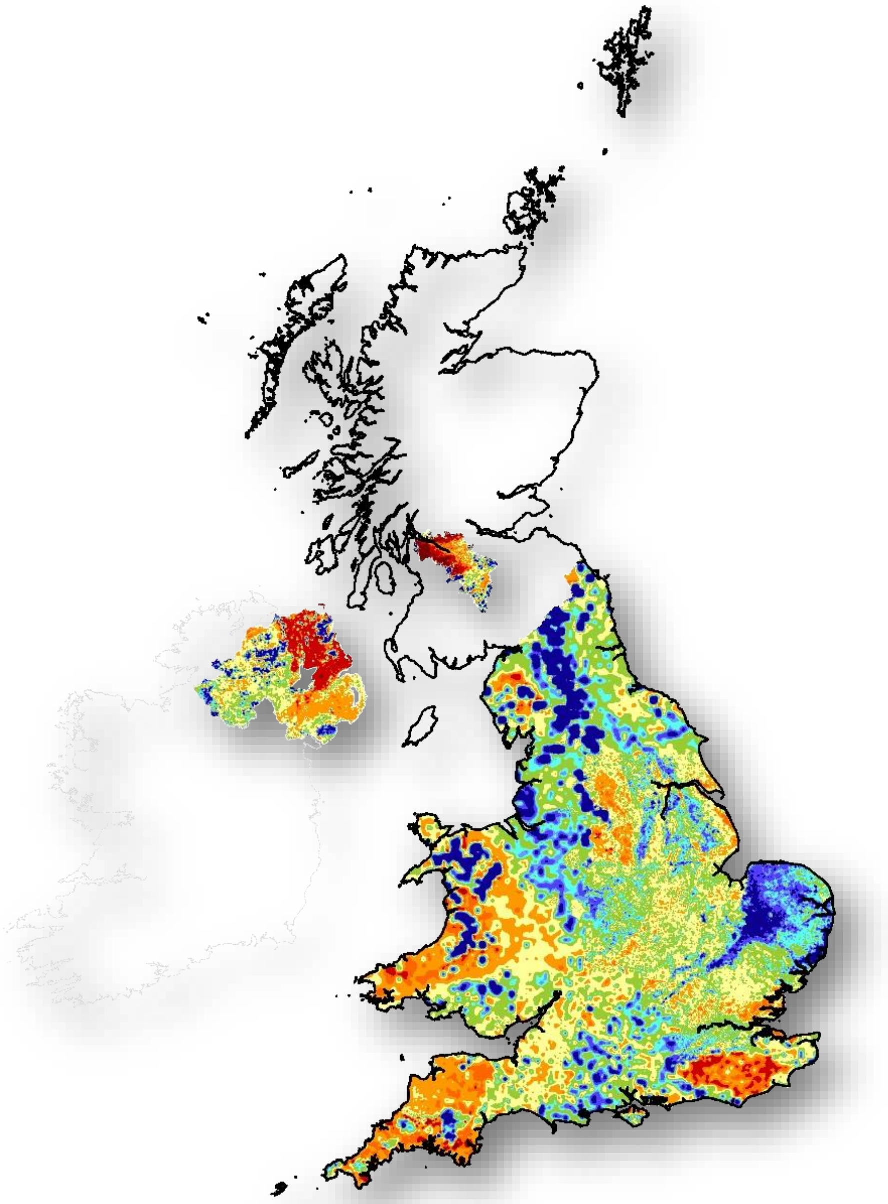

BGS DIGITAL

User Guide: BGS UK Compiled Topsoil raster dataset (UKTSraster)

Open report OR/22/021

British Geological Survey

BRITISH GEOLOGICAL SURVEY

BGS INFORMATICS

OPEN REPORT OR/22/021

User Guide: BGS UK Compiled Topsoil raster dataset (UKTSraster)

British Geological Survey

Keywords

Topsoil, XRF, raster, geoTIFF, png, geochemical.

Front cover

Map showing raster grid of titanium concentrations in UK Topsoil.

Bibliographical reference

BRITISH GEOLOGICAL SURVEY 2023. User Guide: BGS UK Compiled Topsoil raster datas(UKTSraster). *British Geological Survey Open Report*, OR/22/021. 23pp.

Keyworth, Nottingham British Geological Survey 2023

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 UK Research and Innovation.

British Geological Survey offices

Nicker Hill, Keyworth,
Nottingham NG12 5GG

Tel 0115 936 3100

BGS Central Enquiries Desk

Tel 0115 936 3143
email enquiries@bgs.ac.uk

BGS Sales

Tel 0115 936 3241
email sales@bgs.ac.uk

The Lyell Centre, Research Avenue South,
Edinburgh EH14 4AP

Tel 0131 667 1000
email scotsales@bgs.ac.uk

Natural History Museum, Cromwell Road,
London SW7 5BD

Tel 020 7589 4090
Tel 020 7942 5344/45
email bgslondon@bgs.ac.uk

Cardiff University, Main Building, Park Place,
Cardiff CF10 3AT

Tel 029 2167 4280

Maclean Building, Crowmarsh Gifford,
Wallingford OX10 8BB

Tel 01491 838800

Geological Survey of Northern Ireland, Department of Enterprise, Trade & Investment, Dundonald House, Upper Newtownards Road, Ballymiscaw,
Belfast, BT4 3SB

Tel 01232 666595
www.bgs.ac.uk/gsni/

Natural Environment Research Council, Polaris House, North Star Avenue, Swindon SN2 1EU

Tel 01793 411500 Fax 01793 411501
www.nerc.ac.uk

UK Research and Innovation, Polaris House,
Swindon SN2 1FL

Tel 01793 444000
www.ukri.org

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

Foreword

The British Geological Survey (BGS) is a world-leading geological survey, focusing on public-good science for Government, and research to understand earth and environmental processes. We are the UK's premier provider of objective and authoritative geoscientific data, information and knowledge to help society to:

- use its natural resources responsibly
- manage environmental change
- be resilient to environmental hazards

We provide expert services and impartial advice in all areas of geoscience. As a public sector organisation, we are responsible for advising the UK Government on all aspects of geoscience as well as providing geological advice to industry, academia and the public. Our client base is drawn from the public and private sectors both in the UK and internationally.

This publication is a result of the British Geological Survey (BGS) Geochemical Baseline Survey of the Environment (G-BASE) project in collaboration with the Geological Survey Northern Ireland (GSNI) and Rothamsted Research. The BGS is a component of the Natural Environment Research Council (NERC), part of UK Research and Innovation (UKRI). GSNI, our sister organization, is part of the Department for the Economy Northern Ireland. Rothamsted Research receives strategic funding from the Biotechnology and Biological Sciences Research Council (BBSRC).

DATA PRODUCTS

BGS produces a wide range of data products that align to Government policy and stakeholder needs. These include baseline geological, engineering properties, and geohazards datasets. These data products are developed using in-house expertise and are based on the outputs of our research programmes and national data holdings.

Our data products are informed by stakeholder focus groups, identifying gaps in current knowledge and policy assessments. They help to improve understanding and communication of the impact of geo-environmental properties and hazards in Great Britain, thereby improving the resilience of society and enabling people, businesses, and the government to make better-informed decisions.

Acknowledgments

This user guide was written by Antonio M P J Ferreira, T R Lister, Paul Everett, Fiona M Fordyce, and Russell S Lawley. A large number of individuals contributed to the dataset, at all stages of the project. The endeavours of (i) hundreds of voluntary workers who collected the samples, often in extreme weather conditions, (ii) staff who worked on planning, field team management, data quality assurance and interpretation (iii) lab technicians who prepared and analysed over 57,000 samples, are gratefully acknowledged.

The G-BASE project was funded by several UK government departments throughout its lifetime, including the Department of Trade and Industry (DTI), Department of Education and Science (DES), Office of Science and Technology (OST), Department of Business, Innovation and Skills (BIS) and the Department for Business, Energy and Industrial Strategy (BEIS). The TellusNI survey was funded by the Department of Enterprise, Trade and Investment (DETI), now the Department for the Economy (DfE) in Northern Ireland and the INTERREG IVA programme of the European Union (EU) Regional Development Fund.

Acknowledgement is also due to Rothamsted Research and the National Soil Resources Institute, Cranfield University, for facilitating access to the National Soil Inventory of England and Wales sample archive.

Contents

FIGURES

TABLES

Summary

This user guide outlines the background and methods used to produce the British Geological Survey (BGS) UK Compiled Topsoil raster dataset (UKTSraster). This data product consists of 41 interpolated raster grids in Georeferenced TIFF image (GeoTIFF) format, each accompanied by a colour-classified geochemical map in PNG image format, based on the corresponding GeoTIFF. These images display the concentrations of 41 chemical elements in UK topsoil.

The GeoTIFF and PNG files are available for download via the BGS OpenGeoscience web page (http://www.bgs.ac.uk/opengeoscience/downloads.html), and are also displayed in the UK Soil Observatory (UKSO) map portal (http://www.ukso.org).

The UKTS dataset, which has been used to create the UKTSraster, is based on geochemical analyses of 57,966 topsoil samples, which were collected between 1978 and 2014. Sampling covers all of England, Wales, and Northern Ireland, and part of Scotland. The topsoil samples were analysed by X-Ray Fluorescence spectrometry (XRF).

1 Introduction

1.1 BACKGROUND TO THE DATASET

The UK Compiled Topsoil Dataset (UKTS), on which the UKTSraster is based, is the most extensive topsoil geochemistry dataset for the UK available at the time of release (March 2023). The UKTS was brought together by combining data from the following sources (Table 1):

(i) the British Geological Survey's (BGS) Geochemical Baseline Survey of the Environment (G-BASE) rural and urban topsoil dataset (which accounts for 76.4% of the topsoil samples included in the UKTS)

(ii) the Geological Survey of Northern Ireland (GSNI) TellusNI rural and urban topsoil geochemical survey dataset (13.8% of the UKTS samples)

(iii) the BGS-Rothamsted Research X-ray Fluorescence Spectrometry (XRF) rural soil dataset (RR-BGS XRF), based on sub-samples held at Rothamsted Research from the National Soil Inventory (NSI) of England and Wales sample archive, National Soil Resources Institute, Cranfield University (9.8% of the UKTS samples).

The TellusNI rural topsoil dataset was first published by Young, M E and Donald A M (eds) (2013). The RR-BGS XRF dataset was published in 'The advanced soil atlas of England and Wales' (Rawlins et al 2012), which is available on the BGS website (www.bgs.ac.uk/gbase/advsoilatlasEW.html). BGS recently published the 'Rural and Urban Topsoil geochemical atlas for England, Wales, Northern Ireland and part of Scotland' (Everett et. al. 2023), which is also based on the UKTS, and is the first geochemical atlas which has incorporated data from the entire G-BASE topsoil dataset.

The UKTS consists of geochemical data for 57,966 topsoil samples (Table 1) collected between 1978 and 2014 across England, Wales, Northern Ireland and part of Scotland, in both rural and urban settings (Figure 1). Although the collection of these samples involved considerable effort, sample coverage for the whole UK is incomplete, and samples are unevenly distributed (Figure 1). Sampling densities range from 1 sample / 25 km^2 (RR-BGS XRF dataset) to 4 samples / 1 km^2, in urban centres sampled by the G-BASE and TellusNI projects.

The <2 mm size fraction of topsoil samples were analysed by X-ray fluorescence spectrometry (XRF) to determine the total concentration of more than 50 chemical elements; data for 41 of these elements is included in the UKTSraster dataset (Table 2).

Data was acquired with strict quality control procedures in place at all stages from sample collection to analytical data reporting. Accuracy and precision of the analytical data and variability of the sampling and lab procedures have been assessed by including analyses of certified reference materials (CRMs), field duplicates and lab replicates in each analytical batch, with the CRMs used to ensure continuity between different batches and field campaigns. Further detail about the sampling, sample analysis and data quality control procedures followed to produce the UKTS dataset can be found in sections 2, 3 and 4 of the 'Rural and Urban Topsoil geochemical atlas for England, Wales, Northern Ireland and part of Scotland' (Everett et. al. 2023).

The UKTS serves as a baseline soil geochemistry dataset containing valuable information about the geochemistry of the environment around 2000 AD, against which future changes can be compared. This dataset can be used for many applications in different fields including human health, food production, environment and wildlife, mineral resources, social-economic studies and forensics. It can be used for national to local level modelling, planning and decision making. Previous applications of the dataset can be found in the scientific literature, e.g. Cave et al. 2018; Dempster et al. 2016; Ferreira et al. 2018; Wragg et al. 2018; Palmer et al. 2016; Palmer et al. 2015; Barsby et al. 2012.

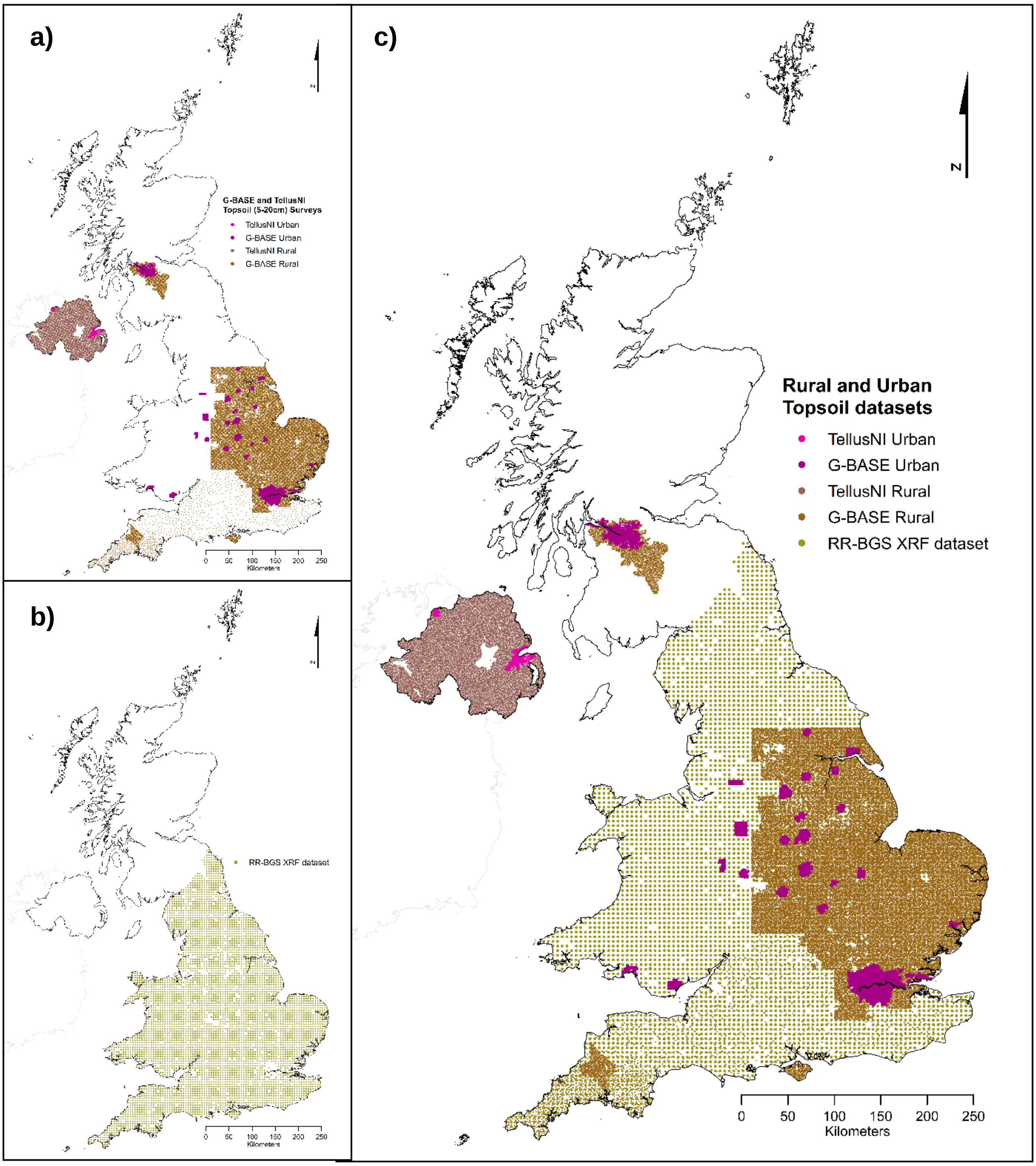

Figure 1. Sample locations for a) the G-BASE and TellusNI surveys; b) the RR-BGS XRF dataset; and c) the UK Compiled Topsoil dataset (UKTS), which combines all three. Ordnance Survey data © Crown Copyright and database rights 2024. Rothamsted Research XRF dataset, based on subsamples held at Rothamsted Research from the National Soil Inventory of England and Wales archive, National Soil Research Institute, Cranfield University.

Table 1. Source of the 57,966 topsoil samples in the UKTS dataset.

Project / dataset	Setting	Number of samples	Depth (m)	Sampling dates	Samples per km^2	Analysis dates
BGS G-BASE	Rural	26 787	0.05 - 0.2	1994 - 2014	0.5 - 0.025[1]	1994 - 2014
	Urban	17 483	0.05 - 0.2	1993 - 2009	4	1993 - 2009
TellusNI	Rural	6 862	0.05 - 0.2	2004 - 2006	0.5	2004 - 2006
	Urban	1 166	0.05 - 0.2	2004 - 2006	4	2004 - 2006

RR-BGS XRF	Rural	5 668	0.0 - 0.15	1978 - 1982	0.04	2008 - 2009

(1) 0.5 in Eastern England, Scotland, Isle of Wight and Tamar Basin, 0.1 in Cornwall and 0.025 in South England

Table 2. Summary of elements presented in the UK Compiled Topsoil raster dataset (UKTSraster).

The units ‘wt%’ and ‘mg/kg’ refer to weight percent and milligrams per kilogram respectively; ‘LLD’ refers to the lower limit of detection of the XRF analysis.

Element		Units	LLD	Element		Units	LLD
Symbol	Name			Symbol	Name		
Al_2O_3	Aluminium	wt%	0.05	**I**	Iodine	mg/kg	0.35
CaO	Calcium	wt%	0.025	**La**	Lanthanum	mg/kg	0.6
Fe_2O_3	Iron	wt%	0.005	**Mo**	Molybdenum	mg/kg	0.2
K2O	Potassium	wt%	0.007	**Nb**	Niobium	mg/kg	0.7
MgO	Magnesium	wt%	0.1	**Nd**	Neodymium	mg/kg	4
MnO	Manganese	wt%	0.005	**Ni**	Nickel	mg/kg	1.3
Na_2O	Sodium	wt%	0.2	**Pb**	Lead	mg/kg	1.2
P_2O_5	Phosphorus	wt%	0.01	**Rb**	Rubidium	mg/kg	0.7
SiO_2	Silicon	wt%	0.1	**Sb**	Antimony	mg/kg	0.25
TiO_2	Titanium	wt%	0.004	**Sc**	Scandium	mg/kg	2.4
As	Arsenic	mg/kg	0.9	**Se**	Selenium	mg/kg	0.1
Ba	Barium	mg/kg	0.5	**Sn**	Tin	mg/kg	0.25
Br	Bromine	mg/kg	0.7	**Sr**	Strontium	mg/kg	0.8
Cd	Cadmium	mg/kg	0.25	**Th**	Thorium	mg/kg	0.7
Ce	Cerium	mg/kg	0.6	**U**	Uranium	mg/kg	0.2
Co	Cobalt	mg/kg	1.5	**V**	Vanadium	mg/kg	2.7
Cr	Chromium	mg/kg	2.8	**W**	Tungsten	mg/kg	0.6
Cs	Caesium	mg/kg	0.5	**Y**	Yttrium	mg/kg	0.8
Cu	Copper	mg/kg	1.2	**Zn**	Zinc	mg/kg	1.1
Ga	Gallium	mg/kg	1	**Zr**	Zirconium	mg/kg	0.8
Hf	Hafnium	mg/kg	1				

1.2 CURRENT VERSION OF THE DATA

This is the first published version of the UK Compiled Topsoil raster (UKTSraster) dataset.

1.3 CITING THIS DATASET

Please cite this dataset as follows:

British Geological Survey. (2024). UK Compiled Topsoil raster dataset (UKTSraster). British Geological Survey. (Dataset). https://doi.org/10.5285/76b69adf-699b-4032-a751-2db0991d55f6

1.4 WHAT THE DATASET SHOWS

The UKTSraster dataset shows predicted spatial variation in element concentrations in topsoil across England, Wales, Northern Ireland and part of Scotland. The data is based on the UKTS topsoil data points shown in Figure 1. The UKTSraster is presented as GeoTIFF raster image layers produced by interpolating point data using ordinary kriging (a method which generates a surface between known data points) into a regular grid with a cell size of 500m x 500m. The GeoTIFF files are spatially referenced to the British National Grid (BNG) coordinate system.

For each element, two raster layers are provided: one containing predicted concentration values (Figure 2), and a second containing prediction standard error values (Figure 3) which provides an indication of the uncertainty associated with the predicted concentrations.

The raster layers can be viewed by simply dragging and dropping the files into GIS software such as QGIS, ArcMap or MapInfo. Any colour classification scheme can be assigned to the grids by the user; the scheme routinely used for predicted concentrations in G-BASE publications is provided as a lookup table, a layer file and a PNG image for each element.

The UKTSraster includes layers for a suite of 41 elements, namely 10 major elements aluminium (Al), calcium (Ca), iron (Fe), potassium (K), magnesium (Mg), manganese (Mn), sodium (Na), phosphorus (P), silicon (Si), and titanium (Ti) presented as oxide weight percent (wt%) – Al_2O_3, CaO, Fe_2O_3, K_2O, MgO, MnO, Na_2O, P_2O_5, SiO_2 and TiO_2 respectively, and 31 trace elements (As, Ba, Br, Cd, Ce, Co, Cr, Cs, Cu, Ga, Hf, I, La, Mo, Nb, Nd, Ni, Pb, Rb, Sb, Sc, Se, Sn, Sr, Th, U, V, W, Y, Zn, Zr) presented in milligrams per kilogram (mg/kg) (Table 2). The lower limit of detection (LLD), that is, the lowest concentration that could be reliably determined by the XRF analysis, is also shown in Table 2.

The source UKTS dataset holds data for other elements, but these have not been included in the UKTSraster dataset due to limited coverage and/or inconsistent or low level of data.

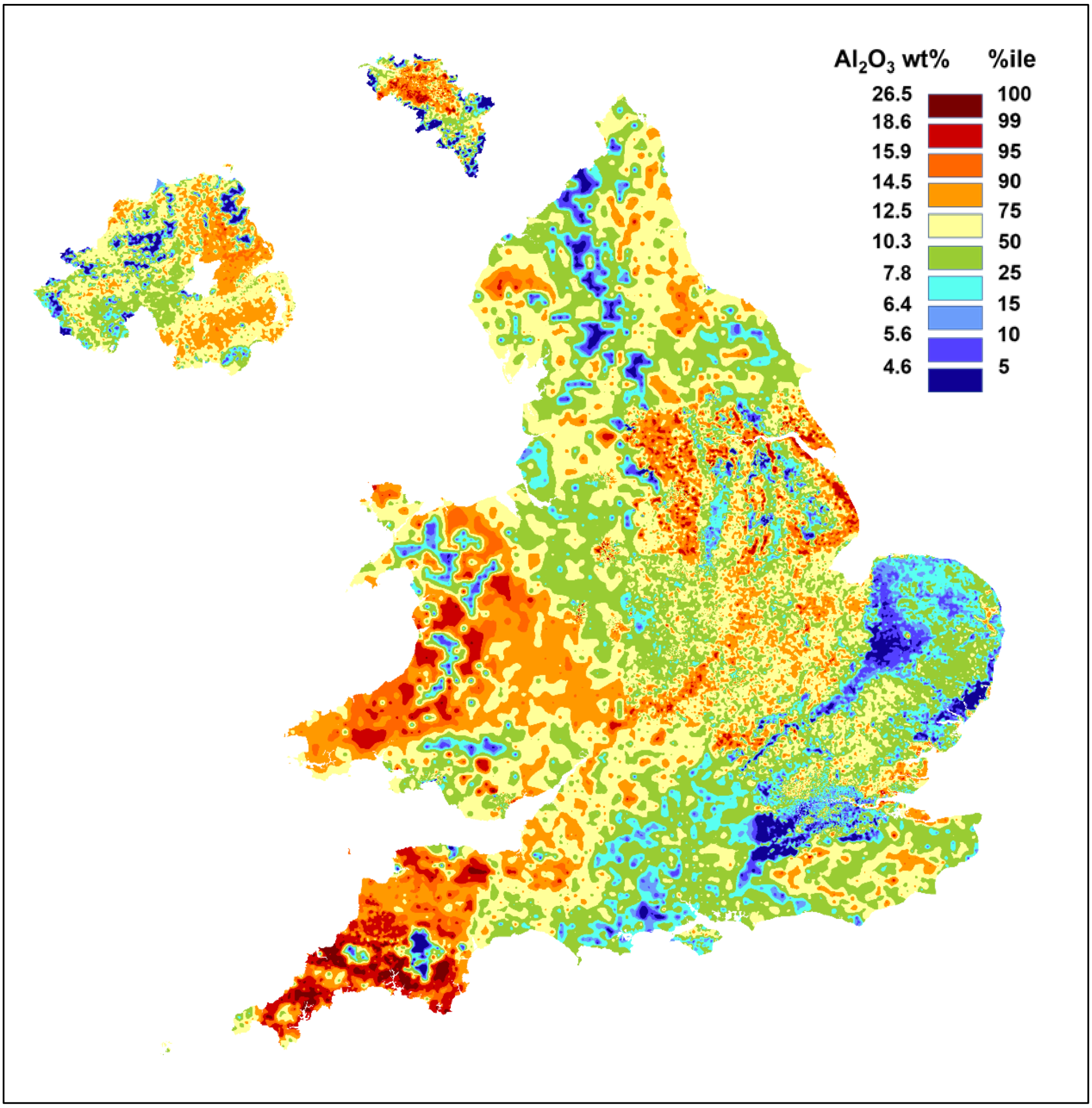

Figure 2. GeoTIFF raster for aluminium (Al_2O_3) concentrations (in wt%) predicted from topsoil samples. The GeoTIFF raster of predicted element concentrations is classified using the 10 colour scheme employed by the G-BASE project. This classification is based on percentiles (%ile) of the source data points.

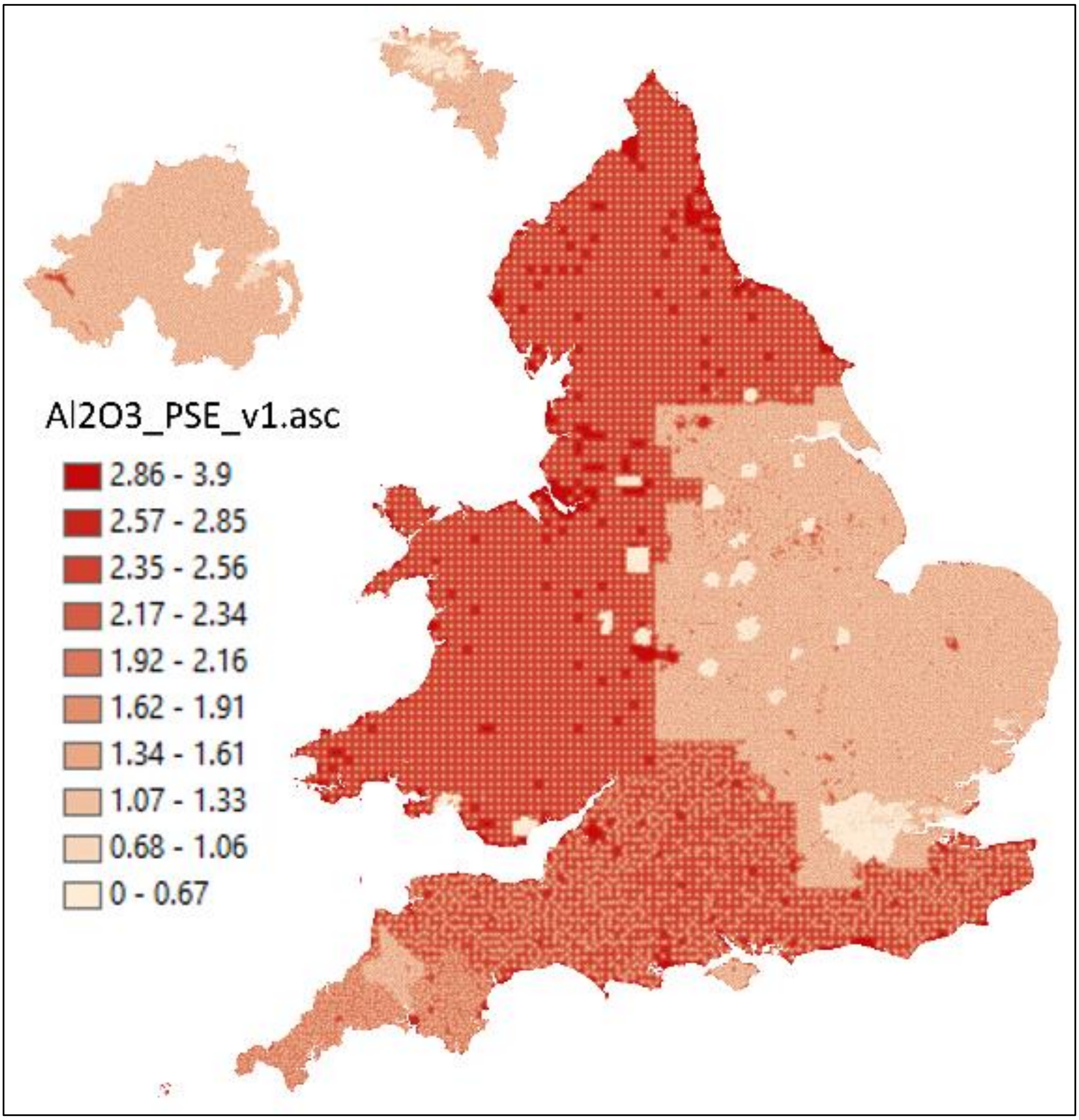

Figure 3. GeoTIFF raster of the Prediction Standard Error (PSE) map for Al_2O_3 (in wt%), which provides an estimate of the error level for the Al_2O_3 predicted concentrations (shown in Figure 2). The PSE maps are shown using 10 classes defined according to natural breaks (jenks). Higher error values are depicted with darker colours. Rothamsted Research XRF dataset, based on subsamples held at Rothamsted Research from the National Soil Inventory of England and Wales archive, National Soil Research Institute, Cranfield University.

2 Case study: cobalt in Northern Ireland topsoil

This section of the user guide presents an example application of the UKTSraster dataset, in the form of a case study on Cobalt (Co) concentrations in Northern Ireland (NI) topsoil. A study of this type could be conducted as part of a review of regional land use and soil quality.

2.1 THE CHALLENGE

The review aims to examine the availability of nutrients in the topsoil, among other geological, land use, chemical and physical environmental characteristics. Cobalt (Co) is one of the nutrients to be screened, as Co is essential for synthesis of vitamin B12, which has a powerful effect on the immune system of grazing cattle. The UKTSraster dataset is to be investigated in order to identify areas where low Co concentrations in topsoil may result in Co deficiency in cattle.

2.2 THE SOLUTION

The UKTSraster **Co.tif** layer was used as a primary estimate for the potential of cobalt deficiency in topsoil. It was imported to a GIS underlain by a topographic base map and viewed at a regional scale for NI (Figure 4). This map shows the main geochemical patterns of predicted Co total concentrations in topsoil across NI. This allows areas of high Co concentrations (e.g. the Eastern half of NI, notably most of the NE sector, and an NNW trend with an ENE extension in the Western half of NI) and areas of low Co concentrations (e.g. the far NE, the Southernmost area and most of the Western half) to be identified.

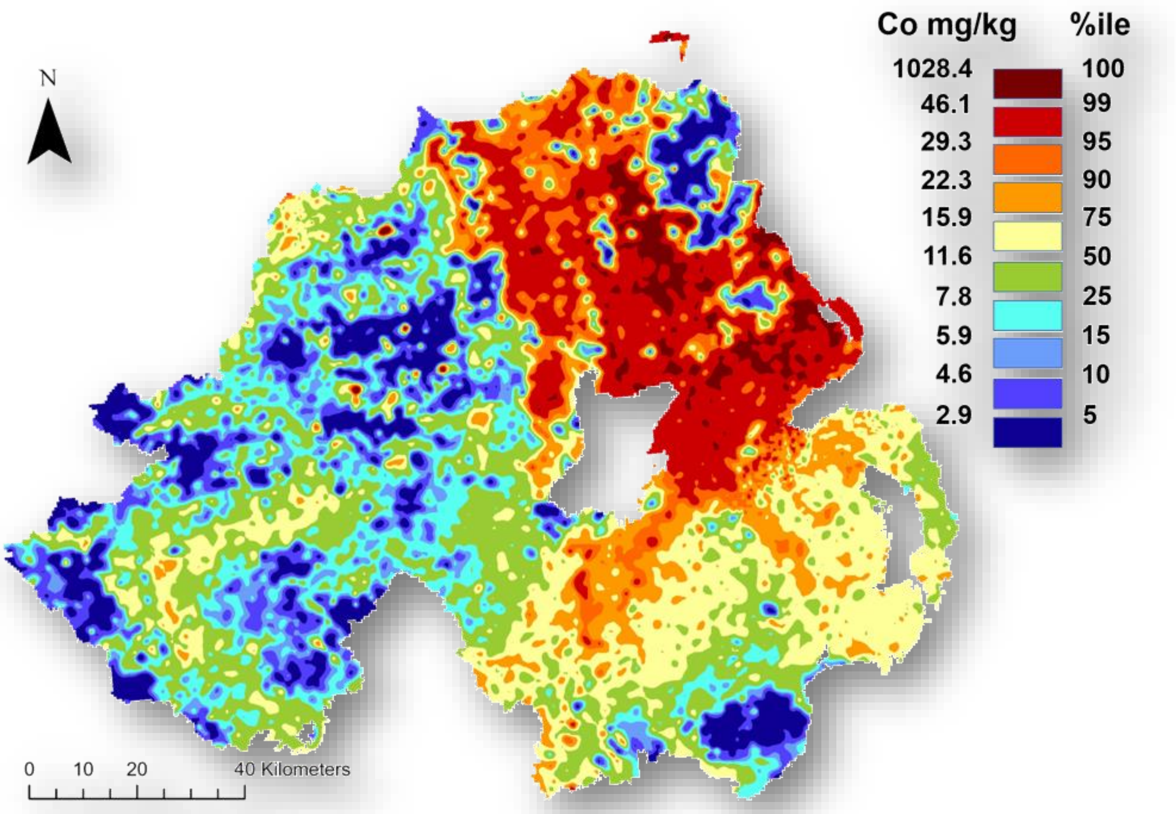

Figure 4. GeoTIFF raster map of predicted cobalt concentrations in topsoil across Northern Ireland.

Only a fraction of the total Co in topsoil will be effectively available to plants and grazing cattle, so the study needs to be extended to better account for the bioavailability of Co, and the consequent effect on Co deficiency.

To achieve this, the Co map can be compared with other map layers that may be relevant for interpreting Co availability in the surface environment. These include other UKTSraster dataset variables (e.g., iron (Fe), manganese (Mn), etc.), soil acidity (pH), the geology (bedrock is the primary source of Co in soil) and other environmental parameters available from other dataset sources.

A modelling approach to identify areas where cattle are at risk of cobalt deficiency can now be designed, which will take into account one or more related environmental factors and/or one or more elements in soil, such as manganese (e.g. Lark et al. 2014).

2.3 THE OUTCOMES AND VALUE

The Co map and other UKTSraster layers of predicted data can be used for various purposes. Examples include:

- nutrient availability (e.g. cobalt, selenium (Se), copper(Cu), zinc (Zn), iodine (I)) for crop production and cattle health;
- soil erosion;
- naturally enriched and / or contaminated soil in elements that may be harmful to humans and wild life (e.g. arsenic (As), cadmium (Cd), lead (Pb), tin (Sn), uranium (U), etc.;
- enrichment in strategic elements (e.g. https://publications.parliament.uk/pa/cm201012/cmselect/cmsctech/726/726.pdf);
- social and economic sciences;
- forensic science;
- resilience to climate change.

The cobalt map and a derived Co deficiency model is an example that can help for identifying naturally favourable areas for certain types of land use, such as crop production and grazing cattle. This can now be used to plan and make decisions seeking a more efficient and sustainable use of the land.

The observation of the Co map (Figure 4) together with other layers of interest allows spatial associations with other features, such as, geology, topography, water bodies, etc., to be identified. Evaluation of the data indicates high Co concentrations are present in the eastern half of NI (orange to red colours). These are associated with Palaeogene (about 60 million years (MA)) volcanic bedrock assigned to the Antrim Lava Group. The low Co concentrations (dark blue to green colours) are associated with the slightly younger granitic rocks of the Mourne mountains and the Neoproterozoic (about 600 – 700 MA) rocks of the Antrim mountains. Intermediate to high Co concentrations (yellow to orange colours) are present in the western half of NI, near Lough Erne and the River Foyle, and are associated with Devonian (c. 360 – 420 MA) sandstones. Most of the lowest Co concentrations (deep blue colour) are observed in mountainous areas, above 150 – 200 m.

In summary, this case study demonstrates how the UKTSraster dataset provides increased knowledge and understanding about soil and the environment in the UK, and strengthens the data available for strategic land use planning. It thereby helps to enable robust planning and decisions based on scientific knowledge.

3 Methodology

The GeoTIFF and PNG images in the UKTSraster dataset were produced in ArcGIS 10.8 (ESRI®) using the Geostatistical Analyst toolbox. For each element, the source point data was interpolated by ordinary kriging (see Appendix 2 for details). The interpolated estimates are computed from the source data points into a regular 500 x 500 m grid, which is shown in the form of a continuous surface map.

4 Technical information

4.1 SCALE

The UKTSraster dataset presents geochemical predictions on a regular 500 x 500 m grid, that is, one prediction per 0.25 km^2 grid node. This scale is equivalent to the sampling density for the sampled urban areas (1 sample point per 0.25 km^2). The sampling density of the remaining point data, however, ranges from 1 sample per 25 km^2 for Wales and Northern England, to approximately 1 sample per 1.9 km^2 on average for Eastern England (this takes into account 1 sample per 2 km^2 from the rural G-BASE survey plus 1 sample per 25 km^2 from the RR-BGS XRF project). The sampling density across most of Southern England is approximately 1 sample per 15.4 km^2 (taking into account 1 sample per 40 km^2 from the rural G-BASE survey plus 1 sample per 25 km^2 from the RR-BGS XRF project) (Figure 1, Table 1).

As a result of the varied sampling density, the prediction modelling has produced less 'smooth' geochemical features in areas of the map where sampling density is higher, such as in sampled urban areas and in Eastern England (Figure 2).

4.2 COVERAGE

The UKTSraster Dataset covers approximately 70% of the UK landmass, including all of England, Wales and Northern Ireland and a small part of Scotland (specifically Glasgow and the drainage basin of the River Clyde, which accounts for about 4% of Scotland).

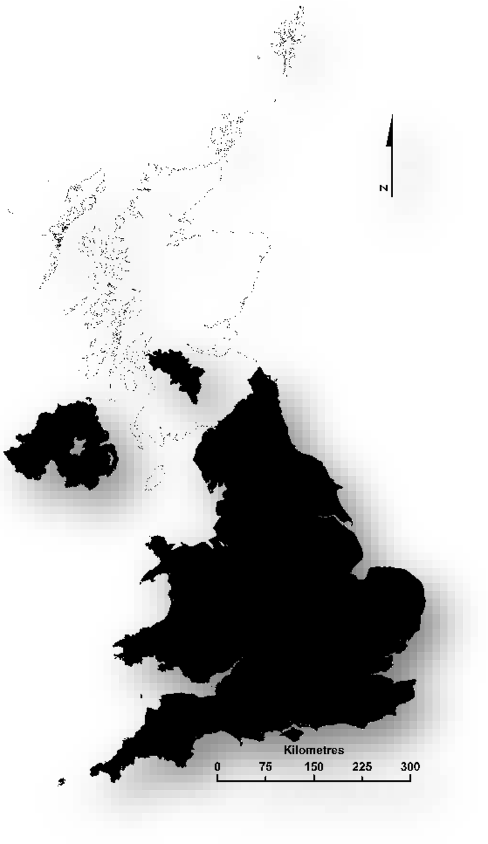

Figure 5. Approximately seventy percent of the United Kingdom (the area in grey) is covered by the UKTSraster dataset. Contains Ordnance Survey data © Crown Copyright and database rights 2023. Contains Ordnance Survey Ireland data © Crown Copyright and database rights 2023. Lough Neagh outline from Flanders Marine Institute (2021).

4.3 ATTRIBUTE DESCRIPTION

For each element, the two GeoTIFF raster images show the predicted concentration (e.g.: Al2O3_v1.tif, Figure 2) and the respective prediction standard error (e.g.: Al2O3_PSE_v1.tif, Figure 3) at the same nodes of the regular 500x500m grid used for the ordinary kriging and in the units reported in Table 2. Data in the grid nodes that fall out with the extent of the sample coverage (Figure 4) are classified as 'no data' ('NoData Value'), for which a low negative value (-9999) is recorded.

The accompanying ArcGIS layer (.lyr) file (e.g.: for Al_2O_3: 'Al2O3_v1.lyr') provides a symbology definition for the GeoTIFF raster maps of predicted concentrations (e.g.: for Al_2O_3: Al2O3_v1.tif, Table 3) based on the G-BASE colour classification scheme (Table 4). The G-BASE classification scheme is also provided as a CSV table and a PNG image.

The Portable Network Graphic (PNG) files are high resolution static images of the predicted concentrations (e.g. for Al_2O_3: Al2O3_raw_v1.png, Al2O3_v1.png and Al2O3_urban_v1.png, Table 3, Figure 6 in Appendix 1).

4.4 DATA FORMAT

The data format types are identified in Table 3, below.

Table 3. Format types available in the UKTSraster dataset.

The two hashes (##) in the file name column correspond to the formula (for the major elements) or symbol (for the trace elements) of the chemical element.

File name	File type	File description
##_v1.tif	GeoTIFF raster with floating point	Georeferenced raster image showing the element concentration in topsoil predicted on the nodes of a 500x500 m grid. Units are wt% or mg/kg (see Table 2).
##_PSE_v1.tif	GeoTIFF raster with floating point	Georeferenced raster image showing the prediction standard error (PSE) of the element concentrations in topsoil predicted on the nodes of a 500x500 m grid. Units are wt% or mg/kg (see Table 2).
##_colour_table_v1.csv	CSV table	Table showing the concentration ranges and colours required to apply the BGS G-BASE colour scheme. The scheme is based on percentiles of the measured values in the source point data. Colours are defined using both HEX and RGB codes.
##_v1.lyr	ArcGIS layer file	An ArcGIS layer file coloured after the CSV table provided, and from which the PNG image maps have been exported from ArcGIS.
##_key_v1.png	PNG image	Colour legend image showing the concentration ranges with the BGS G-BASE colour scheme (as provided in a CSV table), to accompany the static raw map image.
##_raw_v1.png	PNG image	Static raw map image (in PNG format) of the element concentrations, using the BGS G-BASE colour scheme.
##_v1.png	PNG image	Static map image (in PNG format) of the element concentrations with cartographic and general information.
##_urban_v1.png	PNG image	Static map image in PNG format of the element concentrations with cartographic and general information, including the outline of sampled urban areas.

4.5 DATA HISTORY

This is the first published version of the UKTSraster dataset.

4.6 DISPLAYING THE DATA

The colours used to display each element according to the G-BASE classification scheme are reported in a downloadable CSV file, PNG image (Table 3), and are reproduced in Table 4. Concentration ranges (denoted in the CSV file as 'concentration_range') are colour classified according to the BGS G-BASE 10 colour scheme, which uses the following percentiles (denoted as 'percentile_range'): 0, 5, 10, 15, 25, 50, 75, 90, 95, 99 and 100 of the source point data values. The 10 colours are provided as both RGB (from RED, GREEN, BLUE, denoted as 'colour_R' 'colour_G' 'colour_B' respectively) and HEX values (denoted as 'colour_hex').

For some elements, the colour palette is reduced due to technical limitations, such as those related to limits of detection and/or low analytical sensitivity at lower concentrations. Na_2O, Cd and Cs (7 colours), Se and W (8 colours) and Sb (9 colours) are the examples. In these cases, the first colour in Table 4 (dark blue) is always used, while up to the following 3 are successively omitted from the colour classification.

Table 4. BGS G-BASE colour scheme used to display the UKTSraster dataset.

Percentile range	RED	GREEN	BLUE	HEX	LOOKS LIKE
Percentile of input sample values	This is the equivalent red channel colour	This is the equivalent green channel colour	This is the equivalent blue channel colour	This is the equivalent HEXadecimal value	This cell shows the colour as intended
0 - 5	15	0	148	#0E0093	
5 - 10	83	64	255	#5340FE	
10 - 15	107	157	250	#6B9EFA	
15 - 25	89	255	241	#59FFF1	
25 - 50	153	204	51	#99CC33	
50 - 75	255	255	153	#FEFF99	
75 - 90	255	153	0	#FE9900	
90 - 95	255	102	0	#FF6600	
95 - 99	204	0	0	#CC0001	
99 - 100	120	0	0	#780001	

5 Limitations

The UKTSraster dataset has the following limitations, which are related to the input datasets, spatial scale, and methodological uncertainty:

5.1 DATA CONTENT

The UKTSraster dataset is based on and limited to an interpretation of the G-BASE (rural and urban), the TellusNI (rural and urban) and the rural RR-BGS XRF data records at the time when the dataset was assembled.

The values within this dataset are single-element concentrations limited by the 5 surveys on which they are based (e.g. the strategy, the coverage, the sampling density, the quality of the data) and by the methods used to assemble, process and display the data. Given the methodology used to create the UKTSraster dataset (described in Appendix 2), the predicted single-element values provided here are to the best of our knowledge and current data holdings. Each map consists of predictions generated by a geostatistical model only, and does not display the actual sample measurements. Also, whilst the data is compositional in nature, it has not been treated as such in the single-element map production (see the entry for 'compositional data' in the glossary at the end of this user guide for further information).

5.2 SCALE

The UKTSraster dataset is intended to be used at national (1:2,500,000) to regional (1:1,000,000) scales. In all circumstances, any interpretation of the dataset must take into consideration the resolution of the prediction grid (i.e. 500x500m) and the respective original scale of capture (sampling density), which varies across the surveyed land (Table 1, Figure 1).

At the scale of the 25 surveyed urban centres, for example, the UKTSraster dataset can be used with a resolution up to 500 x 500 m for each prediction grid cell (nominal scale: 1:500,000, zoom in up to 1:250 000 scale) as this scale coincides with the respective sampling density.

5.3 REGULAR GRID

The source point data from TellusNI dataset was transformed to the coordinate system (BNG) of the other datasets used (G-BASE and RR-BGS XRF), so as to join the several datasets into a single consistent dataset, and to allow processing the data into a common regular 500 x 500 m grid that consistently covers the area of interest (NI, England, Wales and part of Scotland).

The projection of the UKTSraster dataset into another coordinate system will cause distortion of the cell size and cell shape of the original regular 500 x 500 m grid.

5.4 ACCURACY AND UNCERTAINTY

The topsoil point source data has an inherently undesirable variability related to sampling, sample preparation and analytical procedures. This variability was estimated for each element using a nested analysis of variance (ANOVA) based on several hundred sets of field duplicate and lab replicate samples (Everett et. al. 2023). For the elements presented in this dataset, the natural variation is significantly larger than the undesirable variability. Measurement accuracy and precision were estimated from repeated analyses of certified reference materials (CRMs), using standard procedures.

Each predicted concentration value has an inherent uncertainty related to the interpolation method (Appendix 2). The uncertainty is illustrated in the form of a GeoTIFF map image of the prediction standard error. This should however be interpreted with caution as, to some extent, the estimation of uncertainty is dependent on the sampling density (which varies substantially across the covered land), whether the data was log-transformed or not, on the variogram model used, and on the validity of the underlying assumptions of ordinary kriging. For instance, when the data does not meet the assumptions of ordinary kriging, significant under- or over-estimation of values may occur.

5.5 ARTEFACTS

The raster images were created by clipping a regular grid to the coverage extent (Figure 5), according to whether or not the centre of each grid cell lies on or off the surveyed land. As a result of this process, not all grid cells are contiguous with one-another, for example single offshore cells are present where their cell centre overlies a small land mass, and inland cells may be omitted in situations where the cell centre overlies water.

5.6 DISCLAIMER

The use of any information provided by the British Geological Survey (BGS) is at your own risk. Neither BGS nor the Natural Environment Research Council (NERC) or UK Research and Innovation (UKRI) gives any warranty, condition or representation as to the quality, accuracy or completeness of the information or its suitability for any use or purpose. All implied conditions relating to the quality or suitability of the information, and all liabilities arising from the supply of the information (including any liability arising in negligence) are excluded to the fullest extent permitted by law. No advice or information given by BGS, NERC, UKRI or their respective employees or authorised agents shall create a warranty, condition or representation as to the quality, accuracy or completeness of the information or its suitability for any use or purpose.

6 Frequently asked questions

These questions and answers have been provided to address any potential issues relating to how the product can be used or how it can be interpreted. If you have any additional questions, please contact digitaldata@bgs.ac.uk

Q: What does the BGS UKTSraster dataset show?

A: This dataset is a series of maps showing the spatial variation of 41 element concentrations (Table 2) in the soil cover across 70% of the UK (see figure 5), predicted at the nodes of a 500 x 500 grid (Figure 2) from 57,966 topsoil samples (Figure 1.)

Q: What areas does the BGS UKTSraster dataset cover?
A: This dataset covers England, Wales and Northern Ireland and Glasgow and the watershed of River Clyde in Scotland (see figure 5).

Q: In what data formats can the BGS UKTSraster dataset be provided?
A: The maps in UKTSraster are provided as GeoTIFF raster images of 500 x 500 m cell size, and high resolution static portable network graphic (PNG) images (Table 3). More specialised formats may be available but may incur additional processing costs. Please email iprdigital@bgs.ac.uk to request further information.

Q: At what map scale is the BGS UKTSraster dataset provided?
A: The UKTSraster dataset is provided and to be used at National (1:2,500,000) to Regional (1:1,000,000) scales; in the surveyed 25 urban centres the dataset can be used at larger scale, with up to the resolution of the 500 x 500 m prediction grid cell, that is, a nominal scale of 1:500,000, with a zoom in up to 1:250,000 scale only.

Q: How accurate is the BGS UKTSraster dataset?
A: The accuracy and precision of the data varies from element to element and across the space. Each predicted concentration value at each grid node has an inherent uncertainty related with the interpolation method (Appendix 2) and with the underlying topsoil source point data (Table 1, Figure 1).

The uncertainty is estimated for each element and provided in the form of a Prediction Standard Error map (in GeoTIFF format) and is acceptable for the scales mentioned above only. This, however, should be interpreted with caution as, in some extent, the estimation of uncertainty is dependent on several factors, such as, the sampling density, whether a data transformation was used or not, the variogram model used and the validity of the ordinary kriging underlying assumptions.

Q: How often will the BGS UKTSraster dataset be updated?
A: There are no updates planned for this dataset.

Q: Can I use the BGS UKTSraster dataset as part of a commercial application?
A: To encourage the use and re-use of this data we have made the raster data under the Open Government Licence, subject to the following acknowledgement which should be included in any material that reproduces or uses the BGS materials or data: "Contains British Geological Survey materials ©UKRI [2023]". The underpinning soil chemistry data is sourced from multiple organisations, please refer digitaldata@bgs.ac.uk for further information.

Appendix 1. Example geochemical map

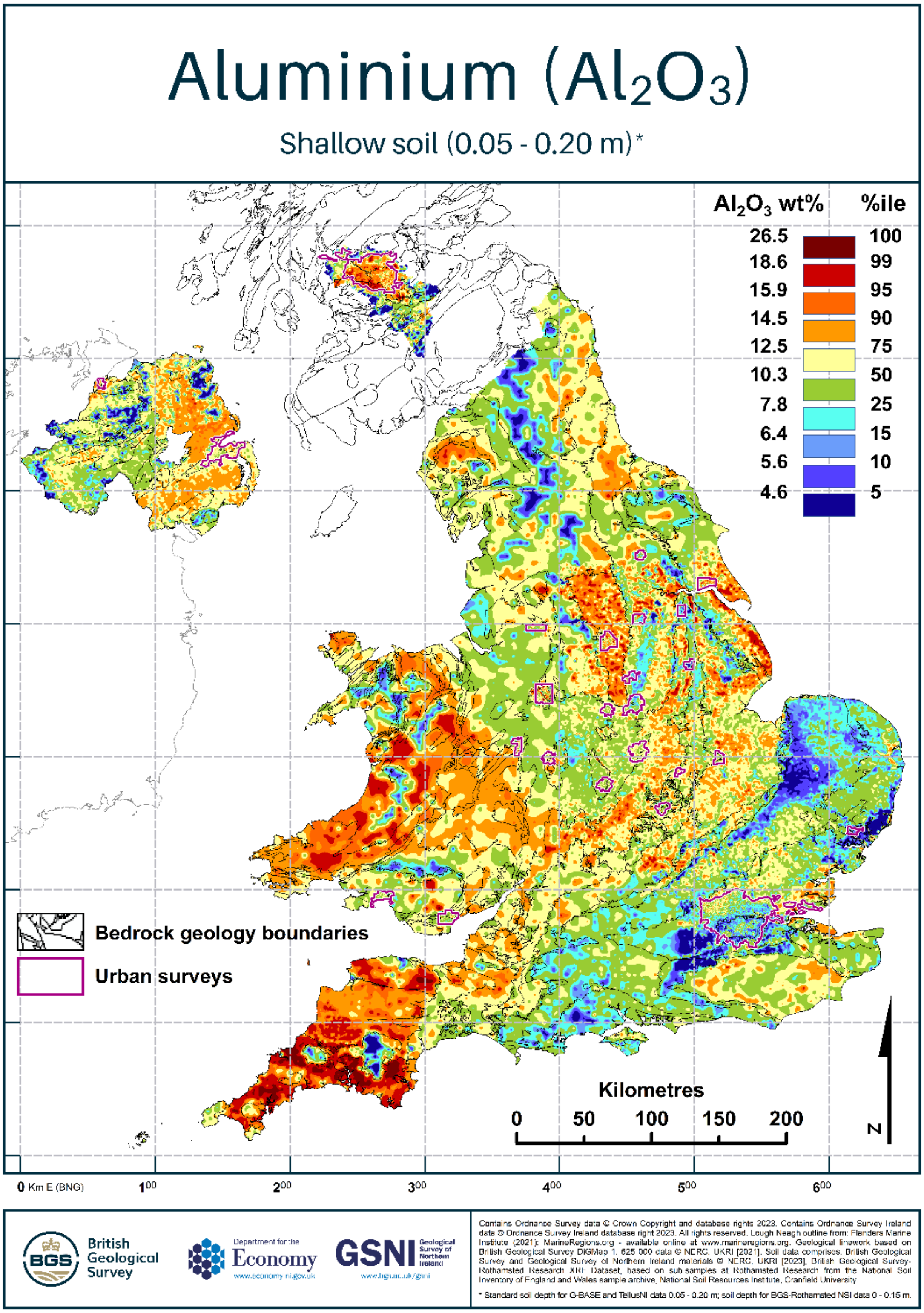

Figure 6. PNG image for aluminium (Al_2O_3) concentrations (in wt%) predicted from topsoil samples collected across most of the UK (Al2O3_urban_v1.png); this map is displayed with the BGS G-BASE 10 colour scheme based on the percentiles (%ile) and cartographic and general information including an outline for the urban surveyed areas. Contains Ordnance Survey data © Crown Copyright and database rights 2023. Contains Ordnance Survey Ireland data © Crown Copyright and database rights 2023. Lough Neagh outline from Flanders Marine Institute (2021). Rothamsted Research XRF dataset, based on subsamples held at Rothamsted Research from the National Soil Inventory of England and Wales archive, National Soil Research Institute, Cranfield University.

Appendix 2. Data interpolation

The GeoTIFF and PNG images in the UKTSraster dataset were produced in ArcGIS 10.8 (ESRI®) using the geostatistical wizard in the geostatistical analyst toolbox. For each element, the source point data was interpolated by ordinary kriging and exported to a GeoTIFF raster image with a regular 500 x 500 m grid. Two GeoTIFF maps are provided for each element. The first shows predicted values, and the second shows prediction standard error (PSE) that displays the level of uncertainty related to the predicted values.

Prior to kriging, the point data from the Tellus NI dataset was re-projected into British National Grid coordinates used by the other datasets (G-BASE and RR-BGS XRF), in order to allow all of the datasets to be processed together.

The ordinary kriging was performed after the log transformation of the source point data when this transformation reduced the skewness of the element distribution. Al_2O_3, K_2O, SiO_2, Ga, Nd, Rb, and Sc did not require the log-transformation (Table 5).

For most elements, an exponential variogram model was derived that excluded the nugget, and the default settings of the geostatistical wizard for lag size and number of lags (12) were accepted. Also, the variogram models were applied to predict concentrations at each grid node by including the nearest 5 to 8 neighbours within a maximum search distance of 10 km. These variogram and kriging options allow the short-scale geochemical variation in areas of higher sampling density to be retained. For eight of the 41 elements presented, namely Pb, Sn, Sb, Zn, Cu, U, Cd and W, the variogram was instead fitted to include the nugget. This was done for several reasons, namely the high number of values below the lower limit of detection (Cd and W); the high within-sample variance (U); and/or the biasing influence of urban samples, which, for some elements (Pb, Sn, Sb, Zn, Cu, U), form a subpopulation that significantly increases the variance within the overall population. For three out of the 41 elements, namely U, Sc and Sb, the lag values were adjusted by a factor of 200, 1000, and 100 respectively as the default lag values were too small (4.77m, 9.99m and 66.7m respectively) for the 10 km maximum search distance implemented. A list of the variogram models and parameters used are provided in Table 5.

The predicted values and respective standard error values generated and back-transformed to the original units (Table 2), were then exported to GeoTIFF raster files.

ORDINARY KRIGING

Ordinary kriging is a geostatistical interpolation method for spatial prediction based on linear probability models. It can be seen as a sophistication of simple inverse distance weighting interpolation, as in ordinary kriging, the weights given to each neighbour are estimated considering the autocorrelation structure of the data as assessed by means of a variogram model (Hengl, 2009). The spatial autocorrelation is the degree to which a variable, e.g. element concentration, correlates with itself through space. Conceptually, it evokes Tobler's first law of geography (Tobler, W R. 1970) which states that "In nature everything is related to everything else, but near things are more related than distant things". In practice, the closer two topsoil sampling sites are, the more similar the element concentrations tend to be.

Ordinary kriging makes certain assumptions, such as, stationarity (i.e. the global mean and autocorrelation structure are similar over the whole area of interest), and that the data is normally distributed.

VARIOGRAM MODELLING

An experimental variogram is a plot that shows the spatial continuity of the data, constituting the most fundamental 'tool' for kriging interpolation. It aims to depict how the data variability changes with distance between the sampled points. Typically, smaller variances at shorter distances and a stable variance (that is equal to global variance) at longer distances are expected to be depicted in the plot (Hengl, 2009). The theoretical variogram fitted to the data plotted in the experimental variogram is a mathematical function (blue line in Figure 7) that is used to optimise the weights given to the nearest neighbours during interpolation, which, in turn, allows optimal predicted concentrations at each of the cell centres of the output grid or raster.

Note that in variography (variogram modelling) a common measure for the variability is half the average squared difference between the concentrations of pair of points at a given distance (lag). This is often described as 'semivariance', which gives rise to the term 'semivariogram' (as used in ArcMap Geostatisical Analyst) used to refer to half the variogram (Figure 7). In this User Guide we use 'variogram modelling' as the basic term and 'semivariance' in specific cases.

NUGGET, SILL, RANGE

Nugget, sill and range (also referred as major range) are the fundamental parameters that define the theoretical variogram (variogram model) from the experimental variogram (the plot showing distance versus variability of source data point pairs, usually binned and averaged at a constant pace (lag) (Figure 7).

The nugget is an optional parameter, corresponding to the value of the variogram model at distance equal to zero, that is, where the variogram function intercepts the y axis (see Figure 7). It represents an un-apportioned mixture of small-scale variance, independent and measurement errors in the dataset which are expected to be encountered in the source data at spatial scales below the shortest point-pair distances. If the variogram model is fitted to include the nugget, the map of predicted values resulting from the interpolation by ordinary kriging tends to show 'smoother' patterns. When the nugget value is excluded from the variogram model, in which case it is fitted to intersect zero variance at zero distance (i.e. a y-axis intercept of 0), overfitting may occur and the patterns in the map will tend to follow more closely the source data points, resulting an image with increased dithering in areas with higher sampling density.

Sill and range are, respectively, the y-axis (semivariance) and x-axis (distance) values at which the variogram model first flattens out. Distances above the range are considered spatially independent, that is, the variance is no longer increasing with distance. Partial sill refers to a part of the variance related to the variogram model, often the variance above the nugget, that is, partial sill = sill - nugget.

UNCERTAINTY

In contrast to inverse distance weighting interpolation methods, ordinary kriging also allows the uncertainty at the grid nodes of the prediction map to be estimated, which, in the ArcGIS geostatistical wizard, is provided as a prediction standard error (PSE) map. The PSE is the square root of the prediction variance, which is the variation associated with the difference between the measured and the predicted value. This prediction variance (standard error squared) is minimized after the particular set of weights given to the nearest samples during the linear estimation by ordinary kriging.

To some extent the PSE estimation is dependent on the variogram model used and the validity of the underlying assumptions, and thus should be interpreted with caution. Generally, 95% of the measured values are within the interval of the predicted value ± 2 times the standard error.

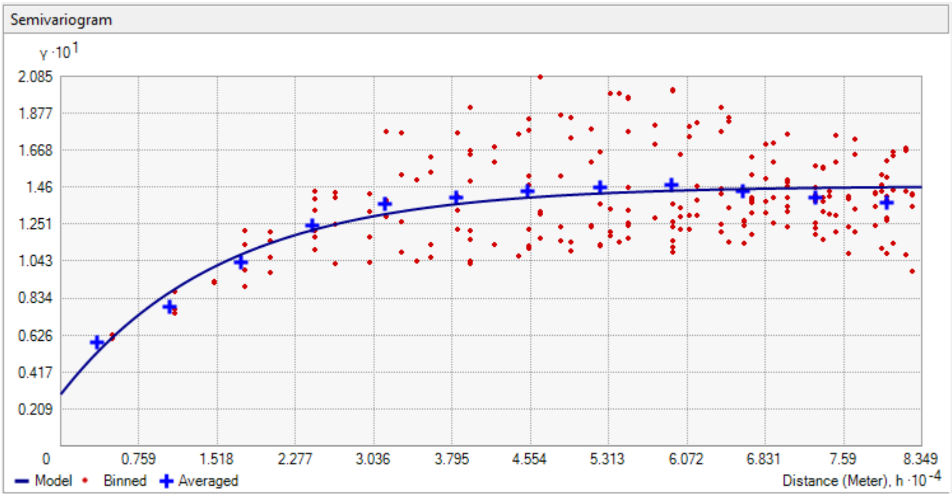

Figure 7. Example of a variogram model (blue line) with a nugget (about 0.03) fitted in ArcGIS geostatistical wizard after semivariances have been binned (red dots) and averaged (blue crosses) to lags of about 7 km. The Sill (about 0.146) is attained at a Range of about 47 km.

Table 5 - Summary table of variogram models used to make predictions at the nodes of a 500 m regular grid by ordinary kriging from the source data.

The variogram model is presented with the formula: Nugget + Partial Sill * Exponential (Major Range); where Nugget and Partial Sill are expressed as proportions of the total variance and Major Range is in metres. The default lag size in metres as calculated by the geostatistical wizard is listed, in cases where this was adjusted, the factor applied to perform the adjustment is given. The number of lags is 12 by default.

Element	Was data log transformed?	Variogram model	Lag Size	Number of Lags
Al_2O_3	No	0 + 9.58 * Exponential (10724)	9512.8	12
CaO	Yes	0 + 1.1125 * Exponential (22943)	13335.6	12
Fe_2O_3	Yes	0 + 0.2798 * Exponential (40081)	27111.8	12
K_20	No	0 + 0.29484*Exponential (13902)	7586.2	12
MgO	Yes	0 + 0.45146*Exponential (33632)	20347.5	12
MnO	Yes	0 + 0.53693 x Exponential (20104)	67209.1	12
Na_2O	Yes	0 + 0.35086 x Exponential (41683)	73128.7	12
P_2O_5	Yes	0 + 0.17709 x Exponential (1006.9)	547.5	12
SiO_2	No	0 + 221.83 x Exponential (37015)	43928.4	12
TiO_2	Yes	0 + 0.23491 x Exponential (80596)	41437.2	12
As	Yes	0 + 0.25149 x Exponential (7238.3)	52150.7	12
Ba	Yes	0 + 0.19485 x Exponential (56490)	33462.6	12
Br	Yes	0 + 0.35143 x Exponential (14864)	10306.8	12
Cd	Yes	0.35297 + 0.1239 x Exponential (3766.2)	318.4	12
Ce	Yes	0 + 0.19188 x Exponential (60736)	69577.4	12
Co	Yes	0 + 0.44394 x Exponential (37813)	57500.1	12
Cr	Yes	0 + 0.12792 x Exponential (1295.5)	520.7	12
Cs	Yes	0 + 0.30326 x Exponential (13841)	45876.1	12
Cu	Yes	0.45788 + 0.28939 x Exponential (555650)	46303.9	12
Ga	No	0 + 17.547 x Exponential (17342)	59756.7	12
Hf	Yes	0 + 0.15762 x Exponential (22373)	40140.3	12
I	Yes	0 + 0.61868 x Exponential (36496)	43236.0	12
La	Yes	0 + 0.2318 x Exponential (26221)	48031.5	12
Mo	Yes	0 + 0.62051 x Exponential (10274)	11288.4	12
Nb	Yes	0 + 0.098904 x Exponential (20794)	12226.5	12
Nd	No	0 + 97.792 x Exponential (18204)	13951.6	12
Ni	Yes	0 + 0.3857 x Exponential (22753)	21856.6	12
Pb	Yes	0.43168 + 0.30376 x Exponential (45519)	6653.2	12
Rb	No	0 + 780.67 x Exponential (52433)	59926.0	12
Sb	Yes	0.50416 + 0.13198 x Exponential (28068)	66.718 x 100	12
Sc	No	0 + 17.22 x Exponential (12436)	9.9948 x 1000	12
Se	Yes	0 + 0.4383 x Exponential (9969)	47186.8	12
Sn	Yes	0.44902 + 0.14151 x Exponential (7889.8)	1043.4	12
Sr	Yes	0 + 0.11475 x Exponential (849.34)	187.1	12
Th	Yes	0 + 0.19276 x Exponential (30475)	12226.5	12
U	Yes	0.18476 + 0.13767 x Exponential (2216.5)	4.7742 x 200	36
V	Yes	0 + 0.21843 x Exponential (23918)	26603.6	12
W	Yes	0.41607 + 0.058955 x Exponential (61573)	5131.1	12
Y	Yes	0 + 0.13464 x Exponential (17886)	5000.2	12
Zn	Yes	0.33169 + 0.12069 x Exponential (10129)	1288.8	12
Zr	Yes	0 + 0.28559 x Exponential (52771)	61548.0	12

Glossary

Jargon	Explanation
ArcGIS	Geographic information system (GIS) software for working with maps and geographic information maintained by the Environmental Systems Research Institute (ESRI).
Attribute	Named property of an entity. Descriptive information about features or elements of a database. For a database feature like census tract, attributes might include many demographic facts including total population, average income, and age. In statistical parlance, an attribute is a variable, whereas the database feature represents an observation of the variable.
Compositional data	When the sum of all parts in a sample sum up to a constant (frequently 1, 100% or 1 000 000 mg/kg), the dataset is called compositional (or closed) data. For closed data the concentration obtained for one part (element) does not vary independently from the others, thus the information is not absolute but only relative. See Ferreira et al. (2018) and references therein for further information on compositional data.
ESRI	Environmental Systems Research Institute (ESRI) is an international supplier of geographic information system (GIS) software, web GIS and geodatabase management applications.
Experimental variogram	A plot showing how (one half the) squared differences between the sampled values - (semi)variance - changes with the distance between the point-pairs. Typically, smaller (semi)variances at shorter distances and then a stable (semi)variance (equal to global variance) at longer distances are expected. The goal is to define a variogram model, a key tool for interpolation by kriging.
Field survey	A strategic campaign to capture observations and measurements of geological features in situ and / or to collect samples for later analysis in a lab.
Geographical Information System (GIS)	Geographic Information Systems (GIS) provides accurate information, assistance, support, and maintains and creates information to aid in the development of maps and data analysis.
Geology	The study or science of the earth, its history, and its life as recorded in the rocks; includes the study of geologic features of an area, such as the geometry of rock formations, weathering and erosion, and sedimentation.
Geospatial data	Data that has a geographic component to it. This means that the records in a dataset have locational information tied to them such as geographic data in the form of coordinates, address, city, or postcode.
Geostatistical prediction	A prediction based on application of quantitative, statistical techniques.
Geostatistics	A subset of statistics specialized in analysis and interpretation of Geospatial (geographically referenced) data (Goovaerts, 1997).
GeoTIFF	A georeferenced tiff (Tag Image File) raster image with a given cell size, 500 x 500 metres for the present case. See Raster for further explanation.
Interpolation	A mathematical technique to estimate a trend in data between known points. The process by which software *invents* new data to fill

	gaps in an image or grid. Inverse distance and kriging are common methods used for interpolation.
Kriging	A geostatistical interpolation technique that considers both the distance and the degree of variation between known data points when estimating values in unknown areas
LYR	ArcGIS Layer Files (. lyr) are a means of preserving a defined symbology for a geospatial data set (usually a raster or a vector).
MA	Millions of years, a common metric used in geological time
Nugget	A parameter of the variogram model. Corresponds to the point of interception of the variogram with the y-axis of the plot, that is, the semivariance at distance zero of the variogram model. It represents a mix of small-scale variability and measurement error of the data.
Ordinary kriging	A type of kriging, which requires certain assumptions, such as, stationarity (mean and autocorrelation structure are similar over the whole area of interest) and normality of the data distribution.
PNG	A raster image of Portable Network Graphics format. See Raster for further explanation.
Range	A parameter of the variogram model. Corresponds to the distance at which the variogram model first flattens out (i.e. reaches the Sill).
Raster	Raster data can be thought of as being similar to a digital photograph. The entire area of the map is subdivided into a grid of tiny cells, or pixels. A value is stored in each of these cells to represent the nature of whatever is present at the corresponding location on the ground. See GeoTIFF and PNG for further explanation.
Resolution	Resolution expresses the size of the smallest object in a spatial data set that can be described. It refers to the amount of detail that can be discerned. It is also known as granularity.
Scale	The relation between the dimensions of features on a map and the geographic objects they represent on the earth, commonly expressed as a fraction or a ratio. A map scale of 1/100,000 or 1:100,000 means that one unit of measure on the map equals 100,000 on the earth.
Sill	A parameter of the variogram model. Corresponds to the semivariance value at which the variogram model first flattens out (i.e. reaches the range)
Source data	Source data is raw data (sometimes-called atomic data) that has not been processed for meaningful use to become Information.
Spatial autocorrelation	The degree to which a variable, e.g. element concentration, correlates with itself through space. As stated in Tobler's first law of geography – "everything is related to everything else, but near things are more related than distant things".
Spatial data	Data describing anything with spatial extent, i.e. size, shape or position. In addition to describing things that are positioned relative to the Earth, spatial data may also describe things using other coordinate systems that are not related to position on the Earth, such as the size, shape and positions of cellular and sub-cellular Spatial Things described using the 2D or 3D Cartesian coordinate system of a specific tissue sample.
Semivariance	One half the squared differences between point values spaced a constant distance apart, that is, half the variance.
Semivariogram	A plot showing the distances versus semivariances of the data point pairs. Same as experimental variogram, from which a variogram model can be derived.
Variogram model	A mathematical formula that defines the curve fitting an experimental variogram.

References

The British Geological Survey holds most of the references listed below and copies may be obtained via the library service subject to copyright legislation (contact libuser@bgs.ac.uk for details). The library catalogue is available at https://envirolib.apps.nerc.ac.uk/olibcgi.

BARSBY A., MCKINLEY J.M., OFTERDINGER U., YOUNG M., CAVE M.R., WRAGG J. (2012). Bioaccessibility of trace elements in soils in Northern Ireland. *Science of The Total Environment*, Volume 433, 2012, Pages 398-417, ISSN 0048-9697, https://doi.org/10.1016/j.scitotenv.2012.05.099.

BRITISH GEOLOGICAL SURVEY (2017) User guide for the British Geological Survey Stream Sediment Geochemistry (500 m grid) dataset. *British Geological Survey Open Report OR/17/004*. Nottingham, UK: British Geological Survey.

CAVE M., WRAGG J., LISTER R. (2018). The effect of lead in soil on crime deprivation in Derby, Leicester and Nottingham, Applied Geochemistry, Volume 88, Part B, Pages 198-212, ISSN 0883-2927, https://doi.org/10.1016/j.apgeochem.2017.06.006. (https://www.sciencedirect.com/science/article/pii/S0883292716305881).

DEMPSTER, M., COOPER, M.R., DUNLOP, P. AND SCHEIB, A.J. (2016) Using soil geochemistry to investigate gold and base metal distribution and dispersal in the glaciated north of Ireland. In M.E. Young (ed.), Unearthed: Impacts of the Tellus surveys of the North of Ireland. Dublin. Royal Irish Academy. DOI:10.3318/ 978-1-908996-88-6.ch7.

EVERETT P A, LISTER T R, FERREIRA A M P J, AND FORDYCE F M. (2023). Rural and Urban Topsoil geochemical atlas for England, Wales, Northern Ireland and part of Scotland. British Geological Survey Open Report, OR/22/037.

EVERETT, P A, LISTER, T R, FORDYCE, F M, FERREIRA, A M P J, DONALD, A W, GOWING, C J B, AND LAWLEY, R S. (2019). Stream sediment geochemical atlas of the United Kingdom. British Geological Survey Open Report OR/18/048. Nottingham, UK: British Geological Survey.

FERREIRA A., DARAKTCHIEVA Z., BEAMISH D., KIRKWOOD C., LISTER R., CAVE M., WRAGG J., LEE K. (2018). Indoor radon measurements in south west England explained by topsoil and stream sediment geochemistry, airborne gamma-ray spectroscopy and geology. *Journal of Environmental Radioactivity, 181:152-171*. https://doi.org/10.1016/j.jenvrad.2016.05.007.

HENGL, T. (2009) A Practical Guide to Geostatistical Mapping. Vol. 52, University of Amsterdam, Amsterdam.

LARK R.M., ANDER E.L., CAVE M.R., KNIGHTS K.V., GLENNON M.M., SCANLON R.P. (2014). Mapping trace element deficiency by cokriging from regional geochemical soil data: A case study on cobalt for grazing sheep in Ireland. Geoderma, Volumes 226–227, 2014, Pages 64-78, ISSN 0016-7061, https://doi.org/10.1016/j.geoderma.2014.03.002.

PALMER S., MCILWAINE R., OFTERDINGER U., COX S.F., MCKINLEY J.M., DOHERTY R., WRAGG J., CAVE M. (2015). The effects of lead sources on oral bioaccessibility in soil and implications for contaminated land risk management, Environmental Pollution, Volume 198, Pages 161-171, ISSN 0269-7491, https://doi.org/10.1016/j.envpol.2015.01.004.

PALMER, S., OFTERDINGER, U. AND MCKINLEY, J.M. (2016) Refining the human health risk assessment process in Northern Ireland through the use of oral bioaccessibility data. in M.E. Young (ed.), Unearthed: impacts of the Tellus surveys of the north of Ireland. Dublin. Royal Irish Academy. DOI:10.3318/978-1-908996-88-6.ch25.

RAWLINS, B G, MCGRATH, S P, SCHEIB, A J, BREWARD, N, CAVE, M, LISTER, T R, INGHAM, M, GOWING, C AND CARTER, S. (2012). The advanced soil geochemical atlas of England and Wales. British Geological Survey, Keyworth. www.bgs.ac.uk/gbase/advsoilatlasEW.html

TOBLER W R (1970) A Computer Movie Simulating Urban Growth in the Detroit Region. Economic Geography. Supplement Proceeding International Geographical Union Commissions on Quantitative Methods 46, 234-240

WRAGG, J.; CAVE, M.; HAMILTON, E.; LISTER, T.R. The Link between Soil Geochemistry in South-West England and Human Exposure to Soil Arsenic. Minerals 2018, 8, 570. https://doi.org/10.3390/min8120570

YOUNG, M E AND DONALD A M (eds) (2013). A guide to the Tellus data. Geological Survey of Northern Ireland, Belfast. 233 pp.