is:

$$[\mathrm{Pb}_{\mathrm{XRFS}}] = 4.42 + (0.982 \times [\mathrm{Pb}_{\textit{aqua regia}}])\ (n=66;\ R^2 = 99\%;\ \text{and } P<0.05),$$

and for the pasture topsoils the result is:

$$[\mathrm{Pb}_{\mathrm{XRFS}}] = 4.67 + (0.953 \times [\mathrm{Pb}_{\textit{aqua regia}}])\ (n=65;\ R^2 = 93\%;\ \text{and } P<0.05).$$

It can be seen that these relationships are very similar and have a far smaller systematic offset in the acid digest data in relation to the XRFS data than is observed in the NSI comparison. This level of systematic offset will have a trivial impact on the setting of typical background concentrations.

When the sources of variance in the data are explored using ANOVA analysis of log(10) transformed concentration, the geochemical variation between sites is found to account for 98% of the variance, confirming the relatively minor concentration variations introduced by the land use and analytical method.

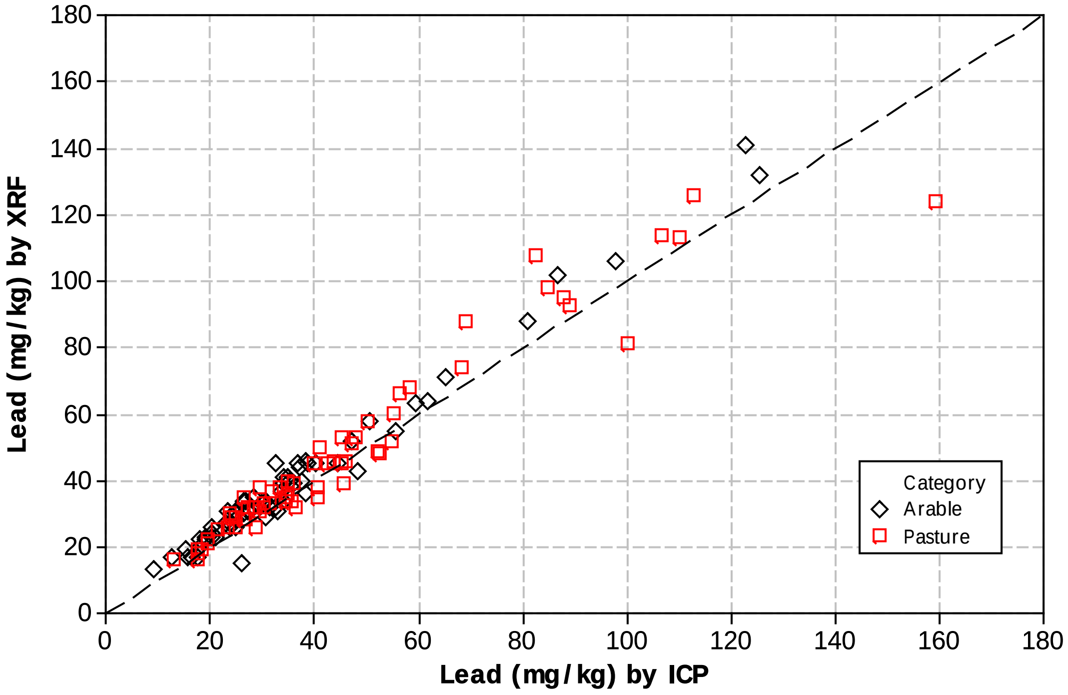


Figure 45: Comparison of GEMAS topsoil Pb data by analytical method and land use (one data point of >1300 mg/kg removed for clarity of comparison)

3.2.3 Benzo[*a*]pyrene

3.2.3.1 INTRODUCTION

Benzo[*a*]pyrene (BaP) is one of the polycyclic aromatic hydrocarbons (PAHs), and a persistent organic pollutant (POP). It is one of four priority PAHs prioritised in environmental monitoring (Creaser *et al.*, 2007). PAHs are produced through the incomplete combustion of organic matter: vegetation, seasoned wood, charcoal, coal, oil, gas and derived products of these (Creaser *et al.*, 2007; Wild and Jones, 1995). Whilst natural fires, or the organic matter itself, provide a natural baseline concentration in soils, and there may be some *in situ* production of PAHs by plants (Wilcke, 2000), these sources are vastly outweighed by anthropogenic production. The diffuse emissions of BaP through domestic activities is both significant and widespread (Wild and Jones, 1995), and coal/wood fires are thought to be the biggest UK source of BaP. It has been postulated that at the European scale differences in total measured PAH (ΣPAHs) (including BaP) between urban areas may partially reflect climate (Morillo *et al.*, 2007) – and thus duration of annual heating and energy requirements; this is also suggested by a weak but significant statistical relationship in ΣPAH in the Countryside Survey data within the UK (Heywood *et al.*, 2006). The data of Cousins *et al.* (1997) does suggest that very low rates of deposition must account for low concentrations in the peats of the Scottish Highlands, which is in contrast to higher concentrations observed in rural peat deposits in Wales, which are closer to more emission sources (Jones *et al.*, 1989b). As is implicitly expected in the sources and persistence of soil BaP, the studies using the Rothamsted archive (Jones *et al.*, 1989a) showed that soil ΣPAH (including BaP) concentrations have increased with time.

Soils have been found to comprise the largest environmental reservoir of this POP in the UK, and losses from soil can be through leaching, degradation or vegetation off-take (Wild and Jones, 1995). BaP is one of the heavier PAHs (molecular weight 252.094), with a very low volatility (boiling point 495°C), which enhances retention in soils, and expected half-life in soils of *c*.8 years (Creaser *et al*., 2007). Degradation rates are typically slow in comparison to rates of deposition, and sorption to organic matter limits leaching and plant uptake (Morillo *et al*., 2007; Wild and Jones, 1995). Organic matter can be both natural organic matter and 'black carbon' from combustion (Morillo *et al*., 2007; Nam *et al*., 2008). There have been some conflicting reports of the relative role of soil properties on PAH/BaP retention. In Wales Jones *et al*. (1989b) found that soil organic matter was significant in ΣPAH concentration, and that increased soil clay content influenced BaP retention; they found urban/industrial land use and peaty/forest soils to have the highest ΣPAH (and generally BaP) concentrations. Cousins *et al*. (1997) noted that the most remote Scottish highlands samples had the lowest PAH (and BaP) concentrations, despite being peats, suggesting that distance from sources was most important in that case.

It can therefore be anticipated that proximity to built-up areas is the likely primary control on BaP concentrations in soil, with a possible secondary influence by soil organic carbon (or matter) concentrations. Jones *et al*. suggest that where ΣPAH >600 µg/kg these are typically urban/industrial proximal soils, whilst lower than that concentration is typically 'rural'; similar calculations for BaP only have not been found during this data gathering phase. An assessment of whether natural PAH sources could contribute significant BaP to soils from parent material weathering is also considered below.

In the absence of large spatially contiguous datasets, such as available for most inorganic contaminants, these considerations have dictated the exploratory data analysis methods employed.

3.2.3.2 DATA AVAILABILITY

The complexity of sample collection, preservation and analysis – resulting in higher costs – associated with organic compounds in soil is much greater than for commonly studied elements such as Pb and there is thus much less data, and little is routinely collected. Also, because the majority of data are from peer-reviewed papers rather than systematic regional/national surveys there is a tendency for those data to be published with summary statistics only. Exceptions to this for British soil samples are Jones *et al*. (1989), and Cousins *et al*. (1997), and these data are used for further analyses in this study; there is sufficiently little data for any part of the UK, that all available data have been used in parts of this analysis, not just that for England. Key BaP data sets for topsoils are summarised below, and shown in Figure 46.

Soil Herbage Survey (UKSHS)

> These data are systematically distributed around the UK (British sample locations are shown in Figure 46), with a common sampling strategy and analytical methodology. Data were collected from settings considered to be rural, urban or industrial. This project is seeking to establish normal background concentrations, so the industrial data have not been used in the assembled data sets. Site locations are downgraded/rounded to ±10 km which places additional uncertainty when capturing point information within polygons in a GIS.

Countryside Survey (CS:2000)

> Organic analyses were undertaken on the 1998 sampling round only (Black *et al*., 2002; Emmett *et al*., 2010) and again have an internally consistent methodology of sampling and analysis. An analysis of the PAH data has been published by Heywood et al (2006); the analyses were preferentially targeted at brown or gley soils, with some additional

peat samples. All samples were collected from sites considered to be rural. As with the UKSHS data, grid references have been downgraded/rounded to ±10km.

Jones et al. and Cousins et al.

These data were both produced in the same laboratory, thus some degree of internal consistency in the data can be expected. Both these peer-reviewed papers have fully published contextual (*e.g*. land use, soil type), location (±100 m) and PAH analytical data for each sample collected. Cousins *et al*. (1997) only collected samples from areas they considered to be rural; Jones *et al*. (1989b) collected a proportion of samples from urban/industrial south Wales valley settings and Cardiff and environs.

3.2.3.3 SUMMARY STATISTICS FROM PEER-REVIEW PUBLICATIONS

Table 15 shows a summary of data from Britain found in the peer-reviewed literature. The data are generally very skewed, with mean concentrations greater than the median, where both are presented. The overall range given is <0.2 - >31,000 µg/kg BaP. For rural areas the median concentrations have a range of 16 - 67 µg/kg, whilst urban median concentrations are up to 714 µg/kg. It is thus immediately apparent from these studies that urban soils typically have higher BaP concentrations than those from rural areas, as would be anticipated.

UK soils are also relatively elevated in relation to those European areas with fewer potential sources, particularly as reflected in lower population density, such as rural Poland and Norway (Jensen *et al*., 2007; Maliszewska-Kordybach *et al*., 2009; Nam *et al*., 2008), and in relation to the expected natural baseline ΣPAH (plant produced, natural fires) which has been estimated as *c*.1-10 µg/kg (Wilcke, 2000), which BaP alone far exceeds in almost all UK soils measured.

3.2.3.4 EXPLORATORY DATA ANALYSIS

The data used in this work are those which could be identified that are available in the published literature, through publication of all results or through national surveys which make their data available.

There are a total of 407 analyses available in Britain from the four key datasets (Table 16a), which are spatially distributed as shown in Figure 46. These data are very skewed, with mean concentration greater than that of the 75th percentile. When data from England, Scotland and Wales are compared it can be seen that the English data is generally higher in concentration (Table 16b). The rural studies of Cousins *et al*. (1997) and the Countryside Survey, have very similar data distributions until the 75th percentile is reached (Table 16c). Notwithstanding the inclusion of *c*.25% urban sites in the data of Jones *et al*. (1989b) from Wales, it can be seen that these concentrations of BaP are much lower than those found in the other studies. The highest summary concentrations are found in the UKSHS data which is a combination of both rural and urban data.

Figure 47 shows how concentrations vary around Britain on a percentile scale. This shows generally higher concentrations of central and southern England, the north-east England conurbation, south Wales and the Edinburgh-Glasgow region. It can also be seen that the range from 5th percentile (6.5 µg/kg) and 95th percentile (577 µg/kg) encompasses almost a hundred-fold increase in concentration. When a k-means cluster analysis is performed on these data a lowest threshold of 96 µg/kg (*c.*70th percentile) is found (Figure 48). This particularly enhances the visibility of the higher concentration samples around large and small urban areas (*e.g.* Merseyside and Norwich). There are noticeably few data points currently available for the London area – the largest urban area in England. Closer geographical examination of many of the data is not appropriate, since the majority are from data with downgraded grid coordinates.

Location type	Place	n	Year collected	BaP [central tendency] (µg/kg)	BaP range (µg/kg)	Reference
Urban	Glasgow	20	2004	971 [mean] 657 [median]	132-3627 992 [sd]	(Morillo *et al*., 2007)
Urban	London	3	1999	1440 [geometric mean] 1652 [median]	*not given*	(Saltiene *et al*., 2002)
Urban	Salford	8	Not specified, ≤1984	731 [mean] 590 [median]	159-2293	(Butler et al., 1984)
Remote, rural & urban	Wales	49$	Not specified, <1989	138 [mean] 16 [median]	3.5-3700	(Jones *et al*., 1989b)
Rural	Hertfordshire	9	1846-1986	n/a – time series data	6.7-120	(Jones *et al*., 1989a)
Rural	UK-wide	45$	1993	134 [mean] 40 [median]	<0.2-1200	(Cousins *et al*., 1997)
Rural	UK-wide	201	Not specified, ≤2006	88.3 [mean] 33.7 [median]	1.11-1440 163 [sd]	(Heywood *et al*., 2006)
Rural	Western UK	27	Not specified, ≤2008	46 [geometric mean]	1.8-1600	(Nam *et al*., 2008)
Rural	UK	366		215 [mean] 46.2 [median]	0.9-24800	(Creaser *et al*., 2007)
Urban	England	42		1590 [mean] 714 [median]	59-31200	(Creaser *et al*., 2007)
Rural	England	183		154 [mean] 67.2 [median]	0.9-1540	(Creaser *et al*., 2007)
Rural		5		4809** [mean]		(Wild and Jones, 1995)
Non-UK European examples						
Rural	Poland	50	Not specified, ≤2008	29 [mean] 20 [geometric mean]	5-116 26 [sd] 37 [75th percentile]	(Maliszewska-Kordybach *et al*., 2009)
Rural	Poland	39	Not specified, ≤2009	22 [mean]	2-248 [95% range]	(Barbara, 1996)
Rural	Norway	26	Not specified, ≤2008	5.3 [geometric mean]	0.5-86	(Nam *et al*., 2008)
Rural-urban	Norway	40$	Not specified, ≤2007	20 [median]	<10-160	(Jensen *et al*., 2007)

$ - reference provides all results of analyses. **Wild & Jones (1995) do not provide any new data; they cite data from Jones *et al* (1989b), but it is not clear from either publication which 40 (Wild and Jones) of the 49 (Jones *et al)* sites they use to derive their summary concentration statistics for urban and rural locations.

Table 15: Literature BaP data summary

When the data for England are categorised according to whether they were reported as rural or urban there is a ten-fold difference in the median with a systematic overall difference, and a reduction in skew of both data sets (Table 16d, Figure 49 and Figure 50). This is consistent with what would be expected from known sources of BaP. When these data are categorised with the GLUD dataset *(i.e*. presented differently to the original publications), for England only, a gradation of increasing concentrations can be seen from rural, semi-urban and urban, as shown in Figure 51. The outliers in Figure 51 (and Table 16e) also show that high concentrations can be found in samples which are associated with 'rural' or 'semi-urban' classes.

Figure 52 appears to confirm a trend for more organic-rich soils to contain higher BaP concentrations (all four studies have also published organic matter data). However, it is also clear that organic matter itself is not a predictor of BaP concentrations, with high concentrations occurring in soils where there are low organic matter concentrations. Figure 53 shows the distribution of 'peat' or 'peat mud' SPMM regions. These may warrant further investigation as a potential domain for BaP (or other POPs), although there are currently only 2 samples from studies with good grid coordinates provided (Cousins *et al*., 1997; Jones *et al*., 1989b) which fall where 'peat' is a mapped parent material.

3.2.3.5 *IN SITU* GEOLOGICAL BAP SOURCES TO SOILS

Geological contribution of PAHs (including BaP) to soils through natural weathering of formations rich in organic matter has not yet been widely considered in the European literature. In England there are marine clays, such as the Kimmeridge Clay and Oxford Clay (both Jurassic; Figure 54) which are widely recognised to be rich in organic matter, with the Kimmeridge Clay acting as a petroleum source rock in Dorset. A small (unpublished) study was undertaken in BGS to look at the organic composition of soils overlying these formations in two soil profiles. The data are presented in Table 17, and study locations are shown in Figure 54. It can be seen that the highest concentrations were found at the surface, also associated with increased organic matter concentrations (Figure 55). There is a lack of systematic increase in BaP with depth in these samples. The Clavell sample shows an increase at >2 m depth to concentrations comparable with the surface soil samples; these samples also have some of the highest concentrations measured (>5 µg/kg). None of the other samples show this increase in the deepest samples. Additionally, a natural source for the outcrop sample of Kimmeridge Clay with a BaP concentration of 1700 µg/kg at Smedmore cannot be ruled out. Information on the soil horizon represented by the depth samples was not available to this study; this could affect the interpretation of the samples in relation to relative proportion of less weathered parent material. All the cores have the highest concentration at surface, and as Figure 55 shows there is also a direct correlation with soil total organic carbon (TOC), which suggests that both atmospheric deposition and the ability of the soil to retain BaP may be the major contributing factors to these data. It would seem possible that outcropping coal seams could be another natural source of PAHs to soil. A similar study to that above, in soils formed over high-carbon Lower Cambrian shales in Norway did not find any relationship to soil PAH concentrations (Jensen *et al*., 2007).

Area	n	Mean	Standard Deviation	Minimum*	25th percentile	Median	75th percentile	Maximum	Skewness
(a)									
All	407	139	325	<DL	14.0	37.7	120	3700	6.4
(b)									
Not known**	32	27	35	<DL	6.7	15.5	34.1	159	2.4
England	178	174	269	<DL	27.3	70.8	178.0	1560	3.0
Scotland	103	100	340	0.1	11.0	24.3	71.6	3140	7.6
Wales	94	153	434	<DL	9.3	30.7	83.6	3700	6.4
(c)									
Cousins *et al* 1997	46	131	219	<DL	13.8	39.5	145	1200	3.1
CS 2000	179	79	147	<DL	12.5	31.4	92	1443	5.6
EA SHS	133	222	404	2.3	23.6	61.6	207	3140	3.9
Jones et al 1989	49	138	534	3.0	9.0	16.0	52	3700	6.5
(d)									
Rural	165	136	205	<DL	24.7	56.8	149	1443	3.6
Urban	13	666	454	72	303	520	1005	1560	0.7
(e)									
Rural	97	110	171	<DL	22.4	47.1	131	1200	3.7
Semi-urban	70	184	238	<DL	29.3	110	206	1150	2.4
Urban	11	684	528	72	260	510	1320	1560	0.7

All concentrations in µg/kg

*<DL – below the detection limit – this is variable and not specified for BaP in every publication, so has not been quantified in the table. ** There are errors in some SHS grid coordinates which will be resolved (corrected or deleted) for final data analysis and publication.

Table 16: Summary statistics for BaP data (a) whole dataset (b) by country (c) by publication (d) English soils categorised as ‘urban’ or ‘rural’ in the publication from which they were taken (e) English soils categorised using the GLUD urbanisation index (UI)

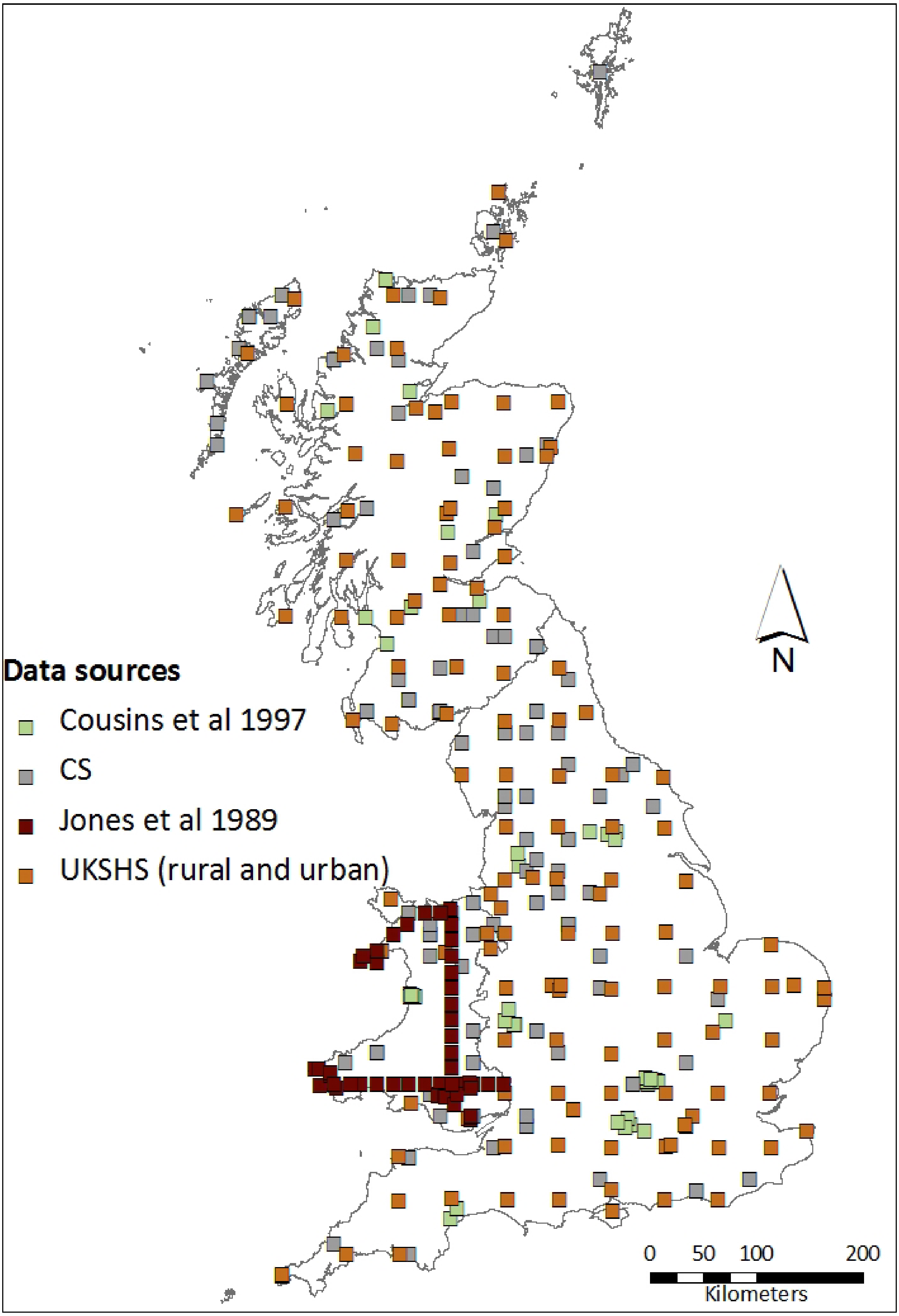


Figure 46: Location of sample sites in Britain with BaP data used in this report

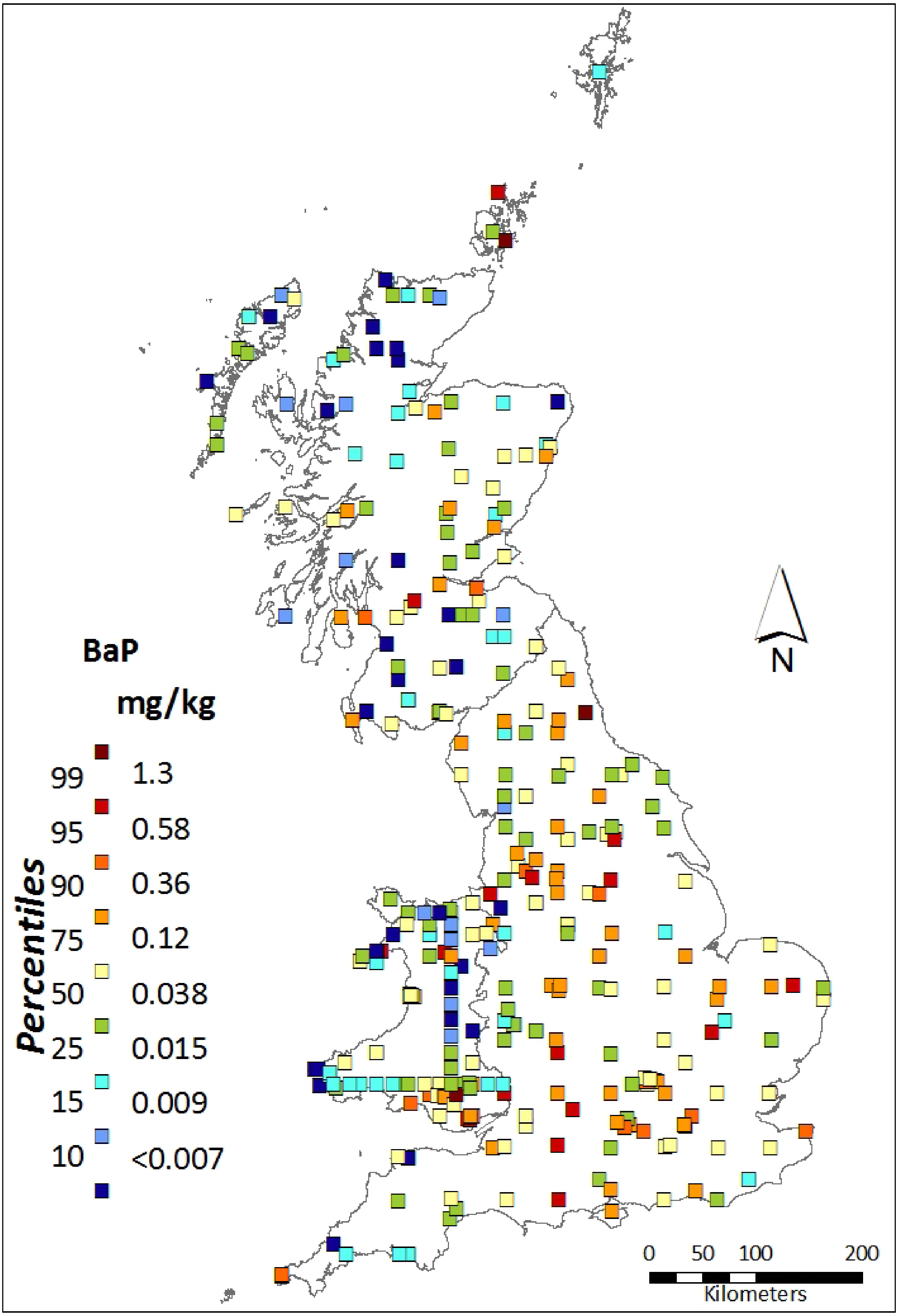


Figure 47: BaP soil concentrations in Britain. (Data sources as shown in Figure 46)

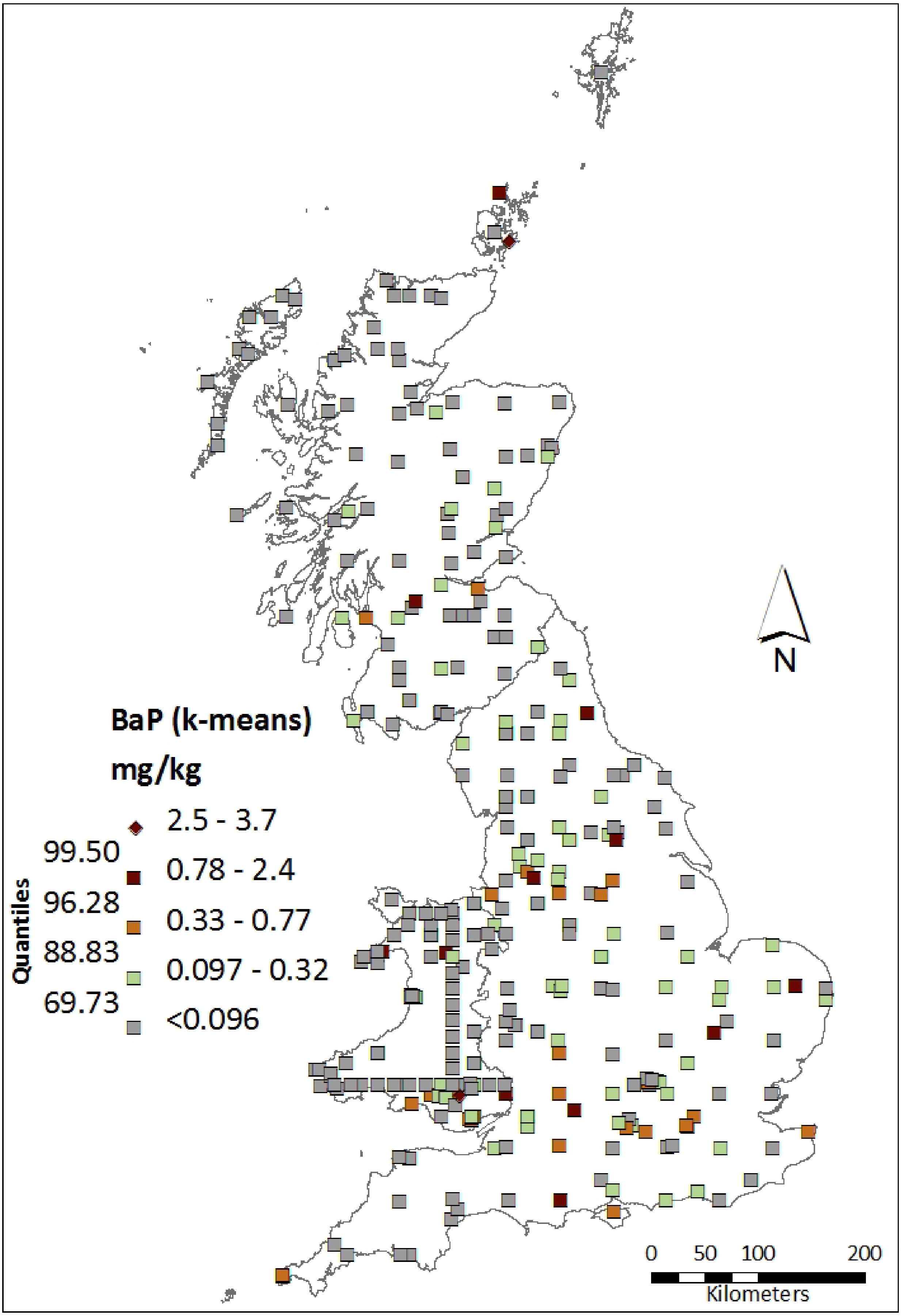


Figure 48: BaP concentration in soils classified by k-means cluster approach

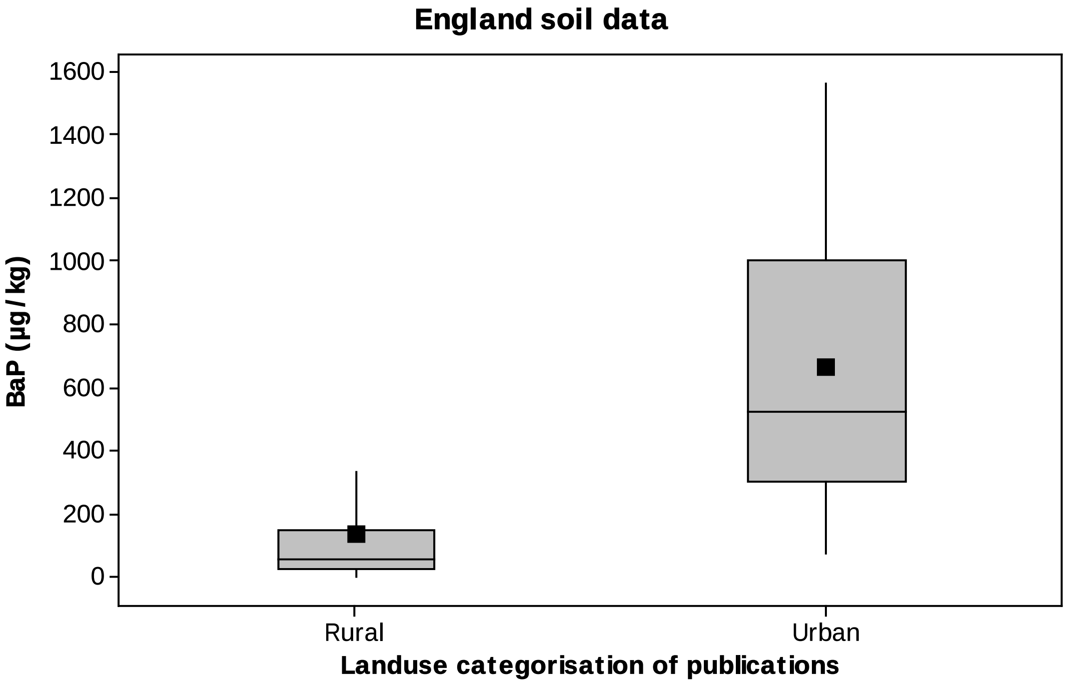


Figure 49: Boxplot of BaP in English soils. Data selected from CS:2000, UKSHS, Cousins *et al.* 1989

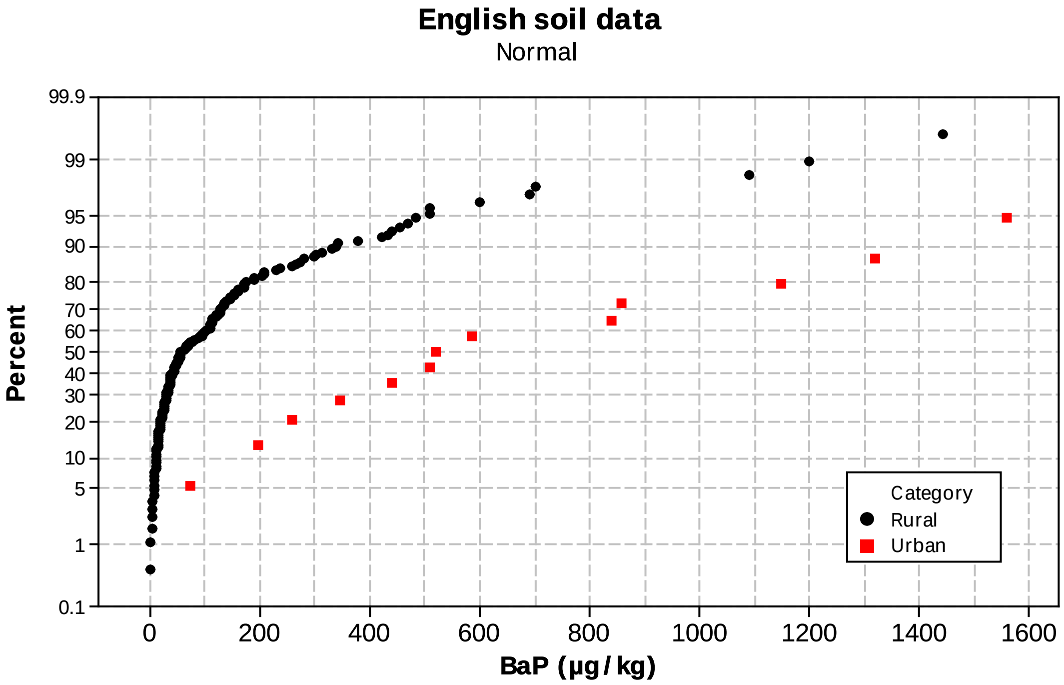


Figure 50: Probability plot of BaP in English soils. Data selected from CS:2000, UKSHS and Cousins *et al*. 1989

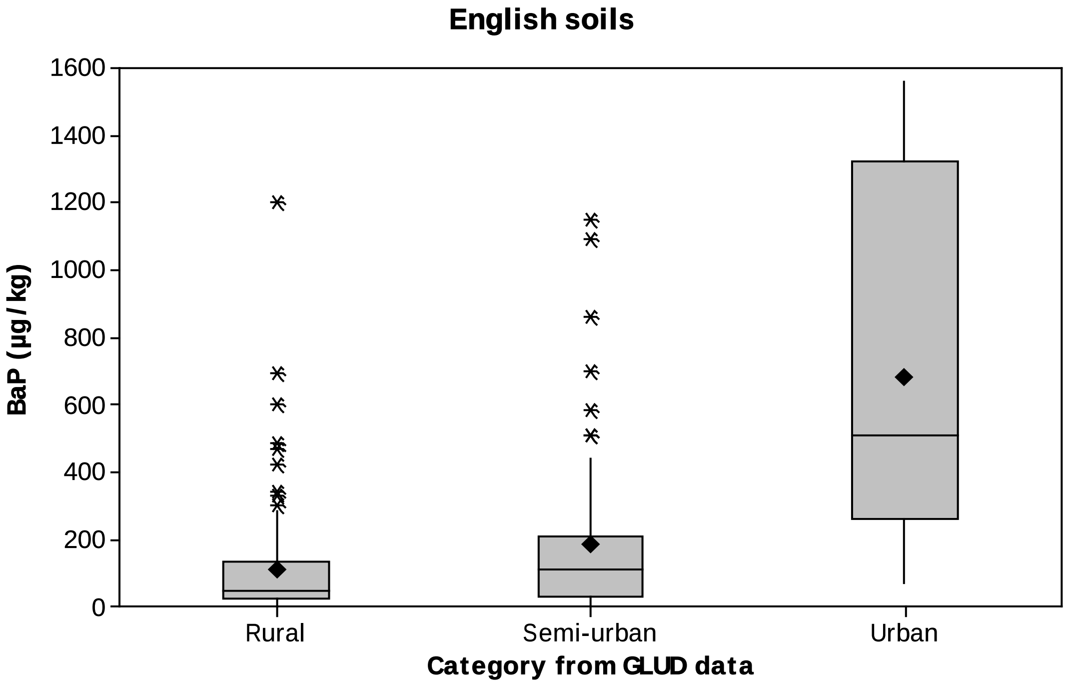


Figure 51: BaP in English soils categorised by the GLUD data. Data from CS:2000, UKSHS and Cousins *et al*. 1989

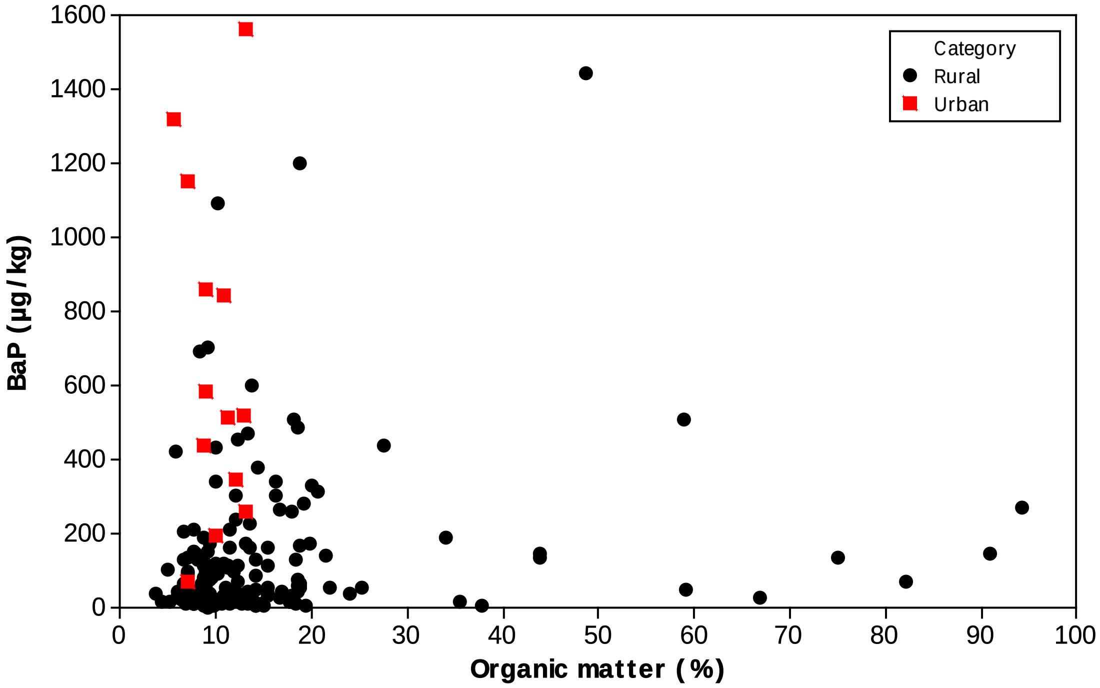


Figure 52: BaP and soil organic matter concentrations in English soils

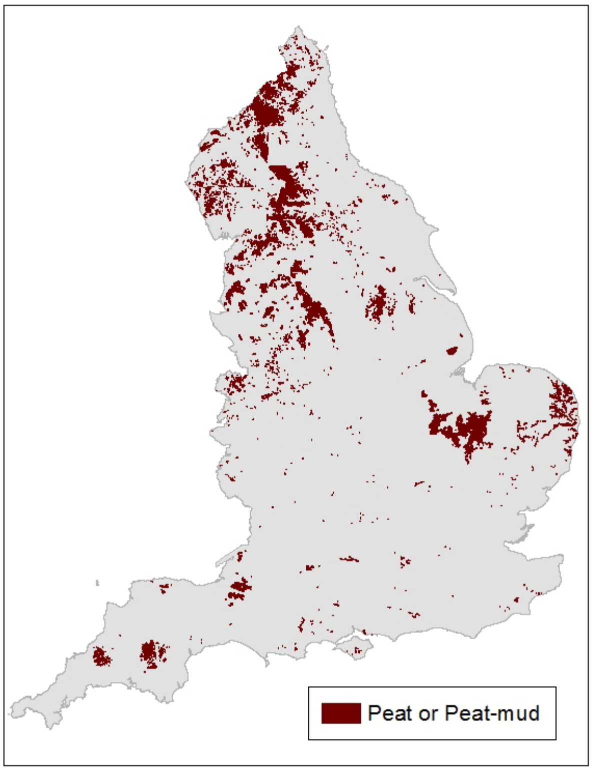


Figure 53: Peat, or 'peat-mud', soil parent material areas, from the BGS Soil-Parent Material Model

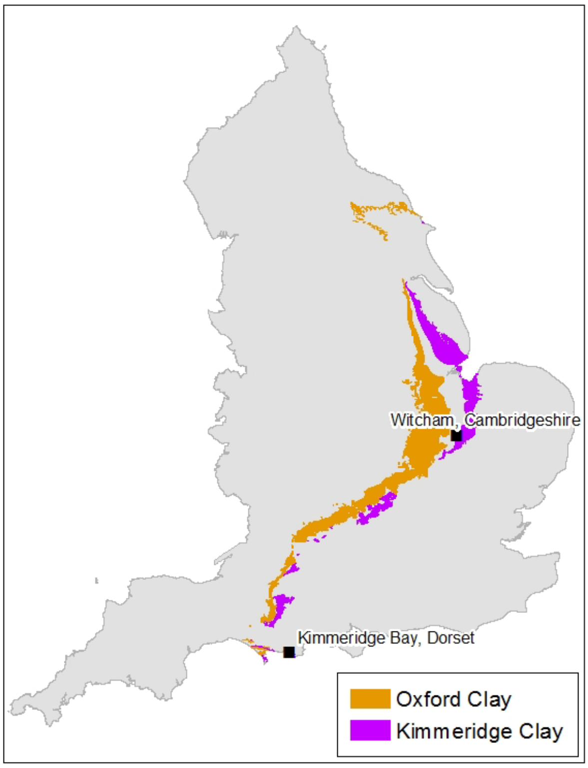


Figure 54: Oxford Clay and Kimmeridge Clay outcrop areas and the location of geological BaP study sites

Site	Clavell	Smedmore	Smedmore2	Witcham 1	Witcham 2	Knowle	Elton 1	Elton 2
Depth (m)	Kimmeridge Clay					London Clay	Oxford Clay	
	BaP (µg/kg)							
0	192*	1701**						
0.025				2471	424		45.0	13.9
0.05	5.6							
0.35	3.7			110.4	10.1	0.1	2.4	1.8
0.45		0.4						
0.55			0.4					
0.65	2.0			1.6	3.0	0.2	0.1	0.0
0.95	0.5			0.2	3.4	0.1	0.1	0.1
1.25	0.1			0.4	0.8	0.1	0.5	0.1
1.55	0.3			0.1	0.2	0.1	1.1	1.4
1.85	0.4			0.2	0.3	0.2	0.6	1.1
2.05				0.2	0.2			
2.15	6.8					0.1	0.6	1.6
2.25	6.4							
2.45				0.7	0.5	0.1	0.5	
2.75				0.8	0.1	0.1	0.2	
3.05						0.1		

* - outcrop sample of Kimmeridge Clay ** - 'turf' sample

Table 17: Summary of soil profile data from organic-rich parent materials in England

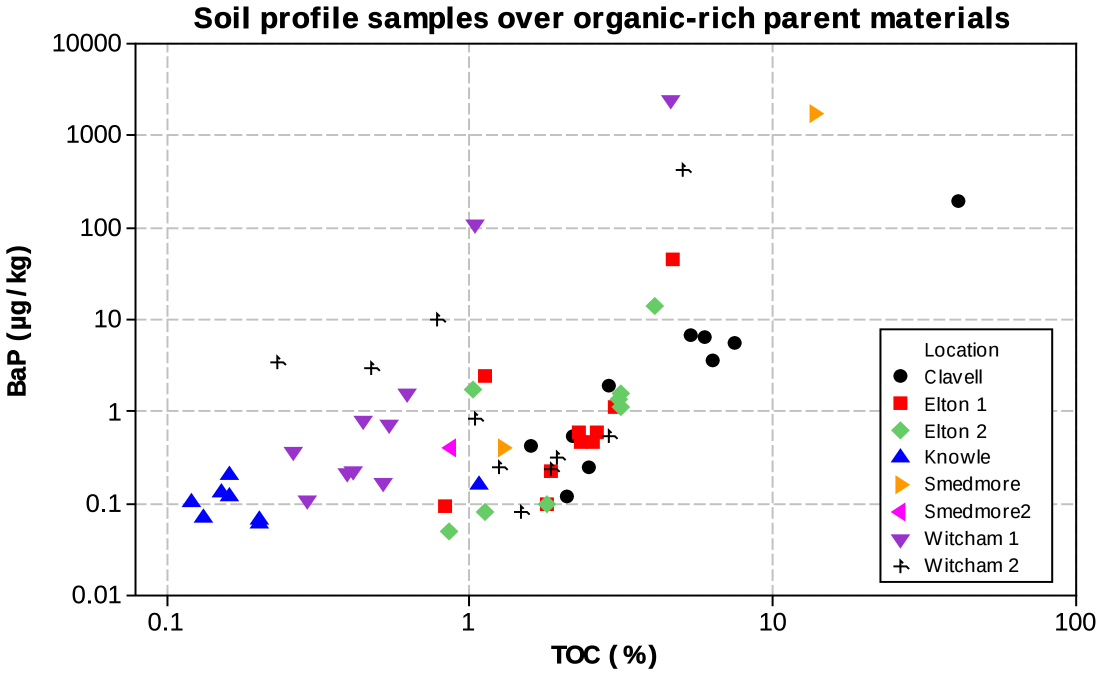


Figure 55: BaP and total organic carbon (TOC) soil data in cores over organic-rich parent materials

3.2.4 Asbestos

Asbestos is a generic term defined by industry referring to certain silicate minerals crystallised in a finely fibrous habit, in bundles of easily separable fibres and with a hair-like elongated shape, thermally stable, with high tensile strengths, smooth faces and with an adamantine or silky lustre (Harper, 2008). Asbestos is a naturally occurring constituent of rocks and soils and includes six asbestiforms minerals from two broad mineral categories of serpentine and amphibole: serpentine – chrysotile (white asbestos); and amphibole - crocidolite (riebeckite) (blue asbestos), amosite (grunerite) (brown asbestos), actinolite, anthophyllite and tremolite (Osinubi *et al*., 2000).

Naturally occurring asbestos (NOA) and the occurrence and concentrations in rocks is largely controlled by the geology. This information can be used to draw attention towards areas with potential to contain NOA (van Gosen, 2007). Asbestos deposits are typically hosted by magnesium-rich rocks (often also iron-rich) that were altered by a metamorphic or magmatic process, including serpentinites, altered ultramafic and some mafic rocks, dolomitic marbles and metamorphosed dolostones, metamorphosed iron formations, and alkalic intrusions and carbonatites (van Gosen, 2007).

Soils developed on serpentinitic bedrock are known to contain significant portions of chrysotile asbestos (Schreier, 1989). These are small in size but have a worldwide distribution. Links between fibers present in soils and rocks and human health problems have been reported in some rural villages in Cyprus, in Greece and in Southwest Turkey and reviewed by Schreier (1989). However, the risk associated with NOA particles in soil remain unclear, as soils can contain aggregating agents such as organic matter, clay, iron oxides and carbonates. These constituents have the potential to bind individual asbestiform particles into larger aggregates and prevent their suspension into the atmosphere as dust. Soil disturbances that destroy the soil aggregates and expose the parent rock to the atmosphere may increase the risk of exposure to NOA particles (Frazell *et al*., 2009).

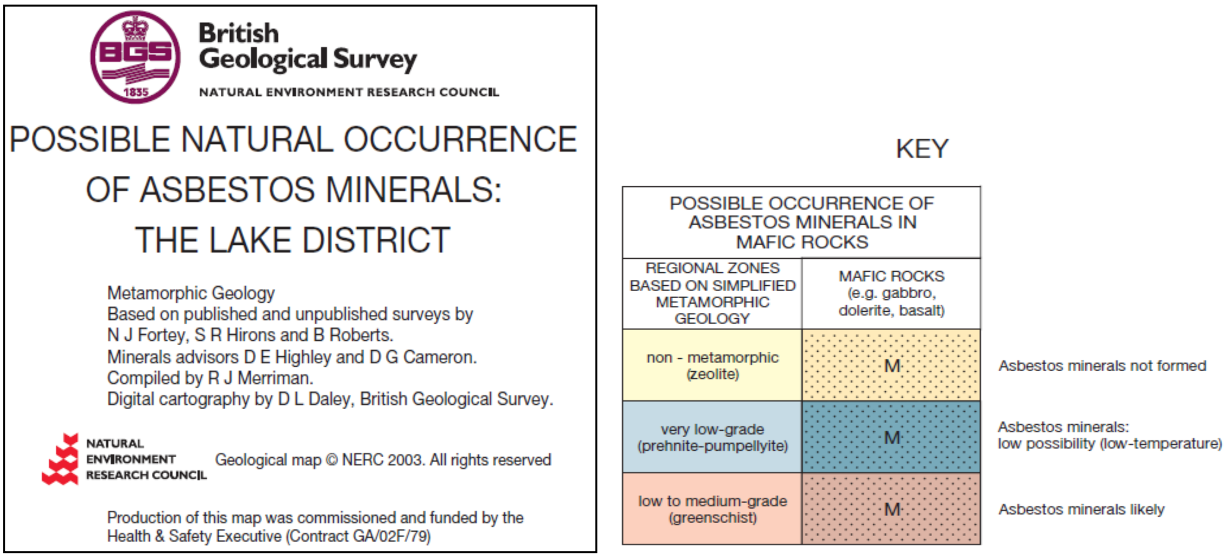
NOA occurrence in UK soils has not been documented. BGS has produced a series of confidential maps for parts of the UK for the UK Health and Safety Executive (HSE) (see BGS, 2003a,b; Figure 56) where asbestos minerals could potentially occur naturally:

1. The Lake District;
2. South West England;
3. Wales;
4. Central and Eastern Scotland;
5. Orkney,Shetland and NE Scotland;
6. Southern Scottish Highlands and the Inner Hebrides;
7. NW Scotland and the Outer Hebrides;
8. Southern Scotland; and
9. Northern Ireland.

The BGS maps of “Possible Natural Occurrence of Asbestos Minerals in Rocks” indicate zones of NOA potential based on the simplified regional distribution of metamorphic zones based on the UK 1:250,000 geological map. Asbestos-bearing rocks may occur in the metamorphic greenschist facies (epizone). In greenschist meta-mafic rocks, asbestos is represented by tremolite-actinolite minerals, which commonly occur as small fibrous crystals. Greenschist facies meta-ultramafic rocks commonly contain serpentine mineral, including fibrous chrysotile (white asbestos). Based on this approach, these maps provide a regional guide to the possible occurrence of natural asbestos minerals in rocks and outcrop. It excludes rock and soil that has been worked and otherwise moved to form man-made ground or similar deposits. Some minor bodies of mafic and ultramafic igneous rocks, with the potential to develop asbestos minerals, are not shown because of the small size of their outcrop. Limited outcrop size may also prevent some of the less common asbestos-bearing rocks being shown. These include calcareous sedimentary rocks metamorphosed within the aureoles of large igneous intrusions, and minor intrusions of alkali microgranites containing riebeckite (blue asbestos). Localised alteration, which has the potential to generate asbestos minerals, was not shown on any of the maps.

An example of contamination arising from the extraction of natural minerals is the vermiculite mining area of Libby, Montana, USA[8].

Unexploited areas of natural asbestos in England can only be considered to be a very minor source contributing to the normal background (Studds, undated). In view of the very limited amount of information on concentrations of asbestos minerals away from contaminated sites, it is inappropriate to pursue NBCs for asbestos.



British Geological Survey
NATURAL ENVIRONMENT RESEARCH COUNCIL

POSSIBLE NATURAL OCCURRENCE
OF ASBESTOS MINERALS:
THE LAKE DISTRICT

Metamorphic Geology
Based on published and unpublished surveys by
N J Fortey, S R Hirons and B Roberts.
Minerals advisors D E Highley and D G Cameron.
Compiled by R J Merriman.
Digital cartography by D L Daley, British Geological Survey.

NATURAL ENVIRONMENT RESEARCH COUNCIL
Geological map © NERC 2003. All rights reserved

Production of this map was commissioned and funded by the Health & Safety Executive (Contract GA/02F/79)

KEY

POSSIBLE OCCURRENCE OF ASBESTOS MINERALS IN MAFIC ROCKS		
REGIONAL ZONES BASED ON SIMPLIFIED METAMORPHIC GEOLOGY	MAFIC ROCKS (e.g. gabbro, dolerite, basalt)	
non - metamorphic (zeolite)	M	Asbestos minerals not formed
very low-grade (prehnite-pumpellyite)	M	Asbestos minerals: low possibility (low-temperature)
low to medium-grade (greenschist)	M	Asbestos minerals likely

Figure 56: Example of map caption and legend from one of the BGS "Possible Natural Occurrence of Asbestos Minerals" map (BGS, 2003a)

---

[8] http://www.epa.gov/region8/superfund/libby/index.html

4 Discussion and Proposals for next phases of the project

1. Work Package 1 has identified soil data that can be used to estimate contaminant NBCs. Data sets collected systematically and with the highest sampling density (*i.e*. covering all land use types) are the most useful. The study of soil chemistry is a multi-disciplinary exercise and as a result information comes from many disparate sources. Technical guidance for contaminant NBCs will need to inform users as to where and how the information can be found. The soil data is a valuable resource for those tasked with assessing contaminated land and needs to be made more readily available. This report, especially Appendix 2 should raise awareness of available data. **The technical guidance (Work Package 4) would be greatly assisted if such information about soil data sets was available from a single web portal.**
2. Much of the soil data has been generated from publicly funded projects, yet there are many aspects that make results difficult and costly to obtain. Complex IPR, data handling and licensing issues stifle the use of soil data for the very sector whose needs are used to justify the funding of such projects. Even within organisations that are the guardians of the data it can be difficult to get easy access to the soil results. An issue highlighted by the work of WP1 is the resolution of site coordinate data. In many instances this has to be degraded to as much as ±10 km to protect land access agreements made at the time of sampling and to ensure the representative nature of sampling sites for the future. Those who plan and manage such projects must be made aware that such restrictions can greatly limit the use of the results for research purposes beyond those of the original survey. The more flexible use of licensing agreements to control access to, and use of, the data must be considered. Land access agreements can stifle legitimate use of the data. **Greater effort needs to be made, where feasible, to ensure soil chemical data is more accessible to those who need to use it.**
3. The data sets used for the data exploration of the inorganic elements in WP2 has primarily been with regard to total element concentrations. For NBCs it is important to have unambiguously total contaminant concentrations to correctly describe spatial variability, therefore such data sets are the data sets of choice. In the contaminated land sector much use is made of partial (*aqua regia*) measurements believed to represent the more extractable fraction of contaminants in soils. Such results, along with other partial extraction techniques, speciation and bioaccessiblity studies are very important in accessing a contaminants potential risk of harm. The relationship between total and *aqua regia* methods has been explored in Section 3.2.1.6 (As) and Section 3.2.2.6 (Pb) and **information on how partial analyses can be used in the context of total results for NBCs needs to be included in the technical guidance.**
4. The proposed approach of this project has been to identify, through exploration of national data sets, areas that can be attributed with NBCs rather than simply assigning a single national value. This method recognises the natural variability of contaminant distributions and how the built environment can impact on these. **The areas to be attributed with NBCs in Work Package 3 are identified as domains and are summarised in Table 18.** The domains are defined using the criteria that are most

influential in determining contaminant distributions in soil – the underlying parent material and areas significantly impacted by anthropogenic activity, *i.e.* the built environment, measured by an urbanisation index.

Contaminant	Domain Name	Defined by
arsenic	Ironstone	SPMM attribute field 'GMPLITH' equals '*Ironstone*'.
	Mineralisation	Mineralisation and mining areas: *1 SW England*; *3 Mendips*; or, *6 Lake District.*
	Principal	Areas not enclosed by 'Ironstone' or 'Mineralisation' domains
lead	Mineralisation	Mineralisation and mining areas: *3 Mendips*; *5 Peak District; 6 Lake District*; or *7 North Pennines.*
	Urban	GLUD log(e) urban:rural ≥ -0.044.
	Principal	Areas not enclosed by 'Mineralisation' or 'Urban'.
BaP	Urban	GLUD log(e) urban:rural ≥ -0.044.
	Principal	Areas not enclosed by 'Urban'.
asbestos	No domain	not applicable

Table 18: A summary of the contaminant domains investigated in Section 3.2. Tables 7 and 11 detail information on mineralised areas

5. In order to correlate the variability of contaminants in soil with the underlying parent material, the BGS Soil Parent Material Model (SPMM), the BGS digital geology maps and available soil maps have been investigated (Section 3.1.1). Although some detailed soil maps are available for England e.g. NSRI NATMAP, there is insufficient coverage at the scale 1:50,000 which is best suited for the more localised definition of domains. Furthermore, the small scale soil maps exhibit similar, if generalised trends, to the mapped geological boundaries. The BGS 1:50,000 digital geology map (DiGMapGB-50) can usefully be used to capture the underlying geogenic sources of contaminants. However, the BGS SPMM, which characterises the near-surface weathered zone, and focuses upon the materials from which topsoils and subsoils develop, is ideally suited to the task of this project. The attribution of some half-a-million polygons with mineralogical, and chemical and physical properties, greatly assists in defining the domains. **It is proposed that the SPMM is used in the methodology of WP3 to determine domain NBCs.** As a result of the data exploration for this project, the BGS maps of "Possible Natural Occurrence of Asbestos Minerals in Rocks" (*i.e*. BGS 2003a,b) will be added to this model.
6. In addition to the geogenic contribution to contaminants in soil, the diffuse anthropogenic component of NBCs also needs to be captured by domains. To investigate this, the CEH Land Cover Map (LCM2000) and the freely available Generalised Land Use Database Statistics for England 2005 (GLUD) were considered (Section 3.1.2). Data exploration using GLUD to define an urbanisation index demonstrated this data set can be applied to define the urban domain without needing

the detail of some of the more comprehensive licensed data sets. **It is proposed that GLUD is used in the methodology of WP3 to determine domain NBCs.**

7. The data exploration for As and Pb demonstrates how metalliferous mineralisation and mining are important to consider as a domain when describing some contaminant NBCs in soil. The DoE sponsored 'Arup' Metalliferous Mineralisation and Mining database, later enhanced by BGS, can be used to capture nature's form of natural contamination. However, at a local scale, the mapping of all historical mine workings, some stretching back several millennia, is incomplete.
8. Investigation of the available soil data sets containing contaminant data shows that generally there is good coverage for most of the inorganic contaminants. The G-BASE urban and rural data gives excellent high resolution soil data for Central and Eastern England and unrivalled coverage for many urban areas. The areas not covered by G-BASE soil sampling are usefully covered by the lower density NSI(XRFS) dataset. Data for Hg and organic contaminants are sparser and **gaps in knowledge for such contaminants will mean that the NBCs determined for these will result in a greater uncertainty that needs to be captured in the methodology of WP3**.
9. Much of the data exploration of WP2 has overlapped with the methodology development for WP3. This data exploration has successfully demonstrated how domains can be derived for three of the selected contaminants, As, Pb and BaP. These contaminants were chosen as they are known to have differing underlying controls and similar exploration can be followed for other contaminants. The fourth contaminant - asbestos, as anticipated, has very little information available to determine NBCs. For such a contaminant the methodology will simply have to be based on the potential for occurrence, derived in urban areas from our knowledge of associated land uses. With respect to natural asbestos mineral occurrence, potential occurrence maps based on the underlying geology is the only approach currently available.
10. Central and eastern England (and Northern Ireland), because of the BGS geochemical baseline mapping programme, are relatively well-covered by high density soil sampling and analysis of inorganic contaminants thus enabling contaminant domains to be well-defined. For those areas where soil sampling has been less dense it is recommended that other surface baseline chemical data, such as that from drainage sediment sampling, can be used along with the soil data to define contaminant domains.
11. The completed work from WP1 and WP2 means the project is now ready to proceed with the definition of some selected contaminant NBCs in WP3. How the NBCs are quantitatively expressed and could relate to derivation of new C4SLs will need to be explored before the technical guidance is prepared (WP4).

Appendix 1 : Contaminant List

The following is a list of inorganic and some selected organic contaminants from Table 1 detailing factors that can contribute to their natural and anthropogenic distribution in soil.

Contaminant	***Factors causing high natural occurrence in soils***	***Anthropogenic sources of soil pollution***
arsenic (As)	High natural contents of As in soils are linked to the presence of underlying As sulphide mineralisation and As-rich parent rocks (*e.g.* ironstones). As is strongly adsorbed in soils by iron oxides and organic matter.	Arsenic can enter the environment through combustion of fossil fuels, incineration of wastes and from use of As-based timber preservatives (chromated copper arsenate) or pesticides.
beryllium (Be)	Granitic and alkaline intrusive rocks are responsible for the high levels of Be in soils. Beryllium is also enriched in micaceous material and secondary minerals weathered from alkaline rocks. Elevated levels of Be are associated with coal deposits.	Beryllium is released into the atmosphere from coal combustion. Many industrial uses of Be can contribute to anthropogenic pollution: in brake systems of airplanes, for neutron monochromatisation, as window material for X-ray tubes, in radiation detectors, and in the electronics industry as a substrate for transistors and silicon chips, coil cores and laser tubes (Salminen *et al*. 2005).
cadmium (Cd)	The main factor determining the Cd content of soil is the chemical composition of the parent material (Kabata-Pendias and Pendias 2001). Cadmium is more concentrated in argillaceous rocks and shale and in metalliferous ore deposits especially associated with Zn.	Anthropogenic sources of Cd include: non-ferrous metallurgical smelting; pyrometallurgical industries; phosphate fertilisers; and sewage sludges. Mostly associated with extensive locations of past mining and smelting of (mainly) Zn and Pb.
chromium (Cr)	The soil Cr is inherited from parent materials and as a result higher contents are found in soils derived from mafic and volcanic rocks. Chromium content is also generally positively correlated with the soil clay fraction.	Main sources of soil pollution are several industrial wastes (electroplating sludges, Cr pigment and tannery wastes, leather manufacturing wastes) and municipal sewage sludges (Kabata-Pendias and Pendias 2001).

Contaminant	*Factors causing high natural occurrence in soils*	*Anthropogenic sources of soil pollution*
copper (Cu)	The concentrations of Cu in soil reflect lithological control and mineralisation. Sulphide minerals (*e.g.* chalcopyrite) are the main detrital Cu phases in mineralised areas and may causes zone of high levels in soils over mineralisation. Copper released during the weathering of these phases is readily adsorbed/complexed by organic matter, clay minerals, hydrous oxides and coprecipitated with carbonates. Zone of high Cu in soil can also be associated to coalfields due to enrichment with organic matter during coal formation resulting from the strong affinity of Cu for organic ligands.	Anthropogenic sources of Cu include mining and smelting, the electrical industry, agriculture, sewage sludge and steel works. Copper compounds are widely used in agriculture (Salminen *et al*. 2005).
lead (Pb)	The main Pb primary mineral in nature is galena (PbS). During weathering Pb sulphides oxidise and can form carbonates or are incorporated in clay minerals, Fe and Mn oxides and in organic matter. UK areas naturally enriched in soil Pb are mainly the Pb-Zn mineralised areas. Organic-rich soils in upland areas can also have some high Pb concentrations.	Long-range atmospheric transport from anthropogenic sources (coal burning, old paint, piping, solders, fuel lead) is a significant source of Pb in soils. Sources of Pb in dust and soil can include Pb from weathering and chipping of Pb-based paint from buildings, bridges, and other structures.
mercury (Hg)	Mercury enters the soil from natural mineralisation or volcanic activities. The accumulation of Hg is related to the C and S content in soils and topsoils have much higher contents than subsoils (Kabata-Pendias and Pendias 2001). Generally organic soils have a higher Hg content than mineral soils due to the high capacity of humus for binding Hg.	Long-range atmospheric transport from anthropogenic sources is a significant source of Hg in humus-rich soils. Hg can enter the environment in atmospheric deposition from combustion of fossil fuels and from ore smelting. Industrial processes, particularly the Hg cell chloralkali process, and the use of Hg fungicides are possible sources of Hg in soils (Ross *et al*. 2007). Hg has also been used in the Au amalgamation process in Au mines. Crematoria furnaces (dental amalgams) are a potential source of Hg in the local environment

Contaminant	***Factors causing high natural occurrence in soils***	***Anthropogenic sources of soil pollution***
molybdenum (Mo)	Mo occurs in soils in sparingly soluble primary minerals (MoS_2, $CaMoO_4$, $PbMoO_4$, $Fe(MoO_4)$ 8H2O), in clay minerals, in iron hydroxides, water soluble Mo and organically-bound Mo. The Mo content of soils usually reflects that of their parent rocks with soils derived from granitic rocks and from some organic rich shales likely to contain large amounts of Mo (Kabata-Pendias and Pendias 2001). Anomalies in Mo contents in UK/English soils are derived from molybdeniferous marine black shale facies, varying in age from Cambrian to Recent (Davies 1980).	Coal combustion is a large atmospheric source of Mo. Mining, smelting, processing of metals and oil refining are other potential sources of Mo in soils (Kabata-Pendias and Pendias 2001).
nickel (Ni)	Nickel is closely related to Fe and is enriched in ultramafic rocks. Nickel content in soils is both dependent on the Ni content of parent rocks and soil forming processes. During weathering Ni is easily mobilised and then co-precipitated with Fe and Mn oxides (Kabata-Pendias and Pendias 2001). Nickel is strongly adsorbed by organic matter.	Nickel is used to make stainless steel and other metal alloys. Anthropogenic sources of Ni in soils include emissions from metal processing operations, from combustion of oil and coal, incinerators, application of sewage sludge, and phosphate fertiliser (Kabata-Pendias and Pendias 2001).
selenium (Se)	Naturally high Se soils are found underlain by organic-rich/black shales, phosphatic rocks and coal. It is also association with sulphide mineralisation and high soil organic matter (Johnson *et al*. 2010).	Important anthropogenic sources of Se include the combustion of coal and petroleum fuels, metal extraction processes and the use of phosphate fertilisers and sewage sludge in agriculture (Johnson *et al*. 2010).
sulphur (S)	Sulphur has a strong affinity for organic matter and as a result soils containing high concentrations of S occur in the peaty areas. Soils overlying organic-rich black shales and coals may contain high S concentrations.	Sulphur is mainly of terrestrial origin, but anthropogenic atmospheric emissions of S from industry have been significant in the past. Sulphur-rich wastes potentially discarded in the environment derive from industries that use sulphates and sulphuric acid, such as mining and smelting operations, Kraft pulp and paper mills, textile mills and tanneries. Fertiliser and pesticide production and use, coal combustion, petrol refining, and vulcanisation (Salminen *et al*. 2005) are additional sources of anthropogenic sulphur.

Contaminant	***Factors causing high natural occurrence in soils***	***Anthropogenic sources of soil pollution***
thallium (Tl)	Thallium occurs in many sulphide ores (Pb and Zn). During weathering, Tl is readily mobilised and fixed in soils by clay minerals and by Mn and Fe oxides. It is enriched in coal.	The largest anthropogenic source is related to coal combustion and smelting and mining. It is also used mostly in manufacturing electronic devices, switches, and closures, primarily for the semiconductor industry. It also has limited use in the manufacture of special glass and for certain medical procedures (ATSDR 1995).
vanadium (V)	Vanadium is concentrated mainly in mafic rocks and shales. It also tends to be associated with organic matter and found enriched in crude oils and coal. In England elevated concentrations occur mainly over the Jurassic strata spatially associated with the ironstones (Breward, 2007).	Vanadium can enter the environment from the burning of fuel oils and is used to make alloys, steel and catalysts.
zinc (Zn)	Zinc most commonly occurs in its sulphide ore mineral sphalerite (ZnS), but is also present in pyroxenes, amphiboles, micas, garnet and magnetite. Soils with high Zn concentrations occur mainly in areas of Pb-Zn mineralisation and mining.	Zinc may enter the environment through atmospheric deposition from mining, the processing of non ferrous metals, the burning of coal and other fossil fuels, the application of sewage sludge to land (Ross *et al*. 2007). In urban areas has a strong association with road corridors (tyre wear) and galvanised structures
cyanide	CN occurs naturally in nature and certain bacteria, fungi, and algae can produce cyanide, and cyanide is found in a number of foods and plants. Cyanide enters air, water, and soil from both natural processes and industrial activities. Cyanides are fairly mobile in soil. Once in soil, cyanide can be removed through several processes. Some cyanide compounds in soil can form hydrogen cyanide and evaporate, whereas some cyanide compounds will be transformed into other chemical forms by microorganisms in soil	The main sources of cyanide releases to the environment are reported to be associated to industrial and municipal wastewater treatments, gas generating sites, aluminium production waste disposal sites, and hydrometallurgical gold mining (Dzombak *et al*., 2006). Releases to soil appear to be primarily from disposal of cyanide wastes in landfills and the use of cyanide-containing road salts.

Contaminant	***Factors causing high natural occurrence in soils***	***Anthropogenic sources of soil pollution***
asbestos	Six naturally occurring asbestiform minerals from two broad mineral categories of serpentine and amphibole. Serpentine – chrysotile (white asbestos); and amphibole - crocidolite (riebeckite) (blue asbestos), amosite (grunerite) (brown asbestos), actinolite, anthophyllite and tremolite (Osinubi *et al*., 2000). Naturally occurring asbestos is referred to as NOA and the occurrence and concentrations in rocks is largely controlled by the geology. Asbestos deposits are typically hosted by magnesium-rich rocks (often also iron-rich) that were altered by a metamorphic or magmatic process (Van Gosen, 2007). Rock types known to host asbestos include serpentinites, altered ultramafic and some mafic rocks, dolomitic marbles and metamorphosed dolostones, metamorphosed iron formations, and alkalic intrusions and carbonatites. NOA in UK is not generally reported though potential natural sources in England have been mapped (BGS, 2003a,b). The risk associated with NOA particles in soil is unclear, as soils can contain aggregating agents such as organic matter, clay, iron oxides and carbonates. These constituents have the potential to bind individual asbestiform particles into larger aggregates and prevent their suspension into the athmosphere as dust. Soil disturbances that destroy the soil aggregates and expose the parent rock to the atmosphere may increase the risk of exposure to NOA particles (Frazell *et al*. 2009)	Natural sources of asbestos in the UK are considered of limited scale. Asbestos contamination in soil is mainly by anthropogenic sources arising from the disposal of asbestos bearing materials (demolished buildings, buried pipes, ducts, fly tipping) and quarrying/tunnelling.
dioxins and furans (PCCD and PCDF)	Whilst the focus for dioxins and furans levels as a normal background will be on anthropogenically derived chemicals there is evidence of natural formation in certain rock types (Jones and Sweetman 2003).	Mainly as the combustion by-products from incineration and bio-fuels, chemical industries, metal processing and paper manufacturing

Contaminant	***Factors causing high natural occurrence in soils***	***Anthropogenic sources of soil pollution***
polychlorinated biphenyls (PCB)	Normal background concentrations of PCBs will be entirely from diffuse anthropogenic pollution	Primarily a human-made technological additive used in a variety of products 1950s – 1980s. Widely used in oils in transformers and capacitors, glues, adhesives, plasticisers in paints and cements, flame retardants, pesticides, hydraulic fluids, lubricating oils and paints (Mielke *et al*. 2011).
polycyclic aromatic hydrocarbons (PAH)	PAH are natural components in oil and hydrocarbons. They would therefore be present in nature associated with rock strata bearing hydrocarbons, *e.g*. coal and organic-rich shales. Natural forest and moorland fires will generate PAHs.	PAHs originate from car exhausts, domestic heating and industrial processes (such as aluminium smelting). Also creosote and asphalt products. Incomplete combustion of wood and oil products is an important anthropogenic source.

Appendix 2 : Summary of soil data sets investigated

This appendix summarises soil data sets investigated for this project and are the reports produced from the Project’s MS ACCESS database (SoilDataSets.accdb) (see Section 2.5).

Organisation abbreviations used in the data set table (Basic Information) are listed below. The numbered list of cited bibliographic references is given at the end of this appendix.

Organisation abbreviations

Origin Organisation	Full name
BGS	British Geological Survey
EGS	EuroGeoSurveys
NSRI	National Soil Resources Institute, Cranfield University
Defra	Department for Environment, Food and Rural Affairs
NERC	Natural Environment Research Council
CEH	Centre for Ecology and Hydrology
EA	Environment Agency
JRC	Joint Research Centre, European Commission
FR	Forest Research
ABER	University of Aberystwyth
UCL	University College of London
ADAS	Agricultural Development and Advisory Science
UEA	University of East Anglia
RECHEM	Rechem International Ltd
LU	Lancaster University
EU	European Union

Available soil data: Basic Information

Data Set	Origin Organisation	Reference	Coverage	Period of collection	No. samples (England)	Sampling strategy	Sampling method
G-BASE (regional)	BGS	1,2,3,5	Central and Eastern England. 1 site per 2 km squares. Limited regional data available from Scotland and Wales. Similar good coverage for all of N Ireland from Tellus Project. Extent of data can be viewed in BGS GeoIndex and Google Earth application	1986 to present	23,686	Systematic sampling outside urban areas targeting multiple land uses in order to create regional geochemical maps. Primarily for inorganic elements (not Hg)	Soils collected with 1m auger. Five sub-samples from corners and centre of 20 m square. Topsoil from 5 - 20 cm. Deep soil from 35 - 50 cm collected but not routinely analysed
G-BASE (urban)	BGS	1,4,5,6,7	Twenty one urban centres from England the biggest being the Greater London Authority area. 4 sites per 1 km square. Other urban centres sampled in other parts of UK. Extent of data can be viewed in BGS GeoIndex and Google Earth application	1992 to present	13,583	Systematic sampling of multiple land uses within urban areas. Primarily for inorganic elements (not Hg)	Soils collected with 1m auger. Five sub-samples from corners and centre of 20 m square. Topsoil from 5 - 20 cm. Deep soil from 35 - 50 cm collected but not routinely analysed. Surface sample (0 - 5 cm) also collected in London area
GEMAS	EGS	8,9,37	Sixty five sites in England representing a density of 1 site per 2,500 sq. km. Survey covered all of Europe	2008	130	Targeting grazing and arable land using 50 km cells. Primarily for metallic elements	Soils collected with spade with a minimum of 5 sub-samples from the corners and centre of 10 m square. Arable sample from 0 - 20 cm and grazed land sample from 0 - 10 cm, with two land use types as close together as possible
FOREGS	EGS	10,11,12,13, 14	Survey covered all of Europe. Thirty three sites in England, average sample density across Europe is 1 site per 4,700 sq. km.	1998	33	Systematic sampling targeting 160 km cells as defined by a global reference network of cells. Primarily for inorganic elements also includes Hg	Soils collected with stainless steel trowel. Composite of 3 -5 sub-samples collected at a minimum distance of 5 m apart. Topsoil from 0 - 25 cm and sub-soil from a 25 cm range from a depth 50 - 200 cm
NSI	NSRI	15,16	England and Wales. 6,127 sites visited and 5,691 samples collected. 1 site per 25 sq. km	1978 - 1983	5,691	Systematic sampling based on 5-km orthogonal grid principally targeting agricultural land. Small number of inorganic elements	Soil collected with a screw auger with 25 sub-samples taken across a 20 m square. Restricted to the upper most 15 cm of mineral soil, i.e. topsoil
NSI (XRFS)	BGS	17	England and Wales. c.5,671 samples available for renanalysis	1978 - 1983	4,864	Systematic sampling based on 5-km orthogonal grid principally targeting agricultural land. Large number of inorganic elements	Soil collected with a screw auger with 25 sub-samples taken across a 20 m square. Restricted to the upper most 15 cm of mineral soil, i.e. topsoil
CS	CEH	18,38,39,67, 68	Great Britain monitoring over time. 1998 survey collected from 366 1-km squares in England.	1978, 1998, 2007	44 - 540	Monitor changes in key soil properties including Hg and organics. Sampling based on plots within 1-km squares based on intersections of 15 km grid. Excludes urban areas so is a study of the rural environment	1998 sampling used pits collecting with a trowel from surface horizon below the organic litter. If the surface horizon was less than 5 cm the soil was collected from the upper 5 cm. The 2007 survey used pipes to collect 0-15 cm sample

Data Set	Origin Organisation	Reference	Coverage	Period of collection	No. samples (England)	Sampling strategy	Sampling method
UKSHS	EA	19,20,21	United Kingdom. There are 61 rural, 13 urban and 22 industrial sites sampled in England. The approximate grid reference for each rural site was determined from the intersection of a 50 km grid	2001-2002	156	To provide at a national scale information about a range of pollutants including PCBs, dioxins, polycyclic aromatic hydrocarbons and trace metals in soils and herbage. Rural, urban and industrial sites were targeted.	Used a 20 m sq. with 3 x 5-cm deep soil cores (Eijkelkamp coring kit) taken at 3 locations within square. The 3 cores were bulked to form 1 sample giving 3 samples for analysis. More complex sampling plan used at industrial sites
ECN	NERC	23, 24, 25, 26,27,28	United Kingdom. Monitoring at 5 and 20 year intervals. Network of 42 freshwater (lakes and rivers) and 11 terrestrial sites - only 6 terrestrial sites in England	1992 to present	5 sites	Monitoring environmental change. Cells (5m x 5m) used for 20-yr and 5-yr sampling (see sampling protocols for detail)	Gouge auger used to collect c.3 kg soil. Samples taken to a maximum depth of 30 cm from each sampled sub-cell. One set is based on depths: 0-5, 5-10, 10-20, and 20-30 cm. The other set corresponds to horizons within the top 30 cm
MRP	BGS	22	Mineral Reconnaissance Programme, Great Britain and Northern Ireland. Targetted metalliferous mineral site and some regional surveys. Extent of data can be viewed in BGS GeoIndex	1973 -1997	c.25,000	Soil sampling targeting known metalliferous mineral areas and some regional exploration. Site-specific investigations usually done along lines with sites at 5 - 100 m intervals. Regional typically at 12 sites per km square	Usually auger (power auger for samples > 1 m). Samples collected at various depths down to 8 m trying to get indication of mineralisation in buried bedrock
LUCAS	JRC	29,30,31	Europe-wide survey. There are 1,373 topsoil samples logged from the UK	2009	1,373 (UK)	Sampling is based on the intersection points of a 2 x 2 km grid covering the EU with topsoil samples colected from 10% of the sites. Variety of landuse covered. CORINE LANDCOVER 2000 used to calculate % area of each land use type	Samples collected by composite sampling (5 sub-samples) and only the mineral topsoil sampled
BIOSOIL (UK)	FR		Europe-wide to give an improved common baseline of forest soils. There are 167 BioSoil plots on 16 x 16 km grid in the UK, of which 125 on private woodland and 42 on Forestry Commission Woodland	2006	167 plots (UK)	Sampled organic soil horizons and the mineral soil layers 0 - 10 cm, 10 - 20 cm, 20 - 40 cm and 40 - 80 cm	Organic layer in 25 x 25 cm area. Humus collected with 8 cm diameter auger. Mineral layer sampled by fixed depth level I: 0-10,10-20 cm; level II: 0-10, 10-20; 20-40 and 40-80 cm, by augering/pits. Samples are composites of 5 (I) or 24 (II) subsamples
FOREST	FR	32,33,66	Europe-wide. 10 long-term intensive monitoring plots covering 3 tree species in the UK established in 1995, with additional plots added in 2002	1995 - 2000	10 plots	sampling design:judgemental and random design or systematic design with a random component; Sampled organic layer and underlying mineral soil	Organic layer in 25 x 25 cm area. Humus collected with 8 cm diameter auger. Mineral layer sampled by fixed depth level I: 0-10,10-20 cm; level II: 0-10, 10-20; 20-40 and 40-80 cm, by augering/pits. Samples are composites of 5 (I) or 24 (II) subsamples
NATURAL ASBESTOS MAP	BGS	40,41	UK-wide. Nine maps for selected parts of the UK showing the "Possible Natural Occurrence of Asbestos Minerals"	2003	2 maps	No samples. Based on an understanding of the underlying geology that is likely to yield asbestos minerals	Geological mapping
TIPPING Hg	CEH	42,43	UK-wide. Reports on Hg data from CS:2000 and UKSHS but also includes further 20 sites from NW England	1998 - 2008	898 (UK)	Mercury data from different surveys brought together to give national coverage	Twenty "new" samples from NW England collected by digging small pits to a depth of 20 - 30 cm

Data Set	Origin Organisation	Reference	Coverage	Period of collection	No. samples (England)	Sampling strategy	Sampling method
LE studies	BGS	unpublished	London Urban area	2005 - 2009	50 - 473	Subset of samples from the G-BASE London Earth survey focusing on Hg, profile line across London for Au+PGEs, subset of samples for organic determinations and bioaccessibility. Sample denisty as for G-BASE (urban)	as for G-BASE (urban). Separate samples collected for organic sample analyses using a sampling procedure specific for organic samples including cooling samples and freezing then soon after collection
MINING Hg	ABER	44	Hg content of soils in Western Britain with special reference to contamination from base metal mining	<1975	51 (E & W)	Targeted historic base metal mining areas of England & Wales with specific interest in Hg	Soils were either surface or from profile pits. Topsoils (0 -15 cm) were sampled, using a mild steel screw auger, from pasture fields. Immediate environs of mines were avoided, as were arable fields. 15 to 20 cores taken at random from fields and bulked
Diss Mere Hg	UCL	45	Diss Mere catchment area, England. 5.3 m sediment core sample and soils	2008	11 sites	5.3 m sediment core sample and soils taken from Diss Mere area to look at historical contamination of the environment by Hg	Soil corer, multiple samples from sites at variable depths down to 55 cm
ADAS	ADAS	46	England and Wales - 16 farms in each ADAS region	1973 - 1980	c. 1,521 (E&W)	Soils collected from 16 farms each year 1973 to 1980 looking at a small number of metallic elements	Soils collected with screw or tubular auger, 0 - 15 cm
SULPHATE	BGS	47	England, Wales and S Scotland		map	digital map 1:625 000 scale of sulphate bearing formations	Geological mapping
MINING Pb	ABER	48	England and Wales	1970s	692 (E & W)	Targeted four Pb mining areas of England and Wales - Central Wales; Halkyn Mt., Wales; Tamar Valley, SW England; and N Somerset, England	Sample depth to 15 cm
JONES et al. PAH	LU	34,35	Wales	before 1989?	49 (Wales)	Various rural and urban locations of varying land use targeted for PAHs	At each site 20 0 - 5 cm cores taken with stainless steel corer
HMIP (1989)	UEA	49,50	UK-wide	1986 - 1987	c. 67 sites	54 urban sites at intersection of regular 50 km grid. Further 13 samples from London and Birmingham looking at PCBs, PCDDs and PCDPs	6-8 core samples to 5 cm depth
HMIP (1995)	UEA	51,52	UK-wide	1988		Samples collected from urban locations and at regular distances from potential dioxin sources	
RECHEM PCB	RECHEM	53,54,55,56, 57	UK-wide	Late 1980s		monitoring for PCB and trace metals in the vicinity of a chemical waste disposal facility	
COUSINS et al. PAH	LU	58	UK-wide	1993	46 UK sites	Returned to sites sampled 1951- 1974 looking at PAHs	Hand-held auger to collect soils from 0 - 2.5 cm and 0 - 25 cm. Soil samples collected, treated and stored specifically for the anlysis of trace organics

Data Set	Origin Organisation	Reference	Coverage	Period of collection	No. samples (England)	Sampling strategy	Sampling method
LEAD et al. PCB	LU	59	UK (and Norway)	1993	46 UK sites	Returned to sites sampled 1951- 1974 (see also Cousins et al., 1997) looking at PCBs	Hand-held auger to collect soils from 0 - 2.5 cm and 0 - 25 cm. Soil samples collected, treated and stored specifically for the anlysis of trace organics
MEIJER et al. PCB	LU	61,62	World-wide	1998 (UK)	7 sites (E & W)	200 "background" surface soils remote from potential sources targeted for PCBs and HCB	Triplicate samples collected with hand held coring device from 0 - 5 cm
URBSOIL	EU	64,65	Six European cites - Aveiro, Glasgow, Ljubljana, Seville, Torino and Uppsala	not given	160	To identify soil quality parameters in urban areas. Sampling based on a grid consisting of at least 25 sites at least 50 m apart	Samples collected with a trowel or plastic-lined corer from 0 - 10, 10 - 20 cm

Available soil data: Analysis Information

Data Set	Sample Preparation	Pre-analysis procedures	Analytical Method	Analytes Determined
G-BASE (regional)	Oven-dried <30ºC, disaggregated and sieved to <2 mm using nylon screen. Pulverised in agate ball mill for 30 mins	Powders pelletised ready for XRFS	XRFS-WD and XRFS-ED	Mg, P, K, Ca, Ti, Mn, Fe, V, Cr, Co, Ba, Ni, Cu, Zn, Ga, As, Se, Rb, Sr, Y, Zr, Nb, Mo, Pb, Bi, Th, U, Ag, Cd, Sn, Sb, Cs, La, Ce, Ge, Sc, Se, Br, Hf, Ta, W, Tl, Te and I plus Loss on Ignition (450ºC for 24 hrs) and pH (calcium chloride) on 50% of samples
G-BASE (urban)	Oven-dried <30ºC, disaggregated and sieved to <2 mm using nylon screen. Pulverised in agate ball mill for 30 mins	Powders pelletised ready for XRFS	XRFS-WD and XRFS-ED	Mg, P, K, Ca, Ti, Mn, Fe, V, Cr, Co, Ba, Ni, Cu, Zn, Ga, As, Se, Rb, Sr, Y, Zr, Nb, Mo, Pb, Bi, Th, U, Ag, Cd, Sn, Sb, Cs, La, Ce, Ge, Sc, Se, Br, Hf, Ta, W, Tl, Te and I plus Loss on Ignition (450ºC for 24 hrs) and pH (calcium chloride)
GEMAS	Oven-dried at max 40ºC, disaggregated and sieved to <2 mm using nylon screen. Homogenised and split into a number of subsamples	One set of subsamples for ICP analysis subjected to hot aqua regia extraction. Samples for XRF milled to <40µm and lithium metaborate/bromide fusion	ICP-MS, ICP-AES, XRF-WD	XRFS: SiO2, TiO2, Al2O3, Fe2O3, MnO, MgO, CaO, Na2O, K2O, P2O5, SO3, Cl, F, As, Ba, Bi, Ce, Co, Cr, Cs, Cu, Ga, Hf, La, Mo, Nb,Ni, Pb, Rb, Sb, Sc, Sn, Sr, Ta, Th, U, V, W, Y, Zn, Zr (41). 53 elements following aqua regia, pH, and LOI (1030°C for 10 mins)
FOREGS	Dried <40ºC, dissagregrated in porcelain bowl, sieved to <2 mm using nylon screen. Pulversised to <63 µm in agate disc mill	Samples prepared at various labs around Europe. Mixed acid extraction before ICP-MS, aqua regia extraction before ICP-AES. Pelletising before XRFS-WD	ICP-MS, ICP-AES, XRFS-WD, Hg analyser	extensive range of elements done by various methods and includes Hg done by Hg analyser and other parameters such as pH and LOI
NSI	Air-dried. Milled in a mild-steel roller mill to pass through 2 mm sieve. A sub-sample of <2 mm fraction ground to <150 µm in agate ball mill	<2 mm fraction to various extraction techniques to determine available/extractable (using ammonium-EDTA) elements. Pulverised (< 150 µm) determined for following aqua regia extraction to give "total" element results	Principally ICP-AES	aqua regia Al, Ba, Cd, Ca, Cr, Co, Fe, Pb, Mg, Mn, Ni, P, K, Na, Sr and Zn. Also pH and organic carbon. Extractable/available K, Mg, P, Cd, Co, Cu, Pb, Ni, Zn
NSI (XRFS)	As per NSI. Excess <2 mm sample powder pulversised in agate ball mill	Pulversised <2 mm fraction pelletised ready for XRFS	XRFS-WD and XRFS-ED	XRFS-WD: Na,Mg,Al,Si,P,S,Cl,K,Ca,Ti,Mn,Fe,Sc,V,Cr,Co,Ni,Cu, Zn,Ga,Ge,As,Sr,Br,Rb,Sr,Y,Zr,Nb,Mo,Nd,Sm,Yb,Hf,Ta,W,Tl,Pb,Bi, Th,U. XRFS-ED:Pd,Ag,Cd,In,Sn,Sb,Te,I,Cs,Ba,La,Ce
CS	Air-dried <2 mm pulversied prior to analysis	1998 & 2007 soils determined following aqua regia extraction	1998&2007: ICP-OES. 2007 ICP-MS. 1998 GC-MS. 1998 Hg cold vapour AAS	ICP-OES:Cd,Cr,Cu,Pb,Ni,V Zn. Hg by cold vapour AAS ICP-MS (2007):Li,Be,Al,Ti,V,Cr,Mn,Fe,Co,Ni,Cu,Zn,As,Rb,Se,Sr,Mo,Cd,Sb, Sn,Cs,Ba,W,Hg,Pb,U. 1998 GC-MS used for <120 POP analyses
UKSHS	No information provided	Heavy metals extracted by aqua regia. Various methodologies (described in Report 3) for organics	ICP-MS, ICP-OES, CV-AAS	ICP-MS:Cd,Cr,Cu,Pb,Ni,Pt,Sn,Ti,Zn. ICP-OES:Mn,V. CV-AAS:Hg,As. Organic methods, 26 PCBs, 22 PAHs, 17 PCDDs/PCDFs
ECN	Samples air-dried and sieved to <2 mm.	Heavy metal determined following aqua regia extraction. As determined after Mg(NO3)2 ashing then sulphuric acid/citrate extraction. Hg nitreic/sulphuric acid digest	various depending on data originator	Pb, Zn, Cd, Cu, Co, Mo, Cr, Ni, As, Hg
MRP	Great variety of sample fractions used, frequently <80 or <60 mesh BSS	Numerous depending on analytical method	mainly AAS or XRFS	predominantly metallic elements. E.g. for Great Britain As data 17,943; Cu data 47,997; Cr data 20156; Ni data 30,159; Pb data 46,768; Hg data 567
LUCAS	Relevant heavy metal analyses not yet published	Relevant heavy metal analyses not yet published	not yet published	the analysis of heavy metalsi proposed

Data Set	Sample Preparation	Pre-analysis procedures	Analytical Method	Analytes Determined
BIOSOIL (UK)	Samples are air dried and living macroscopic roots and all particles, mineral and organic, with a diameter larger than 2 mm removed from the samples by dry sieving	Aqua regia extraction	ICP, AAS (Hg by cold vapour AAS)	Cu, Pb, Cd, Zn, Cr, Ni, Hg (only for organic layer)
FOREST	Samples are air dried and living macroscopic roots and all particles, mineral and organic, with a diameter larger than 2 mm removed from the samples by dry sieving	Aqua regia extraction	ICP, AAS (Hg by cold vapour AAS)	Cu, Pb, Cd, Zn, Cr, Ni, Hg in organic and mineral layer
NATURAL ASBESTOS MAP	None	Not applicable	None	None
TIPPING Hg	New samples - Air-dried, sieved <2 mm and ball milled	New samples - extraction by aqua regia and microwave digestion	ICP-MS	Hg, pH
LE studies	Hg and PGE analyses used samples as per G-BASE urban. Organic samples were done according to a specific organic sample protocol	Organics determined following hexane/acetone extraction	Organics - GC-MS	Hg, Au+PGE, Organics and bioaccessible determinations of some contaminants (additional to G-BASE (urban))
MINING Hg	Dried at 12-15°C, ground in porcelain utensils, sieved through a nylon mesh of 2 mm and stored sealed in polythene bags	(see method of Melton et al. 1971. Soil Sci. Soc. Am. Proc., 35, 850-852)	Cold vapour AAS	Hg (in addition to Pb, Zn, Cu, Cd)
Diss Mere Hg	Soils freeze-dried	Hot nitric acid	Cold vapour AAS	Hg
ADAS	Air-dried, sieved < 2 mm	perchloric/nitric acid digestion. Also some elements done by several less strong extractions	AAS (Se estimated fluorimetrically)	As,Cd,Co,Cr,Cu,Mo,Ni,Pb,Se,Zn
SULPHATE				
MINING Pb	Air-dried, sieved < 2 mm	hot nitric acid digestion	AAS	Pb
JONES et al. PAH	Disaggregrated with pestle and mortar, sieved through 2 mm sieve, bulked and stored in glass jars rinsed with dichloromethane	dichloromethane soxhlet extraction	HPLC	BaP, PAH
HMIP (1989)	Air-dried, sieved <2 mm	extraction into hexane/acetone mixture	HPLC	PCB, PCDD, PCDF
HMIP (1995)				PCB, PCDD, PCDF
COUSINS et al. PAH			HPLC with fluorescence detection	PAH
LEAD et al. PCB				PCB
MEIJER et al. PCB		dichloromethane and toluene soxhlet extraction	GC-MS for PCBs and HCB	PCB, HCB
URBSOIL	Air-dried and sieved <2 mm. A sub-sample ground to <150 µm in agate mill or mortar	microwave aqua regia digest. Hg by pyrolysis AAS with gold amalgamation	AAS, ICP-MS, ICP-AES	Cr,Ni,Pb,Zn,Cu. Also Hg in separate publication on same samples

Available soil data: Data Information

Data Set	Robustness of data	Format	Availability	IPR/Copyright information	Email contact	Web page	Additional Information
G-BASE (regional)	Samples collected to a consistent and documented protocol. Data quality monitored by use of control samples (duplicates, replicates and reference materials,minimum of 6%). Samples prepared and analysed in UKAS accredited laboratories.	Digital maps (pdf), ArcGIS maps, digital data extracted from Oracle database in various formats, reports, publications	Soil results with BNG coordinates are available under licence from the British Geological Survey. Data is free but handling and licencing charge.	IPR detailed in licence agreement	enquiries@bgs.ac.uk	www.bgs.ac.uk/gbase/regional.html http://www.bgs.ac.uk/GeoIndex/geochemistry.htm http://www.bgs.ac.uk/gbase/sampleindexmaps/home.html	G-BASE results are stored in a corporate geochemical database. Requests for data require a complex sql and a thorough understanding of the numerous database tables and their complexity
G-BASE (urban)	Samples collected to a consistent and documented protocol. Data quality monitored by use of control samples (duplicates, replicates and reference materials,minimum of 6%). Samples prepared and analysed in UKAS accredited laboratories.	Digital maps (pdf), ArcGIS maps, digital data extracted from Oracle database in various formats, reports, publications	Soil results with BNG coordinates are available under licence from the British Geological Survey. Data is free but handling and licencing charge. PHE results for urban areas and the London Earth data are available as a BGS digital data product	IPR detailed in licence agreement	enquiries@bgs.ac.uk	www.bgs.ac.uk/gbase/urban.html http://www.bgs.ac.uk/GeoIndex/geochemistry.htm http://www.bgs.ac.uk/gbase/sampleindexmaps/home.html	G-BASE urban data reported in series of urban reports downloadable from G-BASE web page
GEMAS	Samples collected to a consistent and documented protocol across Europe. Data quality monitored and assessed by use of control samples and published as a report	Publications. Data will be made availble as Excel spreadsheets from country representatives	Data not yet published. Current use of data has some restrictions until the full release in 2013	Analytical results for GEMAS samples collected in the UK are owned by BGS and will be freely available from 2013. One restriction is not to publish the 90th percentile value of any element	enquiries@bgs.ac.uk (attention of Andreas Scheib)	In development (Austrian Geological Survey)	Geochemical Mapping of Agricultural Soils (GEMAS) is a cooperation project between the Geochemistry Expert Group of EuroGeoSurveys (EGS) and Eurometaux. The GEMAS project was started to produce soil geochemistry data at the continental scale consistent with REACH (Registration, Evaluation and Authorisation of Chemicals - EC, 2006)
FOREGS	Samples collected and analysed to strict international standards as laid out for Global Geochemical Database requirements report	atlas (maps), digital maps (pdf), digital data (zipped Excel)	Databases, maps and atlases readily and freely available from web pages hosted by the Geological Survey of Finland (GTK)	Data is collectively owned by the component surveys of the EurogeoSurveys. No IPR/Copyright issues	none	www.gsf.fi/publ/foregsatlas/	This was the first initiative arising from the IGCP Project 259 towards completing a global geochemical database. Sponsored by the group of European geological surveys

Data Set	Robustness of data	Format	Availability	IPR/Copyright information	Email contact	Web page	Additional Information
NSI	Strict protocols applied to site location and description, soil sampling strategy, and soil profile description. Considerable effort also went into QC sample preparation and analysis, data recording, error trapping and creation of database	Data normally supplied on CD in an agreed format	Available under licencing agreement through LANDIS. Note that the site coodinates in some distributed data sets are degraded to the nearest 1 km	Arrangements for access to soil data are governed by an agreement between NSRI and Defra acting on behalf of the Crown. Under the LandIS Agreement data are only ever licensed (i.e. not sold) for use over a specified period.	nsridata@cranfield.ac.uk (attention Timothy Farewell)	www.landis.org.uk	Cranfield University's “Land Information System” (LandIS) contains information about the soils of England and Wales, collected first by the Soil Survey of England and Wales between 1939-1987, subsequently by the Soil Survey and Land Research Centre from 1987 to 2006, and then thereafter by the National Soil Resources Institute (NSRI). From 1998 soils there are 540, 44, and 93 samples analysed for heavy metals, Hg & As and POPs, respectively. Extensive exploration of data supported by Defra soil programme R&D project SP0124
NSI (XRFS)	As per NSI data set. XRFS analysis done in BGS UKAS accredited laboratories	Electronic geochemical atlas. Analytical results will be made available as Excel spreadsheets	Analytical data available from BGS with few restrictions. The site coordinates are covered by NSRI IPR restrictions	Site coordinate IPR belongs to NSRI. IPR of reanalysed data shared between BGS and Rothamsted Research	enquiries@bgs.ac.uk (attention Barry Rawlins)	In Development	Samples for reanalysis provided through Prof. Steve McGrath, Rothamsted Research. Note that the site coordinates are estimated to be +/- 10 m or better.
CS	Sample collection and preparation to well-defined documented proceedures. Control samples used in analyses. Some UKAS accredited methods	Digital data, reports, maps	Countryside Survey square level data is now available for licensed users to download for a wide range of geographic regions across Great Britain	Data is licenced through CEH. Because of land access agreements site coordinate data is degraded to nearest 10 km	enquiries@ceh.ac.uk (attention of Claire Wood)	www.countrysidesurvey.org.uk/	Available data is complex covering several phases of sampling with changes to methodology between sampling phases. Emmett et al (2010) gives good overview of project. The poor spatial resolution of grid coordinates limits the value of the data for this project
UKSHS	Analyses carried out at accredited labs (see report 1). Metal and organic results peer-reviewed. Control samples used. Sample collection/lab methods used were accredited to ISO 17025 by UKAS	Digital data available on CD. Reports available in pdf format from various websites	Data CD available through the EA. Grid coordinates to 10 km resolution. Original data not easily accessible	IPR rests with EA on behalf of Defra	enquiries@environment-agency.gov.uk	none	UK Soil and Herbage Pollutant Survey (UKSHS) Defra sponsored project CB01204. All reports available as pdfs. Rural sites are within areas not classified as urban, semi-urban or semi-rural (where semi-urban is more than 25% built up and semi-rural is within small towns or villages greater than 3 km2 in area). Urban sites are defined as being within areas that are more than 90% built up – typically large towns greater than 25–50 km2 in area and cities greater than 50 km2 in area

Data Set	Robustness of data	Format	Availability	IPR/Copyright information	Email contact	Web page	Additional Information
ECN	Common analytical guidelines are provided for each determinand. Each laboratory practises its own internal QC, and most participate in national QA schemes	Digital data, reports, publications	Licensed access to ECN raw data may be provided, through data request form, at the discretion of ECN's sponsoring organisations, except where otherwise legally required	NERC holds Intellectual Property Rights (IPR) for all ECN data jointly with each participants. Data originators retain exclusive rights to ECN data collected at their own sites for applications outside the remit of ECN	ecn@ceh.ac.uk	www.ecn.ac.uk	Since 1992 UK's long-term environmental monitoring programme. Collects, analyses and interprets long-term data from a network of sites. Physical, chemical and biological data are a unique national resource that is improving understanding of how and why environments change.Terrestrial sites: Moor House & Upper Teasdale, Drayton, Wytham, Rothamsted, Alice Holt, Porton and North Wyke
MRP	Samples collected and analysed to a variety of protocols so collectively incoherent	Archived in oracle database. Can be provided in a variety of formats	Can be provided on application to BGS and subject to licencing and data handling charge	BGS licencing agreement	enquiries@bgs.ac.uk	www.bgs.ac.uk/mineralsuk/exploration/potential/mrp.html http://www.bgs.ac.uk/GeoIndex/geochemistry.htm	A large number of soils were collected as part of the Department of Trade and Industry sponsored UK mineral reconnaisance progamme. The great variety of protocols used in sampling and analysis limits the data use on a national scale but Local Authorities where there has been metalliferous mining may find useful supplementary data. The specific targeted nature of the MRP means this data is unsuitable for estimating normal background concentrations
LUCAS	Sampling and field data recording done to strictly defined protocols. No information on analytical results (heavy metal analyses not yet available)	csv database of field information available from Eurostat portal	Soil data from LUCAS are temporarily not available to third parties and general public due to privacy rights of land owners. Legal services are currently evaluating eventual release without precise georeferencing of the data	Limited soil data currently available and is not considered confidential. However, spatial resolution of the data will be limited by land access agreements	estat-user-support@ec.europa.eu	http://epp.eurostat.ec.europa.eu/portal/page/portal/lucas/introduction	LUCAS is the acronym of Land Use and Cover Area frame Survey. It is a pilot project to monitor changes in the management and nature of the land surface of the European Union. In the LUCAS survey 235 000 points were visited by 500 field surveyors. Those sites were selected from a standard 2 km grid with in total around 1 million points all over the EU. Soil samples have been analysed for basic soil properties - particle size, distribution, pH, organic C, carbonates, NPK, CEC and multispectral properties. Heavy metal analysis is planned to be carried out by the laboratories of the Joint Research Centre and expected to be available in 2014

Data Set	Robustness of data	Format	Availability	IPR/Copyright information	Email contact	Web page	Additional Information
BIOSOIL (UK)	The quality of the soil analytical data is controlled by interlaboratory comparisons by the Forest soil co-ordinating Centre	can be made available in digital formats	BIOSOIL data are not yet available (still under QC validation). Georeferencing of the data is not available to third parties due to privacy rights issues. Data available from Programme Co-ordinating Centre (PCC) of ICP Forests in Hamburg, Germany.	There are restrictions on how the data can be used. Requests must include the purpose of a planned study, its time frame, and a declaration that the data will not be submitted to other parties	elena.vanguelova@forestry.gsi.gov.uk, Forest Research	http://www.forestry.gov.uk/fr/INFD-73UDF3	Part of the International Co-operative Programme (ICP) on assessment and monitoring of air pollution effects on forests. Work therefore specifically targets one land use type
FOREST	The quality of the soil analytical data is controlled by interlaboratory comparisons by the Forest soil co-ordinating Centre	can be made available in digital formats	no available data for mineral soils	There are restrictions on how the data can be used. Requests must include the purpose of a planned study, its time frame, and a declaration that the data will not be submitted to other parties	elena.vanguelova@forestry.gsi.gov.uk, Forest Research	http://icp-forests.net/	The purpose of the large-scale soil survey (Level I) is first of all the assessment of basic information on the chemical soil status and its changes over time, and secondly the assessment of soil properties which determine the forest soil's sensitivity to air pollution (e.g. acidification status). Besides providing soil data for the study of atmospheric deposition effects at the broader scale,the soil survey will serve other purposes, as supporting studies related to climate change (e.g.inventory of carbon storage) and sustainable forest management (e.g. nutrient and water balances studies). Monitoring intended for at least 20 years. Soil analysis planned every 10 years. Additional 5 plots established in England in 2002. BIOSOIL project is a project under the FOREST umbrella II plots have been installed in Britain in accordance with EU protocols
NATURAL ASBESTOS MAP	Based on available BGS geology maps. At large scale minor bodies of mafic and ultramafic igneous rocks, with the potential to develop asbestos minerals, are not shown because of small size of outcrops	map (pdf)	Available from BGS	Commissioned project from the Health and Safety Executive. BGS is now seeking to included these maps as a layer in its Parent Material Map. Maps are potentially sensitive and not in the general public domain	enquiries@bgs.ac.uk	none	The study is based on the potential occurrence of asbestos minerals in rocks at outcrop and does not include rock or soil moved to form made ground. Localised alteration, which has the potential to generate asbestos minerals is also not shown on the maps. The work was stimulated by The Asbestos Worker Protection Directive in the context of UK quarrying
TIPPING Hg		Peer reviewed publication data tables	Summary statitics in data tables of publication	Published information	et@ceh.ac.uk	none	

Data Set	Robustness of data	Format	Availability	IPR/Copyright information	Email contact	Web page	Additional Information
LE studies	Sampling protocols as per G-BASE (urban). Samples determined at accredited laboratories with reference materials to indicate precision and accuracy	Unpublished report	Not yet published	as for G-BASE (Urban)	enquiries@bgs.ac.uk	www.bgs.ac.uk/gbase	Pilot studies arising from G-BASE urban samples collected in Greater London Area. Some 50 soils determined for bioaccessibility of contaminants, 70 samples for organic determinations and 473 Hg samples.
MINING Hg		Peer reviewed publication data tables	Summary statitics in data tables of publication	Published information			
Diss Mere Hg	Standard solutions and quality control blanks were measured every five samples to monitor measurement stability and calibrations covering the concentration range of all the digested solutions were made using standard solutions	Peer reviewed publication	Hg data for 20 samples (11 sites) given a table	Published information	hyang@geog.ucl.ac.uk		
ADAS		Peer reviewed publication	Summary statistics presented in tables	Published information			
SULPHATE		Map, Peer reviewed publication					
MINING Pb	Analytical precision cited at 5%	Peer reviewed publication	Summary statistics and cumulative frequency distribution	Published information			Ceredigion area, Wales: samples 119; Halkyn Mountain, Wales: samples 280; Tamar Valley, SW England: samples 184; North Somerset, England: samples 169
JONES et al. PAH	Quality control procedures included the use of reference solutions, blank samples, duplicate samples and the use of standards	Peer reviewed publication	Data available with locations in data table	Published information	k.c.jones@lancaster.ac.uk		
HMIP (1989)		Peer reviewed publication	Data available in table (with locations??)	Published information			Data shows that concentrations of dioxins are higher in urban soils than rural, and show local variations in the vicinity of potential emission sources
HMIP (1995)		Peer reviewed publication					

Data Set	Robustness of data	Format	Availability	IPR/Copyright information	Email contact	Web page	Additional Information
COUSINS et al. PAH	Duplicate sample analyses performed. Reference materials and standards used	Peer reviewed publication	Site locations and data given in tables in the publication	Published information			Grid coordinates to nearest 100m
LEAD et al. PCB		Peer reviewed publication	Site locations and data given in tables in the publication	Published information			
MEIJER et al. PCB		Peer reviewed publications	Summary statistics and plots in publications	Published information	k.c.jones@lancaster.ac.uk		
URBSOIL	Control and reference materials used in multi-lab analyses	eer reviewed publications	Summary statistics and plots in publications	Published information		URBSOIL web site no longer functional. http://www.eugris.info/DisplayProject.asp?P=4452	Data sets of limited use though Hg information on European cities useful for comparison

Reference cited in the data set data base listed in record number order (Appendix 2).

1 Johnson, C.C., Ander, E.L., Lister, T.R. and Flight, D.M.A., 2008. Data conditioning of environmental geochemical data: quality control procedures used in the British Geological Survey's Regional Geochemical Mapping Project. In: B. de Vivo, H.E. Belkin and A. Lima (Editors), Environmental Geochemistry: Site characterisation, data analysis and case histories. Elsevier, Oxford, 93-118.

2 Lister, T.R. and Johnson, C.C. 2005. G-BASE data conditioning procedures for stream sediment and soil chemical analyses. British Geological Survey BGS Internal Report Number IR/05/150. 75pp.

3 Johnson, C.C., Breward, N., Ander, E.L. and Ault, L., 2005. G-BASE: Baseline geochemical mapping of Great Britain and Northern Ireland. Geochemistry: Exploration-Environment-Analysis, 5(4): 347-357.

4 Fordyce, F.M., Brown, S.E., Ander, E.L., Rawlins, B.G., O'Donnell, K.E., Lister, T.R., Breward, N. and Johnson, C.C., 2005. GSUE: Urban geochemical mapping in Great Britain. Geochemistry: Exploration-Environment-Analysis, 5(4): 325-336.

5 Johnson, C.C. 2005. 2005 G-BASE Field Procedures Manual. British Geological Survey, Keyworth, UK Internal Report No. IR/05/097. 130pp.

6 Flight, D.M.A. and Scheib, A., 2011. Soil geochemical baselines in UK urban centres:: the G-BASE project. In: C.C. Johnson, A. Demetriades, J. Locutura and R.T. Ottesen (Editors), Mapping the Chemical Environment of Urban Areas. John Wiley & Sons Ltd., Chichester, UK, 186-206.

7 Johnson, C.C., 2011. Understanding the quality of chemical data from the urban environment - Part 1: Quality Control procedures. In: C.C. Johnson, A. Demetriades, J. Locutura and R.T. Ottesen (Editors), Mapping the Chemical Environment of Urban Areas. John Wiley & Sons Ltd., Chichester, UK, pp. 61-76.

8 Reimann, C., Demetriades, A., Eggen, O.A., Filzmoser, P. and and others 2009. The EuroGeoSurveys geochemical mapping of agricultural and grazing land soils project (GEMAS) - Evaluation of quality control results of *aqua regia* extraction analysis. Geological Survey of Norway (NGU), Trondheim, Norway. Report No. 2009.049. 94pp.

9 EGS Geochemistry Working Group 2008. EuroGeoSurveys Geochemical mapping of agricultural and grazing land soil of Europe (GEMAS) - Field manual. Report for EuroGeosurveys/Eurometaux. Geological Survey of Norway (NGU), Trondheim, Norway. Report No. 2008.038. 46pp.

10 Darnley, A.G., 1997. A global geochemical reference network: the foundation for geochemical baselines. Journal of Geochemical Exploration, 60(1): 1-5.

11 Salminen, R., Batista, M.J., Bidovec, M., Demetriades, A., De Vivo, B., De Vos, W., Duris, M., Gilucis, A., Gregorauskiene, V., Halamic, J., Heitzman, P., Lima, A., Jordan, G., Klaver, G., Klein, P., Lis, J., Locotura, J., Marsina, K., Mazreku, A., O'Connor, P.J., Olsson, S.A., Ottesen, T., Petersell, V., Plant, J.A., Reeder, S., Salpeteur, I., Sandstrom, H., Siewers, U., Steenfelt, A. and Tarvainen, T., 2005. Geochemical atlas of Europe. Part 1, Background information, methodology and maps. Geological Survey of Finland, Espoo, Finland, 526 pp.

12 Darnley, A.G., Bjorklund, A., Bolviken, B., Gustavsson, N., Koval, P.V., Plant, J.A., Steenfelt, A., Tauchid, M., Xuejing, Xie 1995. A global geochemical database for environmental and resource management. UNESCO publishing 19.

13 Salminen, R., Tarvainen, T., Demetriades, A., Duris, M., Fordyce, F.M., Gregorauskiene, V., Kahelin, H., Kivisilla, J., Klaver, G., Klien, H., Larson, J.O., Lis, J., Locutura, J., Marsina, K., Mjartanova, H., Mouvet, C., O'Connor, P., Odor, L., Ottonello, G., Paukola, T., Plant, J.A., Reimann, C., Schermann, O., Siewers, U., Steenfelt, A., Van der Sluys, J., de Vivo, B. and Williams, L. 1998. FOREGS Geochemical Mapping Field Manual. Geological Survey of Finland Guide 47. 36pp.

14 De Vos, W. and Tarvainen, T. (Editors), 2006. Geochemical Atlas of Europe. Part 2. Interpretation of Geochemical Maps, Additional Tables, Figures, Maps, and Related Publications. Geological Survey of Finland, Espoo, Finland, 692 pp.

15 McGrath, S.P. and Loveland, P.J., 1992. The Soil Geochemical Atlas of England and Wales. Blackie Academic and Professional, Glasgow.

16 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. Defra (formerly MAFF) Soil Programme Technical Report Project SP0124.

17 Rawlins, B.G., McGrath, S.P., Scheib, A.J., Cave, M.R., Breward, N., Lister, T.R., Ingham, M., Gowing, C.J.B. and Carter, S., 2012. The Advanced Soil Geochemical Atlas of England and Wales. British Geological Survey, Keyworth, Nottingham. Published as an electronic book. Available on-line at: www.bgs.ac.uk/gbase/advsoilatlasEW.html last accessed 1st May 2012.

18 Black, H.I.J., Garnett, J.S., Ainsworth, G., Coward, P.A., Creamer, R., Ellwood, S., Horne, J., Hornung, M., Kennedy, V.H., Monson, F., Raine, L., Osborn, D., Parekh, N.R., Parrington, J., Poskitt, J.M., Potter, E., Reeves, N., Rowland, A.P., Self, P., Turner, S., Watkins, J., Woods, C. and Wright, J. 2002. MASQ: Monitoring and Assessing Soil Quality in Great Britain. Countryside Survey Module 6: Soils and Pollution.

Centre for Ecology and Hydrogeology (CEH) and Environment Agency (EA) UK. R&D Technical Report E1-063/TR. 207pp.

19 Barraclough, D., 2007. UK Soil and Herbage Pollutant Survey. Introduction and Summary. Environment Agency (EA), UK.

20 Wood, M.D., Copplestone, D. and Crook, P. 2007. UK Soil and Herbage Pollutant Survey (UKSHS). Chemical and radiometric sample collection methods. UK Environment Agency Report No.2. 32pp.

21 Fox, I., Hunter, C. and Hudson, R. 2007. UK Soil and Herbage Pollutant Survey (UKSHS). Chemical analytical methods. UK Environment Agency Report No.3. 38pp.

22 Haslam, H.W., Cameron, D.G. and Evans, A.D. 1990. The Mineral Reconnaissance Programme in the United Kingdom. A continuing programme of mineral exploration in the United Kingdom carried out by the British Geological survey on behalf of the Department of Trade & Industry. Mineral Reconnaissance Programme Report No. 114. British Geological Survey, Keyworth, Nottingham, UK.

23 NERC 1999. NERC Data Policy Handbook Version 2.1. NERC Scientific Services: Swindon, UK.

24 Hornung, M., Beard, G.R., Sykes, J.M. and Wilson, J.M. (Editors), 1996. The United Kingdom Environmental Change Network: Protocols for Standard Measurements at Terrestrial Sites. NERC (version 1.1 updated Dec 2001).

25 Lane, A.M.J., 1996. The ECN database, The United Kingdom Environmental Change Network: Protocols for standard measurements at terrestrial sites. HMSO London, pp. 7.

26 Lane, A.M.J., 1996. The ECN data reporting. In: J.M. Sykes and A.M.J. Lane (Editors), The UK Environmental Change Network Protocols for Standard Measurements at Terrestrial Sites. HMSO London, pp. 37.

27 Sykes, J.M., 1996. Introduction. In: J.M. Sykes and A.M.J. Lane (Editors), The UK Environmental Change Network Protocols for Standard Measurements at Terrestrial Sites. HMSO London, pp. 20.

28 Sykes, J.M. and Lane, A.M.J. (Editors), 1996. The United Kingdom Environmental Change Network: protocols for standard measurements at terrestrial site. Stationery Office, London.

29 Montanarella, L., Tóth, G. and Jones, D.G., 2011. Soil component in the 2009 LUCAS Survey. In: G. Tóth and T. Németh (Editors), The international conference on ’Land Quality and Land Use Information in the European Union’ 26-27 May 2011 in Keszthely, Hungary. European Commission, Joint Research Centre, Institute for Environment and Sustainability, Keszthely, Hungary, pp. 209-218.

30 Eurostat 2009. LUCAS 2009: Quality Control Procedures. European Commission, Eurostat, Directorate E: Agriculture and environment statistics Technical reference document C-4. 10pp.

31 Eurostat 2009. LUCAS 2009: Quality Assurance. European Commission, Eurostat, Directorate E: sectoral and regional statistics Technical reference document C-4. 8pp.

32 Durrant, D., 2000. Environmental monitoring in British forests Information Note 37. Forestry Commission, Edinburgh.

33 Vanguelova, E., Barsoum, N., Benham, S., Broadmeadow, M., Moffat, A.J., Nisbet, T.R. and Pitman, R. , 2007. Ten years of long-term environmental monitoring in British forests. Forestry Commission Information Note 88, Edinburgh.

34 Jones, K.C., Stratford, J.A., Waterhouse, K.S. and Vogt, N.B., 1989. Organic contaminants in Welsh soils - polynuclear aromatic hydrocarbons. Environmental Science & Technology, 23(5): 540-550.

35 Jones, K.C., Stratford, J.A., Waterhouse, K.S., Furlong, E.T., Giger, W., Hites, R.A., Schaffner, C. and Johnston, A.E., 1989. Increases in the polynuclear aromatic hydrocarbon content of an agricultural soil over the last century. Environmental Science & Technology, 23(1): 95-101.

37 Reimann, C., de Caritat, P., GEMAS Project Team and NGSA Project Team. 2012. New soil composition data for Europe and Australia: demonstrating comparability, identifying continental-scale processes and learning lessons for global geochemical mapping. Science of The Total Environment, 416, 239-252.

38 Emmett, B.A., Frogbrook, Z.L., Chamberlain, P.M., Griffiths, R., Pickup, R., Poskitt, J.M., Reynolds, B., Rowe, E., Spurgeon, D.J., Brittain, S.A., Rowland, P., Wilson, J. and Wood, C.M. 2008. Countryside Survey. Soils Manual. Centre for Ecology & Hydrology CS Technical Report No. 3/07. 180pp.

39 Emmett, B.A., Reynolds, B., Chamberlain, P.M., Rowe, E., Spurgeon, D.J., Brittain, S.A., Frogbrook, Z.L., Hughes, S., Lawlor, A.J., Poskitt, J.M., Potter, E., Robinson, D.A., Scott, A., Wood, C.M. and Woods, C. 2010. Countryside Survey. Soils Report from 2007. Centre for Ecology & Hydrology CS Technical Report No. 9/07. 194pp.

40 BGS, 2003. Possible natural occurrence of asabestos minerals: The Lake District. In: R.J. Merriman (Editor). British Geological Survey, Keyworth, Nottingham, UK, pp. Commissioned Project funded by the Health & Safety Executive (Contract GA/02/79).

41 BGS, 2003. Possible natural occurrence of asabestos minerals: Southwest England. In: R.J. Merriman (Editor). British Geological Survey, Keyworth, Nottingham, UK, pp. Commissioned Project funded by the Health & Safety Executive (Contract GA/02F/79).

42 Tipping, E., Poskitt, J.M., Lawlor, A.J., Wadsworth, R.A., Norris, D.A. and Hall, J.R., 2011. Mercury in United Kingdom topsoils; concentrations, pools, and Critical Limit exceedances. Environmental Pollution, 159: 3721 - 3729.

43 Tipping, E., Wadsworth, R.A., Norris, D.A., Hall, J.R. and Ilyin, I., 2011. Long-term mercury dynamics in UK soils. Environmental Pollution, 159(12): 3474-3483.

44 Davies, B.E., 1976. Mercury content of soils in Western Britain with special reference to contamination from base metal mining. Geoderma, 16(3): 183-192.

45 Yang, H.D., 2010. Historical mercury contamination in sediments and catchment soils of Diss Mere, UK. Environmental Pollution, 158(7): 2504 - 2510.

46 Archer, F.C. and Hodgson, I.H., 1987. Total and extractable trace element contents of soils in England and Wales. Journal of Soil Science, 38: 421-432.

47 Forster, A., Culshaw, M.G., Bell, F.G, 1995. Regional distribution of sulphate in rocks and soils of Britain. In: M. Eddleston, Walthall, S, Cripps, J.C., Culshaw, M.G. (Editor), Engineering geology of construction. Geological Society of London London, pp. 95-104.

48 Davies, B.E., 1983. A graphical estimation of the normal lead content of some British soil. Geoderma, 29(1): 67 - 75.

49 Creaser, C.S., Fernandes, A.R., Harrad, S.J., Hurst, T. and Cox, E.A., 1989. Background levels of polychlorinated biphenyls in British soils - II. Chemosphere, 19(8-9): 1457-1466.

50 HMIP 1989. *Determination of polychlorinated biphenyls, polychlorinated dibenzo*-p-*dioxins and polychlorinated dibenzofurans in UK soils. . Technical report*. London: HMSO. Her Majesty's Inspectorate of Pollution (HMIP)

51 Cox, E.A. and Creaser, C.S. 1995. Determination of polychlorinated biphenyls, polychlorinated dibenzo-p-dioxins and polychlorinated dibenzofurans in UK soils. Second Technical Report to Her Majesty's Inspectorate of Pollution (HMIP).

52 HMIP 1995. *Determination of polychlorinated biphenyls, polychlorinated dibenzo*-p-*dioxins and polychlorinated dibenzofurans in UK soils. . Second technical report*. London: HMSO. Her Majesty's Inspectorate of Pollution (HMIP)

53 Badsha, K. and Eduljee, G., 1986. PCB in the U.K. environment - A preliminary survey. Chemosphere, 15(2): 211-215.

54 Badsha, K., Eduljee, G. and Scudamore, N., 1986. Environmental monitoring for PCB and trace metals in the vicinity of a chemical waste disposal facility - III. Chemosphere, 15(7): 947-957.

55 Eduljee, G., Badsha, K. and Price, L., 1985. Environmental monitoring for PCB and heavy metals in the vicinity of a chemical waste disposal facility - I. Chemosphere, 14(9): 1371-1382.

56 Eduljee, G., Badsha, K. and Scudamore, N., 1986. Environmental monitoring for PCB and trace metals in the vicinity of a chemical waste disposal facility - II. Chemosphere, 15(1): 81-93.

57 Eduljee, G.H., Badsha, K.S. and Mundy, K.J., 1987. PCB concentrations in soil from central and southern Wales. Chemosphere, 16(7): 1583-1598.

58 Cousins, I.T., Kreibich, H., Hudson, L.E., Lead, W.A. and Jones, K.C., 1997. PAHs in soils: contemporary UK data and evidence for potential contamination problems caused by exposure of samples to laboratory air. Science of The Total Environment, 203(2): 141-156.

59 Lead, W.A., Steinnes, E., Bacon, J.R. and Jones, K.C., 1997. Polychlorinated biphenyls in UK and Norwegian soils: spatial and temporal trends. Science of The Total Environment, 193(3): 229-236.

60 Hassanin, A., Lee, R.G.M., Steinnes, E. and Jones, K.C., 2005. PCDD/Fs in Norwegian and U.K. Soils: Implications for Sources and Environmental Cycling. Environmental Science & Technology, 39(13): 4784 - 4792.

61 Meijer, S.N., Ockenden, W.A., Sweetman, A., Breivik, K., Grimalt, J.O. and Jones, K.C., 2003. Global Distribution and Budget of PCBs and HCB in Background Surface Soils: Implications for Sources and Environmental Processes. Environmental Science & Technology, 37(4): 667-672.

62 Meijer, S.N., Steinnes, E., Ockenden, W.A. and Jones, K.C., 2002. Influence of Environmental Variables on the Spatial Distribution of PCBs in Norwegian and U.K. Soils: Implications for Global Cycling. Environmental Science & Technology, 36(10): 2146 - 2153.

63 Nam, J.J., Thomas, G.O., Jaward, F.M., Steinnes, E., Gustafsson, O. and Jones, K.C., 2008. PAHs in background soils from Western Europe: Influence of atmospheric deposition and soil organic matter. Chemosphere, 70(9): 1596 - 1602.

64 Rodrigues, S., Pereira, M.E., Duarte, A.C., Ajmone-Marsan, F., Davidson, C.M., Grcman, H., Hossack, I., Hursthouse, A.S., Ljung, K., Martini, C., Otabbong, E., Reinoso, R., Ruiz-Cortes, E., Urquhart, G.J. and Vrscaj, B., 2006. Mercury in urban soils: A comparison of local spatial variability in six European cities. Science of the Total Environment, 368(2-3): 926-936.

65 Madrid, L., Diaz-Barrientos, E., Ruiz-Cortés, E., Reinoso, R., Biasioli, M., Davidson, C.M., Duarte, A.C., Greman, H., Hossack, I., Hursthouse, A.S., Kralj, T., Ljung, K., Otabbong, E., Rodrigues, S., Urquhart, G.J. and

Ajmone-Marsan, F., 2006. Variability in concentrations of potentially toxic elements in urban parks from six European cities. Journal of Environmental Monitoring, 8: 1158-1165.

66 Cools, N. and de Vos, B., 2010. Sampling and Analysis of Soil. Manual Part X Manual on methods and criteria for harmonized sampling, assessment, monitoring and analysis of the effects of air pollution on forests. UNECE, ICP Forests, Hamburg. , pp. 208.

67 Spurgeon, D.J., Rowland, P., Ainsworth, G., Rothery, P., Long, S. and Black, H.I.J., 2008. Geographical and pedological drivers of distribution and risks to soil fauna of seven metals (Cd, Cu, Cr, Ni, Pb, V and Zn) in British soils. Environmental Pollution, 153(2): 273-283.

68 Heywood, E., Wright, J., Wienburg, C.L., Black, H.I.J., Long, S.M., Osborn, D. and Spurgeon, D.J., 2006. Factors influencing the national distribution of polycyclic aromatic hydrocarbons and polychlorinated biphenyls in British soils. Environmental Science & Technology, 40(24): 7629-7635.

Glossary and Abbreviations

Glossary of some of the terms used in this report (alphabetical order). Those marked with an '*' are defined in the draft Statutory Guidance (Defra, 2011b).

alkaline rock	rocks formed from a magma that are enriched with sodium and potassium bearing minerals (relative to silica)
ambient background	concentration levels which are predominantly geogenic but with some anthropogenic input
anthropogenic	having an origin associated with human activity
argillaceous rock	a sedimentary rock formed from clay deposits
asbestos	a generic term defined by industry referring to certain silicate minerals crystallised in a finely fibrous habit, in bundles of easily separable fibres and with a hair-like elongated shape, thermally stable, with high tensile strengths, smooth faces and with an adamantine or silky lustre
background	see "normal background"
baseline	in geochemistry used to describe the spatial distribution of an element. It is usually defined by extrapolating between sample sites to model the varying spatial distribution of a chemical element
basic rock	a quartz (silica) free igneous rock
boxplot	a graphical representation of a range of values where the length of the box represents the spread of values. Sometimes called a box-and-whisker plot as lines (whiskers) are added to the box to further demonstrate a range of results
bioaccessibility	is the fraction of a substance that is released into solution from the soil during digestion making it available for absorption
bioavailability	can be defined as the fraction of an ingested contaminant from the soil that is released into solution and subsequently absorbed which reaches the systemic circulation where it may then cause adverse effects on human health
built environment	human-made environment associated with activity from the scale of personal shelter and buildings to neighbourhoods and cities including their supporting infrastructure, such as transport networks
contaminant*	a substance relevant to the Part 2A regime (DETR, 2000) which is in, on or under the land and which has the potential to cause significant harm to a relevant receptor, or to cause significant pollution of controlled waters. Synonymous with terms "pollutant" and "substance"
contaminated land*	is any land which appears to the local authority in whose area it is situated to be in such a condition, by reason of substances in, on or under the land that – (a) significant harm is being caused or there is a significant possibility of such harm being caused; or (b) significant pollution of controlled waters is being caused, or there is a significant possibility of such pollution being caused

deep soil	generally a soil from below the ploughing depth (> 30cm) and is more likely to be influenced with contributions to its physical and chemical properties from the underlying parent material
domain	a region defined by a boundary derived from a soil's underlying parent material, an urbanisation index, or an area of metalliferous mineralisation/mining and is used in this report to nationally characterise significantly different areas of NBCs
diffuse pollution	pollution from widespread human activities with no one discrete source. These activities may be recent or have been carried out in the past but cannot be tied down to a specific location or source. Examples of diffuse pollution include atmospheric deposition of contaminants arising from industry, domestic and industrial coal combustion and traffic exhaust, and disposal of domestic coal ash.
dioxins	polychlorinated dibenzodioxins and dibenzofurans
duplicate	a sample that is collected from the same site as another sample. A duplicate is a control sample that can be used to monitor the quality of data produced from a sampling and analytical strategy
Gaussian	a term used in statistics, the Gaussian (or normal) distribution is a continuous probability distribution that has a bell-shaped probability density function, known informally as the bell curve
geochemistry	is the study of the distribution of chemical elements in and at the earth's surface
geogenic	pertaining to a geological origin
granitic rock	is an intrusive igneous rock characterised by large mineral crystals and is felsic in nature, *i.e*. consisting of >69% of silicate minerals, like quartz, alkali feldspars, and micas.
Guidance*	see Statutory Guidance
interquartile range	a statistical term being a measure of dispersion equal to the difference between the upper and lower quartiles
intrusive rock	an igneous rock formed from magma underground meaning that it will have cooled slowly resulting in a rock with large visible mineral crystals
lithology	is a geological term referring to the general characteristics of a rock formation focusing on macroscopic hand-sample or outcrop-scale description of rocks
mafic rock	is an igneous rock that is dominated by the silicate minerals pyroxene, amphibole, olivine, and mica. These are high in Mg and Fe oxides, and their presence gives the rock a characteristic dark colour
mean	a term used in statistics to quantify the average of a group of numbers determined by calculating the total of all the numbers and dividing this by the number of values used in the calculation

median	a term used in statistics for the middle result in a sorted list of results. There are therefore 50% of the results below this value
micaceous	an adjective used to describe something that contains mica minerals which are characteristically flat in shape and often glitter
normal*	a term used to describe contaminant concentrations when they are seen as typical and encompasses contributions both from geogenic sources and diffuse anthropogenic pollution. Normal Background Concentrations (NBCs) are an expression of normal contaminant levels and should not be considered to cause land to qualify as contaminated land, unless there is a particular reason to consider otherwise
Part 2A*	means Part 2A of the Environmental Protection Act 1990 (as amended)
pedological	relating to soils and the processes associated with them
percentile	a statistical term based on ranking numbers in order and assigning a value based on below which a certain percentage of observations fall. It is therefore a non-parametric measure
pollutant	see "contaminant"
quantile	quantiles are values which divide the distribution of an ordered list of results such that there is a given proportion of results below the quantile. For example, the median is a quantile
quartile	quartiles are a set of values being the three points that divide a data set into four equal groups, each representing a fourth of the population being sampled
replicate	a sample that is created by subdividing another sample following a strictly defined protocol. A replicate is a control sample that can be used to monitor the quality analytical methodology
rural	areas not categorised as industrial or urban. Predominantly agricultural land or undeveloped countryside. In this report quantitatively defined using an urbanisation index
semi-urban	A classification of land use defined in this report by the urbanisation index and is intermediary between urban and rural land use
skewness	a statistical term that is a measure of the asymmetry of the probability distribution of a population of results
speciation	chemical elements can exists in differing chemical forms or species which can be described as speciation. Different forms of an element can be associated with differing potential to cause harm, *e.g*. Cr(III) and Cr (VI)
standard deviation	a statistical term that is a measure of how spread out numbers are. It is quantified by taking the square root of the variance

Statutory Guidance*	is issued by the Secretary of State for Environment, Food and Rural Affairs in accordance with section 78YA of the Environmental Protection Act 1990 ("the 1990 Act"). It is intended to explain how local authorities should implement the regime, including how they should go about deciding whether land is contaminated land in the legal sense of the term.
substance	see "contaminant"
surface soil	a surface soil is generally a sample of soil from the top several cm of a soil profile and so will greatly be influenced by additions of substances to the surface (*e.g.* atmospheric deposition) and are likely have a higher organic content than soils lower in the profile
topsoil	is generally a sample of mineral soil collected from the top 30 cm of a soil profile and so represents a combination of both geogenic and anthropogenic processes that will influence the soil's chemistry
ultrabasic rock	see basic rock. 'Ultra' implies that it is very basic
ultramafic rock	see mafic rock. 'Ultra' implies that it is very mafic
unacceptable risk*	a risk of such a nature that it would give grounds to be considered contaminated land under Part 2A
urban area	from a physical scientist's perspective an urban area is a built environment (Johnson and Ander, 2008), *i.e.* an area with significant anthropogenic modification. The UK Soil and Herbage Pollutant Survey (Wood *et al*., 2007), for example, defines urban as being an area which is ≥90% built-up
urbanisation	see "urban area"
urbanisation index	in this report it is an estimate used to indicate the degree to which an area has been subjected to anthropogenic influence and is derived from the GLUD statistics and is the ratio of built space to open space using the Census Area Statistical Wards (CASW)
variance	is a statistical term calculated for a spread of results by taking the average of the squared differences from the mean

Abbreviations listed in alphabetical order.

AAS	atomic adsorption spectrometry
ABER	Aberystwyth University
BaP	Benzo[*a*]pyrene
BARGE	Bioaccessibility Research Group of Europe
BGR	Bundesanstalt für Geowissenschaften und Rohstoffe (Institute of Geosciences and Natural Resources, Germany)
BGS	British Geological Survey
BNG	British National Grid
C4SL	category 4 screening level
CASW	Census Area Statistical Wards
CD	compact disk
CEH	Centre for Ecology and Hydrology
CRM	certified reference material
CS:2000	Countryside Survey 2000
Defra	Department for Environment, Food and Rural Affairs
DETR	Department of the Environment, Transport & the Regions
DoE	Department of Environment UK (now included in Defra role)
EA	Environment Agency (UK)
EC	European Commission
ECN	Environmental Change Network
ED-XRFS	energy dispersive X-ray fluorescence spectrometry
EDTA	ethylene diamine tetra acetic acid
EGS	EuroGeoSurveys
EU	European Union
FOREGS	FORum of European Geological Surveys
FR	Forest Research (UK)
G-BASE	Geochemical BAseline Survey of the Environment
GC	gas chromatography
GC-MS	gas chromatography mass spectrometry
GEMAS	GEochemical Mapping of Agricultural and grazing Soils
GLA	Greater London Authority
GLUD	Generalised Land Use Database
GPC	gel permeation chromatography
GTK	Geologian Tutkimuskeskuksen (Geological Survey of Finland)
HCB	hexachlorobenzene
HPLC	high-performance liquid chromotography
HRGC-MS	high resolution gas chromatography mass spectrometry
HSE	Health & Safety Executive (UK)
idw	inverse distance weighted
ICP-AES	inductively coupled plasma atomic emission spectrometry
ICP-MS	inductively coupled plasma mass spectrometry
ICP-OES	inductively coupled plasma optical emission spectrometry
IPR	intellectual property rights
ISO	International Standards Organisation
JRC	Joint Research Centre (EU)
LandIS	Land Information System
LE	London Earth (G-BASE project)
LOI	loss on ignition
LRMS	low resolution mass spectrometry
LUCAS	Land Use and Cover Area frame Survey
MASQ	Monitoring and Assessing Soil Quality
MRP	Mineral Reconnaissance Programme
NAA	neutron activation analysis
NBC	normal background concentration
NERC	Natural Environment Research Council (UK)
NLS	National Laboratory Service
NLUD	National Land Use Database
NOA	Naturally Occurring Asbestos
NSI	National Soil Inventory
NSRI	National Soil Resources Institute
ONS	Office of National Statistics (UK)

OS	Ordnance Survey
OSHA	Occupational Safety & Health Administration (USA)
PAH	Polynuclear(cyclic) aromatic hydrocarbon
PCB	polychlorinated biphenyl
PCDD	polychlorinated dibenzo-*p*-dioxin
PCDF	polychlorinated dibenzofuran
POP	persistent organic pollutant
ppb	parts per billion (equivalent to mg/g)
ppm	parts per million (equivalent to mg/kg)
PPM	parent material map (BGS)
PTV	programmed temperature vaporisation
QA	quality assurance
QC	quality control
REACH	Registration, Evaluation and Authorisation of Chemicals
SGV	soil guideline value
SI	Système international [d'unités]
SPE	solid phase extraction
SPPM	soil-parent material model
SQL	standard query language
SRM	secondary reference material
TOC	total organic carbon
UEA	University of East Anglia
UCL	University College London
UI	urbanisation index
UK	United Kingdom
UKAS	United Kingdom Accreditation Service
USEPA	United States Environmental Protection Agency
UKSHS	UK Soil and Herbage pollutant Survey
WD-XRFS	wavelength-dispersive X-ray fluorescence spectrometry
WP	work package
XRFS	X-ray fluorescence spectrometry

References

Abrahams, P.W. and Thornton, I. 1987. Distribution and extent of land contamination by arsenic and associated metals in mining regions of southwestern England. Transactions of the Institution of Mining and Metallurgy Section B-Applied Earth Science, 96, B1-B8.

Andersson, M., Holt, Y. and Eggen, O.A., 2011. Polychlorinated dibenzo-*p*-dioxins and dibenzofurans (PCDDs/PCDFs) in urban surface soil in Norway. *In*: C.C. Johnson, A. Demetriades, J. Locutura and R.T. Ottesen (Editors), *Mapping the Chemical Environment of Urban Areas.* John Wiley & Sons Ltd., Chichester, UK, 473-486.

Appleton, J. D. 1995. Potentially harmful elements from natural sources and mining areas: characteristics, extent and relevance to planning and development in Great Britain. British Geological Survey. Technical Report WP/95/3. Commission Report for the Department of Environment (DoE), UK.

Aston, S.R., Thornton, I., Webb, J.S., Milford, B.L., and Purves, J.B. 1975. Arsenic in stream sediments and waters of southwest England. Science of The Total Environment, 4(4), 347-358.

ATSDR. 1995. Thallium ToxGuide. CAS7440-28-0. September 1995. Agency for Toxic Substances and Disease Registry, Atlanta, GA: U.S. Department of Health and Human Services, Public Health Services. http://www.atsdr.cdc.gov Last accessed 12 December 2011.

Avery, B.W. 1990. *Soils of the British Isles*. Wallingford, Oxfordshire, CAB International.

Bacon, J.R., Jones, K.C., McGrath, S.P. and Johnston, A.E. 1996. Isotopic character of lead deposited from the atmosphere at a grassland site in the United Kingdom since 1860. Environmental Science & Technology, 30(8), 2511-2518.

Barbara, M.-K. 1996. Polycyclic aromatic hydrocarbons in agricultural soils in Poland: preliminary proposals for criteria to evaluate the level of soil contamination. Applied Geochemistry, 11(1-2), 121-127.

Bellis, D.J., Satake, K. and McLeod, C.W. 2004. A comparison of lead isotope ratios in the bark pockets and annual rings of two beech trees collected in Derbyshire and South Yorkshire, UK. Science of The Total Environment, 321(1-3), 105-113.

Butler, J.D., Butterworth, V., Kellow, S.C. and Robinson, H.G. 1984. Some observations on the polycyclic aromatic hydrocarbon (PAH) content of surface soils in urban areas. Science of The Total Environment, 33, 75-85.

BGS, 2003a. Possible natural occurrence of asbestos minerals: The Lake District. In: R.J. Merriman (Editor). British Geological Survey, Keyworth, Nottingham, UK, Commissioned Project funded by the Health & Safety Executive (Contract GA/02/79). Unpublished. Confidential.

BGS, 2003b. Possible natural occurrence of asbestos minerals: Southwest England. In: R.J. Merriman (Editor). British Geological Survey, Keyworth, Nottingham, UK, Commissioned Project funded by the Health & Safety Executive (Contract GA/02F/79). Unpublished. Confidential.

BGS. 2010. Environmental Geochemical Atlas of Central and Eastern England. British Geological Survey Electronic Atlas, available from http://www.bgs.ac.uk/downloads/start.cfm?id=1633 . Last accessed 12 December 2011.

BGS, 2011. Proposal Form (EVID2) for "Establishing data on normal/background levels of soil contamination in England". Defra Research Project Proposal call CTE1118, British Geological Survey, Keyworth, Nottingham, UK. 5th September 2011. *Commercial in Confidence*.

Biasiolia, M. and Ajmone-Marsan, F., 2007. Organic and inorganic diffuse contamination in urban soils: the case of Torino (Italy). Journal of Environmental Monitoring, 9, 862-868.

Biasiolia, M., Barberis, R. and Ajmone-Marsan, F. 2006. The influence of a large city on some soil properties and metals content. Science of The Total Environment, 356, 154-164.

Black, H.I.J., Garnett, J.S., Ainsworth, G., Coward, P.A., Creamer, R., Ellwood, S., Horne, J., Hornung, M., Kennedy, V.H., Monson, F., Raine, L., Osborn, D., Parekh, N.R., Parrington, J., Poskitt, J.M., Potter, E., Reeves, N., Rowland, A.P., Self, P., Turner, S., Watkins, J., Woods, C. and Wright, J. 2002. MASQ: Monitoring and Assessing Soil Quality in Great Britain. Countryside Survey Module 6: Soils and Pollution. Centre for Ecology and Hydrogeology (CEH) and Environment Agency (EA) UK. R&D Technical Report E1-063/TR. 207pp.

Bowman, A.W. and Azzalini, A. 1997. *Applied Smoothing Techniques for Data Analysis - The Kernel Approach with S-Plus Illustrations*. Clarendon Press, Oxford.

Breward, N. 2007. Arsenic and presumed resistate trace element geochemistry of the Lincolnshire (UK) sedimentary ironstones, revealed by a regional geochemical survey using soil, water and stream sediment sampling. Applied Geochemistry, 22 (9), 1970-1993.

Cave, M.R., Wragg, J., Palumbo, B. and Klinck, B.A. 2002. Measurement of the bioaccessibility of Arsenic in UK Soils. BGS-EA, P5-062/TR02.

Clarke, M.C. and Linkson, P.B. 1987. The toxicity and degradation of cyanide in mine tailings. 87 - 94. In: *Mining and environment- a professional approach*. The Australian Institute of Mining and Metallurgy. Victoria. Australia.

Colbourn, P. and Thornton, I. 1978. Lead pollution in agricultural soils. Journal of Soil Science, 29(4), 513-526.

Cole, S. and Jeffries, J. 2009. Using Soil Guideline Values. Environment Agency, UK SC050021/SGV Introduction. 32pp.

Communities and Local Government. 2007. Generalised Land Use Database Statistics for England 2005. Department for Communities and Local Government. Product Code 06CSRG04342. February 2007. Available online: http://www.communities.gov.uk/publications/planningandbuilding/generalisedlanduse last accessed 29 March 2012.

Cools, N. and de Vos, B., 2010. Sampling and Analysis of Soil. Manual Part X Manual on methods and criteria for harmonized sampling, assessment, monitoring and analysis of the effects of air pollution on forests. UNECE, ICP Forests, Hamburg. , pp. 208.

Cousins, I.T., Kreibich, H., Hudson, L.E., Lead, W.A. and Jones, K.C., 1997. PAHs in soils: contemporary UK data and evidence for potential contamination problems caused by exposure of samples to laboratory air. Science of The Total Environment, 203(2), 141-156.

Creaser, C.S., Fernandes, A.R., Harrad, S.J., Hurst, T. and Cox, E.A. 1989. Background levels of polychlorinated biphenyls in British soils - II. Chemosphere, 19(8-9), 1457-1466.

Creaser, C. S., Wood, M. D., Alcock, R. E., Copplestone, D. and Crook, P. J. 2007, UK Soil and Herbage Pollutant Survey. Environmental concentrations of polycyclic aromatic hydrocarbons in UK soil and herbage. UKSHS Report No.9. Environment Agency. UK

Culbard, E.B., Thornton, I., Watt, J.M., Wheatley, M., Moorcroft, S. and Thompson, M. 1988. Metal contamination in British urban dusts and soils. Journal of Environmental Quality, 17, 226-234.

D'Arcy, B.J., Ellis, B., Ferrier, R., Jenkins, J. and Dils, R. (Editors), 2000. *Diffuse Pollution Impacts: the Environmental and Economic Impacts of Diffuse Pollution in the UK*. Chartered Institute of Water and Environmental Management (CIWEM), London.

Darnley, A.G., Bjorklund, A., Bolviken, B., Gustavsson, N., Koval, P.V., Plant, J.A., Steenfelt, A., Tauchid, M. and Xuejing, X. 1995. A *global geochemical database for environmental and resource management*. UNESCO publishing 19.

Davies, B.E. 1980. *Applied Soil Trace Elements*. John Wiley & Sons.

Davies, B.E. 1983. A graphical estimation of the normal lead content of some British soil. Geoderma, 29(1), 67 - 75.

Defra-EA 2002. Potential contaminants for the assessment of land. Department for Environment, Food and Rural affairs and Environment Agency (UK). 34pp.

Defra, 2008. Guidance on the Legal Definition of Contaminated Land, Department for Environment Food and Rural Affairs (Defra, UK)

Defra, 2011a. Simplification of contaminated land statutory guidance. Impact Assessment. IA No: Defra1133, Department for Environment, Food and rural Affairs (Defra)/Welsh Assembly Government, UK. 6th October 2011.

Defra, 2011b. Draft Contaminated Land Statutory Guidance – end of 2011 version. Laid before Parliament 8th February 2012 .

DETR, 2000. Environmental Protection Act 1990: Part IIA Contaminated Land. DETR Circular 02/2000, Department of the Environment, Transport & the Regions (DETR), UK. 20th March 2000. 162pp. London: HMSO.

Dzombak, D.A., Ghosh, R.S. and Wong-Chong, G.M. 2006. Cyanide in water and soil: chemistry, risk, and management. CRC Taylor & Francis, 602 pp.

Emmett, B.A., Reynolds, B., Chamberlain, P.M., Rowe, E., Spurgeon, D.J., Brittain, S.A., Frogbrook, Z.L., Hughes, S., Lawlor, A.J., Poskitt, J.M., Potter, E., Robinson, D.A., Scott, A., Wood, C.M. and Woods, C. 2010. Countryside Survey. Soils Report from 2007. Centre for Ecology & Hydrology CS Technical Report No. 9/07. 194pp.

Finnish Government. 2007. Government Decree on the Assessment of Soil Contamination and Remediation Needs. 1st March 2007. Decree 214/2007. Ministry of the Environment. (In Finnish – Unofficial translation available).

Fordyce, F.M., 2005. Selenium deficiency and toxicity in the environment. *In*: O. Selinus et al. (Editors), *Essentials of Medical Geology*. Elsevier, London, 373-415.

Fordyce, F.M. and Ander, E.L. 2003. Urban soils geochemistry and GIS-aided interpretation – a case study from Stoke-on-Trent. British Geological Survey Research Report IR/01/35 (unpublished). Available online at http://nora.nerc.ac.uk/7018/ Last accessed 12th December 2011.

Fordyce, F.M., Brown, S.E., Ander, E.L., Rawlins, B.G., O'Donnell, K.E., Lister, T.R., Breward, N. and Johnson, C.C. 2005. GSUE: Urban geochemical mapping in Great Britain: Geochemistry: Exploration-Environment-Analysis, 5(4), 325-336.

Forster, A., Culshaw, M.G. and Bell, F.G. 1995. Regional distribution of sulphate in rocks and soils of Britain. 95-104. *In*: Eddleston, M., Walthall, S., Cripps, J.C. and Culshaw, M.G. (Editors*), Engineering geology of construction*. Geological Society of London, London,

Frazell, J., Elkins, R., O'Green, A.T., Reynolds,R. and Meyers, J. 2009. Facts about serpentine rocks and soil containing asbestos in California. University of California Division of Agriculture and Natural Resources Publication 8399.

Fuge, R., Glover, S.P., Pearce, N.J.G. and Perkins, W.T. 1991. Some observations on heavy-metal concentrations in soils of the Mendip Region of North Somerset. Environmental Geochemistry and Health, 13(4), 193-196.

Harper, M., 2008. 10th Anniversary Critical Review: Naturally occurring asbestos. Journal of Environmental Monitoring, 10, 1394-1408.

Hartigan, J. A., and Wong, M. A. 1979. A K-means clustering algorithm. Applied Statistics, 28, 100-108.

Heywood, E., Wright, J., Wienburg, C.L., Black, H.I.J., Long, S.M., Osborn, D. and Spurgeon, D. J. 2006. Factors influencing the national distribution of polycyclic aromatic hydrocarbons and polychlorinated biphenyls in British soils. Environmental Science & Technology, 40(24), 7629-7635.

ISO: 10381-2. 2002. Soil quality: Sampling – Part 2: Guidance on sampling techniques. ISO 10381-2:2002. International Organisation for Standardisation.

ISO 19258:2011. 2011. Soil quality: Guidance on the determination of background values. ISO 19258:2005. International Organisation for Standardisation. Corrigendum 31st August 2011.

Jensen, H., Reimann, C., Finne, T.E., Ottesen, R.T. and Arnoldussen, A. 2007. PAH-concentrations and compositions in the top 2 cm of forest soils along a 120 km long transect through agricultural areas, forests and the city of Oslo, Norway. Environmental Pollution, 145(3), 829-838.

Johnson, C.C. 2005. 2005 G-BASE Field Procedures Manual. British Geological Survey, Keyworth, UK Internal Report No. IR/05/097. 130pp. Available online at http://nora.nerc.ac.uk/5190/ . Last accessed 12th December 2011.

Johnson, C.C. and Ander, E.L. 2008. Urban Geochemical Mapping Studies: How and why we do them. Environmental Geochemistry and Health, 30, 511-530.

Johnson, C.C., Breward, N., Ander, E.L. and Ault, L. 2005. G-BASE: Baseline geochemical mapping of Great Britain and Northern Ireland. Geochemistry: Exploration, Environment, Analysis, 5(4), 347-357.

Johnson, C.C., Fordyce, F.M. and Rayman, M.P., 2010. Factors controlling the distribution of selenium in the environment and their impact on health and nutrition. Proceedings of the Nutrition Society, 69, 119-132.

Jones, K.C. and Sweetman, A.J., 2003. Research priorities for dioxins and polychlorinated biphenyls (PCBs). Defra Project CB01035, Environmental Science Department, Lancaster Univeriesty.

Jones, K. C., Stratford, J. A., Waterhouse, K. S., Furlong, E.T., Giger, W., Hites, R.A., Schaffner, C. and Johnston, A.E. 1989a. Increases in the polynuclear aromatic hydrocarbon content of an agricultural soil over the last century. Environmental Science & Technology, 23(1), 95-101.

Jones, K.C., Stratford, J.A., Waterhouse, K.S., and Vogt, N.B. 1989b. Organic contaminants in Welsh soils - polynuclear aromatic hydrocarbons. Environmental Science & Technology, 23(5), 540-550.

Kabata-Pendias, A. and Pendias, H., 2001. *Trace elements in soils and plants*. CRC Press LLC.

Kelly, J., Thornton, I. and Simpson, P.R. 1996. Urban Geochemistry: a study of the influence of anthropogenic activity on the heavy metal content of soils in traditionally industrial and non-industrial areas of Britain. Applied Geochemistry, 11, 363-370.

Lawley, R. 2011. The Soil-Parent Material Database: A User Guide. *British Geological Survey Open Report*, OR/08/034. 53pp.

Li, X. D. and Thornton, I. 1993. Mulitelement contamination of soils and plants in old mining areas, UK. Applied Geochemistry, Supplementary Issue No 2, January 1993: Environmental Geochemistry, 51-56.

Maliszewska-Kordybach, B., Smreczak, B. and Klimkowicz-Pawlas, A. 2009. Effects of anthropopressure and soil properties on the accumulation of polycyclic aromatic hydrocarbons in the upper layer of soils in selected regions of Poland. Applied Geochemistry, 24(10), 1918-1926.

Martin, I. and Cowie, C. 2008. Compilation of data for priority organic pollutants for derivation of Soil Guideline Values. Environment Agency, UK SC050021/SR7. 171pp.

Marchant, B.P., Tye, A.M. and Rawlins, B.G. 2011. The assessment of point-source and diffuse soil metal pollution using robust geostatistical methods: a case study in Swansea (Wales, UK). European Journal of Soil Science, 62(3), 346-358.

Matschullat, J., Ottenstein, R. and Reimann, C. 2000. Geochemical background - can we calculate it? Environmental Geology, 39, 990 - 1000.

McGrath, S.P. and Loveland, P.J. 1992. *The Soil Geochemical Atlas of England and Wales*. Glasgow, Blackie Academic and Professional.

Mielke, H.W., Alexander, J., Langedal, M. and Ottesen, R.T., 2011. Children, Soils, and Health: How do polluted soils influence children's health. *In*: C.C. Johnson, A. Demetriades, J. Locutura and R.T. Ottesen (Editors), *Mapping the Chemical Environment of Urban Areas*. John Wiley & Sons Ltd., Chichester, UK, pp. 134-150.

Morillo, E., Romero, A.S., Maqueda, C., Madrid, L., Ajmone-Marsan, F., Grcman, H., Davidson, C.M., Hursthouse, A.S. and Villaverde, J. 2007. Soil pollution by PAHs in urban soils: a comparison of three European cities. Journal of Environmental Monitoring, 9(9), 1001-1008.

Nam, J.J., Thomas, G.O., Jaward, F.M., Steinnes, E., Gustafsson, O. and Jones, K.C. 2008. PAHs in background soils from Western Europe: Influence of atmospheric deposition and soil organic matter. Chemosphere, 70(9), 1596-1602.

Napiera, F., D’Arcy, B. and Jefferies, C., 2008. A review of vehicle related metals and polycyclic aromatic hydrocarbons in the UK environment. Desalination, 226, 143-150.

Office for National Statistics, 2011. Census 2001 Key Statistics - Local Authorities KS01 Usual resident population. Census 2001 Key Statistics - Local Authorities KS01 Usual resident population.

Ordnance Survey, 2011. OS MasterMap Topography Layer. OS MasterMap® Topography Layer. http://www.ordnancesurvey.co.uk/oswebsite/products/os-mastermap/index.html . Last accessed 16th December 2011.

Osinubi, O., Gochfeld, M. and Kipen, H. 2000. Health effects of asbestos and non-asbestos fibers. Environmental Health Perspectives, 108(supplement 4), 665-674.

Ove Arup. 1990. Mining Instability in Britain. Department of Environment contract. Hardcopy maps digitised and adapted by the British Geological Survey.

Palumbo-Roe, B. and Klinck, B. 2007. Bioaccessibility of arsenic in mine waste-contaminated soils: A case study from an abandoned arsenic mine in SW England (UK). Journal of Environmental Science and Health, Part a-Toxic/Hazardous Substances & Environmental Engineering, 42(9), 1251-1261.

Palumbo-Roe, B., Cave, M. R., Klinck, B. A., Wragg, J., Taylor, H., O'Donnell, K. and Shaw, R. A. 2005. Bioaccessibility of arsenic in soils developed over Jurassic ironstones in eastern England. Environmental Geochemistry and Health, 27(2), 121-130.

Rawlins, B.G., Lister, T.R. and Mackenzie, A.C. 2002. Trace metal pollution of soils in northern England. Environmental Geology, 42, 612-620.

Rawlins, B.G., O'Donnell, K.E. and Ingham, M. 2003a. Geochemical survey of the Tamar catchment (southwest England): British Geological Suvey. Commissioned Report CR/03/027.

Rawlins, B.G., Webster, R. and Lister, T.R. 2003b. The influence of parent material on top soil geochemistry in eastern England: Earth Surface Processes and Landforms, 28, 1389-1409.

Rawlins, B.G., Lark, R.M., O'Donnell, K.E., Tye, A.M. and Lister, T.R. 2005. The assessment of point and diffuse metal pollution of soils from an urban geochemical survey of Sheffield, England. Soil Use and Management, 21(4), 353-362.

Rawlins, B.G., McGrath, S.P., Scheib, A.J., Cave, M.R., Breward, N., Lister, T.R., Ingham, M., Gowing, C.J.B., and Carter, S. 2012. The Advanced Soil Geochemical Atlas of England and Wales. British Geological Survey, Keyworth, Nottingham. Published as an electronic book. Published as an electronic book. Available on-line at: www.bgs.ac.uk/gbase/advsoilatlasEW.html last accessed 1st May 2012.

R Development Core Team. 2011. R: A language and environment for statistical computing. R Foundation for Statistical Computing. http://www.r-project.org/ Last accessed 16th December 2011.

Reaves, G.A., and Berrow, M.L. 1984. Total lead concentrations in Scottish soils. Geoderma, 32(1), 1-8.

Reimann, C., Matschullat, J., Birke, M. and Salminen, R. 2009. Arsenic distribution in the environment: The effects of scale. Applied Geochemistry, 24(7), 1147-1167.

Reimann, C., de Caritat, P., GEMAS Project Team and NGSA Project Team. 2012. New soil composition data for Europe and Australia: demonstrating comparability, identifying continental-scale processes and learning lessons for global geochemical mapping. Science of The Total Environment, 416, 239-252.

Ross, S.M., Wood M.D., Copplestone, D., Warriner, M. and and Crook, P. 2007. UK Soil and Herbage Pollutant Survey. Environmental Concentrations of heavy metals in UK soil and herbage. UKSHS Report No.7 Environment Agency. UK

Runnells, D.D., Shepherd, T.A. and Angino, E.E. 1992. Metals in Water. Determining natural background concentrations in mineralized areas. Environmental Science & Technology, 26(12), 2316-2323.

Salminen, R., Batista, M. J., Bidovec, M., Demetriades, A., De Vivo, B., De Vos, W., Duris, M., Gilucis, A., Gregorauskiene, V., Halamic, J., Heitzman, P., Lima, A., Jordan, G., Klaver, G., Klein, P., Lis, J., Locotura, J., Marsina, K., Mazreku, A., O'Connor, P. J., Olsson, S. A., Ottesen, T., Petersell, V., Plant, J. A., Reeder, S., Salpeteur, I., Sandstrom, H., Siewers, U., Steenfelt, A., and Tarvainen, T., 2005. *Geochemical atlas of Europe. Part 1, Background information, methodology and maps*. Geological Survey of Finland, Espoo, Finland, 526pp.

Saltiene, Z., Brukstiene, D. and Ruzgyte, A. 2002. Contamination of soil by polycyclic aromatic hydrocarbons in some urban areas. Polycyclic Aromatic Compounds, 22(1), 23-35.

Self, S., Entwistle, D.C and Northmore, K.J. 2011. The structure and operation of the BGS National Geotechnical Properties Database. *British Geological Survey Internal Report*.

Schreier, H., 1989. Asbestos Fibers in the Soil Environment. Chapter 4, 68-94. *In*: H. Schreier (Editor), *Asbestos in the natural environment.* Studies in environmental science. Volume 37. Elsevier, Amsterdam.

Shepherd, T.J., Chenery, S.R.N., Pashley, V., Lord, R.A., Ander, L.E., Breward, N., Hobbs, S. F., Horstwood, M., Klinck, B.A. and Worrall, F. 2009. Regional lead isotope study of a polluted river catchment: River Wear, Northern England, UK. Science of The Total Environment, 407(17), 4882-4893.

Sinclair, A. J. 1974. Selection of threshold values in geochemical data using probability graphs. Journal of Geochemical Exploration, 3, 129-149.

Smith, A. 1994. The geochemistry of Cyanide in Mill Tailings. *In*: *Short Course Handbook on Environmental Geochemistry of Sulfide Mine-Wastes*, J.L. Jambor and D.W. Blowes (eds.), Waterloo, Ontario, May 1994, 293-332.

Smith, B., Rawlins, B.G., Ferguson, A.J., Fordyce, F.M., Hutchins, M.G., Finnamore, J.R. and Barr, D.M. 1999. Information on Land Quality in the UK. Land Contamination in England, Environment Agency R&D Technical Report P291, pp. 95.

Smyth, D. 2007. Methods used in the Tellus Geochemical Mapping of Northern Ireland. British Geological Survey Open Report OR/07/022.

Spurgeon, D.J., Rowland, P., Ainsworth, G., Rothery, P., Long, S. and Black, H.I.J. 2008. Geographical and pedological drivers of distribution and risks to soil fauna of seven metals (Cd, Cu, Cr, Ni, Pb, V and Zn) in British soils. Environmental Pollution, 153(2), 273-283.

Studds, P. (undated). Asbestos in soils. A pdf presentation available from: http://www.ideal-homes.org.uk/__data/assets/pdf_file/0008/596267/ . Last accessed 12 December 2011.

Tipping, E., Poskitt, J.M., Lawlor, A.J., Wadsworth, R.A., Norris, D.A. and Hall, J.R., 2011. Mercury in United Kingdom topsoils; concentrations, pools, and Critical Limit exceedances. Environmental Pollution, 159, 3721 - 3729.

van Gosen, B.S., 2007. The geology of asbestos in the United States and its practical applications. Environmental & Engineering Geoscience, 13(1), 55-68.

Wilcke, W. 2000. Polycyclic aromatic hydrocarbons (PAHs) in soil - a review. Journal of Plant Nutrition and Soil Science-Zeitschrift Fur Pflanzenernahrung Und Bodenkunde, 163(3), 229-248.

Wild, S.R. and Jones, K.C. 1995. Polynuclear aromatic hydrocarbons in the United Kingdom environment - a preliminary source inventory and budget. Environmental Pollution, 88, (1), 91-108.

Wood, M.D., Copplestone, D. and Crook, P. 2007. UK Soil and Herbage Pollutant Survey (UKSHS). Chemical and radiometric sample collection methods. Report No.2, UK Environment Agency. UK

Wragg, J., Cave, M.R., Basta, N., Brandon, E., Casteel, S., Denys, S.e.b., Gron, C., Oomen, A., Reimer, K., Tack, K. and Van de Wiele, T., 2011. An Inter-laboratory Trial of the Unified BARGE Bioaccessibility Method for Arsenic, Cadmium and Lead in Soil. Science of the total Environment 409: 4016-4030.

Young, D.S., Hunter, D.R., Chauveau, D. and Benaglia, T., 2009. An R Package for Analyzing Mixture Models. Journal of Statistical Software, 32(6), 1-29.

Zhao, F.J., Adams, M.L., Dumont, C., McGrath, S.P., Chaudri, A.M., Nicholson, F.A., Chambers, B.J. and Sinclair, A.H. 2004. Factors affecting the concentrations of lead in British wheat and barley grain. Environmental Pollution, 131(3), 461-468.