Part 2A, Environmental Protection Act 1990

Technical Guidance Sheet (TGS) on normal levels of contaminants in English soils

Supplementary Information

COPPER (Cu)

Technical Guidance Sheet Supplementary Information TGS03s, July2012.

Contents

defra
Department for Environment Food and Rural Affairs

British Geological Survey
NATURAL ENVIRONMENT RESEARCH COUNCIL

List of Figures

List of Tables

Acknowledgments

This supplementary information for the copper Technical Guidance Sheet (TGS) is compiled with information derived mainly from the reports prepared for the Department for Environment Food and Rural Affairs (Defra) soil R&D project SP1008 by the British Geological Survey. This work has been led by Chris Johnson with assistance from Louise Ander, Mark Cave and Barbara Palumbo-Roe (all BGS, Keyworth) with additional contributions and comments from Murray Lark, Barry Rawlins, Don Appleton and Chris Vane (BGS Keyworth); Stephen Lofts (CEH Lancaster); and Paul Nathanial Land Quality Management Group, Nottingham. The authors also thank the Defra Soils Policy Team, the Project Steering Group and several Local Authority contaminated land officers who have given valuable advice to improve the content of this information sheet.

When referring to this document the following bibliographic reference should be made:

Defra, 2012. Technical Guidance Sheet on normal levels of contaminants in English soils: Copper – supplementary information. Technical Guidance Sheet No. TGS03s, July 2012. Department for Environment, Food and Rural Affairs (Defra), Soils R&D Project SP1008. Available on-line from Defra project SP1008 web page.

The copper Technical Guidance Sheet which this document supplements:

Defra, 2012. Technical Guidance Sheet on normal levels of contaminants in English soils: Copper. Technical Guidance Sheet No. TGS03, July 2012. Department for Environment, Food and Rural Affairs (Defra), Soils R&D Project SP1008. Available on-line from Defra project SP1008 web page.

Supplementary Information

Important soil sample and analytical information

Aggregate samples, soil depth and fraction size

Both the NSI (XRFS) and G-BASE data sets are derived from a soil sample that has been aggregated (composited) from a number of subsamples collected over the area of a site, rather than a single point sample. In the case of NSI this is 25 cores (subsamples) from a 20-m square (McGrath and Loveland 1992) whereas G-BASE is 5 cores, also from a 20-m square (Johnson *et al.* 2005; Fordyce *et al.* 2005). If a sample is collected as a single core, and the result is compared to the NBC, it is important to be aware that short-range variation (which can be substantial) for the single core sample will be potentially much greater than for the samples from which the NBC values are derived (Lark, 2012).

Soil samples used to calculate the Cu NBCs have been collected from the top 15 cm of the mineral soil profile (hence they are referred to as topsoils). When the sample is collected from a site covered with vegetation the surface organic layers (leaf litter) do not form part of the sample collected. Any recently deposited airborne particulates that have not yet migrated into the soil profile will not be sampled and surface organic material, which has the capacity to fix some contaminants from atmospheric deposition, is not included as part of the sample. In urban areas the top 15 cm will be expected to have been modified by historical urban land uses and, in rural agricultural areas, where relevant, will be within the ploughed horizon. Surveys targeting recent airborne pollution added to the soil will generally only collect from the top 2 cm of the profile in order to bias the soil results toward the airborne pollutant inputs. Such data has not been used in the NBC calculations.

Another consideration is the soil size fraction to be submitted for chemical analysis. The <2 mm fraction is widely used for soil analyses. However, other fractions are sometimes reported (*e.g.* <150 µm) in order to enhance some chemical contrasts and to reduce variability in the chemical results – coarser grains mean that a single "nugget" will give rise to greater variability in the analyses than will occur with a finer more homogenous material. The NBCs calculated here are exclusively based on the <2 mm soil fraction.

Total and partial analytical determinations of element concentrations

There are established international procedures and standards for the determination of naturally occurring elements in the Earth's surface environment (Darnley *et al.* 1995). These procedures have been set up in order to develop a global database of chemical results that is compatible and of sufficient quality to be used for environmental and resource management. The analytical requirements to realise this objective includes: "*The total amount of each element present is the most fundamental (and reproducible) quantity in any sample, therefore direct measurement techniques, e.g. XRFS or neutron activation analysis (NAA), or total extraction procedures should be employed as a first priority*." The British Geological Survey has been one of the leading organisations in the development of this global geochemical database. Therefore, the vast majority of systematically collected soil sample data that is available for NBC calculations for English soils are total element concentrations determined by laboratory-based XRFS. Other analytical techniques that do not give total element concentrations are used to determine the nature of occurrence and speciation of an element within a sample.

When using NBCs a common question will be "how should I interpret NBCs in the context of non-total analyses"? This was investigated as part of the data exploration phase of this project (Ander *et al.* 2011; 2012). Figure 1 shows a plot of Cu in the NSI topsoils which have been analysed by both a total (XRFS) (Rawlins *et al.* 2012) and partial (*aqua regia* followed by ICP-AES) (McGrath and Loveland 1992). There is a close linear relationship between the two analytical methods with a systematic bias to higher concentrations by XRFS; this would be expected from this total measurement, unlike the acid digest which will leave a quantity of trace element bearing, residual material. There would also be a systematic bias expected between two analytical measurement techniques.

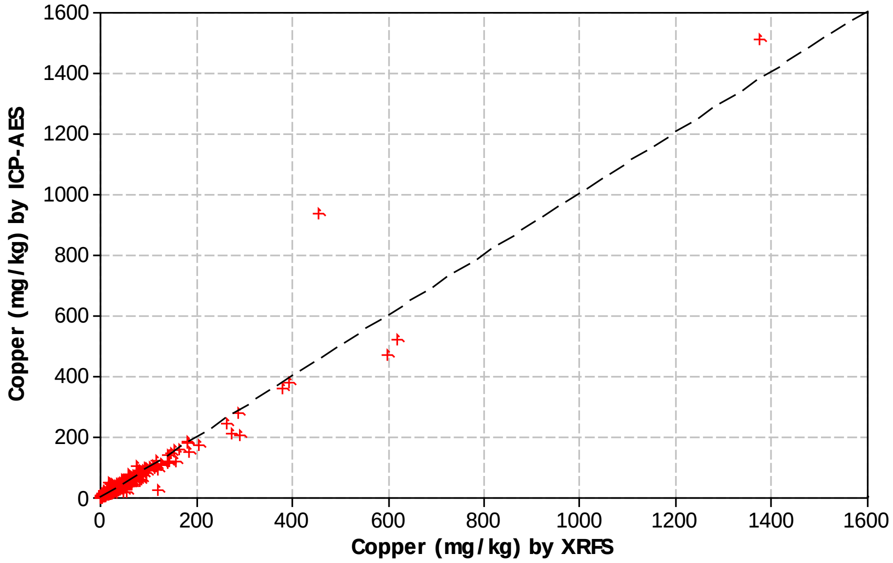


Figure 1: Comparison of topsoil Cu concentrations in NSI samples by XRFS and ICP-AES following *aqua regia* acid digest.

The regression equation for the NSI data comparison:

$[Cu_{XRFS}] = 2.89 + (0.911 \times [Cu_{aqua\ regia}])$ (n=3956; R^2 = 93%; and $P<0.05$)……*Equation 1*

A similar exercise has been done with the recently collected GEMAS project samples (see Ander *et al.* 2012). The arable and pasture topsoils samples from England were analysed by both XRFS and *aqua regia* digest followed by ICP-MS analysis (Figure 2).

The regression equation for the GEMAS data comparison is:

$[Cu_{XRFS}] = -1.00 + (0.972 \times [Cu_{aqua\ regia}])$ (n=130; R^2 = 97%; and $P<0.05$)……*Equation 2*

A further comparison of Cu results determined by XRFS and ICP-MS (following an *aqua regia* extraction) can be made using the Tellus Project (Smyth 2007) topsoil results (Figure 3). The regression equation for these data is:

$[Cu_{XRFS}] = -0.381 + (1.18 \times [Cu_{aqua\ regia}])$ (n=3956; R^2 = 93%; and $P<0.05$)……*Equation 3*

A soil sample determined to have 100 mg/kg Cu following an *aqua regia* extraction and measurement by ICP-MS, from equations 1, 2 and 3 will have an estimated total Cu of 94, 96 and 118 mg/kg, respectively, rounding to the nearest 1 mg/kg. The regression equations can be used as a tool to estimate total concentrations of Cu for *aqua regia*/ICP-MS determined samples, though its application must be done with an awareness of the analytical error range, particularly at higher concentrations.

Supplementary Information

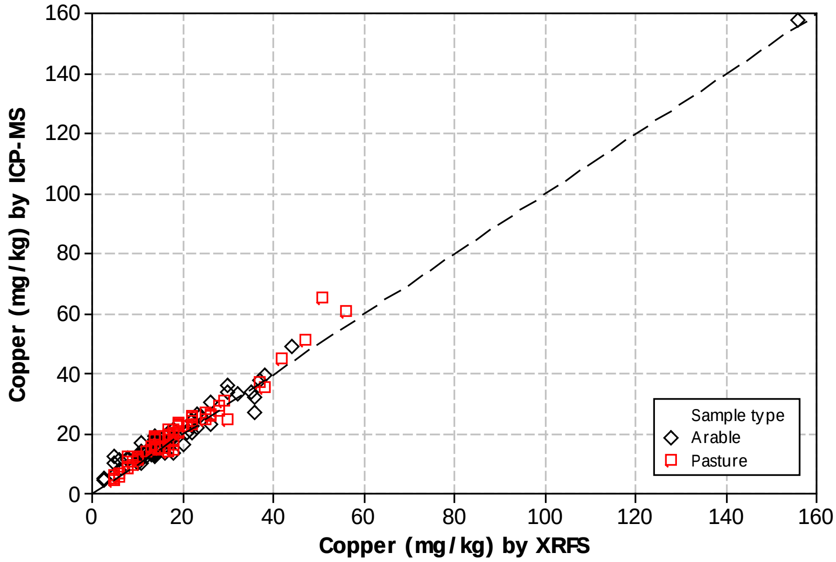


Figure 2: A comparison of GEMAS project topsoil Cu data by analytical method and categorised by land use type.

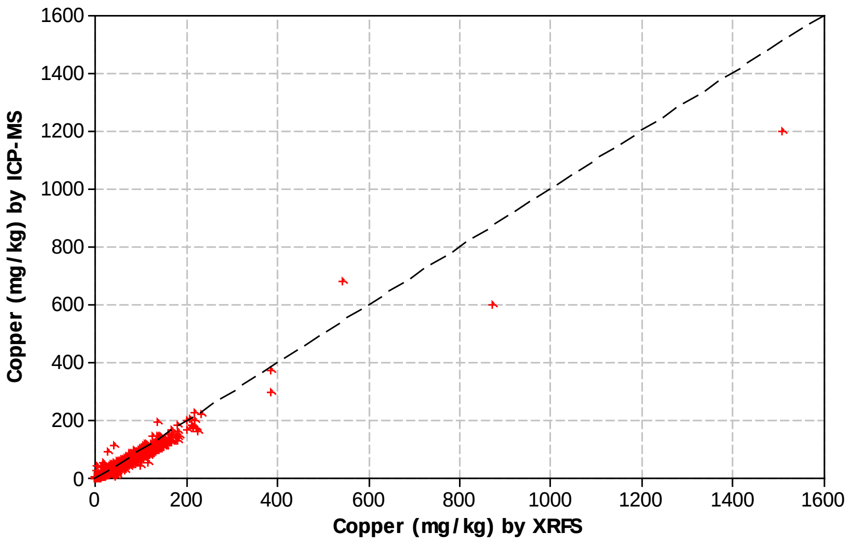


Figure 3: A comparison of topsoil Cu concentrations in Northern Ireland (Tellus Project) by XRFS and *aqua regia* digest ICP-MS.

Scale and use of Normal Background Concentrations

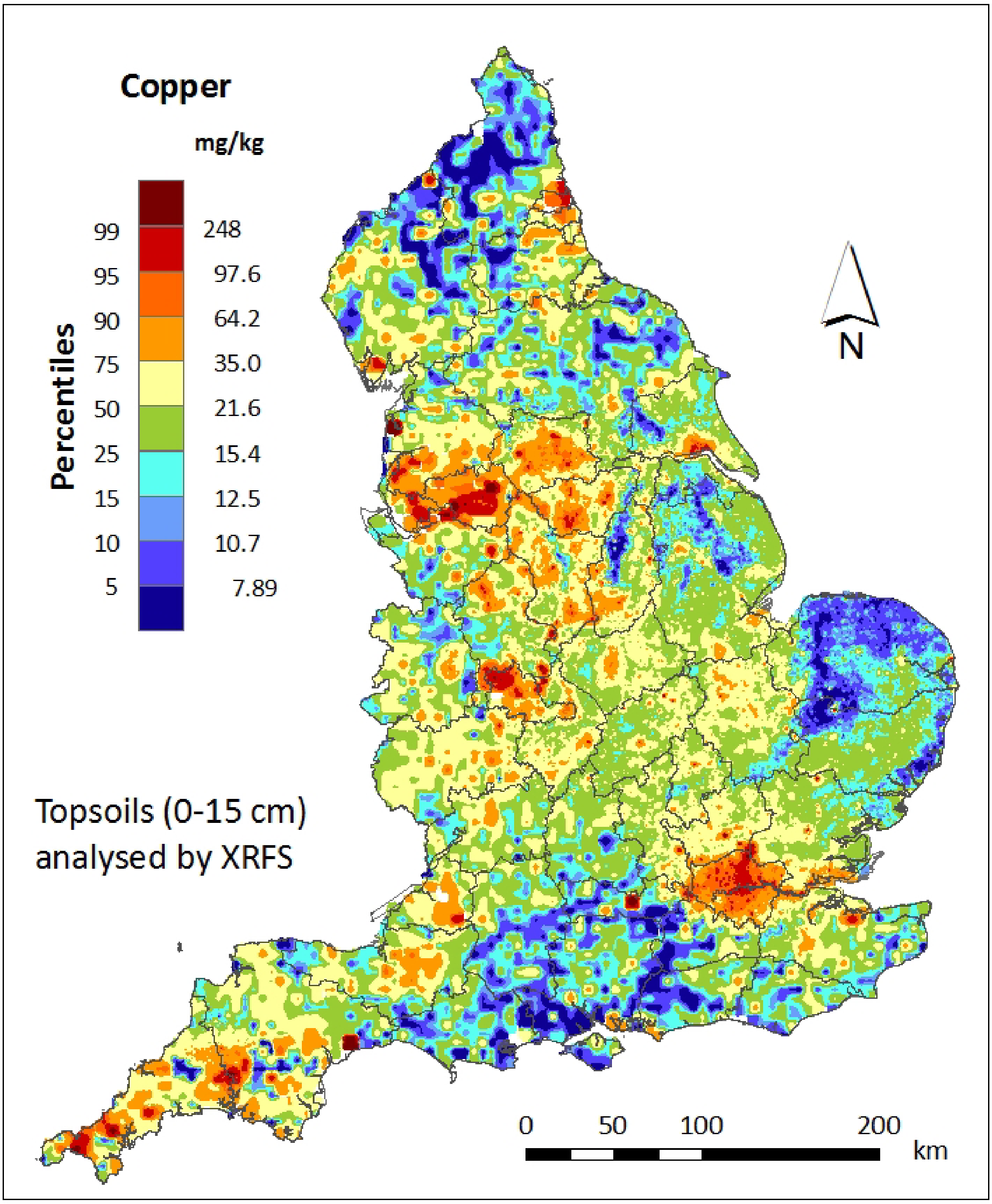
NBCs have been determined for Cu using soils collected at a range of sampling densities, from 1 sample per 0.25 km^2 (G-BASE urban) through to 1 sample per 25 km^2 (NSI XRFS). The G-BASE urban samples provide a definition of the chemical surface environment to a much higher resolution than do the NSI (XRFS) samples. Thus G-BASE rural samples (collected at 1 per 2 km^2 sampling density) can show contaminant variability at a local area scale (1:50,000). When investigating a sample result in the context of a NBC, it is important to ask whether localised variability (scales at less than 1:50,000), say within the Principal Domain, has been truly captured during the determination of the NBCs. Ander *et al*. 2011 describe that at a local scale mineralisation (potentially Cu veins) may not be mapped. Therefore, a high contaminant result should be attributed to a domain taking account of the localised underlying parent material feature, even where this has an extent which is very discontinuous.

Use of variograms

The domains that are defined for a particular contaminant correspond to major sources of variation in concentrations of that contaminant in soil, such as urbanisation, mining or mineralisation. Concentrations of the contaminant vary within the domains, the procedure to define normal background concentrations (NBCs) quantifies this variation with robust statistics, from which the NBCs are computed. The spatial variation of a contaminant within a domain can be quantified by the variogram (Matheron, 1962). The variogram is a function that shows how the variation between observations of a variable at two sites depends on the distance in space between the sites. The variogram is half the mean squared difference between two observations plotted against the distance between them for all the results in a data set. Typically the variogram increases with distance until a plateau in the plot is reached at a value called the sill variance, which it reaches at a distance called the range. If the range is very short then this shows that the spatial variation is very intricate. If the range is longer then it may be feasible to map spatial variations from sample observations on a grid

Supplementary Information

National map showing the distribution of copper in topsoils



The distribution of samples used in this interpolated map is shown in Figure 1 of the Cu technical guidance sheet.

Figure 4: National map of copper distribution in topsoils with county boundaries (using G-BASE and NSI (XRFS) results).

The national map of Cu distribution in topsoils (Figure 4) is shown here along with county boundaries to help with location at a regional scale. This map is given to demonstrate the variability in Cu across England and is also available to view on-line at the BGS project web page. The map has been generated from G-BASE and NSI (XRFS) topsoil data using 42,133 samples. Because central and eastern England have been sampled at a much higher density (by G-BASE), resolution of information in these areas is much higher. Figure 4 has been produced in ArcGIS v9.3 using the IDW option of the Spatial Analyst tool, cell size 1000 m and search radius 5000 m (inverse square option selected). The percentile classification is based on **all data** and differs from the domain data sets in which results are modelled to fit a normal distribution and the effect of outliers (representing point rather than diffuse pollution) have been reduced by normalisation of the data.

The map shown in Figure 4 uses soils to represent the geochemical baseline. Other national/regional scale geochemical atlases for soils are those of McGrath and Loveland (1992) (NSI *aqua regia* data) and Rawlins *et al.* (2012) (NSI XRFS data). A preferred way of representing the geochemical baseline at a national/regional scale is to use stream sediments. The fine stream sediment in a drainage channel is representative of material washed down the drainage catchment to the sampling site in the stream and so gives a much better regional average of the chemical environment than is given by soils. The G-BASE project also collects stream sediments at a sampling density of approximately one sample per 2 km^2 and results for England have been presented in a series of atlases (*e.g.* Lake District (BGS 1992) and NE England (BGS 1996)) and these can be used to further demonstrate element variability across the surface environment of England. For the more recently sampled parts of England, the G-BASE project has also determined a large range of elements in stream waters (*e.g.* Environmental Geochemical Atlas of Central and Eastern England). Comparing the element concentrations and distributions of different sample types collected from the same locality can provide useful information about the mobility of a chemical element in that area.

A stream sediment atlas for England and Wales was also completed by Webb *et al.* (1978) (Wolfson Geochemical Atlas). More recently, low density sampling has produced continental scale geochemical baselines for Europe based on a number of sampling media, including stream sediments, stream waters and soils, namely the FOREGS atlas project (Salminen *et al.* 2005) and the ongoing GEMAS project (Reimann *et al.* 2012).

The importance of using resources other than the available soil maps to identify areas of high natural background concentrations is demonstrated by the fact that the high density stream sediment sampling of G-BASE delineates the Lake District copper mineralisation (see BGS 1992) which is not shown by the lower density soil sampling of the NSI (Rawlins *et al.* 2012).

Supplementary Information

Descriptive statistics for copper in topsoil data

Copper Domain percentile classifications

Copper data for soils has been gathered from data sets as described in the Cu TGS and classified according to the most important domains as detailed by Ander *et al.* (2011). A percentile of a data distribution (in this case the distribution of Cu in soil for a given domain) is the value of a variable below which a certain percentage of observations fall. The 95th percentile, for example, is the value below which 95% of the observations may be found, *i.e.* it encompasses the majority of the data. The contaminant concentrations in the soil for a given domain are a subset of the total population of all possible soil concentrations and therefore any percentile calculation will only be an approximation of the true value. The uncertainty on the percentile increases as the number of samples used to calculate it decreases. Lower and upper limits can be statistically estimated for each percentile giving a confidence interval for that percentile. **The Cu NBC for each domain is defined as the upper 95% confidence limit of the 95th percentile for the Cu topsoil concentrations that fall within that domain** (Cave *et al.* 2012). A summary of domain percentiles with their upper and lower limits is given in Table 1.

Percentile	Urban Domain (7,475)			Mineralisation Domain (153)			Principal Domain (34,504)		
	lower	middle	upper	lower	middle	upper	lower	middle	upper
50	50	51	52	42	49	57	20	20	21
55	55	56	57	48	56	65	22	22	22
60	61	62	63	54	63	75	24	24	24
65	67	69	70	60	72	85	26	26	27
70	75	77	78	68	82	99	29	29	29
75	84	86	88	78	95	120	32	32	32
80	95	98	100	90	110	140	35	36	36
85	110	110	120	110	140	170	40	41	41
90	130	140	140	130	170	220	47	48	49
95	180	180	**190**	180	250	**340**	60	61	**62**

Figure in brackets represents the number of samples used in the domain calculation

Table 1: A summary of the copper domain percentile classifications. Domain NBCs shown in bold red. Concentrations in mg/kg.

Descriptive statistics copper topsoil data set

Table 2 shows descriptive statistics for all the topsoil Cu results from the G-BASE and NSI (XRFS) data sets. The cities and towns in Table 2(c) are those that have been systematically sampled by the G-BASE project. Some of these data sets have associated reports that can be downloaded by clicking on the location place marker on the map at http://www.bgs.ac.uk/gbase/urban.html. Other data sets for other English cities may exist but they are not made publicly available and are not sampled and analysed to a nationally consistent standard.

(a) All data	Number	Mean	Minimum	25th percentile	Median	75th percentile	Maximum	Skewness
G-BASE(urban + rural) + NSI (XRFS)	42132	36	<0.5	15.4	21.7	35	5330	26
(b) Data set type	**Number**	**Mean**	**Minimum**	**25th percentile**	**Median**	**75th percentile**	**Maximum**	**Skewness**
All NSI(XRFS)	4864	23.9	<0.5	12.9	18.6	25.9	1380	19
G-BASE (rural)	23685	22.4	<0.5	13.7	18.5	24.4	2770	36
G-BASE (urban)	13583	64.2	1.64	25.1	39.4	67.3	5330	18
Eastern England (G-BASE)	23221	21.9	<0.5	13.6	18.4	24	1470	22
Tamar catchment (G-BASE)	464	47	2.06	19.8	30.5	44	2770	17
(c) Urban (G-BASE)	**Number**	**Mean**	**Minimum**	**25th percentile**	**Median**	**75th percentile**	**Maximum**	**Skewness**
Corby	133	32	11.4	17.8	20.8	24.7	908	10
Coventry	390	48	9.98	22.5	31.9	53.7	464	4
Derby	275	56.2	16.1	30.4	41.3	57.7	659	6
Doncaster	279	53.5	7.89	22.5	31.9	53.7	1280	10
Hull	407	76.9	5.81	27.7	41.2	80.8	1170	5
Leicester	652	38.9	10.7	22.4	29.6	45.1	508	6
Lincoln	215	32.4	3.72	11	17.3	33.9	362	4
London (GLA area)	6494	72.6	3.24	29.1	46.2	76.6	5330	19
Manchester (part of)	300	125	7.06	59	89.1	134	2160	8
Mansfield	257	41.8	2.68	15.2	24.6	40.7	1800	13
Northampton	275	33.9	7.47	18.3	24.7	34.7	1070	14
Nottingham	636	49.6	8.93	26.7	37.1	53.7	1010	9
Peterborough	272	34.9	11.5	20.4	26.1	35.6	270	4
Scunthorpe	196	22.9	1.64	9.98	15.2	24.6	451	9
Sheffield	575	81	12.1	40.2	52.7	85	1640	8
South Essex Towns	715	50.1	4.71	21.8	30.9	49.2	2590	16
Stoke-on-Trent	745	52.1	6.85	22.5	33.9	53.7	1800	12
Telford	292	38.3	7.89	19.6	26.7	38.1	434	5
Wolverhampton	284	146	14.1	50.6	81.9	153	3180	8
York	191	37.7	5.81	21.4	26.7	44.4	236	3

Table 2: Descriptive statistics of underlying primary data sets for Cu in all topsoils. These are classified by various data set subgroups of the original projects (total concentrations (XRFS) in mg/kg) (from Ander *et al*. 2012, with results cited to three significant figures).

Supplementary Information

Data distributions

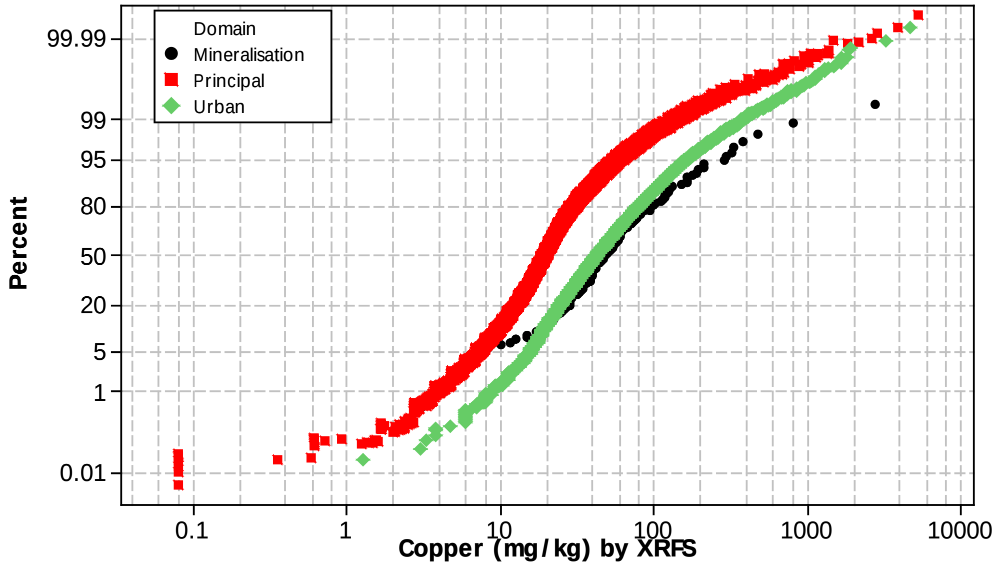


Figure 5: Probability plot of topsoil Cu results categorised by domains.

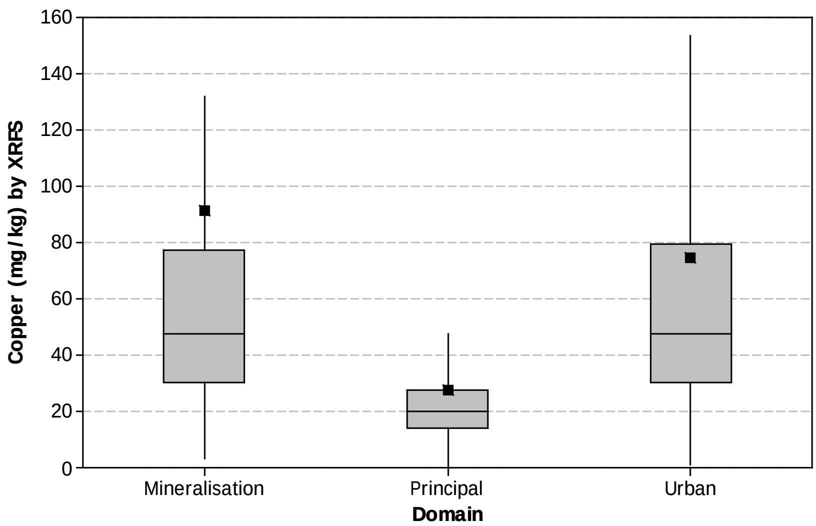


Figure 6: Boxplot of Cu topsoil results attributed to domains.

Figure 5 and Figure 6 show the frequency distribution of results for soils over the three domains defined for Cu using the G-BASE urban and rural data sets and the NSI(XRFS) data. These plots can be used in conjunction with any new results plotted in a similar way to compare distributions with the defined domains. The box of the boxplot represents the interquartile range (Q1, Q3), with the median (Q2) as a line within the box. The point symbol shows the mean value. The upper whisker = Q3 + 1.5(Q3-Q1); lower whisker = Q1-1.5(Q3-Q1).

Archer and Hodgson (1987) carried out a study of total and extractable trace element contents of agricultural soils (from a depth of 15 cm) in England and Wales, including Cu. "Total" Cu analyses were done by AAS following a digestion using perchloric and nitric acids. They defined the normal range for trace element contents as that between twice the log-derived standard deviation above and below the mean; approximately 95% of the data range. For 1,468 agricultural topsoils they determined a Cu median of 18.4 mg/kg and a "normal" range of 5.8-62 mg/kg.

Paterson *et al*. (2003) reporting on background levels of contaminants in Scottish soils report a range of Cu concentrations for mineral soils from 0.2 – 63.9 mg/kg with Q1, Q2 and Q3 values of 4.7, 7.4 and 11.5 mg/kg, respectively. This is much lower compared to the English soils (all data - Table 2) of 15.4, 21.7 and 35.0 mg/kg (Q1,Q2 and Q3, respectively), which would be expected as there are no urban soils included in with the Scottish data.

Supplementary Information

Landscape data used to define contaminant domains

Rather than seeking to define a single Cu NBC for the whole of England, the project has, through its data exploration (Ander *et al.* 2012), determined the most significant domains that can be defined in order to capture the most significant controls on Cu distribution in soils. For Cu these have been identified as soils in urban areas and a mineralised/metalliferous mining area. These domains have been defined using some key datasets within a GIS environment, namely: the BGS Soil-Parent Material Model (SPMM) (Lawley, 2009) and a revised and digitally updated version of the Ove Arup (1990) Department of the Environment (DoE) Metalliferous Mining and Mineralisation data set; and an urbanisation index derived from the Generalised Land Use Database (GLUD) Statistics for England 2005 (Communities and Local Government 2007).

Soil parent material

The Soil-Parent Material Model[1] (SPMM) 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[2]), 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. 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 for the soil parent material is also relevant – 1:50,000 is the scale at which much of the systematic geochemical soil sampling has been undertaken, and gives the user a reasonable feel for the degree of uncertainty on the data. Where geographical information 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 series mapping is available at a national-scale (see *e.g.* NSRI NATMAP[3]) 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.

[1] http://www.bgs.ac.uk/products/onshore/soilPMM.html
[2] http://www.bgs.ac.uk/products/digitalmaps/digmapgb_50.html
[3] http://www.landis.org.uk/data/natmap.cfm

Metalliferous mining and mineralisation

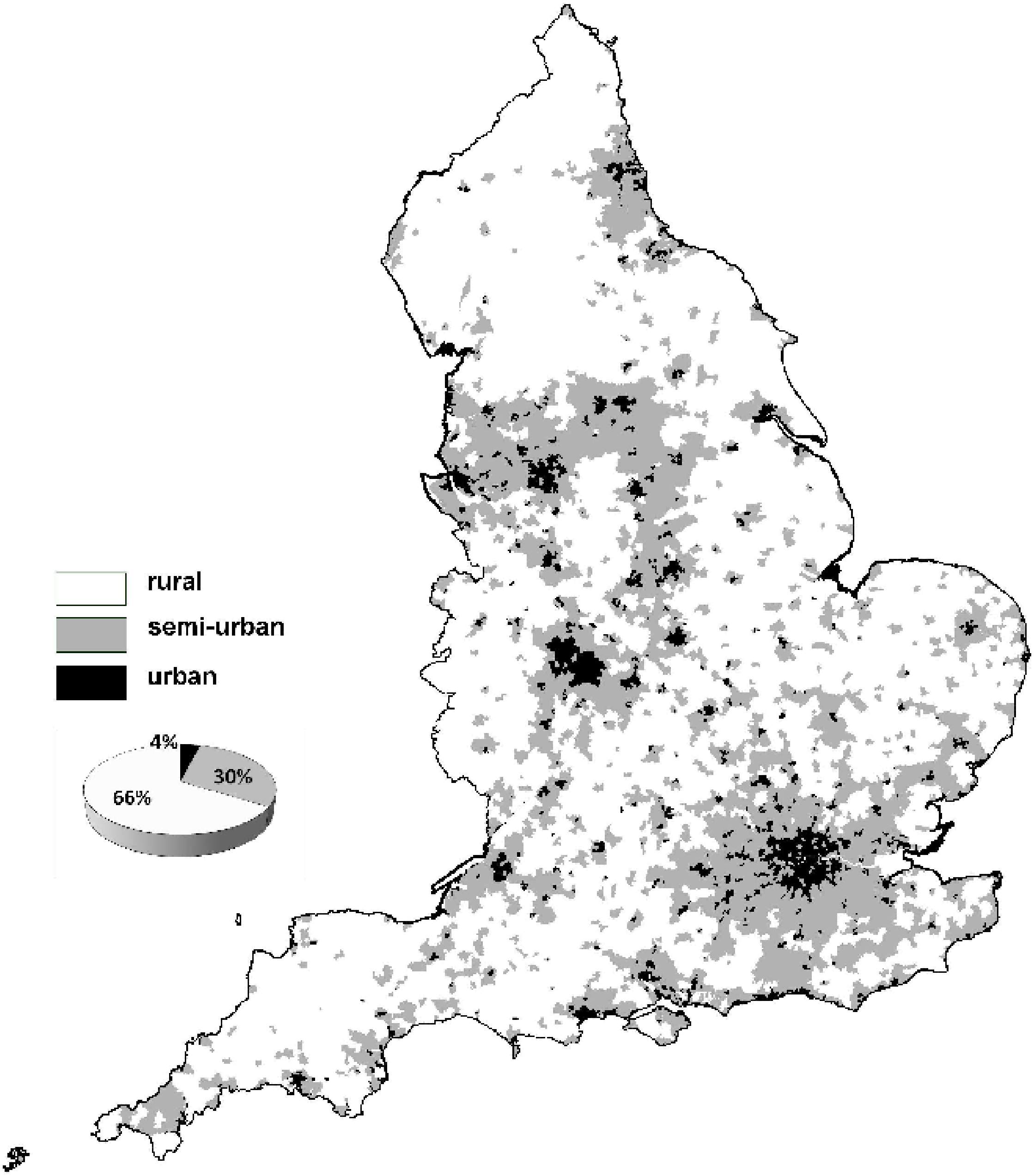
The data set which has been examined in this project is that of non-ferrous Metalliferous Mineralisation and Mining database, originally produced in hard-copy by Ove Arup (1990) for DoE (Department of Environment), but which has been 'cleaned' and turned into a polygon layer by BGS. The data for England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

Definition of urban areas

The definition of normal levels of contaminant concentrations in soils includes the contribution from diffuse pollution. As much diffuse pollution is associated with built-up regions, defining areas of urbanisation to create an urban domain is important in the attribution of NBCs. The definitive database for land use in 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[4], 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. Using the land use data from the 8850 Census Area Statistical Wards (CASW) an urbanisation index can be determined as described in Ander *et al*. (2011). This index can be used to create the map used to define urban domains (Figure 7). The urban classification map of England is available as a GIS layer from the BGS project web page.

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

Supplementary Information



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

Figure 7: A map of England showing urban, semi-urban and rural areas of England defined from an urbanisation index using the GLUD database.

Summary of statistical procedure to determine NBCs

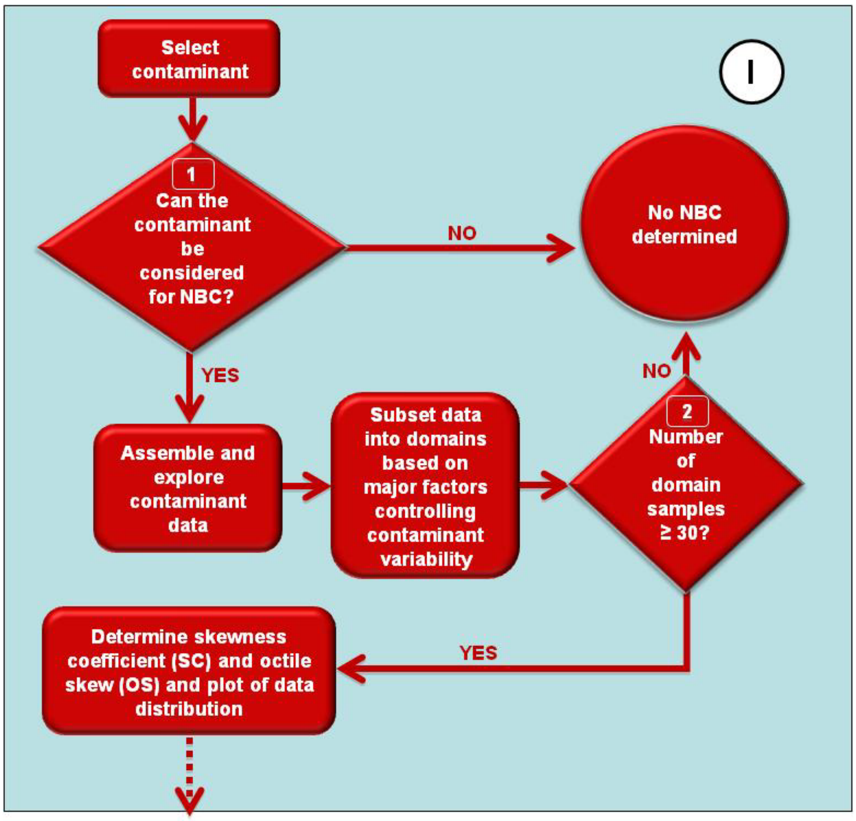


Figure 8: Flow chart for the calculation of the NBC for a given contaminant domain (OS and SC are octile skew and skewness coefficient, respectively. MAD = median absolute deviation). See text for explanation, continued overleaf.

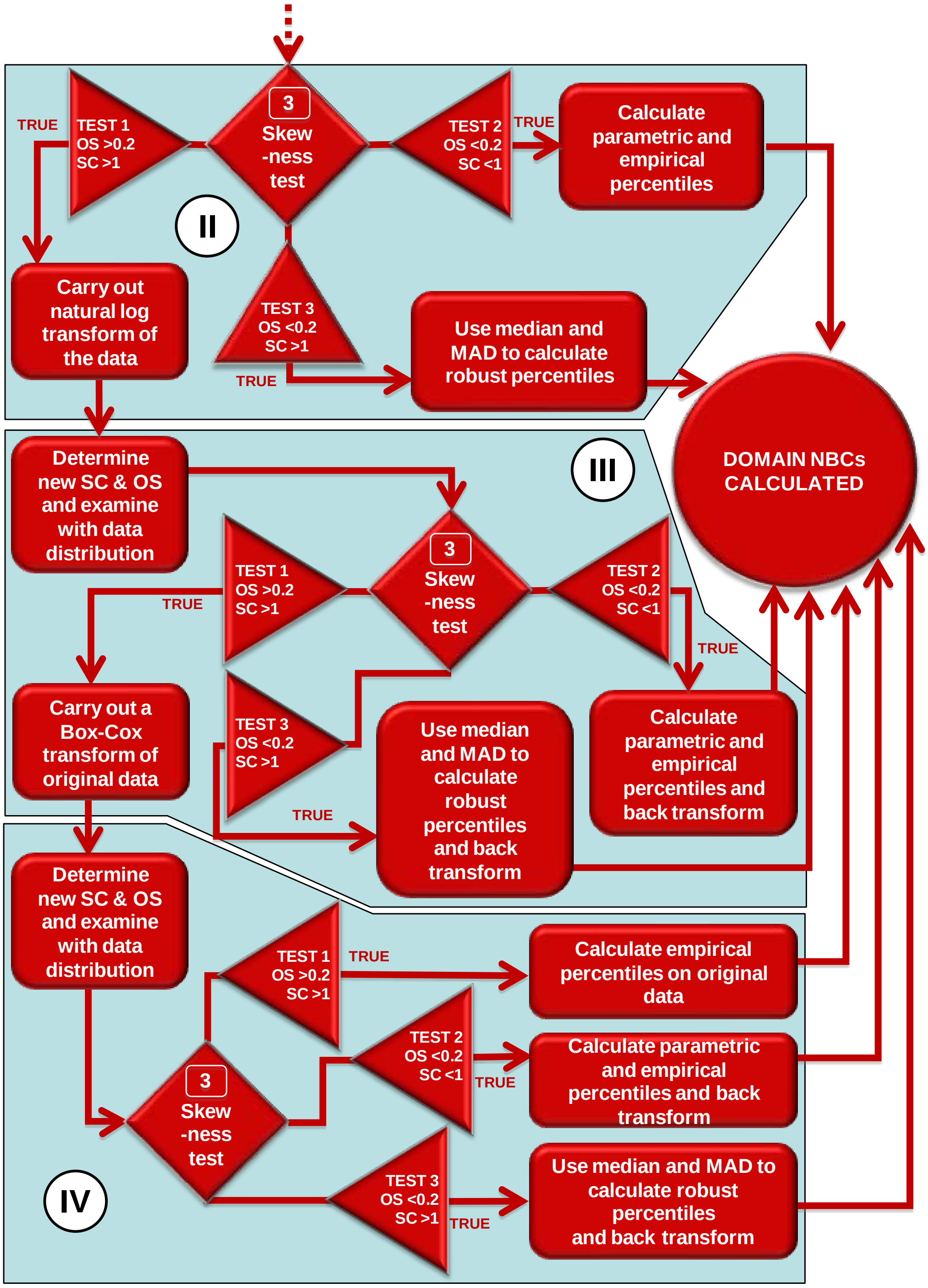


Figure 8 continued. Flow chart for the calculation of the NBC for a given contaminant domain (OS and SC are octile skew and skewness coefficient, respectively. MAD = median absolute deviation). See text for explanation.

Figure 8 summarises the statistical procedure used to determine contaminant NBCs (see Cave *et al.* 2012). Part I essentially represents the data gathering and exploration phase of the project (WP1&2) in which domain areas are identified. Question 1 asks if the contaminant is suitable for a NBC. Asbestos and manufactured organic contaminants with no natural origin, for example, fail this question. The data exploration (Ander *et al.*, 2011) identifies the areas (domains) where there are clearly identifiable controls on high concentrations of a specified contaminant. The contaminant data set is then subdivided into domain data sets. In question 2 a minimum of 30 results are considered necessary to determine a NBC (see Cave *et al*., 2012). Once the data has been subsetted into domains, then skewness testing and inspection of frequency distribution plots can be done to select the appropriate data transform and method of calculating percentiles (Parts II – IV). Question 3, the skewness test, has three possible outcomes. TEST 1 (OS > 0.2 and SC >1) is true if the data distribution is skewed and not suitable for fitting to a Gaussian model and the data need to be transformed to using either a logarithmic or Box-Cox transform. If TEST 2 (OS < 0.2 and SC <1) is true then the data are consistent with the assumption of a Gaussian distribution and the parametric percentiles are fitted based on the mean and standard deviation of the data. Finally, TEST 3 (OS < 0.2 and SC >1) means the data show a mostly symmetrical distribution but with potential outliers. Here the data are consistent with the assumption of a Gaussian distribution and the parametric percentiles are fitted using median and the median absolute deviation (MAD) in place of the mean and standard deviation as these measures are robust to outliers.

Supplementary Information

Access to data and information resources used to calculate NBCs

Project Reports and information

These resources are available from the BGS project web page [5] and include:

Data Exploration Reports (BGS reports Nos. CR/11/145 and CR/012/041); Methodology Report (BGS report No. CR/12/003); Final Project Report (BGS report No. CR/12/035); Technical Guidance Sheets and supplementary information; MS Access Database summary of available data; Project Bibliography (Endnote bibliography); R code scripts used to determine NBCs; and GIS Resources served as WMS files (Domain polygons; the urbanisation index polygons defined from GLUD database; and the national contaminant interpolated image maps).

Web map services (WMS) are an industry standard protocol for serving georeferenced images across the web. They were developed and first published by the Open Geospatial Consortium (OGC) in 2000. Since this date WMS have had a steady uptake and are being increasingly used in traditional desktop based GIS, web-based GIS systems (including Google Earth), and Smartphone 'apps'. BGS holds the data on their servers and publish it openly via the BGS project web page.

Principal contaminant data sets for England

Intellectual Property Rights for the raw soil data sets resides with the organisations responsible for those data sets. In the case of the G-BASE and NSI (XRFS) data is made freely available subject to certain licensing terms and conditions. For large data sets there will also be a data handling fee. Further information regarding access to the G-BASE and NSI (XRFS) soil data is given at the BGS project web page and enquiries should be sent to enquiries@bgs.ac.uk.

Other data sets providing information on soil chemistry are summarised in Appendix 2 of Ander *et al.* (2011) and this includes contact and web site links.

Soil parent material

The BGS Soil-Parent Material Model is described on a BGS web page (SPPM)[6] and this contains information regarding further information and pricing.

Land use data including metalliferous mining and mineralisation

The Generalised Land Use Database (GLUD) Statistics for England 2005 is available for free from the Communities and Local Government website.[7] Users interested in the detailed maps at land parcel level who hold the appropriate public sector licence to use OS MasterMap® can request to see the GLUD data at this large scale level (gis@communities.gsi.gov.uk).
The Ove Arup Mineralisation and mines data updated and modified by BGS is available from BGS subject to terms and conditions (see the BGS project web page).

[5] http://www.bgs.ac.uk/gbase/NBCDefraProject.html
[6] http://www.bgs.ac.uk/products/onshore/soilPMM.html
[7] http://www.communities.gov.uk/publications/planningandbuilding/generalisedlanduse

Further Reading

The following is a list of bibliographic references that provide more detailed information regarding the distribution and behaviour of copper in the surface environment. Some of these references are referred to in this supplementary information section.

Ander, E.L., Cave, M.R., Johnson, C.C. 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.

Ander, E.L., Cave, M.R., Johnson, C.C. and Palumbo-Roe, B. 2012. Normal background concentrations of contaminants in the soils of England. Results of the data exploration for Cu, Ni, Cd and Hg. *British Geological Survey Commissioned Report*, CR/12/041. 88pp.

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, 421-431.

Bevins, R.E., Young, B., Mason, J.S., Manning, D.A.C. and Symes, R.F. (2010), *Mineralization of England and Wales*, Geological Conservation Review Series, No. 36, Joint Nature Conservation Committee, Peterborough, 598 pages, illustrations, A4 hardback, ISBN 978 1 86107 566 6. Available on-line at: http://www.jncc.gov.uk/page-3298

BGS. 1992. Regional geochemistry of the Lake District and adjacent areas. British Geological Survey Regional Geochemistry Atlas Series. BGS, Keyworth, Nottingham, UK. ISBN 0 85272 225 7.

BGS. 1996. Regional geochemistry of north-east England. British Geological Survey Regional Geochemistry Atlas Series. BGS, Keyworth, Nottingham, UK. ISBN 0 85272 255 9.

BGS. 1998. Minerals in Britain. Past production … Future potential: Copper. A British Geological Survey Report for the Department of Trade and Industry, UK. Available on –line at: http://www.bgs.ac.uk/downloads/start.cfm?id=1324 . Last access 21st March 2012.

Breward, N. 2003. Heavy-Metal Contaminated Soils Associated with Drained Fenland in Lancashire, England, UK, Revealed by BGS Soil Geochemical Survey. *Applied Geochemistry*, 18 11), 1663-1670.

Cave, M.R., Johnson, C.C., Ander, E.L. and Palumbo-Roe, B. 2012. Methodology for the determination of normal background contaminant concentrations in English soils. *British Geological Survey Commissioned Report*, CR/12/003. 56pp.

Colbourn, P., Alloway, B.J. and Thornton, I. 1975. Arsenic and Heavy Metals in Soils Associated with Regional Geochemical Anomalies in South-West England. *Science of The Total Environment*, 4(4), 359-363.

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.

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. 1978. Plant-Available Lead and Other Metals in British Garden Soils. *Science of The Total Environment*, 9(3), 243-262.

Davies, B.E. 1995. Lead *In: Heavy Metals in Soil*. Editor Alloway, B.J. Blackie Academic & Professional, London. Chapter 9, 206-223.

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.

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., Ander, E.L., Cave, M.R. and Palumbo-Roe, B. 2012. Normal Background Concentrations of contaminants in English soil: Final project report. *British Geological Survey Commissioned Report*, CR/12/035. 40pp.

Supplementary Information

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(1–2), 363-370.

Lark, R.M. 2012. Some considerations on aggregate sample supports for soil inventory and monitoring. *European Journal of Soil Science*, February 2012, 63, 86-95.

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

Lepp, N.W., Hartley, J.,Toti, M. and Dickinson, N.M. 1997. Patterns of Soil Copper Contamination and Temporal Changes in Vegetation in the Victinity of a Copper Rod Rolling Factory. *Environmental Pollution*, 95(3), 363-369.

Matheron, G. 1962. Traité de géostatistique appliquée. Tome 1, Editions Technip, Paris, 334pp.

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

Mighall, T.M., Abrahams, P.W., Grattan, J.P., Hayes, D., Timberlake, S. and Forsyth, S. 2002. Geochemical Evidence for Atmospheric Pollution Derived from Prehistoric Copper Mining at Copa Hill, Cwmystwyth, Mid-Wales, UK. *Science of The Total Environment*, 292(1–2), 69-80.

Ordnance Survey, 2011. OS MasterMap Topography Layer. OS MasterMap® Topography Layer. http://www.ordnancesurvey.co.uk/oswebsite/products/os-mastermap/index.html. Last accessed 7th March 2012.

Ove Arup (1990). Mining Instabililty in Britain. Unpublished reports by Ove Arup for the Department of the Environment (DoE), UK. Associated data sets modified by BGS into GIS format (see http://www.bgs.ac.uk/gbase/NBCDefraProject.html).

Paterson, E., Towers, W., Bacon, J.R. and Jones, M. 2003. Background levels of contaminants in Scottish soils. The Macaulay Institute. *Final Contract Report to SEPA, 60pp.*

Rawlins, B.G., Lark, R.M., Webster, R. and O'Donnell, K.E. 2006. The Use of Soil Survey Data to Determine the Magnitude and Extent of Historic Metal Deposition Related to Atmospheric Smelter Emissions Across Humberside, UK. *Environmental Pollution*, 143(3), 416-426.

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. Copper, pages 70-73.

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.

Salminen, R. (chief ed.) *et al*. 2005. Geochemical Atlas of Europe. Part 1 - Background Information, Methodology and Maps. Geological Survey of Finland, Otamedia Oy, Espoo, 525 pp.

Sanders, J.R., McGrath, S.P. and Adams, T.McM. 1987. Zinc, Copper and Nickel Concentrations in Soil Extracts and Crops Grown on Four Soils Treated with Metal-Loaded Sewage Sludges. *Environmental Pollution*, 44 (3), 193-210.

Smyth, D. 2007. Methods used in the Tellus Geochemical Mapping of Northern Ireland. British Geological Survey Open Report OR/07/022. 89pp. Available on-line from NORA.

Thornton, I. and Abrahams, P. 1993. Soil ingestion — a major pathway of heavy metals into livestock grazing contaminated land. *Science of The Total Environment*, 28(1–3), 287-294.

Tipping, E., Rieuwerts, J., Pan, G., Ashmore, M.R., Lofts, S., Hill, M.T.R., Farago, M.E. and Thornton, I. 2003. The Solid–Solution Partitioning of Heavy Metals (Cu, Zn, Cd, Pb) in Upland Soils of England and Wales. *Environmental Pollution*, 125(2), 213-225.

Tye, A.M., Young, S., Crout, N.M.J., Zhang, H., Preston, S., Zhao, F.J., McGrath, S.P. 2004. Speciation and solubility of Cu, Ni and Pb in contaminated soils. European Journal of Soil Science, 55, 579-590.

Webb, J.S, Thornton, I, Thompson, M, Howarth, R.J, and Lowenstein, P.L. 1978. *The Wolfson geochemical atlas of England and Wales*. Clarendon Press, Oxford.