

Historical development and its effect on soil contamination in urban areas

Rebekka McIlwaine¹, Siobhan Cox¹, Rory Doherty¹ and Mark Cave²

Corresponding author: rmcilwaine05@qub.ac.uk

¹ Civil Engineering Research Centre
School of Planning, Architecture and Civil Engineering
Queen's University Belfast
Belfast BT7 1NN
Northern Ireland, UK

² British Geological Survey
Keyworth
Nottingham NG12 5GG
England, UK

Research published in; McIlwaine, R., Doherty, R. and Cox, S.F., 2016. The relationship between historical development and potentially toxic element concentrations in urban soils. *Environmental Pollution*, Available online, doi 10.1016/j.envpol.2016.11.040.

Contents

Executive Summary

Acknowledgements

1	Introduction.....	1
1.1	Research hypothesis and objectives	1
2	Methodology	3
2.1	Research approach.....	4
	Stage 1 – Defining the urban boundaries	4
	Stage 2 – Identification of geogenic and historical urban controls	4
	Stage 3 – Source identification via depth comparison and PTE groups	5
	Stage 4 – Consider element enrichment across city development zones.....	6
	Stage 5 – Identification of urban subdomains.....	6
	Stage 6 – Calculate background concentrations within urban subdomains.....	7
3	Results and discussion	9
3.1	Study areas.....	9
3.1.1	Bedrock and superficial geology	9
3.1.2	City development	12
3.2	Statistical distribution.....	14
3.3	Spatial distribution	19
3.3.1	Calculation of Moran’s I	23
3.4	Grouping PTEs of similar sources	24
3.4.1	Depth comparison	24
3.4.2	Cluster analysis	27
3.4.3	Principal component analysis	29

3.4.4	Pollution indices	36
3.5	Urban subdomains.....	39
3.6	Background concentration calculations.....	39
4	Conclusions	44

References

Appendices

Appendix A	Boxplots and histograms
Appendix B	PTE concentration maps
Appendix C	Sample R scripts
Appendix D	PCA with all available elements

Executive Summary

This project creates a new methodology for identifying local sources of contamination and relationships between historical urban development and groups of potentially toxic elements (PTEs) in urban soils. The Belfast Tellus data is used to create the methodology, which is then tested on Sheffield using G-BASE data. Eleven PTEs are considered within this research; arsenic (As), cobalt (Co), chromium (Cr), copper (Cu), molybdenum (Mo), nickel (Ni), lead (Pb), antimony (Sb), tin (Sn), vanadium (V) and zinc (Zn). These elements are expected to be related to different geogenic and anthropogenic sources within the study areas.

The statistical distribution of the PTEs in each of the study areas is used to determine anticipated source controls. Simple interpolated maps are compared with statistical distributions to gain an understanding of how the sources vary spatially. Shallow and deep geochemical data (only available for Belfast) is compared to gain an understanding of geogenic and anthropogenic contributions to PTE soil concentrations. Multivariate statistical analysis in the form of cluster analysis and principal component analysis is used to group PTEs anticipated to be from similar sources.

Potentially toxic elements with a controlling geogenic source in Belfast are identified as Co, V, Cr and Ni; the Tertiary basalts are found to have the largest control over their concentrations. Elements with predominantly anthropogenic sources in Belfast are split into three groups; 1) Sn, Pb and Sb, 2) Cu and Zn and 3) As and Mo. Copper and Zn receive some geogenic contribution to their concentrations from the Tertiary basalts while increasing anthropogenic contributions to both As and Mo see them grouped together. The greater anthropogenic contribution is noted for Pb, Sb and Sn.

All of the PTEs investigated are found to be under some anthropogenic influence in Sheffield. Nickel, Co, Cr and V are shown to occur at elevated concentrations in a large area along the north-east boundary of Sheffield; many of Sheffield's iron and steel works were/are found here. Therefore, elevated PTE concentrations could be related to the industrial use (in these various factories) of the coal. Similarly to Belfast, Pb, Sb and Sn are shown to form a halo around the oldest area of the city. The widespread nature of these PTEs on their total concentration maps suggests an atmospheric deposition anthropogenic source. A geogenic contribution to Zn, Cu, Ni and Cr concentrations is identified in the form of coal outcrops in Sheffield.

The pollution indices (PIs) calculated show that different historical zones are related to different levels of PTE accumulation. The relationship between historical development and differing PTEs is a novel finding from this research, suggesting PTEs have the potential for use as 'urbanisation tracers', and allowing each of the historical zones to be identified as urban subdomains.

Background values are calculated for the PTEs deemed to have some form of anthropogenic input within each of the city's development zones. These PTEs demonstrate varying historical sources within the development zones which result in varying background values. The background values for Belfast are generally highest in the 1919-1939 development zone, suggesting that contamination was at its greatest in Belfast between 1901 and 1939. The background values calculated for Sheffield vary more widely across the different development zones considered.

The scale of PTE concentration data available within the study areas considered allows for a thorough examination of the effects of historical development on soil PTE concentrations. The combination of statistical and spatial techniques used provided a thorough methodology for examining the geochemical data. The techniques must be used side by side to gain an understanding of the PTE distributions from both statistical and spatial perspectives. Simply considering statistical distributions and spatial patterns alone provides a lot of information which can then be augmented by multivariate techniques. The investigative methodology employed within this research may be useful for application within other urban environments.

Acknowledgements

This research has been funded by The Society of Brownfield Risk Assessment's Annual Scholarship. The Tellus data was provided by Geological Survey of Northern Ireland and the G-BASE data by British Geological Survey.

1 Introduction

Reuse of brownfield sites is a key aspect of government policy to promote sustainable land use; the UK government has pledged to ensure that 90% of brownfield sites have planning permission for housing by 2020 (Conservative Party 2015). In 2005 it was estimated that up to 300,000 hectares of UK brownfield land could be affected by industrial contamination, with an even greater area affected by diffuse anthropogenic contamination. Between £140 to £280 million per annum is estimated to have been spent on unnecessary remediation in England and Wales due to an inability to identify 'contaminated' land (Defra & WAG 2011).

Many potential redevelopment sites in urban areas are on brownfield or potentially contaminated land, with little or no, spatial assessment of pollution by potentially toxic elements (PTEs) systematically performed. The proportion of the population living in urban areas is expected to continue to rise, resulting in an increased pressure to develop and reuse urban soils. An improved understanding of PTE sources, concentrations and distribution in urban soils will help to reduce this cost to the economy and facilitate sustainable development of lasting benefit to the UK.

In particular, 'background concentrations' of elements are useful in determining the typical concentration of a PTE within a defined area. Previous research has utilised a variety of methods to calculate background concentrations of PTEs in both regional (McIlwaine et al. 2014; Ander et al. 2013) and, more recently, urban environments (Rothwell & Cooke 2015). They usually aim to differentiate between concentrations of elements that result from geogenic and diffuse anthropogenic sources and concentrations generated by point sources, fitting with current statutory guidance.

1.1 Research hypothesis and objectives

This research develops robust tools to identify sources of soil contamination and relationships between historical development and PTEs, using the wealth of environmental data already collected in urban geochemical surveys. PTEs are considered spatially and compared with historical development zones to develop a greater understanding of the resulting environmental change. From this, urban subdomains are developed, within which background concentrations of PTEs in soils are calculated, helping to inform on risks posed in urban environments. Elements of emerging environmental concern (e.g. antimony (Sb) and molybdenum (Mo)), which have recently been added to recognised "urban elements" (e.g. lead (Pb), copper (Cu) and zinc (Zn)), are included within this study alongside arsenic (As), cobalt (Co), chromium (Cr), nickel (Ni), tin (Sn) and vanadium (V).

The outcomes of this research will enable improved risk assessment, and provide key scientific knowledge of urban processes and tools that will inform transparent decision making in wider planning and public health issues. Where environmental forensic techniques have previously been applied to datasets to apportion sources of contamination in urban soils, studies were generally undertaken with fewer sample locations than are available in this research.

2 Methodology

Geochemical data from Belfast (Tellus survey) and Sheffield (Geochemical Baseline Survey of the Environment (G-BASE)) were used to complete this study.

The Tellus survey included geochemical surveying of all of Northern Ireland following the G-BASE methodology (Johnson 2005). Four soil samples were recovered every 1km² in urban areas at depths of 5 to 20cm (shallow) and 35 to 50cm (deep) and analysed for major, minor and trace elements by x-ray fluorescence (XRF) (shallow samples only) and inductively coupled plasma (ICP) following an aqua regia extraction (shallow and deep samples). Similar sampling density and analytical techniques were employed during G-BASE sampling of 26 other urban areas in the UK. After discussions with Mark Cave (BGS) and consultation with the SoBRA committee, Sheffield was selected as a suitable G-BASE sampled city for comparison with Belfast. It was felt to be an appropriate location due to the;

- existence of G-BASE data,
- similar (though greater and slightly more varied) industrial history to Belfast,
- additional inputs from its coal mining legacy,
- similar type of bedrock geology to Belfast.

This project aimed to create a new methodology for identifying local diffuse sources of contamination and relationships between historical urban development and groups of PTEs, and to calculate background concentrations of these PTEs in urban soils. The Belfast Tellus data was used to create the methodology, which was then tested on Sheffield using G-BASE data. This methodology allowed for an assessment of land quality by completing five objectives:

- (1) Identifying groups of elements thought to be controlled by similar sources of contamination.
- (2) Considering element enrichment across city development zones to gain an understanding of how historical industries have influenced soil quality.
- (3) Identifying the geogenic and anthropogenic contributions to PTE concentrations in soil.
- (4) Developing urban subdomains – using the findings from (1), (2) and (3) to define where particular sources are expected to govern PTE concentrations.
- (5) Calculating background concentrations for grouped anthropogenic PTEs in urban subdomains.

Eleven PTEs were considered within this research; As, Co, Cr, Cu, Mo, Ni, Pb, Sb, Sn, V and Zn. These elements were selected as they were expected to be related to numerous different geogenic and

anthropogenic sources within the study areas; in particular they are likely to represent contamination from a variety of industrial processes. PTEs such as As, Cr, Cu, Ni, Pb and Zn commonly feature in urban geochemical studies (Johnson & Ander 2008) while Carrero et al. (2013) demonstrate a relationship between a variety of the chosen PTEs, including Mo, Sb and Sn, and soils heavily impacted by traffic. Previous research has demonstrated concentrations of various PTEs, including Cr, Cu, Ni and V, which are controlled by geogenic sources in the form of bedrock geology (McIlwaine et al. 2014).

2.1 Research approach

In order to address our specific objectives a novel investigative methodology was employed utilising depth ratios, a range of multivariate statistical techniques and pollution indices (PIs). The project was addressed in six stages to allow clear definition of deliverables.

Stage 1 – Defining the urban boundaries

Stage 1 consisted of selecting a second UK city for consideration, collating all urban data for both cities to ensure non-detects and duplicates were handled appropriately, and preparation of initial concentration maps for all PTEs of interest. This provided information on spatial trends in contaminant concentrations, allowing for identification of controlling factors over element concentrations. The Moran's I statistic was calculated to allow interpretation of spatial autocorrelation in both study areas. Moran's I is widely used for this purpose (Fu et al. 2014; Helbich et al. 2012; Liu et al. 2012; Goyal et al. 2015) and identifies if there are "relationships between location and attribute values" (Helbich et al. 2012). The results of Moran's I indicate perfect positive spatial autocorrelation at value "1" (i.e. high values or low values cluster together), while "-1" suggests perfect negative spatial autocorrelation (a checkerboard pattern). A value of "0" implies perfect spatial randomness (Fu et al. 2014).

Data was plotted in histograms, box plots and cumulative distribution plots so that careful consideration could be given to what methods were appropriate to normalise the data. Data was interrogated alongside land use data from the Corine land cover dataset to determine urban boundaries. The deliverables of stage 1 are normalised datasets for all elements of interest and defined urban study boundaries.

Stage 2 – Identification of geogenic and historical urban controls

In order to differentiate between geogenic and anthropogenic contributions it was vital to have a clear understanding of bedrock and superficial geology, in addition to any other obvious geogenic controls. In order to assess likely anthropogenic contributions to soil PTE concentrations, a summary of the

historical development in each city was completed. This involved using historical maps to define the urban boundary at distinct times in the cities' development; these were used to produce city development zones which spatially demonstrate how the city has developed and changed over time. The deliverables of stage 2 are summaries of geogenic controls in Belfast and Sheffield and maps of city development zones.

Stage 3 – Source identification via depth comparison and PTE groups

Geogenic contributions are believed to result in elevated soil PTE concentrations at depth, whereas elevated shallow soil PTE concentrations are often related to anthropogenic contributions (Chiprés et al. 2009a; Galán et al. 2008). As Belfast data exists at both depths, a boxplot comparison of the deep (35-50cm) and shallow (5-20cm) concentrations was undertaken to define the main controls over the PTE concentrations.

Geochemical data is compositional in nature, i.e. all the elements analysed in a sample sum to a constant value and therefore are reported in mg/kg or parts per million. Therefore, prior to completion of multivariate statistical analysis (cluster analysis and principal component analysis (PCA)), the data was centred logratio (clr) transformed in order to 'open' it. The clr transformation uses the geometric mean to divide all elements measured before completing a log-transformation (Aitchison 1982).

Cluster analysis and PCA were used in both study areas to split the datasets into groups of PTEs which are similar in their characteristics or behaviour. Cluster analysis aims to split the data under consideration into a number of similar groups. Ward's minimum variance method was used within this cluster analysis - it is a commonly used procedure for determining hierarchical clusters (Ward 1963). This method forms "hierarchical groups of mutually exclusive subsets on the basis of their similarity with respect to specified characteristics" (Ward 1963). It is therefore an agglomerative technique whereby initially each sample is considered a cluster and this is reduced until all observations are included in one cluster.

PCA is a popular multivariate statistical technique used to eliminate 'noise' from multivariate data and present variability in a reduced number of components (Candeias et al. 2011; Yunker et al. 2012; Reimann et al. 2007). The PCA method plots geochemical data in multivariate space, searching for the direction that contains maximum variability. This axis becomes principal component (PC) one, which explains the most variance, with each component after that orthogonal to the previous component and explaining the maximum amount of remaining variability. By reducing the data down to a few components which contain the maximum variance of all the variables, it should become easier to determine the main factors affecting the data (Reimann et al. 2008).

Within PCA, the biplot is a common image used to visualise the scores and loadings of the PCs simultaneously. The loadings describe the relationship between the original variables and the PCs, while the scores describe the relationship between each of the observations and the PCs (Reimann et al. 2008). Biplots are particularly useful in helping to understand the controlling variables over each PC; the distance between the variables is a demonstration of their correlation while their distance in the direction of the PCs demonstrates their effect on that PC. In order to understand the spatial distribution of the PCs the score for each observation can be plotted spatially, forming a map of how the PC is distributed. This allows a comparison between the identified controlling variables (from the biplot) and geographical distribution, which should allow a full interpretation of the PCA (Reimann et al. 2008).

These methods aided in the identification of the main controls over the PTEs being investigated. The deliverable of stage 3 is R-scripts of the developed chemometric methods.

Stage 4 – Consider element enrichment across city development zones

In order to gain an understanding of how the soil PTE concentrations within the city's development zones show various levels of enrichment, PIs were calculated using;

$$PI = U_c/R_c$$

where U_c is the median element concentration within the development zone under consideration and R_c is the median rural element concentration. Previous research (Biasioli et al. 2006) has used PIs in a similar manner in order to estimate the enrichment of a city with certain elements. The median rural element concentrations were easy to calculate for the Belfast study area, as the remaining Tellus rural data was available. The median rural element concentrations for the Sheffield study area were taken from Freestone et al. (2004), which provided the 'median concentrations in regional surface soil samples overlying Carboniferous Coal Measures, Humber-Trent atlas areas'. The deliverable of stage 4 is a summary of the PIs calculated within each of the city's development zones.

Stage 5 – Identification of urban subdomains

The results from stages 1-4 were collated at this stage and used to produce the deliverable of stage 5; urban subdomains. A domain is an area where a readily identifiable factor can be shown to control the concentration of a PTE, therefore the various source apportionment techniques applied allowed for subdomain identification.

Stage 6 – Calculate background concentrations within urban subdomains

A variety of methods for calculating background concentrations were utilised, and the results contrasted and compared to gain an understanding of their strengths. As previously suggested for use at a regional scale in Northern Ireland, the typical threshold value (TTV) (McIlwaine et al. 2014) methodology which utilises the Finnish upper limit of geochemical baseline variation (ULBL) (Jarva et al. 2010) to define background concentrations was employed. The ULBLs are based on the upper limit of the upper whisker line of the box and whisker plots, which can be calculated using:

$$\text{ULBL} = P_{75} + 1.5 \times (P_{75} - P_{25})$$

P_{75} and P_{25} are the 75th and 25th percentiles of the element concentrations respectively (Jarva et al. 2010). Logarithmic transformed data were not used to plot the box and whisker plots, as the untransformed data led to the highest amount of outliers and therefore gives a more conservative value. This is a very straightforward method for calculating TTVs, solely requiring the calculation of the 75th and 25th percentiles.

In addition, the Normal Background Concentration (NBC) method employed for use alongside Part 2A of the Environmental Protection Act in England and Wales was used. Within the NBC methodology, it is recommended that the domains (areas where a readily identifiable factor can be shown to control the concentration of an element) are based on at least 30 values. The NBC is calculated for each domain using a statistical methodology that (1) assesses the skewness of the geochemical data by observing a histogram and calculating the skewness and octile skewness of the distribution. Based on the results of that assessment, the method (2) performs either a log transformation or a box-cox transformation on the data if necessary and then (3) computes percentiles using either parametric, robust or empirical methods depending on the results of the transformation applied.

The NBC is then taken to be the upper 95% confidence limit of the 95th percentile. A detailed explanation of how the methodology was constructed and how it should be applied is given in Cave et al. (2012). The project outputs included R code scripts which can be used to determine NBCs, as discussed by Johnson et al. (2012). These R scripts have been utilised by the author to calculate NBCs within this research.

The median + 2MAD (median absolute deviation) (Reimann et al. 2005) method was also utilised. The median + 2MAD, boxplot upper whisker and English NBC methods were compared as methods by which to calculate background concentrations in urban environments in a study by Rothwell & Cooke (2015). The lack of systematically collected geochemical data (no G-BASE data was available) meant that a different approach had to be taken within this study in Gateshead; site investigation data collected during the planning process was instead used. The local authority determined that the

median + 2MAD method provided the preferred NBCs within this study as it consistently gave the most conservative values i.e. the lowest NBC (Rothwell & Cooke 2015). The deliverables of stage 6 are a summary of the background concentrations calculated.

3 Results and discussion

3.1 Study areas

The section aims to provide background to both the Belfast and Sheffield study areas and summarise the main anticipated controls over PTE concentrations in soils. The Corine land cover data (European Environment Agency 2012) was utilised to define both the Belfast and Sheffield study area boundaries. This data is split into 44 different land uses based on the interpretation of satellite images. The majority of land uses were simplified by the author into either the urban or rural category; areas of urban fabric and industrial units were obviously urban while areas of pasture and forest were clearly rural. Where any inconclusive land uses were identified, the surrounding land use was utilised as an indicator of land use type on a site by site basis. If the surrounding land use was majority urban then the area was also classed as urban, and vice versa where the surrounding land uses were mainly rural the area was classed as rural. This allowed identification of study area boundaries for both Belfast (Figure 3-1) and Sheffield (Figure 3-2). Within Sheffield, this had to be correlated with the spatial extent of the G-BASE data to ensure a suitable boundary was defined. These boundaries incorporated 782 samples within Belfast from the Tellus survey and 495 in Sheffield from its G-BASE survey.

3.1.1 Bedrock and superficial geology

Belfast

As shown in Figure 3-1, the bedrock geology in the Belfast urban area is diverse. Silurian greywacke and Silurian shale are the two oldest rock formations, forming part of the Southern Uplands-Down-Longford Terrane, followed by the Permo-triassic sandstones and mudstones (known locally as Sherwood Sandstone and Mercia Mudstone respectively). This is covered in the west of the city by Cretaceous sandstone and chalk and finally by the most recent Tertiary basalts which run along the north-west boundary of the city (Mitchell 2004). A number of Palaeogene intrusions occur within the study area.

Geology has been previously identified as a control over element concentrations in Northern Irish soil, with geogenic controls from underlying basalt or sandstone defining domains in a regional study by McIlwaine et al. (2014). From this perspective, this study area is of great interest due to the expected geogenic controls within an urban environment.

Superficial geology is found in the form of till, glacial sands and gravels, and alluvium within the vicinity of the River Lagan (Bazley et al. 1984).

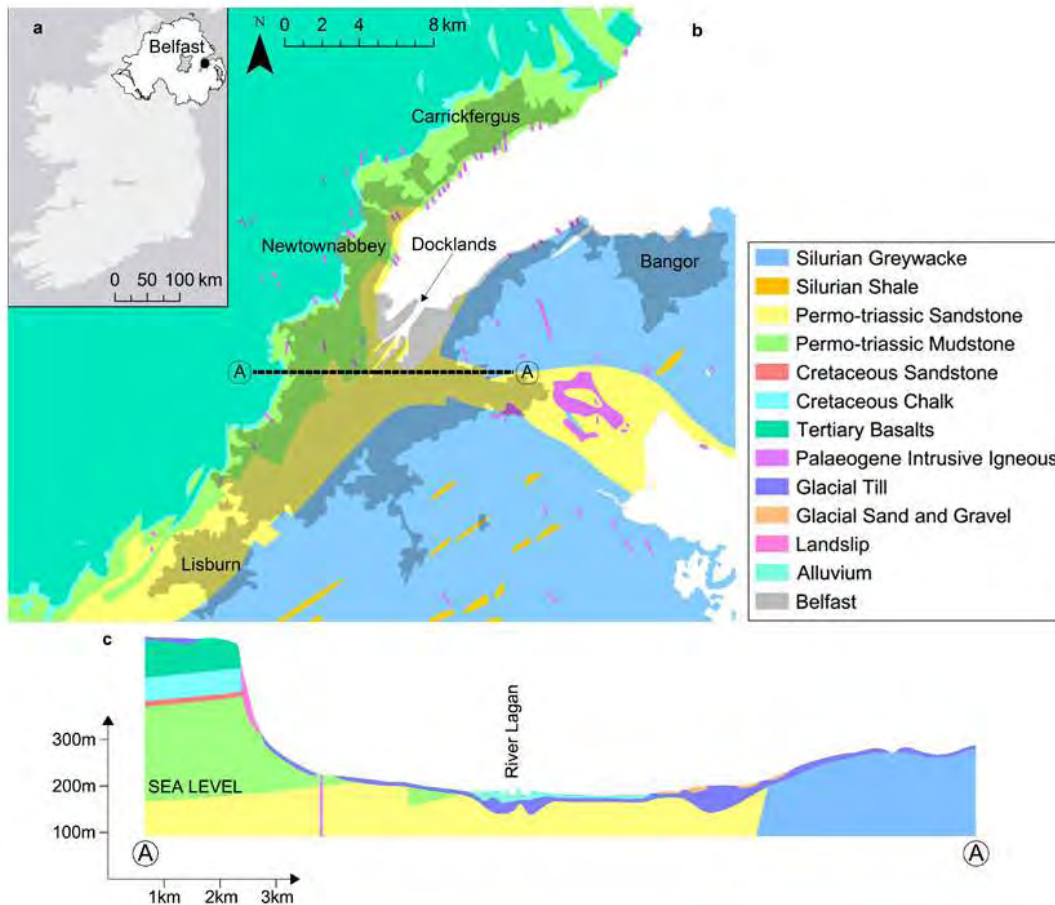


Figure 3-1 Maps showing a) location of Belfast within Northern Ireland, b) simplified bedrock geology and c) topographical cross section showing bedrock and superficial geology within study area (Bedrock geology derived from data provided by GSNI (Crown Copyright))

Sheffield

Sheffield is underlain by Carboniferous deposits of Westphalian and Namurian age (Freestone et al. 2004). As shown on Figure 3-2, the rocks are highly faulted and folded creating many discontinuous outcrops. The Middle Coal Measures Formation outcrops to the east of the city centre; it is Westphalian in age and composed of sandstone. The Lower Coal Measures Formation, also composed of sandstone and Westphalian in age, underlies most of the city centre. The Rossendale and Marsden Formations are present to the west of the study area comprising of mudstone and siltstones.

As shown on Figure 3-2, quaternary deposits cover approximately 10% of Sheffield. Silt alluvium is located in the river valleys around and in the city of Sheffield. Some river terrace deposits in the form of sand and gravel also infill these valleys (Freestone et al. 2004).



Figure 3-2 Map showing a) location of Sheffield within the United Kingdom, b) bedrock geology (taken from BGS GeoIndex) and c) superficial geology (taken from BGS GeoIndex) within Sheffield

3.1.2 City development

Belfast

Belfast is located on the eastern coast of Northern Ireland and is its largest city. The city's origin dates back to 1613 when it was given 'town' status; this was later updated to 'city' in 1888. The city is historically recognised for both linen production and ship-building; the early 18th century saw the introduction of the linen industry involving bleaching, weaving and spinning processes while ship-building was introduced later in the 18th century (Beckett & Glasscock 1967; Crawford 1986). It was this ship-building industry that created one of Belfast's most famous exports; the Titanic. The city was an important manufacturing centre during the industrial revolution with other common businesses including rope works, bleachers, glass manufacturers, tobacco factories and distilleries (Royle 2007). A contamination legacy remains from historical gasworks sites in Belfast; Belfast was one of the first towns in Ireland to use coal gas for lighting purposes (Niskavaara et al. 1997).

Sectarian and political unrest between the Protestant and Catholic populations in Northern Ireland came to a head in 1968 and the well-known conflict, 'The Troubles', lasted until 1998 when the Good Friday Agreement was signed. 'The Troubles' took over the latter half of the 20th century resulting in the loss of much industry and commerce. Present-day Belfast is a prosperous and lively city, much more reliant on service provision related occupations (82% in Northern Ireland in the 2013 Census of Employment (NISRA 2014)) than the historical industrial and manufacturing employment.

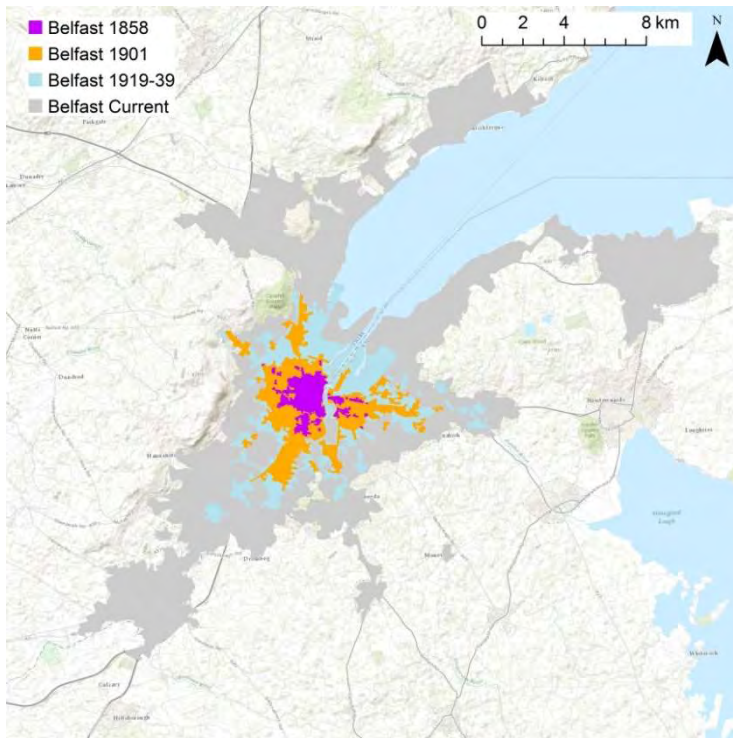


Figure 3-3 Development of Belfast over time (1858, 1901, 1901-1939 to the current boundary)

Historical maps of Belfast have been used to produce historical study area boundaries (Figure 3-3). These historical boundaries were constructed by the author using historical 6 inch county maps (Land and Property Services 1858; Land and Property Services 1901; Land and Property Services 1919); the boundaries were constructed for 1858, 1901 and 1919-1939. The last boundary is inclusive of the period between 1919 and 1939 as detailed mapping for a more specific year was not available. Although towns such as Carrickfergus, Bangor and Newtownabbey (see Figure 3-1) also witnessed development prior to the current boundary, their historical development was not incorporated as they could only be considered part of the larger Belfast area within the current boundary.

Sheffield

Sheffield is located in South Yorkshire in England and is believed to have been originally founded in the 8th century. It was given town charter in 1297 and city status in 1893. Coal has played an important role in the city's history, being mined and burnt for "space heating and industrial purposes in Sheffield since Roman times" (Rawlins et al. 2005). By the 1750s, more than 150 firms were dedicated to steel manufacture within the city. High quality cutlery, an export for which Sheffield is recognised, has been produced in the city since that time (Gilbertson et al. 1997). Industrial expansion continued until the late 1960s, when British Steel opened their Tinsley Park Works in the north east of the city. The steel and cutlery industry in Sheffield began to decline in the late 1970s and 1980s when cheaper alternatives

were being produced in other areas of the world. This also affected the coal used to fuel industry in Sheffield, with coal use declining dramatically from the mid-1980s (Gilbertson et al. 1997).

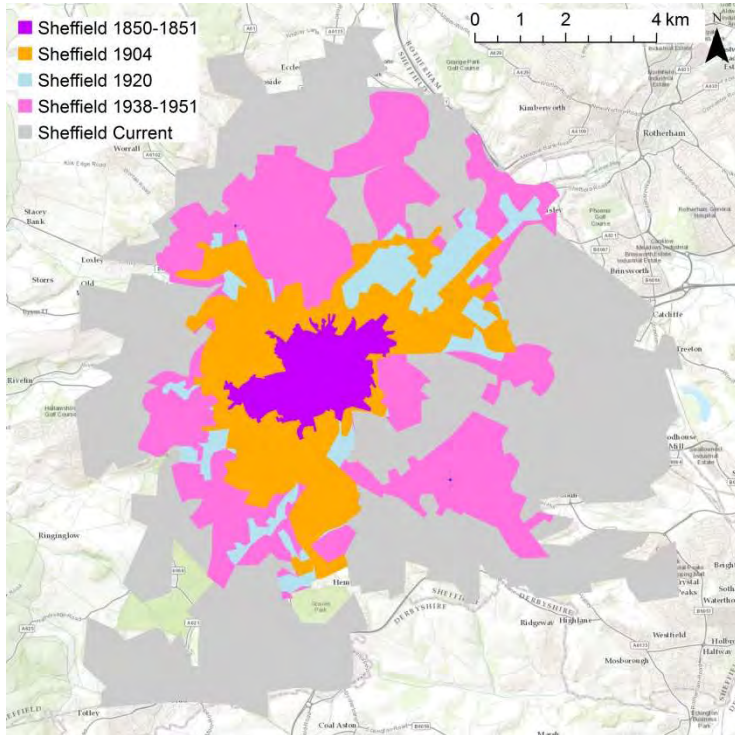


Figure 3-4 Development of Sheffield over time (1850-1851, 1904, 1920, 1938-1951 to the current boundary)

Historical maps of Sheffield have also been used to produce its historical study area boundaries (Figure 3-4). The maps came from a variety of sources and are listed below:

- Yorkshire 294 (includes: Sheffield), surveyed 1850-1851, published 1855 from OS Six-inch England and Wales, 1842-1952.
- Sheet 9 – Sheffield, published 1904 from Bartholomew’s “Half Inch Maps” of England and Wales 1902-1906.
- Sheet 9 – Sheffield, published 1920 from Bartholomew’s Revised “Half Inch Maps”.
- SK38, SK39, SK48 and SK49, published 1951 to 1953 from 1:25,000 maps of Great Britain, 1937-1961

3.2 Statistical distribution

Median concentrations of the PTEs in each city were compared to gain an understanding of their anticipated differences in scale. As the median is robust to extreme values, it was used to give an average concentration of the PTEs in each city.

Table 3-1 Summary of median concentrations of each PTE in Belfast and Sheffield

PTE	Belfast median (mg/kg)	Sheffield median (mg/kg)
As	10.0	22.0
Co	19.8	16.4
Cr	171.7	88.7
Cu	53.6	54.5
Mo	1.3	3.9
Ni	65.1	35.8
Pb	73.4	166.5
Sb	1.4	2.9
Sn	5	12.6
V	110.7	84.8
Zn	120.3	158.8

From Table 3-1 it is immediately obvious that very different median concentrations of many of the PTEs exist in Belfast and Sheffield. Sheffield has higher median concentrations of As, Mo, Pb, Sb, Sn and Zn, while Belfast has higher median concentrations of Cr, Ni and V. The median concentrations of Co and Cu are relatively similar in both cities. Source identification will help to determine the causes of these differences in scale i.e. are the PTEs controlled by geogenic or anthropogenic sources in the two cities.

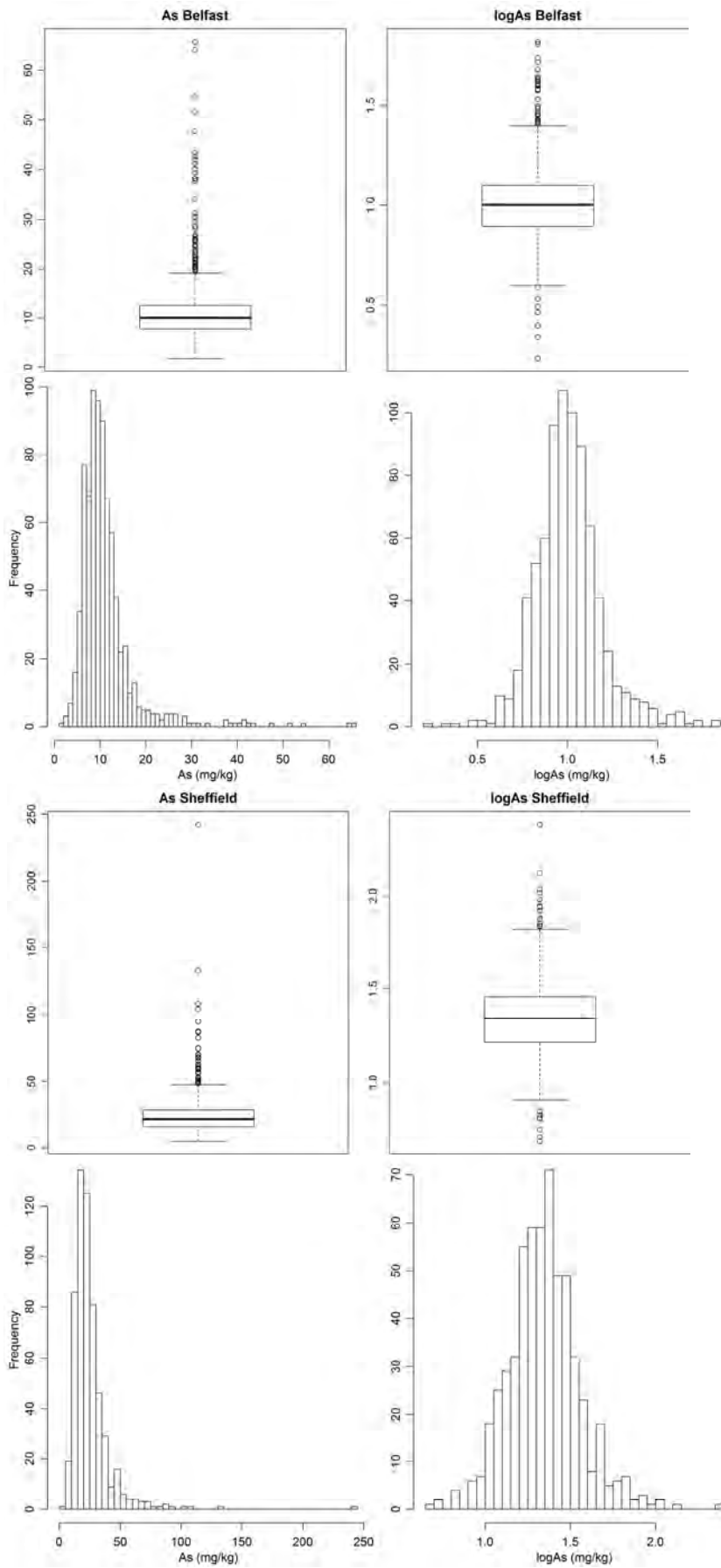


Figure 3-5 Boxplots and histograms used to examine the distribution of the data and the effectiveness of a log transformation for As in Belfast and Sheffield

Figure 3-5 provides an example of the coupled boxplot and histogram approach taken to examine the distribution of the data in Belfast and Sheffield. The remainder of the boxplots and histograms are included in Appendix A.

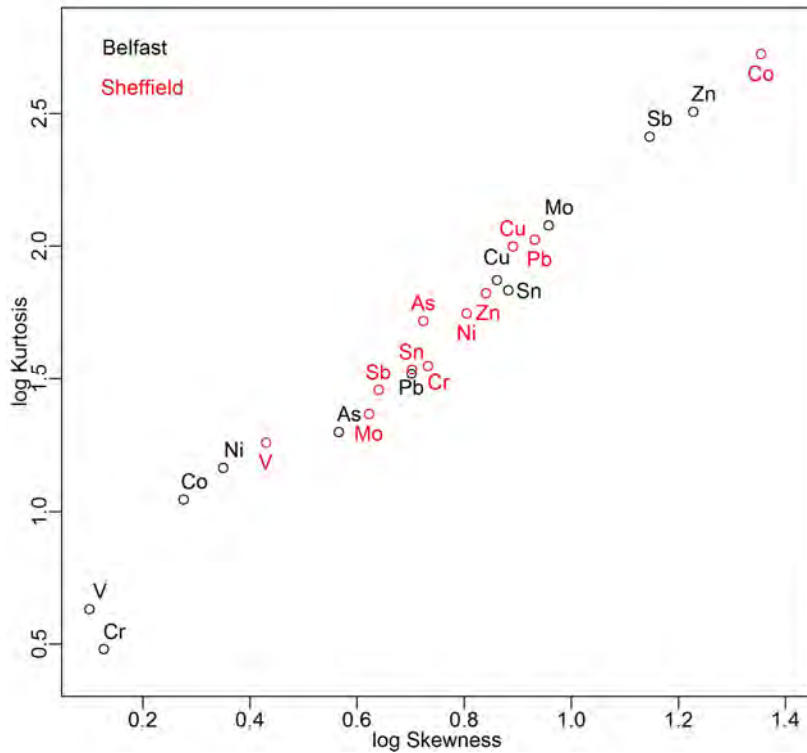


Figure 3-6 Plot of the (log) skewness against the (log) kurtosis for each of the PTEs in Belfast (black) and Sheffield (red)

Figure 3-6 provides a comparison of the skewness and kurtosis of the data in each of the cities. Skewness is a measure of the symmetry of the data while kurtosis indicates whether the data is heavy-tailed relative to the normal distribution. All the skewness and kurtosis values calculated are positive, demonstrating the right-skewed and highly peaked nature of geochemical data. Figure 3-6 provides an indication of which PTEs are more likely to be governed by geogenic sources and therefore have a distribution closer to the normal distribution. These PTEs have smaller skewness and kurtosis values, for example Cr, Co and Ni in Belfast and V in both cities. PTEs with greater skewness and kurtosis values are likely to be governed by anthropogenic sources of both diffuse and point nature.

The concentrations of Cr are much more right-skewed in Sheffield than they are in Belfast. The Belfast distribution of Cr is relatively normal with the presence of some upper outliers. In contrast, in Sheffield the Cr distribution is highly skewed suggesting a greater anthropogenic control over Cr in Sheffield. Ni also demonstrates very different distributions in both areas. The Sheffield distribution of Ni is much

more highly skewed than the Belfast distribution, suggesting a greater anthropogenic contribution to Ni concentrations in Sheffield. The distribution in Belfast is relatively normal with the presence of some upper outliers. As shown in Table 3-1, the median concentrations are higher for Cr and Ni in Belfast than they are in Sheffield.

The distribution of V in both areas is probably the closest to a normal distribution, suggesting controlling geogenic contributions in both areas. The distribution in Belfast is slightly closer to a normal distribution than the distribution in Sheffield but low skewness and kurtosis values are calculated for this PTE in both cities.

Cobalt shows very different statistical distributions in Belfast and Sheffield. In Belfast the data is relatively normally distributed with only a few upper outliers. A log transformation is still required to allow the data to closely approximate the normal distribution. In Sheffield the data is highly skewed; the presence of one extreme outlier (>1000 mg/kg) has a large effect over its distribution and is generating the high skewness value shown in Figure 3-6. This suggests a greater geogenic control over Co concentrations in Belfast compared to greater anthropogenic contributions to Co concentrations in Sheffield.

The statistical distribution of arsenic in both Belfast and Sheffield is quite similar. Prior to log transformation both datasets are right-skewed with the presence of a reasonable amount of outliers. The Belfast data is slightly less skewed than the Sheffield data with a lower median value, suggesting more of an anthropogenic input to the Sheffield As concentrations.

Copper demonstrates quite similar distributions in both urban areas, with similar skewness and kurtosis values calculated. The untransformed distributions demonstrate highly right-skewed data suggesting anthropogenic controls over this element. Molybdenum also demonstrates quite similar, right-skewed distributions in both areas. The data in Sheffield is less skewed than the data in Belfast. These distributions would suggest an anthropogenic control over Mo concentrations.

Lead concentrations in Belfast and Sheffield are both highly skewed suggesting an anthropogenic control over this element. A lot of upper outliers are present in both datasets with the main difference between the two urban areas being a number of extremely elevated concentrations in Sheffield (>1500 mg/kg) that are well above the maximum Belfast concentrations. These elevated concentrations result in the calculation of higher skewness and kurtosis values in Sheffield than in Belfast.

The presence of two extreme outliers creates a highly skewed distribution of Sb in Belfast. The distributions are actually highly skewed in both cities with the added presence of two extreme outliers in Belfast, suggesting an anthropogenic control over Sb in both urban areas. Similarly to Sb, the

distribution of Sn is highly skewed in both urban areas. The presence of a number of extreme outliers in Belfast means that it appears initially to be more skewed than Sheffield. Both these PTEs appear to be governed by anthropogenic controls in both cities.

Anthropogenic controls over Zn concentrations are anticipated based on the distribution of Zn in both areas. The Belfast data appears to be even more skewed due to the presence of two extreme outliers (>6000 mg/kg), resulting in much higher skewness and kurtosis values in Belfast than in Sheffield.

3.3 Spatial distribution

Soil concentration maps were produced for the PTEs in Belfast and Sheffield using the inverse distance weighting (IDW) interpolation technique in the ArcMap (ESRI 2010) package of ArcGIS. The IDW technique was given an output cell size of 200 m, power of 2 and a search radius of 500 m.

Belfast

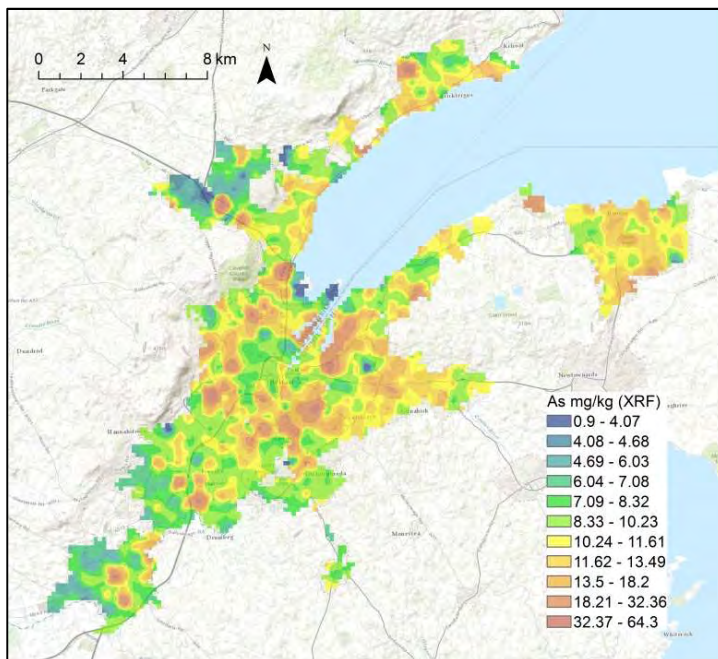


Figure 3-7 Interpolated map showing As concentrations (mg/kg) across Belfast

The soil concentration maps for Belfast, shown in Figure 3-7 and Appendix B, can be broadly split into two groups of PTEs which demonstrate similar patterns. Cobalt, V, Cr and Ni all show very similar distributions that are generally elevated in the north-west of the Belfast study area. This is related to the presence of the Tertiary basalts, which have previously been identified as a control over these PTEs (McIlwaine et al. 2014). Other less widespread elevated concentrations of these PTEs do occur within the Belfast study area, which are likely to be related to anthropogenic sources - however, the

overwhelming control over Co, Cr, Ni and V concentrations seems to be bedrock geology in the form of the Tertiary basalts.

The other group of PTEs consists of Cu, Zn, Sn, Pb, Sb, As and Mo. As discussed in Section 3.1, all of these PTEs would be anticipated to have some anthropogenic contribution. They are likely to be anthropogenically controlled, by a number of sources, within the Belfast study area. All these PTEs demonstrate different distributions, with As and Pb appearing to show greater contamination spread than more recent 'elements of concern' such as Mo and Sb. Generally, the docklands area of Belfast (Figure 3-1) appears to be relatively highly contaminated, with many of the PTEs highlighting elevated concentrations in this area. Many of these anthropogenic PTEs show elevated concentrations in a halo surrounding the oldest part of the city centre.

Sheffield

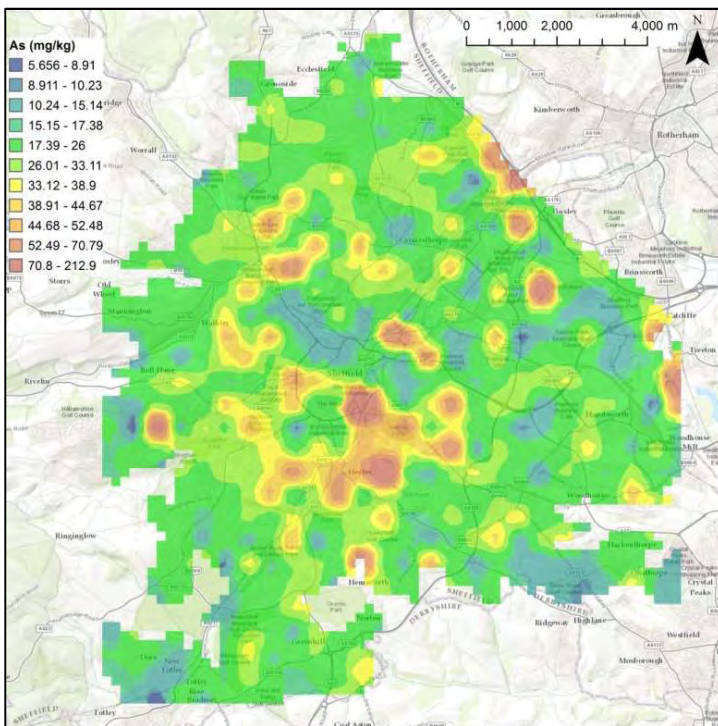


Figure 3-8 Interpolated map showing As concentrations (mg/kg) across Sheffield

The maps for Sheffield (Figure 3-8 and Appendix B) show quite different patterns. One of the most obvious features for Ni, Cr, Co, Mo, Zn and V is that they all demonstrate elevated concentrations in an area to the north-east of the study area. The map for Cr has been included in Figure 3-9 to demonstrate this area of elevated concentrations (outlined approximately in a dashed black line). Some of these PTEs show a more widespread pattern of elevated concentrations in this north-east area than others.

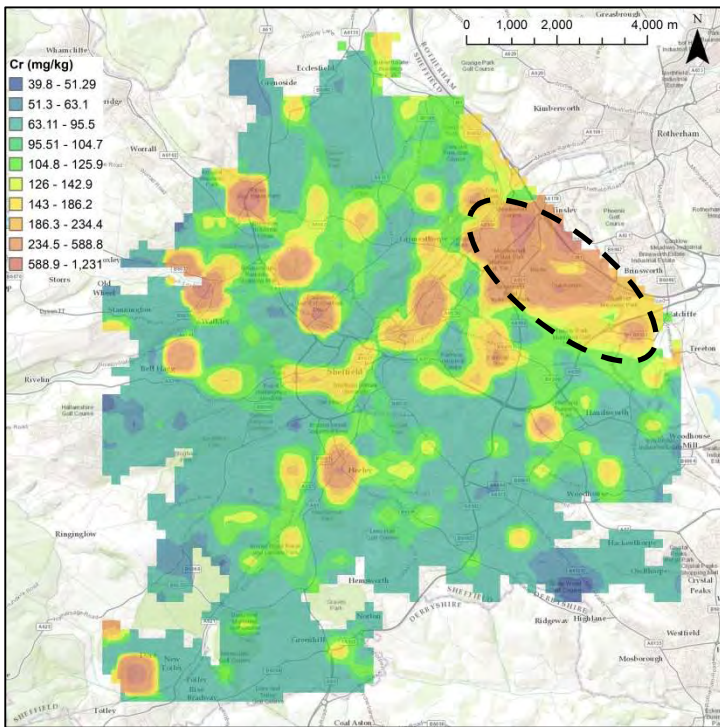


Figure 3-9 Interpolated map showing Cr concentrations (mg/kg) across Sheffield, dashed line shows area of elevated concentration noted in the text

Nickel and Cr show very similar spatial distributions, with their largest area of elevated concentrations located in the north-east of the study area. This area is shown as partially developed in the 1920 Sheffield map (Figure 3-4) with further development within the 1938-1951 development boundary. This area has obviously expanded further since then. As shown on Figure 3-10, a concentrated area of past and current surface coal mining is located in this north-east location in Sheffield, along with many of the large steel factories for which Sheffield is famous (outlined approximately in dashed red line). It is likely that the elevated PTE concentrations identified in this area are linked to the geogenic presence of coal, its mining and/or its industrial use in the steel factories. According to Rawlins et al. (2005), the disposal of slag from locally produced stainless steel may have led to generally elevated concentrations of Cr and Ni in soils in Sheffield.

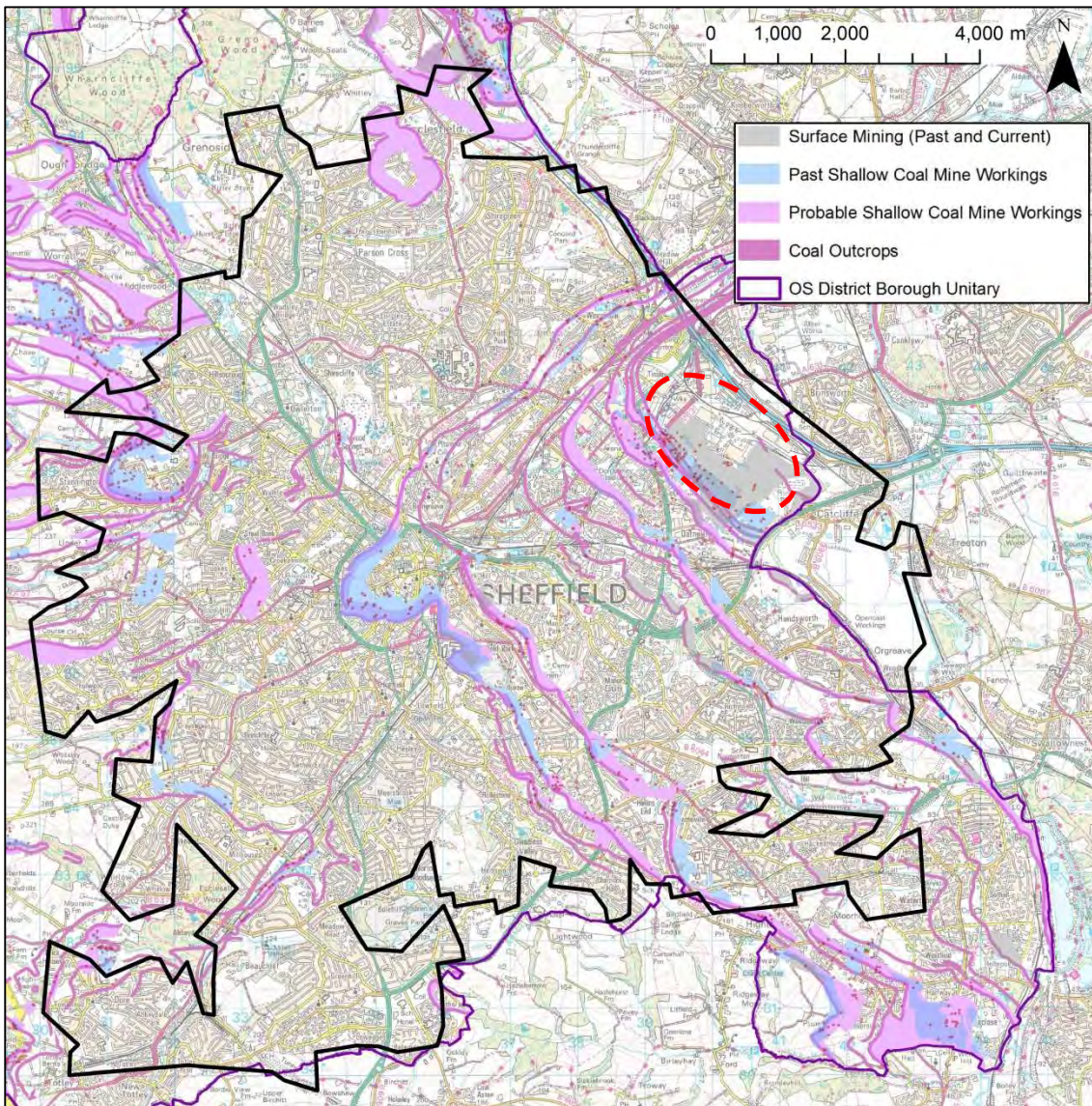


Figure 3-10 Map summarising specific coal mining legacy plan for Sheffield created by the Coal Authority (The Coal Authority 2015), dashed red line encircles the area of large steel factories as well as past and current surface coal mining

The map for Co also demonstrates elevated concentrations in the same location to the north-east of the city, with a small number of other point sources across the study area contributing to elevated Co concentrations. In contrast, although Zn concentrations are shown to be elevated in the north-east area, more widespread elevated concentrations of Zn are obvious across Sheffield's city centre. The concentrations of Mo are elevated in the same north-east area of Sheffield, with other sources generating elevated concentrations of Mo in dispersed locations across the city.

Likewise V concentrations are elevated in a number of locations across the city, with elevated concentrations visible in the same north-east location. It should be noted that concentrations of V are

substantially lower in Sheffield than they are in Belfast. This is likely to be because V is elevated in the Tertiary basalts in Belfast.

The concentration maps for As and Cu are similar for Sheffield. Elevated concentrations of these PTEs are generally related to point sources across the study area, with the largest areas of elevated concentrations surrounding the oldest area of the city (Figure 3-4).

Elevated concentrations of Pb, Sb and Sn are more widespread across the Sheffield study area. All of these PTEs show a halo effect surrounding the oldest area of Sheffield, with their most elevated concentrations surrounding this area. The widespread nature of their elevated concentrations perhaps suggests a controlling atmospheric deposition source of these PTEs. Although the use of Pb in petrol was phased out in the UK in 2000, elevated Pb concentrations resulting from traffic remains a possibility in both Sheffield and Belfast. The high density road network in Sheffield and its associated heavy traffic (Rawlins et al. 2005) may have had an effect over Pb concentrations in soils. Elevated concentrations of Sn are also quite widespread, with less spatial distribution shown for Sb, acknowledging its more recent contribution to anthropogenic concentrations of PTEs in soils (Tian et al. 2012).

3.3.1 Calculation of Moran's I

In order to assess patterns of spatial correlation the Moran's I method was applied to the various PTEs in both Belfast and Sheffield. The calculated statistics for Moran's I provide a measure of global spatial autocorrelation across the entire study area. The results of Moran's I indicate perfect positive spatial autocorrelation at value "1" (i.e. high values or low values cluster together), while "-1" suggests perfect negative spatial autocorrelation (a checkerboard pattern). A value of "0" implies perfect spatial randomness (Fu et al. 2014).

The Moran's I calculations for both Belfast and Sheffield generated significant ($p=0.05$) positive statistics for all the PTEs. This tells us that all the distributions are more spatially clustered than randomly distributed. The Moran's I statistics are higher in Belfast for all the PTEs apart from Mo and Sn, for which they are higher in Sheffield. This suggests more heterogeneous patterns in Sheffield for the majority of the PTEs, while the distributions in Belfast show greater spatial autocorrelation.

Table 3-2 Moran's I statistics calculated for all PTEs in Belfast and Sheffield to indicate spatial autocorrelation

	Observed Moran's I	
	Belfast	Sheffield
As	0.01791	0.01185
Co	0.15913	0.00001
Cr	0.14950	0.02949
Cu	0.04868	0.01863
Mo	0.03826	0.04055
Ni	0.10599	0.04092
Pb	0.02528	0.02191
Sb	0.01645	0.01461
Sn	0.02444	0.02504
V	0.18792	0.04058
Zn	0.02264	0.01071

From Table 3-2, the Moran's I statistics in Belfast are highest for V, Co, Cr and Ni (0.1060 to 0.1879) demonstrating the geogenic Tertiary basalt influence over these PTEs which generates a higher spatial correlation. The lowest spatial correlation is noted for Sb (0.01645), followed by As (0.0179), Zn, Sn, Pb, Mo and Cu (0.0487) respectively, suggesting more contributions from anthropogenic point sources to these PTEs.

In Sheffield, the highest Moran's I statistic is calculated for Ni (0.0409), followed closely by V (0.0406) and Mo (0.0406), suggesting some spatial autocorrelation for these PTEs. Chromium (0.0295), Sn, Pb, Cu, Sb, As and Zn (0.0107) gradually show a reduction in the statistic calculated, suggesting less spatial autocorrelation for these PTEs in Sheffield. The lowest statistic is calculated for Co, suggesting the greatest contribution from anthropogenic point sources to this PTE.

3.4 Grouping PTEs of similar sources

Sample R scripts detailing the packages and functions used to complete the depth comparison, cluster analysis and PCA are included in Appendix C.

3.4.1 Depth comparison

Belfast

Both shallow (0-20cm) and deep (35-50cm) ICP (following an aqua regia digestion) data are available for Belfast, however data at depth is not available for Sheffield. Therefore a comparison of the results

from the two depths was completed for Belfast only. Depth ratio boxplots allow for the evaluation of geogenic and anthropogenic influences over the PTEs to be assessed. By comparing the geochemical concentrations measured at different depths, an understanding of the controlling sources over PTEs can be gained. If the shallow concentrations are more elevated this suggests an anthropogenic control over the PTE (Chiprés et al. 2009b). Generally a geogenic control will result in elevated concentrations in the deep samples (Galán et al. 2008). There will obviously be exceptions to these assumptions, for example the presence of deep anthropogenic point sources, such as buried tanks and pipes. However, across the urban study areas considered in this chapter, where geochemical samples have been taken at high densities, these assumptions are believed to be appropriate for this hypothesis regarding depth comparisons. The R script used to complete this comparison is provided in Appendix C.

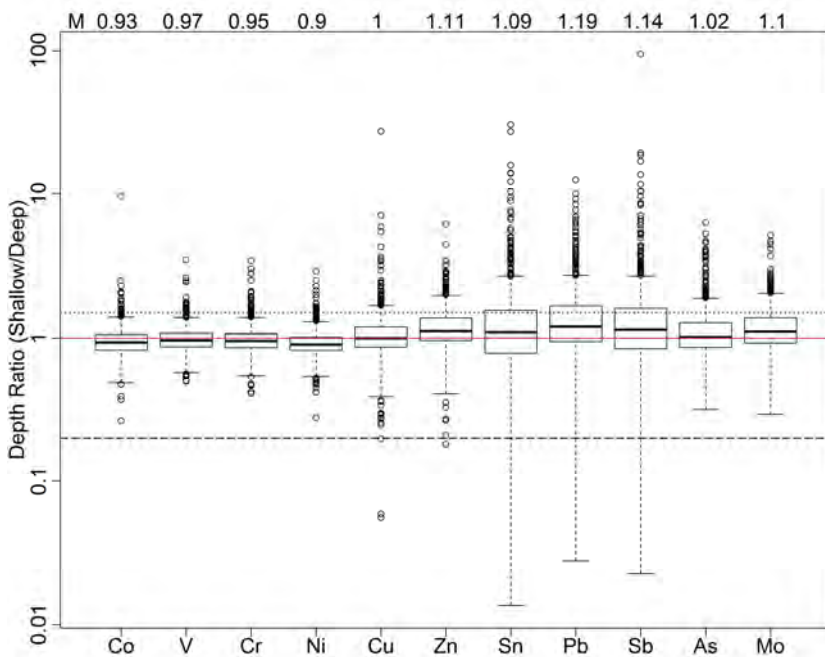


Figure 3-11 Boxplots of the shallow/deep PTE concentrations (depth ratio) using ICP following an aqua regia digestion data (solid red line shows where the shallow and deep concentrations are equal, dashed line shows where the depth ratio is equal to 0.2, dotted line shows where the depth ratio is equal to 1.5 and the M values represent the median depth ratio)

Similar boxplot characteristics are shown for Co, V, Cr and Ni; the majority of the boxplot falls below a ratio of one i.e. the median is less than one, representing more elevated concentrations in deep soils. A number of lower outliers are identified, as well as some upper outliers. The lower outliers are related to elevated concentrations at depth, while the upper outliers are likely to be related to anthropogenic sources of these PTEs within shallow soils. This pattern suggests a predominantly geogenic control

over these PTEs within Belfast. These upper outliers may be a useful way of spatially assessing anthropogenic point sources of these PTEs in the study area. These outliers show a great deal of consistency across the Co, V, Cr and Ni distributions; the depth ratio falls above 1.5 (Figure 3-11) for at least 3 of the PTEs at 18 sample locations and for all 4 PTEs at 8 sample locations.

Similarly, geogenic controls are also exerted over Cu and Zn, however the overall pattern for these PTEs is a little different. A larger variance of the ratio is obvious, with a wider dispersion of the boxplot's whiskers. A number of lower outliers represent the geogenic influence over these PTEs, however the increased amount of upper outliers suggest a more substantial anthropogenic contribution to Cu and Zn concentrations. The geogenic contribution to soil PTE concentrations is more obvious overlying the Tertiary basalts for Cu than Zn in Figure 3-11.

The remainder of the PTEs appear to be controlled by anthropogenic processes as the medians of the depth ratio are generally high and only upper outliers are present. They can be split into two groups; As and Mo behave in a similar manner as both show a relatively small variance with only upper outliers. Although Pb, Sb and Sn again only have upper outliers they demonstrate a much larger variance. This is related to a small number of samples where the deep concentration is much greater than the shallow concentration, for example the number of samples where the depth ratio is below 0.2 (dotted line on Figure 3-11) is 7, 6 and 15 for Pb, Sb and Sn respectively. Most of these samples are dispersed across the study area and are potentially related to sites where the deeper soil has been disturbed or replaced (perhaps with waste materials) during development leading to higher concentrations of these PTEs at depth. An alternative possible explanation for this pattern is that these PTEs are more easily leached from shallow to deeper soils at these locations.

In summary, these boxplots suggest the highest levels of anthropogenic input for Pb, Sb and Sn followed by As and Mo in Belfast, whilst geogenic inputs control the concentrations of Ni, Co, Cr and V, and Cu and Zn are influenced by both anthropogenic and geogenic inputs. These results correlate well with the spatial assessment completed in Section 3.3.

3.4.2 Cluster analysis

Belfast

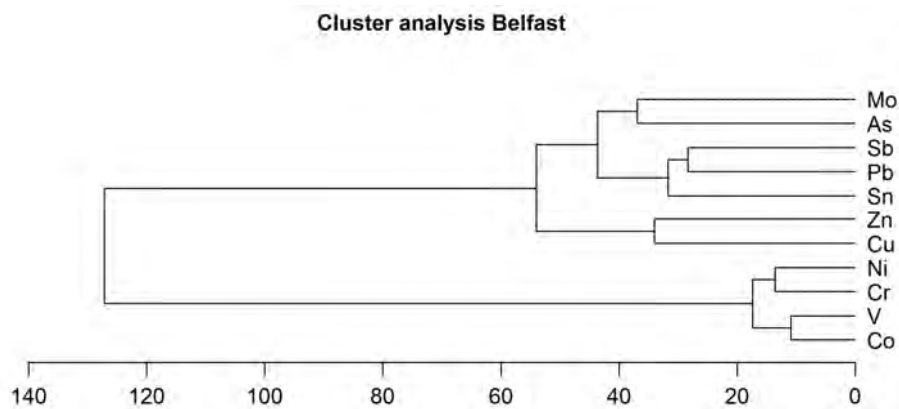


Figure 3-12 Results of cluster analysis completed for Belfast using shallow XRF data

Figure 3-12 provides the results of a cluster analysis completed for the shallow XRF data in Belfast; more detail on this analysis can be found in the R scripts in Appendix C. The results of the cluster analysis for Belfast groups the PTEs in a very similar pattern to that previously determined by the depth ratio analysis (Figure 3-11), thereby validating both sets of results. The results from both techniques also align closely with the results from Sections 3.2 and 3.3. The PTEs are first split into two main groups which can be explained by controlling geogenic (Co, V, Cr and Ni) and anthropogenic (Cu, Zn, Sn, Pb, Sb, As and Mo) factors.

Within the geogenic cluster two separate groups of Co and V, and Cr and Ni are present. Previous research has shown strong correlations between all these PTEs in a Northern Ireland context (Barsby et al. 2012), related to the stark control areas of Tertiary basalt exert over these PTEs. The total concentration maps of Co, V, Cr and Ni (Appendix B) show similar spatial distributions. Although the main control over these PTEs in Belfast is the Tertiary basalts, a few anthropogenic hot spots (many of them similar for these elements) are dispersed across Belfast. It could be these point sources, and their influence over the PTEs separately, that creates the difference between the Co and V, and Cr and Ni groups.

Three smaller groups make up the anthropogenic cluster; firstly Cu and Zn, secondly As and Mo and finally Sn, Pb and Sb (Pb and Sb are most closely related within this cluster). As previously discussed in Section 3.4.1, these separate groupings are related to the different source contributions for these PTEs. Although anthropogenic controls govern these concentrations, Cu and Zn are probably grouped because they also have geogenic contributions from the Tertiary basalts (Young & Donald 2013). From

an anthropogenic perspective, the close grouping of Cu and Zn could also be explained by their role in the production of brass (Herting et al. 2008). Arsenic and Mo are anthropogenically controlled, but a geogenic influence from the Silurian greywackes may contribute to elevated concentrations of As and Mo in overlying soils (Young & Donald 2013) and could cause them to cluster together within the overall anthropogenic cluster. No geogenic contributions to Pb, Sb or Sn could be identified within the study area suggesting a sole anthropogenic control over these PTEs, creating their separate grouping (Figure 3-12).

Sheffield

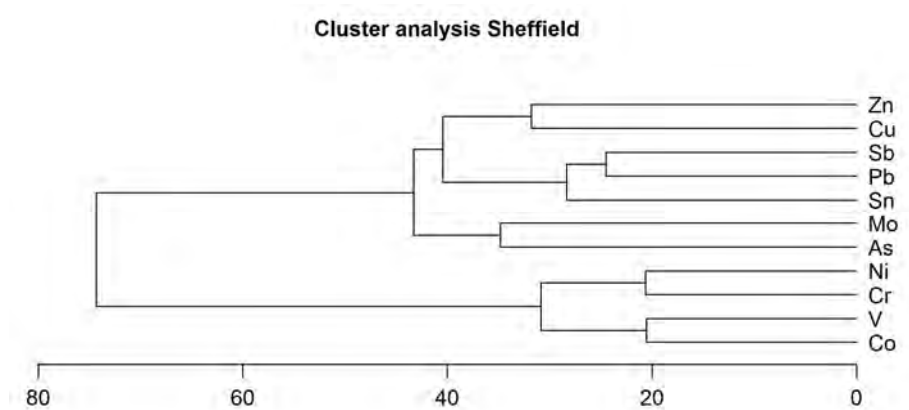


Figure 3-13 Results of cluster analysis completed for Sheffield using shallow XRF data

The results of the cluster analysis for Sheffield (Figure 3-13) are strikingly similar to those presented for Belfast (Figure 3-12). The Sheffield results also align quite well with the previous discussions in Sections 3.2 and 3.3. Two main separate groupings are noted; Zn, Cu, Sb, Pb, Sn, Mo and As are grouped separately from Ni, Cr, V and Co. As demonstrated on the concentration maps, Ni, Cr, V and Co all show elevated concentrations to the north-east of the study area. Although they are all affected by the presence of other point sources in Sheffield, the anthropogenic source to the north-east is probably their most obvious characteristic, resulting in them being grouped together.

In contrast, the other PTEs all have other factors affecting their concentration distributions. This cluster analysis suggests similar sources controlling three groups of PTEs, 1) Zn and Cu, 2) Sb, Pb and Sn and 3) As and Mo. Although these elements are anthropogenically controlled it is difficult to narrow down their specific sources. As noted for Belfast, Zn and Cu could be grouped together due to their role in the production of brass. The Pb, Sb and Sn concentrations are again likely to be solely anthropogenically controlled; their distributions (Section 3.3) would suggest a controlling atmospheric deposition source.

The difference in the scale on the cluster analysis from the two cities should be noted. The overall difference between the two groups, (Ni, Cr, V and Co are grouped separately from Mo, As, Sb, Pb, Sn, Zn and Cu) is much greater for Belfast than for Sheffield. In addition, the PTEs identified as being of geogenic origin in Belfast (Ni, Cr, V and Co) have a much smaller difference in height (between the two groups and the adjacent PTEs), than is noted for these PTEs in Sheffield. These two findings suggest a stronger similarity between these PTEs in Belfast, defending the identification of a geogenic control over these PTEs in Belfast and an anthropogenic control over these PTEs in Sheffield.

Although the other group of PTEs (Mo, As, Sb, Pb, Sn, Zn and Cu) show similar height (a measure of difference between either the individual data point or clusters) differences in Belfast and Sheffield between most of the groups, the overall difference between the Zn and Cu group, and the Mo, As, Sb, Pb and Sn group is much greater for Belfast than it is for Sheffield. This could be explained by the geogenic contributions to Zn and Cu in Belfast compared to the governing anthropogenic controls over Mo, As, Sb, Pb and Sn.

These results are useful for providing preliminary information on links between PTEs; in order to ascertain the underlying associations between the anthropogenic PTEs more detailed multivariate analysis is required, along with an understanding of how their sources vary spatially.

3.4.3 Principal component analysis

An initial PCA utilising all the elements included in the G-BASE and Tellus analysis of Sheffield and Belfast respectively was completed to gain an understanding of the general controls governing geochemistry in both cities. The results of this PCA are included in Appendix D.

Within this PCA utilising all the elements, PC1 in Belfast seems to be controlled by geology with areas influenced by the Tertiary basalts shown in red on the PC1 map (Appendix D) and areas mainly overlying the Silurian Greywacke shown in blue. As would be anticipated, the Tertiary basalts are associated with elements such as Co, Ni, Fe₂O₃, CaO, Cr, V and Cu while the samples overlying the Silurian Greywackes show correlation between Rb, K₂O, SiO₂, Zr, Hf, Na₂O and Ba. Principal Component 2 within Belfast highlights PTEs under anthropogenic influence in a red/orange halo surrounding the oldest part of the city. This area is associated with Pb, Sb, Sn, Mo, Zn and As and as such is similar to the trends previously identified for Belfast.

For Sheffield, PC1 from the PCA utilising all the elements demonstrates a relationship between a number of anthropogenically controlled elements and the area surrounding the oldest part of Sheffield city centre. It should be noted that this area is more spatially dispersed than the anthropogenic halo

defined for Belfast. This area is related to Sb, Cu, Pb, Sn, Mo and Zn which are therefore all anticipated to be under anthropogenic influence in Sheffield. The PC2 results for Sheffield demonstrate a relationship between Ni, Co, Cr and V and an area mainly along the north-east of the study area. It is difficult to determine the source of these PTEs in this area.

This overall PCA is useful for determining the anticipated main geogenic and anthropogenic controls over element concentrations, and gaining an idea of their spatial patterns. In order to further elucidate the controls over the PTEs under investigation in this report, a PCA has been completed in the following section with solely As, Co, Cr, Cu, Mo, Ni, Pb, Sb, Sn, V and Zn.

Belfast

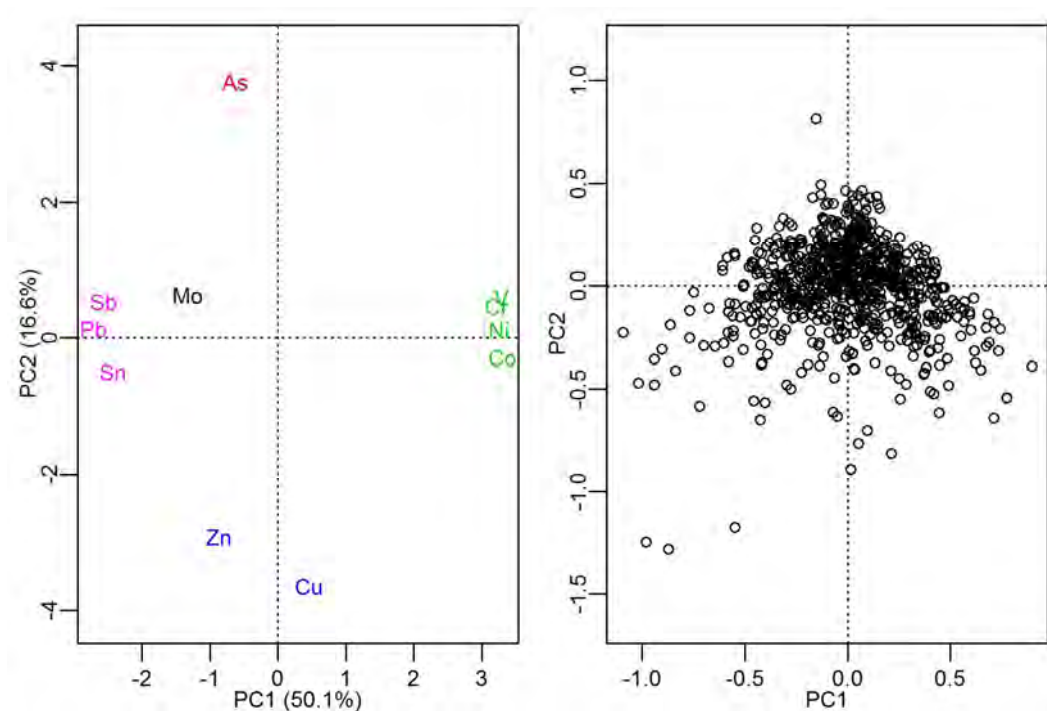


Figure 3-14 Results of PCA completed for Belfast using shallow XRF data

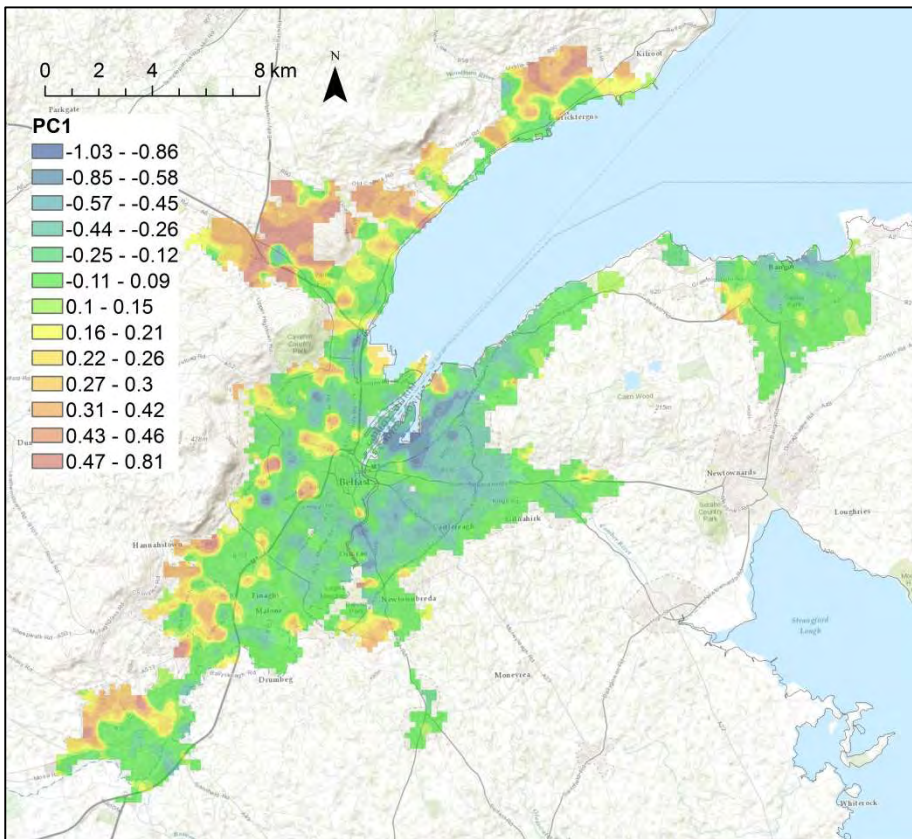


Figure 3-15 Map of PC1 from PCA completed for Belfast using shallow XRF data

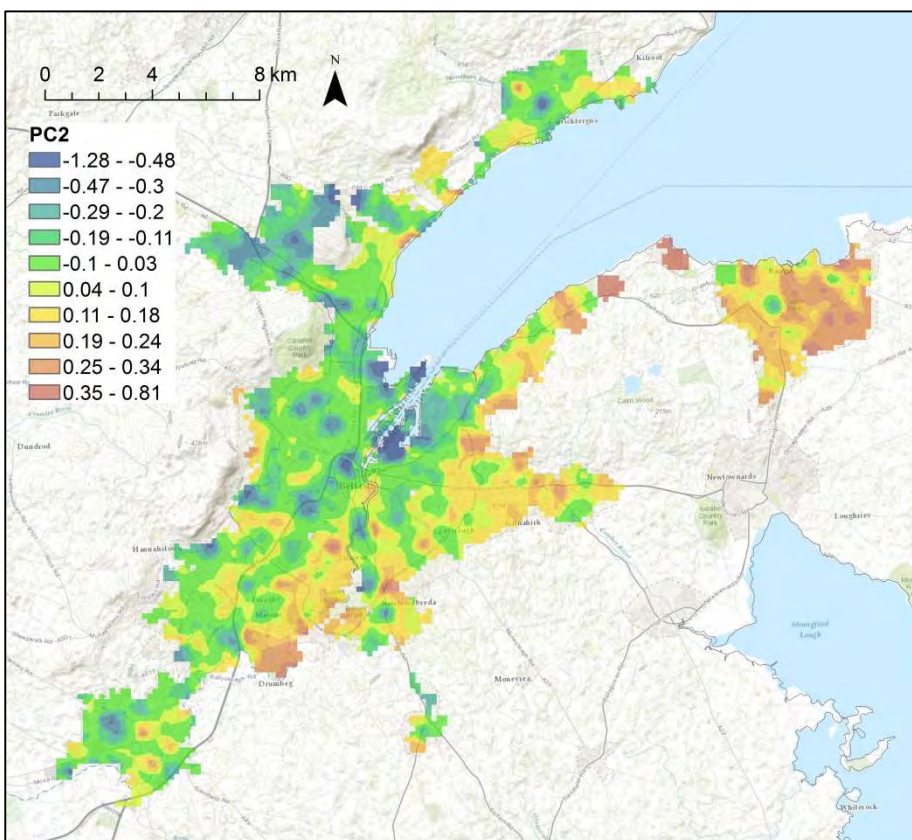


Figure 3-16 Map of PC2 from PCA completed for Belfast using shallow XRF data

The first two identified PCs explain 67 % of the total variance within the results of the Belfast PCA.

Table 3-3 Main controls over PC1 and PC2 for Belfast

	Controlling PTEs
PC1	Co, V, Cr & Ni vs. Pb, Sb & Sn
PC2	Cu & Zn vs. As

PC1 separates the geogenic controls over Co, V, Cr and Ni from the anthropogenic PTEs. Cobalt, V, Cr and Ni, shown in green, cluster closely on the biplot (Figure 3-14) reemphasising their almost identical spatial distributions related to their similar sources. The strong control exerted by the Tertiary basalts over these PTEs is represented in the red areas in the PC1 map (Figure 3-15).

The PC1 results distinguish Pb, Sb and Sn (shown in pink) as the most clearly defined anthropogenic PTEs. The map of PC1 shows the anthropogenic group to create a halo effect around the oldest part of the city, with a stronger presence towards the east of the city. The shipbuilding industry has been based within this area of this city for many years, with George Best Belfast City Airport now also located here. This pattern is particularly obvious from the 1901 zone out to the current Belfast zone, suggesting a long-term pattern of contamination within these soils.

PC2 appears to be explained by different contaminant sources contributing to As (red), and Cu and Zn (blue). The widespread As contributions suggest a domestic source such as coal combustion (Duan & Tan 2013), whereas the dark areas related to Cu and Zn suggest point sources across the city centre and also highlight the area of Tertiary basalts in the west of the city. A geogenic contribution to As in areas overlying Silurian greywacke is also possible; although the concentrations of As in this type of bedrock would not be expected to be particularly elevated, we would expect them to be higher than those found over the Tertiary basalts (Young & Donald 2013).

Sheffield

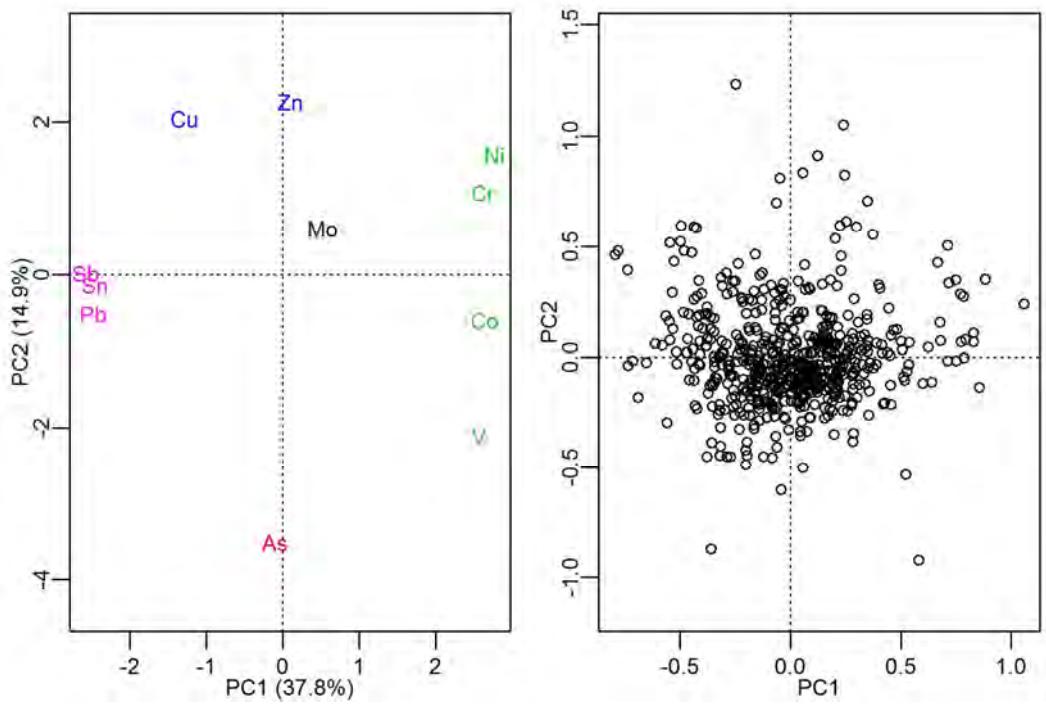


Figure 3-17 Results of PCA completed for Sheffield using shallow XRF data

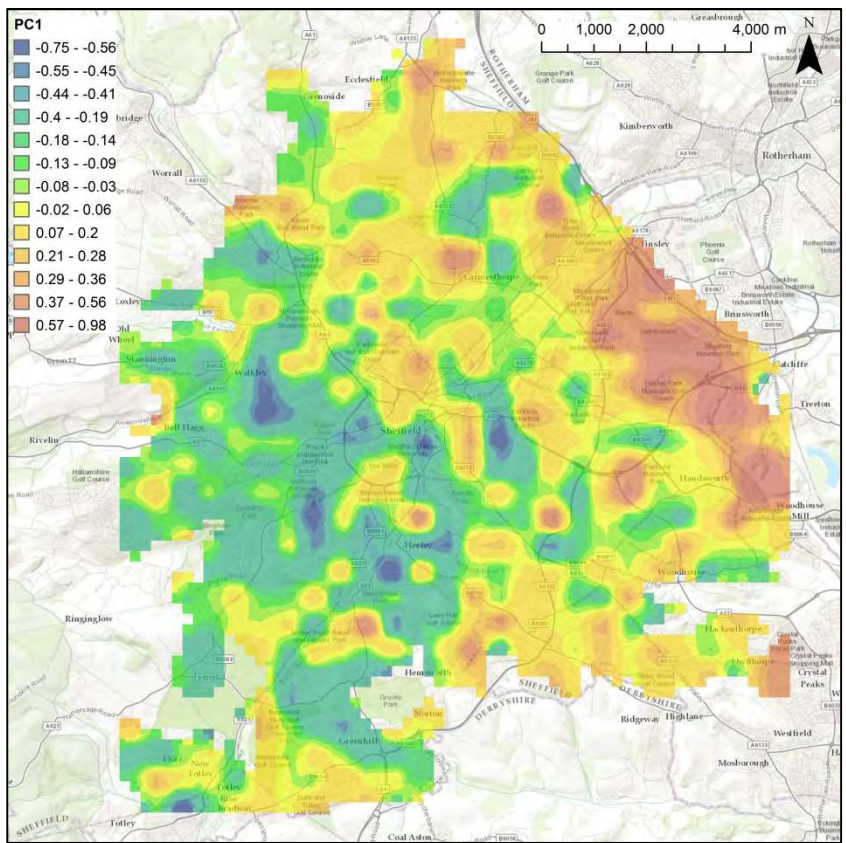


Figure 3-18 Map of PC1 from PCA completed for Sheffield using shallow XRF data

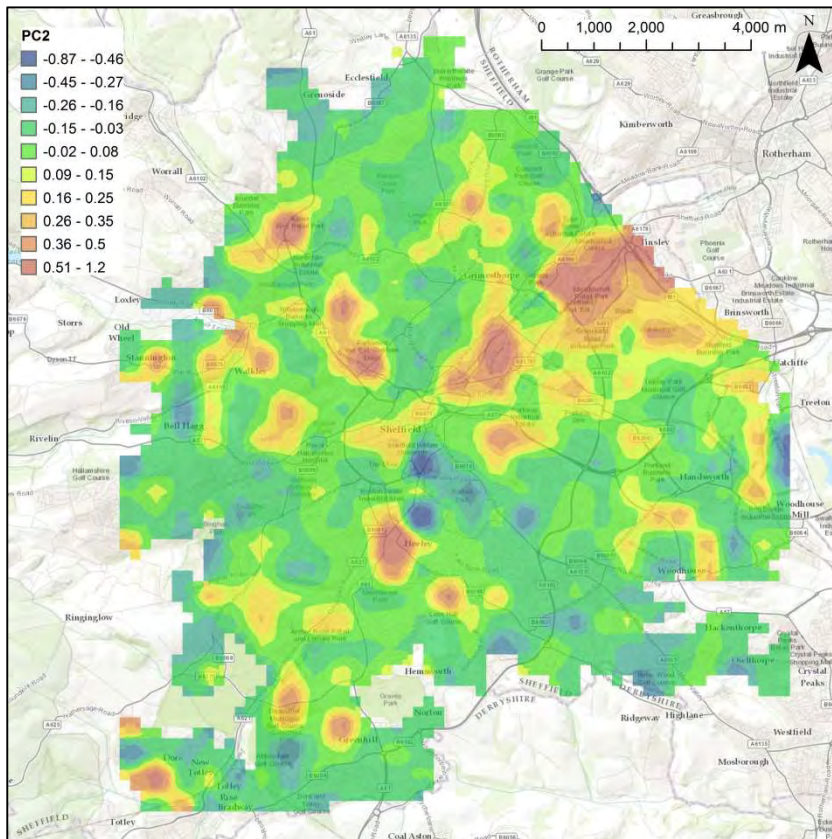


Figure 3-19 Map of PC2 from PCA completed for Sheffield using shallow XRF data

The PCA outputs for Sheffield are again similar to those for Belfast (Figure 3-17). In total PC1 and PC2 account for 53% of the total variance within the geochemical dataset for Sheffield. This is 14% less than the PCA completed for Belfast demonstrating a greater complexity associated with soil geochemistry in Sheffield.

Within PC1 the results again clearly separate Ni, Cr, Co and V (green) from Sb, Pb and Sn (pink). In contrast to the Belfast results, Ni, Cr, Co and V are distributed much more widely on the biplot (Figure 3-17), demonstrating that although they are still anticipated to be from the same or similar sources, they are not as closely correlated in Sheffield as they are in Belfast. The map for PC1 demonstrates that these PTEs exert the greatest control over a large area along the north-east boundary of the study area where these PTEs were previously shown to occur at elevated concentrations (Appendix B). This pattern is thought to be related to the various industrial land uses located in this part of the city; many of Sheffield's iron and steel works were/are found here. The elevated concentrations could be related to the industrial use (in these various factories) of the coal which also occurs naturally in this area.

The darker blue colours on the PC1 map represent the areas controlled by Pb, Sb and Sn, and similarly to Belfast these PTEs are shown to form a halo around the oldest area of the city. The widespread

nature of these PTEs on their total concentration maps (Appendix B) suggests an atmospheric deposition anthropogenic source.

For Sheffield, PC2 separates As (red), from Cu and Zn (blue). In contrast to the Belfast PCA results, Ni and Cr also seem to have an influence over PC2, falling towards the same side as Zn and Cu on the biplot. The blue areas on the PC2 map are related to elevated concentrations of As, suggesting a particular point source of As in these areas. The red areas are related to Zn, Cu, Ni and Cr and fall in a similar, though smaller, area to the north-east of the location identified in PC1. As shown in Figure 3-20, PC2 seems to be influenced by the geogenic presence of coal with many of the red areas on the map aligning with coal outcrops in Sheffield. Therefore, PC2 seems to identify a geogenic or mining contribution to PTE concentrations (Zn, Cu, Ni and Cr) in Sheffield in the form of coal outcrops.

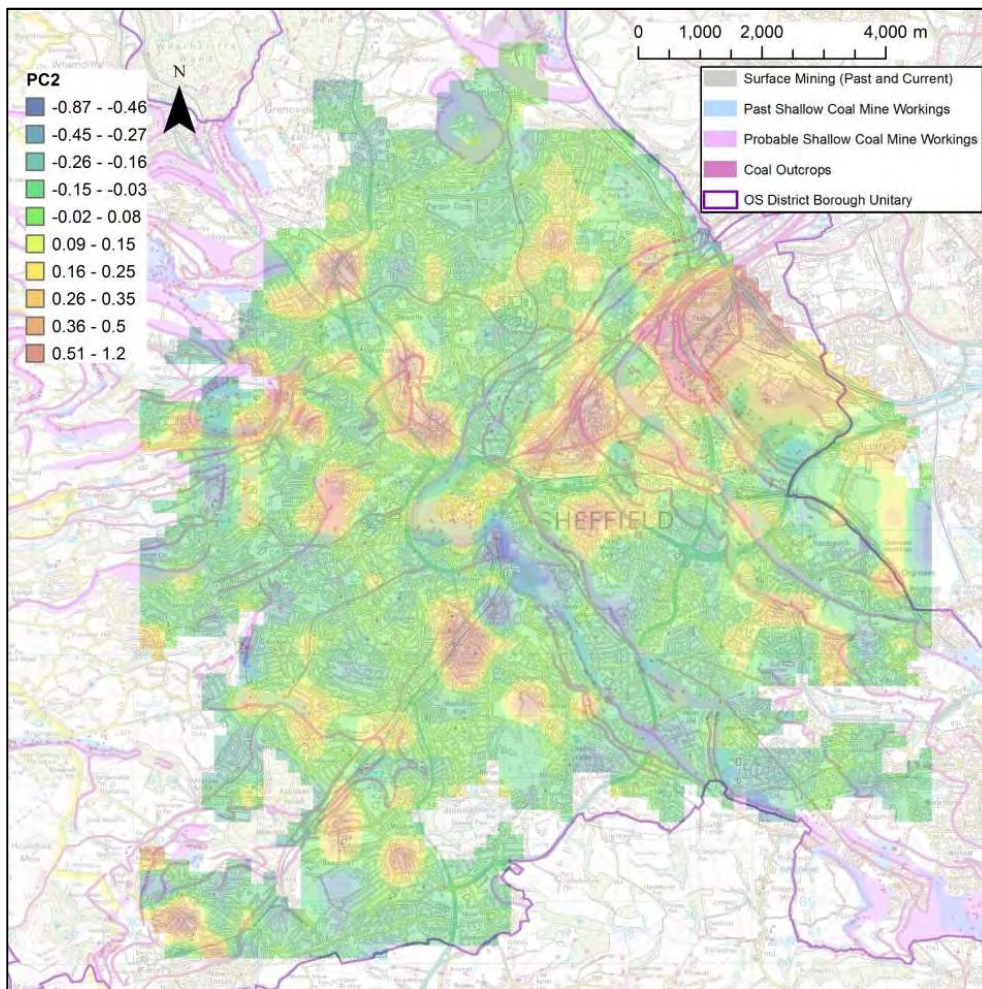


Figure 3-20 Map of PC2 from PCA completed for Sheffield using shallow XRF data compared with the specific coal mining legacy plan for Sheffield created by the Coal Authority (The Coal Authority 2015)

Compared to the overall PCA completed with all available elements (Appendix D), the more refined PCA completed with the eleven PTEs of specific interest to this study provides a better approach for identifying specific sources of PTEs.

3.4.4 Pollution indices

Belfast

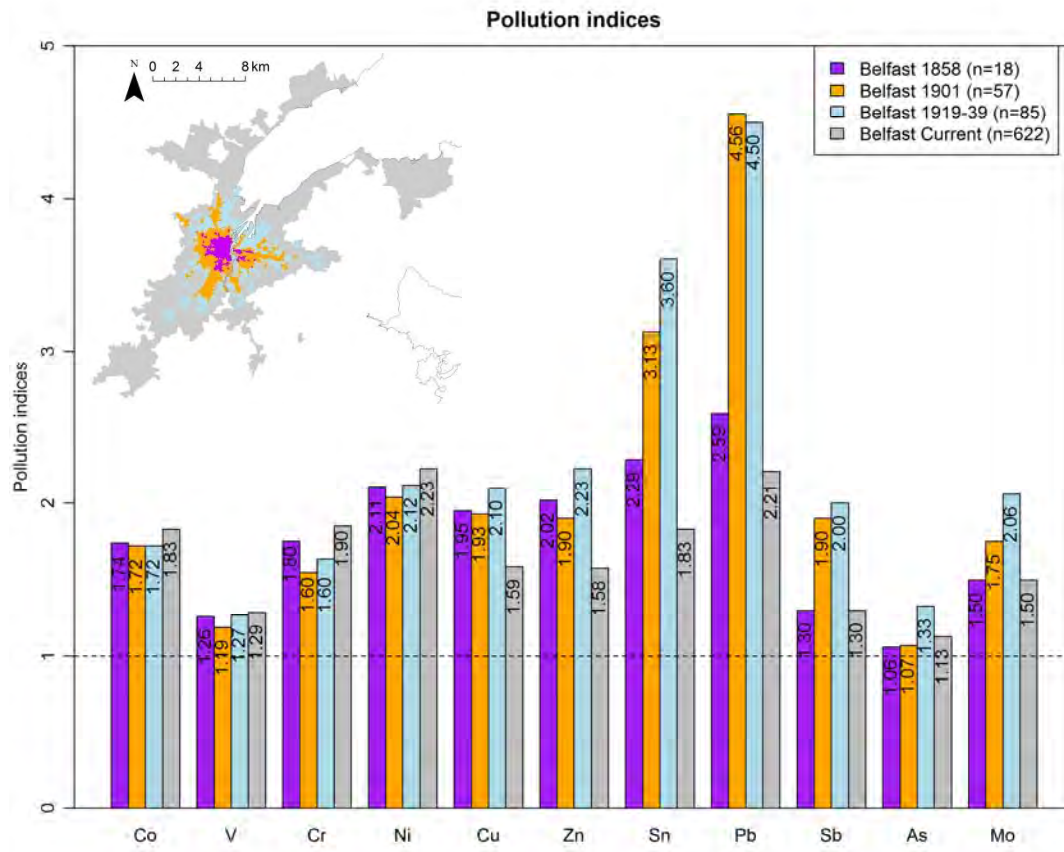


Figure 3-21 Bar chart showing the PI for the PTEs within each of Belfast’s development zones (dashed line where PI = 1)

Pollution indices are used to assess the enrichment of the PTEs in the different historical development zones of both cities. In Belfast, the PIs are calculated using rural median concentrations, thereby demonstrating enrichment of these PTEs in the different Belfast zones compared to rural Northern Ireland. The PIs suggest that certain groups of PTEs are related to different development zones of Belfast (Figure 3-21).

For Co, V, Cr and Ni the highest PI is within the current Belfast zone, due to the fact that a greater proportion of the current Belfast area overlies basalts, which have a significant control over their concentrations, than in the remainder of the development zones. These PIs are much reduced when the

median concentration of the PTEs in soils overlying areas of Tertiary basalts replaces the rural median concentration (Table 3-4). For example, the PIs for V, Co, Cr and Ni within the current development zone reduce from 1.29, 1.83, 1.85 and 2.23 to 0.59, 0.69, 0.62 and 0.66 respectively.

Table 3-4 PIs generated for V, Cr, Co and Ni using the median basalt concentration in place of the overall rural median concentration

	V	Cr	Co	Ni
Belfast 1858 median / Basalt median	0.58	0.65	0.59	0.62
Belfast 1901 median / Basalt median	0.55	0.58	0.58	0.60
Belfast 1919-1939 median / Basalt median	0.58	0.61	0.58	0.62
Belfast current median / Basalt median	0.59	0.69	0.62	0.66

Pb concentrations are at their highest within the 1901 zone, although the PI is only slightly higher at 4.56 than that within the 1919-1939 zone (4.50). For Cu, Zn, Sn, Sb, As and Mo the highest PIs are located within the 1919-1939 zone suggesting this development zone is the most enriched in the majority of the anthropogenically controlled PTEs. However, these PTEs do differ in the zones containing the second highest PI; for Cu and Zn these fall within the 1858 zone, for Sn, Sb and Mo within the 1901 zone and for As within the current Belfast zone. This suggests that rapid growth associated with development of heavy industry in the city between 1901 and 1919-1939 may be responsible for the elevated concentrations of a number of PTEs however contamination from all of these PTEs is likely to have begun before this period of time. The similar pattern shown for Cu and Zn again could be related to their role as constituents of brass.

Overall, Pb and Sn have the highest pollution indices suggesting the biggest anthropogenic enrichment of these PTEs in Belfast. These are also two of the oldest and best recognised urban contaminants. Although Sb and Mo also show anthropogenic enrichment, it is at a much lower level with PIs ranging between 1.30 and 2.00 for Sb and 1.50 and 2.06 for Mo. Geogenic contributions to As in the form of natural mineralisation (McIlwaine et al. 2014) in rural areas of Northern Ireland probably results in lower PIs for As.

Sheffield

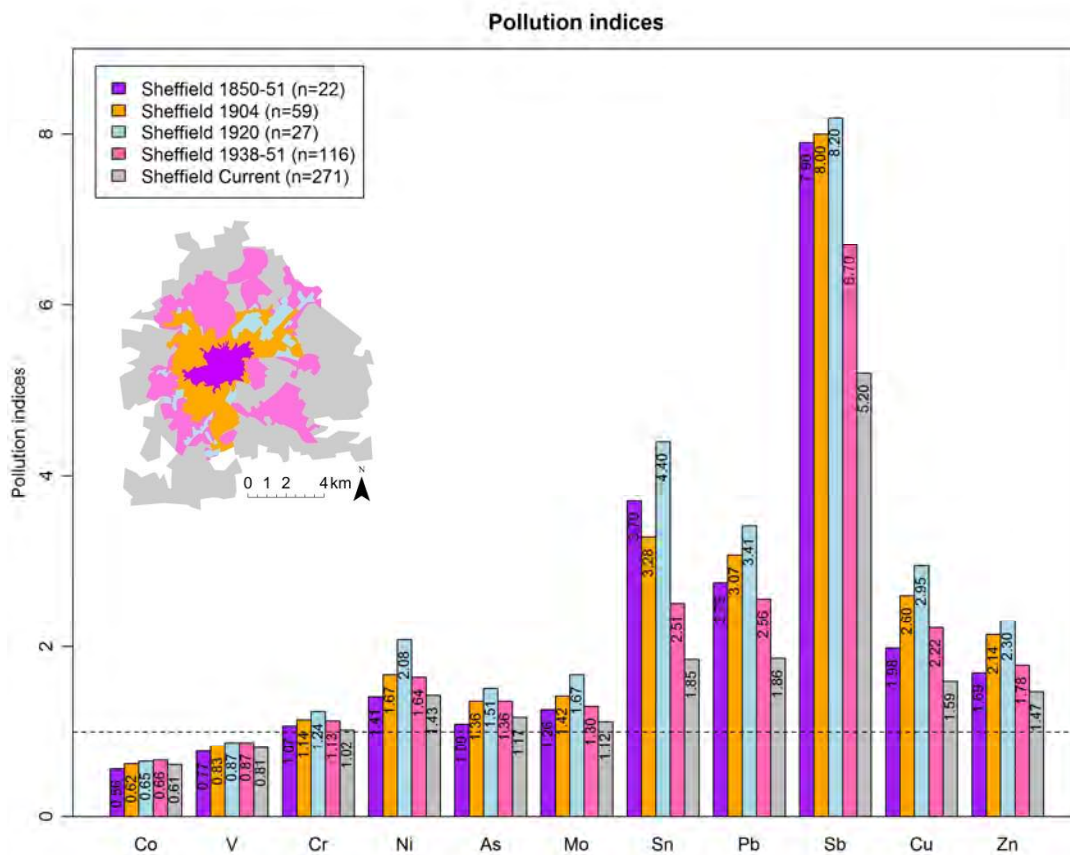


Figure 3-22 Bar chart showing the PI for the PTEs within each of Sheffield’s development zones (dashed line where PI = 1)

As opposed to the PIs calculated for Belfast, the PIs for Sheffield are calculated using a rural median solely from the Carboniferous Coal Measures, rather than the median concentrations from the entire surrounding rural area. When interpreting the PIs calculated for Sheffield, it should be noted that the results therefore tell us about enrichment of these PTEs in the urban area compared to one type of bedrock geology, rather than enrichment compared to the rural area as a whole.

The PIs calculated for Sheffield are consistently highest within the 1920 zone. This suggests that industry in Sheffield between 1904 and 1920 may be responsible for the most elevated PTE concentrations in soils. For Sn, Pb and Sb the PIs fall considerably within the 1938-1951 zone, below all three of the previous historical zones. The PIs are lowest for Cr, Mo, Sn, Pb, Sb, Cu and Zn in the Current zone, and lowest for Co, V, Ni and As in the 1850-1851 zone.

Lead, Sb and Sn show the greatest accumulation in urban soils (highest PIs), followed by Cu and Zn. The PIs for Sb are so high due to its low median concentration in the Carboniferous Coal Measures (0.5

mg/kg). Cobalt and V are actually depleted in urban soils when compared to the rural median, while Cr, Ni, As and Mo show minimal accumulation.

3.5 Urban subdomains

In summary, a mixture of geogenic and anthropogenic influences were identified over the PTE concentrations in both cities. In Belfast, Co, Cr, Ni and V were shown to be geogenically controlled by the Tertiary basalts. This geogenic influence was also evident for Cu and Zn, however both these PTEs were also under anthropogenic influence, showing elevated concentrations across Belfast city centre. Potential geogenic influences over As and Mo were identified in the form of the Silurian Greywackes, with additional anthropogenic influences over both PTEs across Belfast. Solely anthropogenic influences were identified over Pb, Sb and Sn in Belfast, most likely in the form of atmospheric deposition.

An industrial area to the north-east of Sheffield was identified to be contributing to elevated concentrations of Ni, Co, Cr and V. Areas of open cast mining as well as iron and steel works were identified in this area, suggesting an anthropogenic source of these PTEs related to the industrial use of the mined coal. The natural presence of coal was also found to be contributing geogenically to concentrations of Zn, Cu, Ni and Cr in Sheffield. Similarly to Belfast, a sole anthropogenic source of Pb, Sb and Sn in the form of atmospheric deposition was identified. Both As and Mo were also considered to be anthropogenically controlled in Sheffield.

Although the evidence is more conclusive for Belfast, the different historical zones within both cities have been demonstrated to be related to different sources of the PTEs. This is demonstrated by the PIs where differing levels of enrichment are shown across the varying historical boundaries and is clearly shown for a number of the PTEs on their spatial maps where halos surrounding the oldest areas of the cities are visible (Section 3.3). To allow calculation of background values, the different historical zones will be considered as urban subdomains, and background values calculated within each of them for any anthropogenically influenced PTEs.

3.6 Background concentration calculations

Within the Belfast study area, four PTEs (Co, Cr, Ni and V) have been clearly identified as elements that are geogenically controlled. The remainder of the PTEs in Belfast (As, Cu, Mo, Pb, Sb, Sn and Zn) are thought to be affected in some way by anthropogenic inputs. In Sheffield, all the PTEs are thought to be under some anthropogenic influence. From these results, different methods for calculating

background concentrations have been applied to assess the concentrations of the anthropogenically controlled PTEs within the city development zones, which have been shown to act as urban subdomains. These background values indicate what a 'typical' or background concentration of these PTEs would be within the defined domain; aiming to differentiate between concentrations related to geogenic and diffuse anthropogenic sources, and concentrations generated by point sources.

Belfast

Table 3-5 Summary of ULBLs calculated for the PTEs in Belfast regarded as having some anthropogenic input within the separate development zones (highest shown in bold and lowest in italics)

	ULBLs (mg/kg)			
	1858	1901	1919-1939	Current
As	<i>18</i>	21	26	19
Cu	<i>120</i>	160	200	<i>120</i>
Mo	2.8	3.5	5.1	<i>2.7</i>
Pb	<i>190</i>	430	620	200
Sb	<i>2.7</i>	4.3	7.2	3.0
Sn	16	20	51	<i>14</i>
Zn	<i>240</i>	310	510	240

Table 3-6 Summary of Median + 2MAD background values calculated for the PTEs in Belfast regarded as having some anthropogenic input within the separate development zones (highest shown in bold and lowest in italics)

	Median + 2MAD (mg/kg)			
	1858	1901	1919-1939	Current
As	<i>13</i>	15	17	14
Cu	95	100	120	<i>80</i>
Mo	<i>1.8</i>	2.4	3.0	<i>1.8</i>
Pb	140	280	270	<i>120</i>
Sb	<i>1.9</i>	3.1	3.7	2.1
Sn	11	14	18	<i>7.7</i>
Zn	<i>210</i>	220	290	<i>170</i>

Table 3-7 Summary of NBC values calculated for the PTEs in Belfast regarded as having some anthropogenic input within the separate development zones (BC = box-cox transformation, Log = log transformation and Emp = empirical percentiles utilised)

	NBC (mg/kg)					
	1901		1919-1939		Current	
As	BC	37	Emp	52	Log	21
Cu	BC	210	BC	640	Log	130
Mo	Log	4.7	Emp	18	Log	3.1
Pb	Log	490	BC	1300	Log	260
Sb	Log	10	BC	33	Log	4.3
Sn	Log	33	BC	1000	BC	24
Zn	Log	470	BC	2100	Log	290

NBCs cannot be calculated for the 1858 Belfast domain as there are only 18 samples available within this area. With the exception of the Median + 2MAD method for Pb, all the background values calculated are highest within the 1919-1939 zone. This would be expected for all the PTEs apart from Pb, where the calculated PI is highest within the 1901 zone (Figure 3-21) suggesting the greatest enrichment of Pb in this zone. Upon further investigation this was found to be related to the distribution of the data. As the PI is based on the median of the dataset it is less affected by skew than the background value calculations via the ULBL and NBC methods; the 1919-1939 Pb data is more highly skewed than the 1901 Pb data. This suggests a more homogenous source of Pb, such as atmospheric deposition, within the 1901 zone, whereas the 1919-1939 zone is possibly witnessing atmospheric deposition as well as more independent point sources of Pb. This is possibly associated with increased development in the east of the city within this 1919-1939 zone, as identified and discussed within the PC1 results.

Sheffield

Table 3-8 Summary of ULBLs calculated for the PTEs in Sheffield regarded as having some anthropogenic input within the separate development zones (highest in bold and lowest in italics)

	ULBLs (mg/kg)				
	1850-1851	1904	1920	1938-1951	Current
As	62	70	54	50	<i>41</i>
Co	31	29	36	29	29
Cr	220	190	270	190	<i>160</i>
Cu	210	300	250	180	<i>120</i>
Mo	12	12	15	9.1	8.3
Ni	97	82	120	75	<i>68</i>
Pb	940	930	800	560	<i>370</i>
Sb	16	14	12	11	7.2
Sn	100	77	69	46	33
V	130	140	130	140	140
Zn	420	620	520	420	330

Table 3-9 Summary of Median + 2MAD background values calculated for the PTEs in Sheffield regarded as having some anthropogenic input within the separate development zones (highest in bold and lowest in italics)

	Median + 2MAD (mg/kg)				
	1850-1851	1904	1920	1938-1951	Current
As	34	44	39	38	<i>31</i>
Co	22	23	24	23	23
Cr	130	130	140	130	<i>120</i>
Cu	110	160	150	120	<i>80</i>
Mo	7.4	8.3	8.3	6.7	5.9
Ni	57	62	82	58	<i>51</i>
Pb	440	530	460	360	<i>250</i>
Sb	8.0	8.2	6.7	6.2	4.6
Sn	47	45	48	28	21
V	110	110	110	110	110
Zn	290	400	380	280	230

Table 3-10 Summary of NBC values calculated for the PTEs in Sheffield regarded as having some anthropogenic input within the separate development zones (None = no transformation required, BC = box-cox transformation, Log = log transformation and Emp = empirical percentiles utilised)

	NBC (mg/kg)					
	1904		1938-1951		Current	
As	Log	77	Log	61	Log	50
Co	Log	40	Emp	41	Emp	28
Cr	Emp	780	Emp	600	Emp	370
Cu	Log	590	Log	260	BC	200
Mo	Log	23	BC	19	Log	13
Ni	BC	240	BC	140	Log	89
Pb	Log	1400	Log	790	Log	600
Sb	Log	20	BC	22	Log	9.2
Sn	Log	170	Log	64	Log	46
V	None	140	Emp	150	Log	130
Zn	Log	920	BC	880	Log	390

NBCs cannot be calculated for the 1850-1851 or 1920 Sheffield domain as there are only 22 and 27 samples available within these areas respectively. The domains with the most elevated background values are reasonably consistent across the ULBL and Median + 2MAD calculations; it is difficult to draw a comparison with the NBCs due to the missing subdomains. Where these differences occur it is likely to be related to the distribution of the data, as explained for Pb in the Belfast results.

The main differences between the ULBL and Median + 2MAD concentrations are noted for Pb, Sb and Sn. The highest concentrations via the ULBL method are in the 1850-1851 zone for all three PTEs, whereas via the Median + 2MAD method the highest concentrations are in the 1904 zone for Pb and Sb and the 1920 zone for Sn. In conjunction with the spatial distribution previously discussed, the most elevated concentrations of all of these PTEs seem to form a halo surrounding the oldest part of the city, suggesting a stronger alignment with the ULBLs than the Median + 2MAD values.

The ULBL method highlights the highest concentrations for Co, Cr, Mo and Ni in the 1920 zone, while the Median + 2MAD method has the highest concentrations for Co, Cr, Mo, Ni, Sn and V in this zone. The concentrations of V do not vary much across the development zones for any the methods used to calculate background values, suggesting this PTE is not influenced by different periods of historical development to the same extent as the other PTEs.

Table 3-11 Summary of the percentage of concentrations exceeding the various background concentrations using Sheffield's Current zone as an example

	Percentage of concentrations exceeding various background concentrations within Sheffield's Current zone										
	As	Co	Cr	Cu	Mo	Ni	Pb	Sb	Sn	V	Zn
ULBL	6.3	3.0	11.1	8.9	11.1	8.9	11.4	7.0	8.1	2.6	10.0
Median + 2MAD	14.8	7.7	20.3	19.6	17.3	17.0	18.5	19.2	19.2	13.7	19.6
NBC	3.0	3.0	3.0	4.4	5.5	4.4	3.3	3.7	3.7	5.2	5.9

As can be seen in Table 3-11, the most conservative background concentrations (i.e. the lowest concentrations) are calculated by the Median + 2MAD method, while the least conservative (i.e. the highest concentrations) are calculated via the NBC method, for all the PTEs except V. This is true for the majority of the background concentrations calculated, not just for Sheffield's Current zone (Table 3-11).

Rothwell & Cooke (2015) suggested the median + 2MAD method for use because it consistently calculated the most conservative background concentrations in their study in Gateshead. Although this is a sensible precaution from a risk perspective, it may not be realistic to state that further investigation may be required at 19% of sampled sites within Sheffield's Current zone where the concentration of Pb is above the calculated Median + 2MAD value. The ULBLs provide concentrations between the most conservative Median + 2MAD concentrations and the least conservative NBCs; they may therefore be more appropriate values to use in gaining an understanding of background concentrations of different PTEs in these studies. In addition, the NBC methodology is only applicable to domains with more than 30 samples and so it can't be applied in a number of the urban subdomains identified.

The ULBL methodology has previously been applied on a Northern Ireland regional scale to generate TTVs (McIlwaine et al. 2014) as it was identified as the most appropriate method for calculating background values.

4 Conclusions

The scale of PTE concentration data available within the study areas considered allowed for a thorough examination of the effects of historical development on soil PTE concentrations. Clear groups of PTEs were identified within the study areas investigated via depth ratios, a range of multivariate statistical techniques and PIs. Controlling sources and links between historical development and PTE concentrations were identified.

The combination of statistical and spatial techniques used provided a thorough methodology for examining the geochemical data. The techniques must be used side by side to gain an understanding of the PTE distributions from both statistical and spatial perspectives. Simply considering statistical distributions and spatial patterns alone provides a lot of information which can then be augmented by multivariate techniques. The investigative methodology employed within this research may be useful for application within other urban environments.

Potentially toxic elements with a controlling geogenic source in Belfast were identified as Co, V, Cr and Ni. The marked similarity in the spatial distributions of these PTEs clearly demonstrates the control that the Tertiary basalts have over their concentrations, and suggests similar point sources of anthropogenic contributions perhaps due to importing of 'clean' topsoil, which originally overlaid the Antrim basalts, in these areas. PTEs under predominantly anthropogenic sources in Belfast can be split into three groups; 1) Sn, Pb and Sb, 2) Cu and Zn and 3) As and Mo. Cu and Zn receive some geogenic contribution to their concentrations from the Tertiary basalts but similar anthropogenic contributions are also obvious. Increasing anthropogenic contributions to both As and Mo see them grouped similarly, with Pb, Sb and Sn noted for the greatest anthropogenic contribution.

All of the PTEs investigated were found to be under some anthropogenic influence in Sheffield. Nickel, Co, Cr and V were found to align well in a large area along the north-east boundary of Sheffield where they were shown to occur at elevated concentrations. This pattern is thought to be related to the various industrial land uses located in this part of the city; many of Sheffield's iron and steel works were/are found here. The elevated concentrations could be related to the industrial use (in these various factories) of the coal which also occurs naturally in this area. Similarly to Belfast, Pb, Sb and Sn were shown to form a halo around the oldest area of the city. The widespread nature of these PTEs on their total concentration maps suggests an atmospheric deposition anthropogenic source. A geogenic contribution to Zn, Cu, Ni and Cr concentrations was identified in the form of coal outcrops in Sheffield.

The PIs calculated show that different historical zones are related to different levels of PTE accumulation. The relationship between historical development and differing PTEs is a novel finding from this research. This suggests that PTEs have the potential for use as 'urbanisation tracers' as different PTEs have been shown to be associated with different historical anthropogenic sources.

Background values were calculated for the PTEs deemed to have some form of anthropogenic input within each of the city's development zones. These PTEs demonstrated the varying historical sources within the development zones which result in varying background values. The background values for

Belfast were generally highest in the 1919-1939 development zone, suggesting that contamination was at its greatest in Belfast between 1901 and 1939. The background values calculated for Sheffield varied more widely across the different development zones considered.

Point sources have not been considered in great detail in this study. Although they are likely to cause problems on a small scale, this research suggests that diffuse contamination from larger scale issues such as traffic, domestic combustion and more widespread industrial processes contribute a large part of the contamination within urban environments. Current statutory guidance for Part 2A of the Environmental Protection Act states that widespread geogenic or diffuse anthropogenic pollution in soil should not be regarded as contaminated land unless other evidence demonstrates that they pose a risk. These findings suggest that further research into the risk posed by these elevated concentrations may be required.

References

- Aitchison, J., 1982. The Statistical Analysis of Compositional Data. *Journal of the Royal Statistical Society*, 44(2), pp.139–177.
- Ander, E.L., Johnson, C.C., Cave, M.R., Palumbo-Roe, B., Nathanail, C.P. & Lark, R.M., 2013. Methodology for the determination of normal background concentrations of contaminants in English soil. *The Science of the total environment*, 454-455, pp.604–618. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/23583985> [Accessed May 7, 2013].
- Barsby, A., McKinley, J.M., Offerdinger, U., Young, M., Cave, M.R. & Wragg, J., 2012. Bioaccessibility of trace elements in soils in Northern Ireland. *The Science of the total environment*, 433, pp.398–417. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/22819891> [Accessed November 29, 2012].
- Bazley, R.A.B., Manning, P.I., Wilson, H.E. & Dunham, K.C., 1984. *Geology of Belfast and District*, Southampton: Director General of the Ordnance Survey.
- Beckett, J.C. & Glasscock, R.E. eds., 1967. *Belfast: Original & growth of an industrial city*, British Broadcasting Corporation.
- Biasioli, M., Barberis, R. & Ajmone-Marsan, F., 2006. The influence of a large city on some soil properties and metals content. *The Science of the total environment*, 356, pp.154–164. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/15941578> [Accessed December 18, 2014].
- Candeias, C., Ferreira da Silva, E., Salgueiro, A.R., Pereira, H.G., Reis, A.P., Patinha, C., Matos, J.X. & Ávila, P.H., 2011. The use of multivariate statistical analysis of geochemical data for assessing the spatial distribution of soil contamination by potentially toxic elements in the Aljustrel mining area (Iberian Pyrite Belt, Portugal). *Environmental Earth Sciences*, 62, pp.1461–1479. Available at: <http://www.springerlink.com/index/10.1007/s12665-010-0631-2> [Accessed November 22, 2012].
- Carrero, J.A., Arrizabalaga, I., Bustamante, J., Goienaga, N., Arana, G. & Madariaga, J.M., 2013. Diagnosing the traffic impact on roadside soils through a multianalytical data analysis of the concentration profiles of traffic-related elements. *The Science of the total environment*, 458-460, pp.427–434. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/23685368> [Accessed February 4, 2014].
- Cave, M.R., Johnson, C.C., Ander, E. & Palumbo-Roe, B., 2012. Methodology for the determination of normal background contaminant concentrations in English soils. *British Geological Survey Commissioned Report*, CR/12/003, p.42pp.

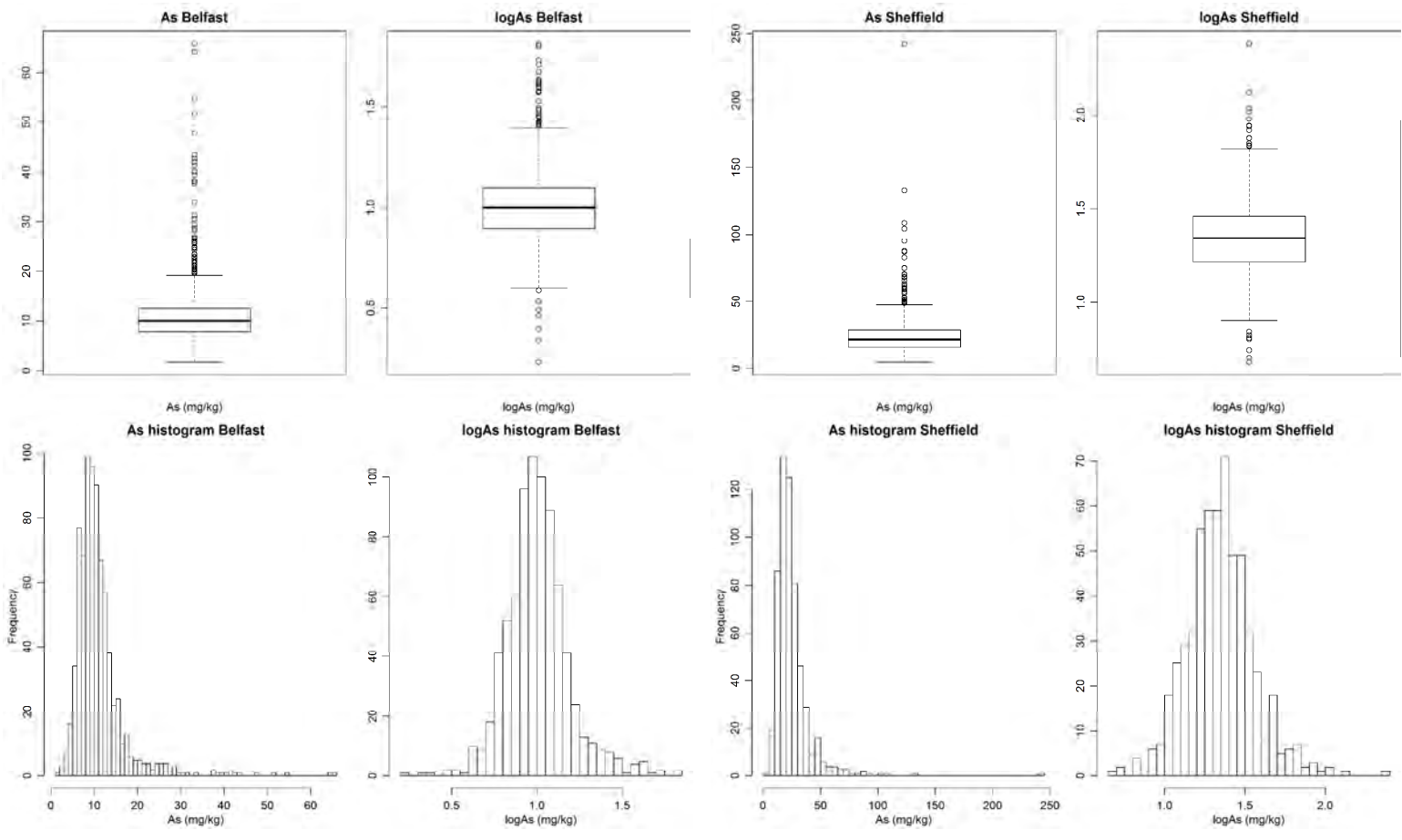
- Chiprés, J.A., de la Calleja, A., Tellez, J.I., Jiménez, F., Cruz, C., Guerrero, E.G., Castro, J., Monroy, M.G. & Salinas, J.C., 2009a. Geochemistry of soils along a transect from Central Mexico to the Pacific Coast: A pilot study for continental-scale geochemical mapping. *Applied Geochemistry*, 24, pp.1416–1428. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S0883292709001231> [Accessed February 20, 2013].
- Chiprés, J.A., de la Calleja, A., Tellez, J.I., Jiménez, F., Cruz, C., Guerrero, E.G., Castro, J., Monroy, M.G. & Salinas, J.C., 2009b. Geochemistry of soils along a transect from Central Mexico to the Pacific Coast: A pilot study for continental-scale geochemical mapping. *Applied Geochemistry*, 24, pp.1416–1428. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S0883292709001231> [Accessed February 20, 2013].
- Conservative Party, 2015. *The Conservative Party Manifesto 2015*, Available at: <https://s3-eu-west-1.amazonaws.com/manifesto2015/ConservativeManifesto2015.pdf>.
- Crawford, W.H., 1986. *Industries of the North one hundred years ago: Industrial and Commercial Life in the North of Ireland 1888-91* 2nd ed., Belfast: Universities Press Ltd.
- Department for Environment Food and Rural Affairs (Defra) and Welsh Assembly Government (WAG), 2011. *Simplification of the contaminated land regime: Impact assessment*, London.
- Duan, J. & Tan, J., 2013. Atmospheric heavy metals and Arsenic in China: Situation, sources and control policies. *Atmospheric Environment*, 74, pp.93–101. Available at: <http://dx.doi.org/10.1016/j.atmosenv.2013.03.031>.
- ESRI (Environmental Systems Resource Institute), 2010. *ArcMap 10.0*, Redlands, California: ESRI.
- European Environment Agency, 2012. Corine Land Cover 2006 seamless vector data. , p.Version 16 (04/2012). Available at: <http://www.eea.europa.eu/data-and-maps/data/clc-2006-vector-data-version-2#tab-additional-information>.
- Freestone, S.E., O'Donnell, K.E. & Brown, S.E., 2004. Geochemical baseline data for the urban area of Sheffield. *British Geological Survey Internal Report*, IR/02/084, p.71. Available at: <http://nora.nerc.ac.uk/7019/>.
- Fu, W.J., Jiang, P.K., Zhou, G.M. & Zhao, K.L., 2014. Using Moran's I and GIS to study the spatial pattern of forest litter carbon density in a subtropical region of southeastern China. *Biogeosciences*, 11(8), pp.2401–2409.

- Galán, E., Fernández-Caliani, J., González, I., Aparicio, P. & Romero, A., 2008. Influence of geological setting on geochemical baselines of trace elements in soils. Application to soils of South–West Spain. *Journal of Geochemical Exploration*, 98, pp.89–106. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S0375674208000071> [Accessed March 26, 2013].
- Gilbertson, D., Grattan, J., Cressey, M. & Pyatt, F., 1997. An air-pollution history of metallurgical innovation in iron and steel-making: a geochemical archive of Sheffield. *Water, Air and Soil Pollution*1, 100, pp.327–341.
- Goyal, N., Camacho, F., Mangano, J. & Goldenberg, D., 2015. Evaluating for a geospatial relationship between radon levels and thyroid cancer in Pennsylvania. *Laryngoscope*, 125(1), pp.45–49.
- Helbich, M., Leitner, M. & Kapusta, N.D., 2012. Geospatial examination of lithium in drinking water and suicide mortality. *International journal of health geographics*, 11, p.19. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3441892&tool=pmcentrez&rendertype=abstract>.
- Herting, G., Goidanich, S., Odnevall Wallinder, I. & Leygraf, C., 2008. Corrosion-induced release of Cu and Zn into rainwater from brass, bronze and their pure metals. A 2-year field study. *Environmental Monitoring and Assessment*, 144(1-3), pp.455–461.
- Jarva, J., Tarvainen, T., Reinikainen, J. & Eklund, M., 2010. TAPIR - Finnish national geochemical baseline database. *Science of the Total Environment*, 408, pp.4385–4395.
- Johnson, C.C., 2005. 2005 G-BASE Field Procedures Manual. *British Geological Survey Internal Report*, IR/05/097, p.130.
- Johnson, C.C., Ander, E., Cave, M.R. & Palumbo-Roe, B., 2012. Normal background concentrations (NBCs) of contaminants in English soils: Final project report. *British Geological Survey Commissioned Report*, CR/12/035, p.40pp.
- Johnson, C.C. & Ander, E.L., 2008. Urban geochemical mapping studies: how and why we do them. *Environmental geochemistry and health*, 30, pp.511–530. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/18563589> [Accessed February 11, 2014].
- Land and Property Services, 1858. *Ordnance Survey 6 Inch County Maps*, © Crown copyright and database rights MOU203.
- Land and Property Services, 1901. *Ordnance Survey 6 Inch County Maps*, © Crown copyright and database rights MOU203.
- Land and Property Services, 1919. *Ordnance Survey 6 Inch County Maps*, © Crown copyright and database rights MOU203.

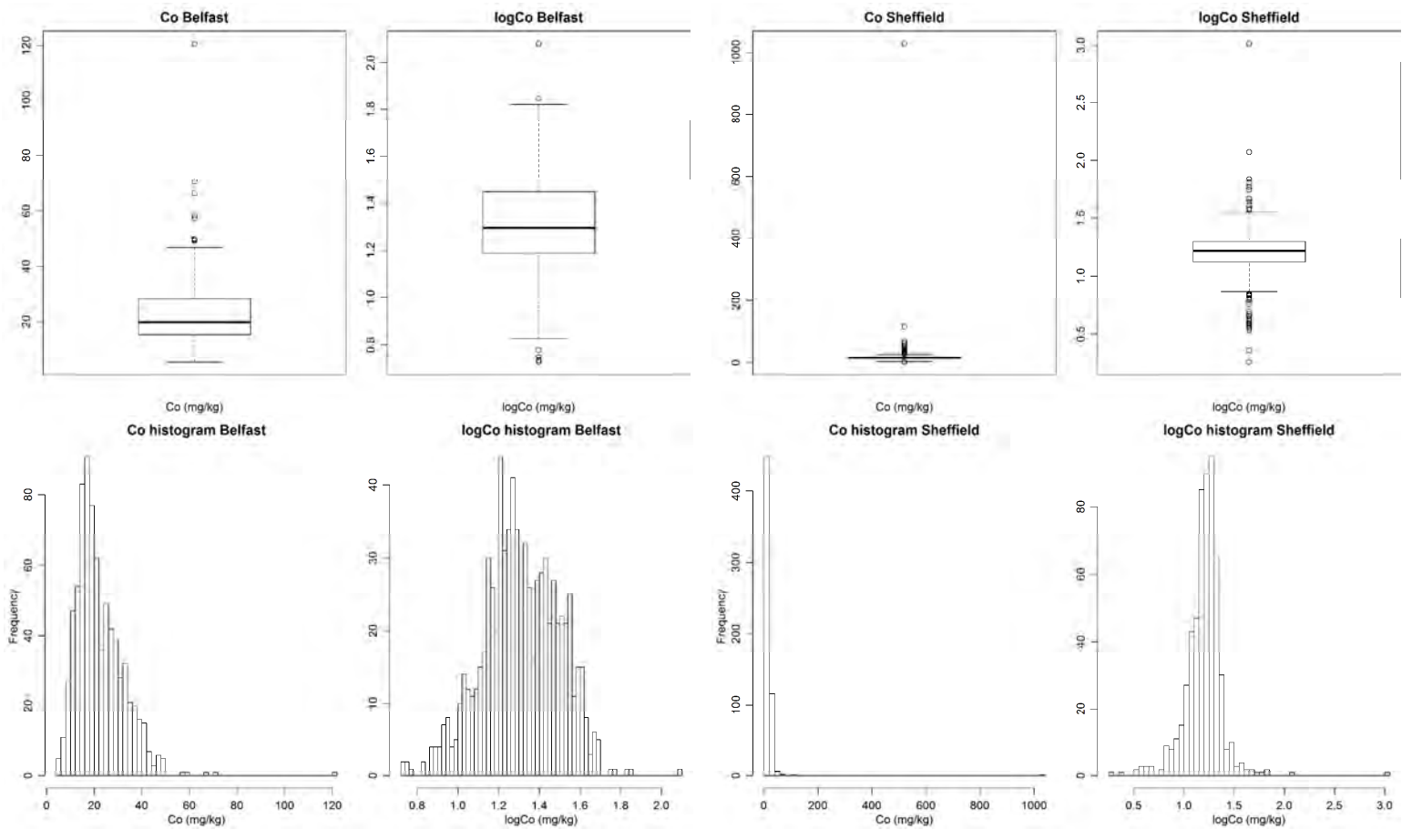
- Liu, Y., Li, X., Wang, W., Li, Z., Hou, M., He, Y., Wu, W., Wang, H., Liang, H. & Guo, X., 2012. Investigation of space-time clusters and geospatial hot spots for the occurrence of tuberculosis in Beijing. *International Journal of Tuberculosis and Lung Disease*, 16(4), pp.486–491.
- McIlwaine, R., Cox, S.F., Doherty, R., Palmer, S., Ofterdinger, U. & McKinley, J., 2014. Comparison of methods used to calculate typical threshold values for potentially toxic elements in soil. *Environmental Geochemistry and Health*, 36(5), pp.953–971.
- Mitchell, W. ed., 2004. *The Geology of Northern Ireland* Second Edi., Belfast: Geological Survey of Northern Ireland.
- Niskavaara, H., Reimann, C., Chekushin, V. & Kashulina, G., 1997. Seasonal variability of total and easily leachable element contents in topsoils (0-5 cm) from eight catchments in the European Arctic (Finland, Norway and Russia). *Environmental pollution*, 96(2), pp.261–74. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/15093425>.
- Northern Ireland Statistics and Research Agency, 2014. Northern Ireland Census of Employment 2013. *Labour Market Statistics Bulletin*, (July).
- Rawlins, B.G., Lark, R.M., O'Donnell, K.E., Tye, A.M. & Lister, T.R., 2005. The assessment of point and diffuse metal pollution of soils from an urban geochemical survey of Sheffield, England. *Soil Use and Management*, 21, pp.353–362. Available at: <http://doi.wiley.com/10.1079/SUM2005335> [Accessed October 19, 2012].
- Reimann, C., Arnoldussen, A., Englmaier, P., Filzmoser, P., Finne, T.E., Garrett, R.G., Koller, F. & Nordgulen, Ø., 2007. Element concentrations and variations along a 120-km transect in southern Norway – Anthropogenic vs. geogenic vs. biogenic element sources and cycles. *Applied Geochemistry*, 22, pp.851–871. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S0883292707000170> [Accessed February 21, 2013].
- Reimann, C., Filzmoser, P. & Garrett, R.G., 2005. Background and threshold: critical comparison of methods of determination. *The Science of the total environment*, 346, pp.1–16. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/15993678> [Accessed March 12, 2012].
- Reimann, C., Filzmoser, P., Garrett, R.G. & Dutter, R., 2008. *Statistical Data Analysis Explained: Applied Environmental Statistics with R*, Chichester: John Wiley & Sons Ltd.
- Rothwell, K.A. & Cooke, M.P., 2015. A comparison of methods used to calculate normal background concentrations of potentially toxic elements for urban soil. *Science of The Total Environment*, 532, pp.625–634. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S0048969715302783>.

- Royle, S.A., 2007. *Belfast Part II, 1840 to 1900* A. Simms, H. B. Clarke, & R. Gillespie, eds., Royal Irish Academy in association with Belfast City Council: Dublin University Press.
- The Coal Authority, 2015. *Sheffield District (B) - Specific Risk*, Available at: <https://www.gov.uk/government/publications/coalfield-plans-sheffield-city-council-area>.
- Tian, H., Zhao, D., Cheng, K., Lu, L., He, M. & Hao, J., 2012. Anthropogenic atmospheric emissions of antimony and its spatial distribution characteristics in China. *Environmental Science & Technology*, 46, pp.3973–3980. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/22397522>.
- Ward, J.H., 1963. Hierarchical Grouping to Optimize an Objective Function. *Journal of the American Statistical Association*, 58(301), pp.236–244.
- Young, M. & Donald, A. eds., 2013. *A guide to the Tellus data*, Belfast: Geological Survey of Northern Ireland.
- Yunker, M.B., Perreault, A. & Lowe, C.J., 2012. Source apportionment of elevated PAH concentrations in sediments near deep marine outfalls in Esquimalt and Victoria, BC, Canada: Is coal from an 1891 shipwreck the source? *Organic Geochemistry*, 46, pp.12–37. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S0146638012000071> [Accessed October 3, 2012].

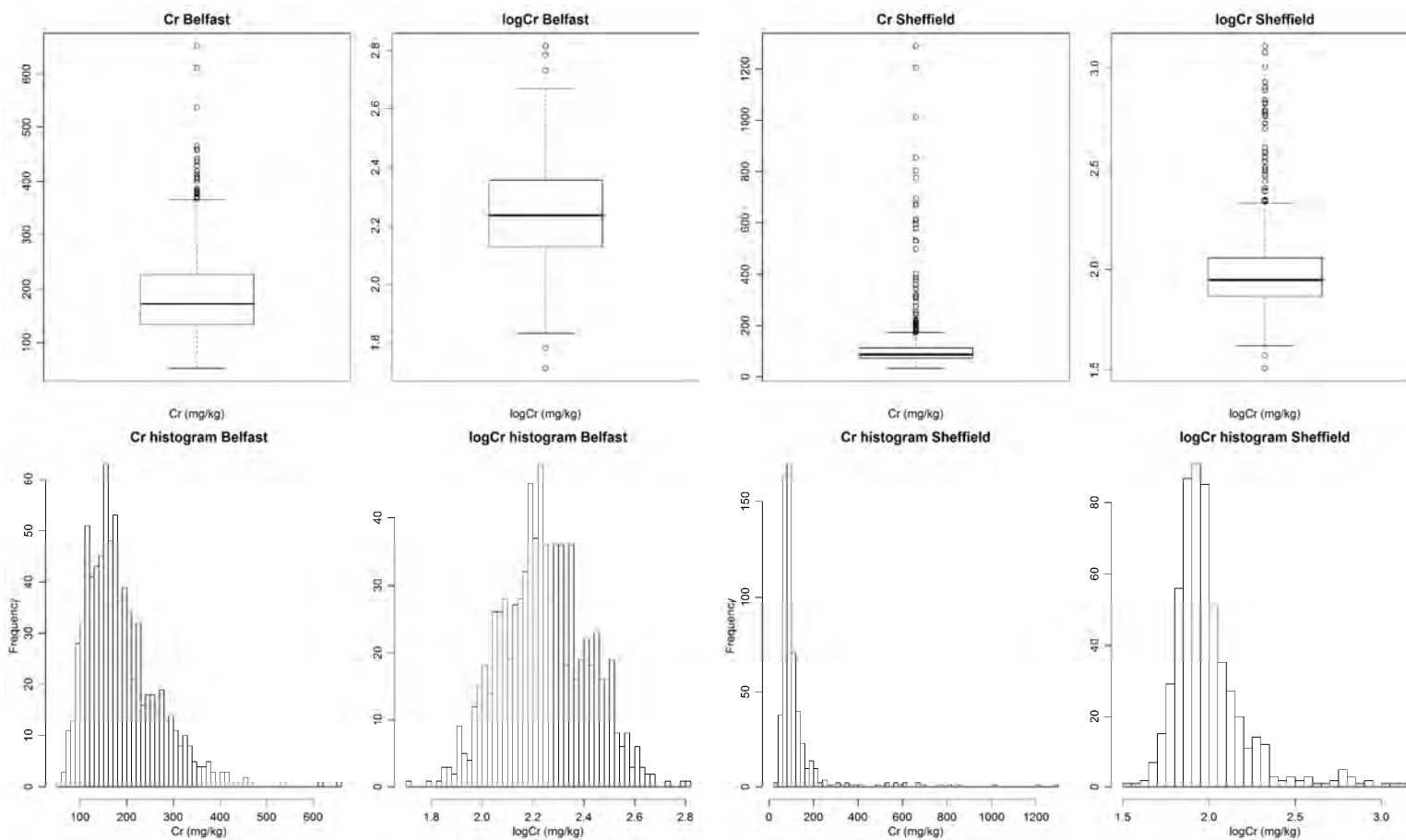
Appendix A – Boxplots and histograms



Appendix A – Boxplots and histograms

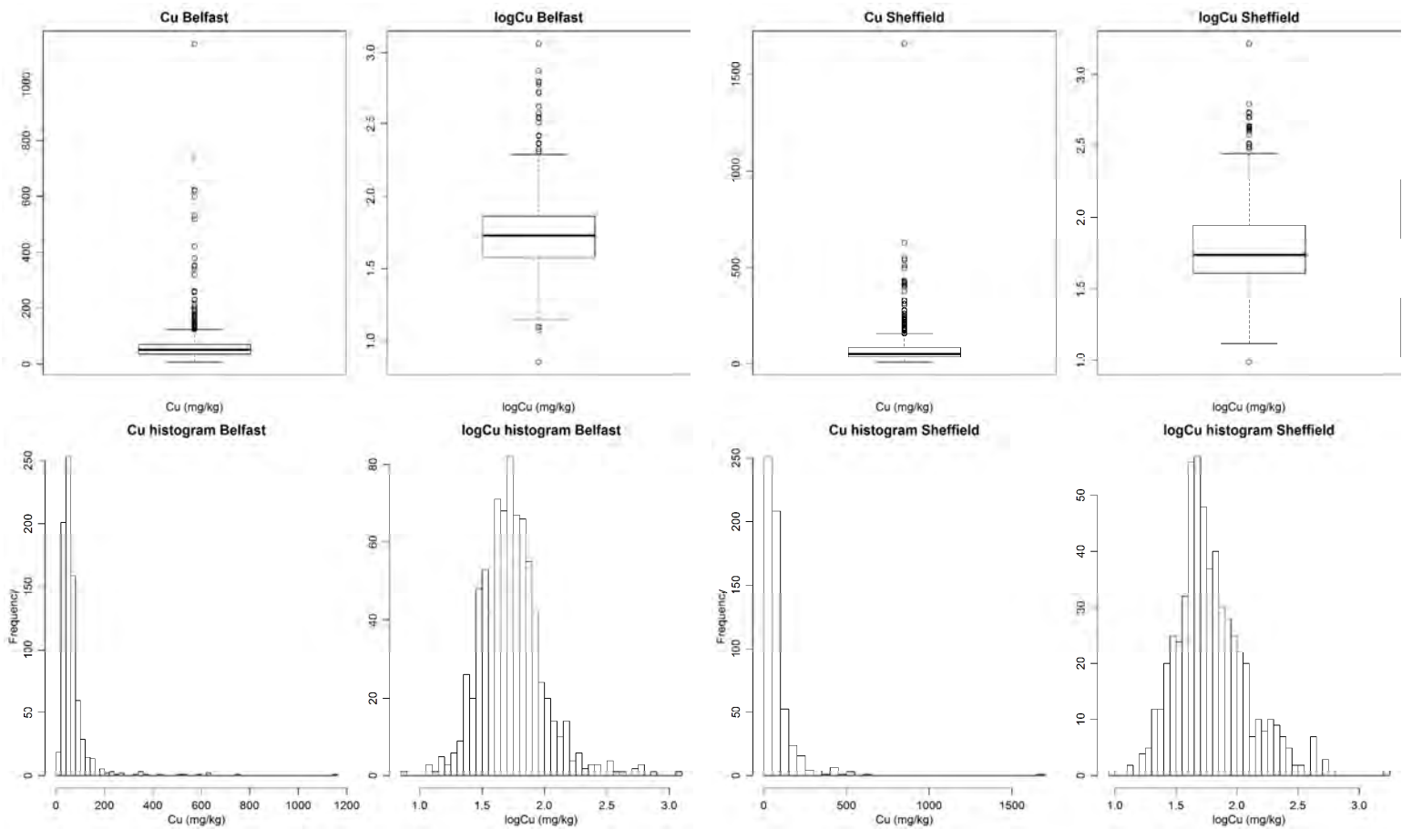


Appendix A – Boxplots and histograms

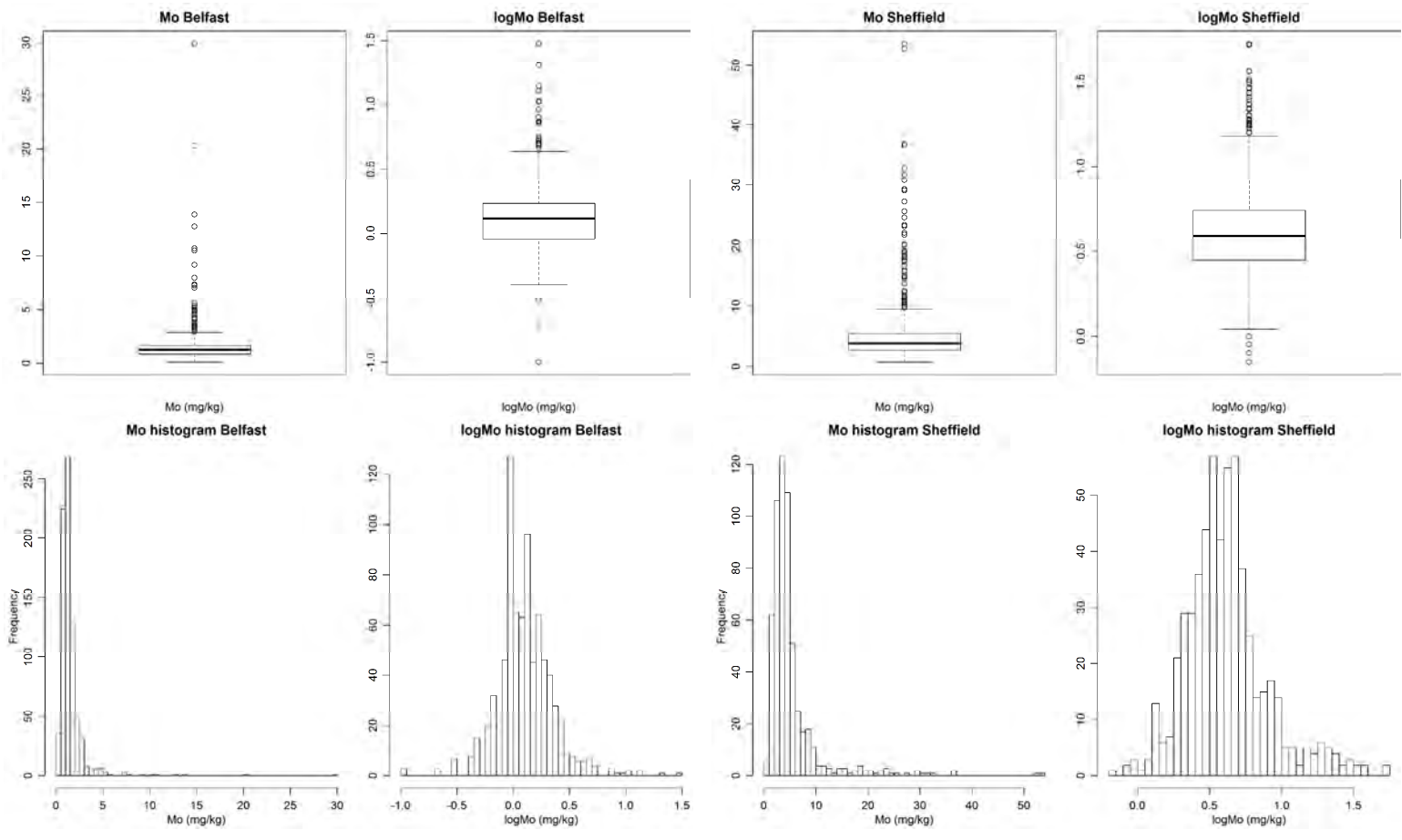


Historical development and its effect on soil contamination in urban areas

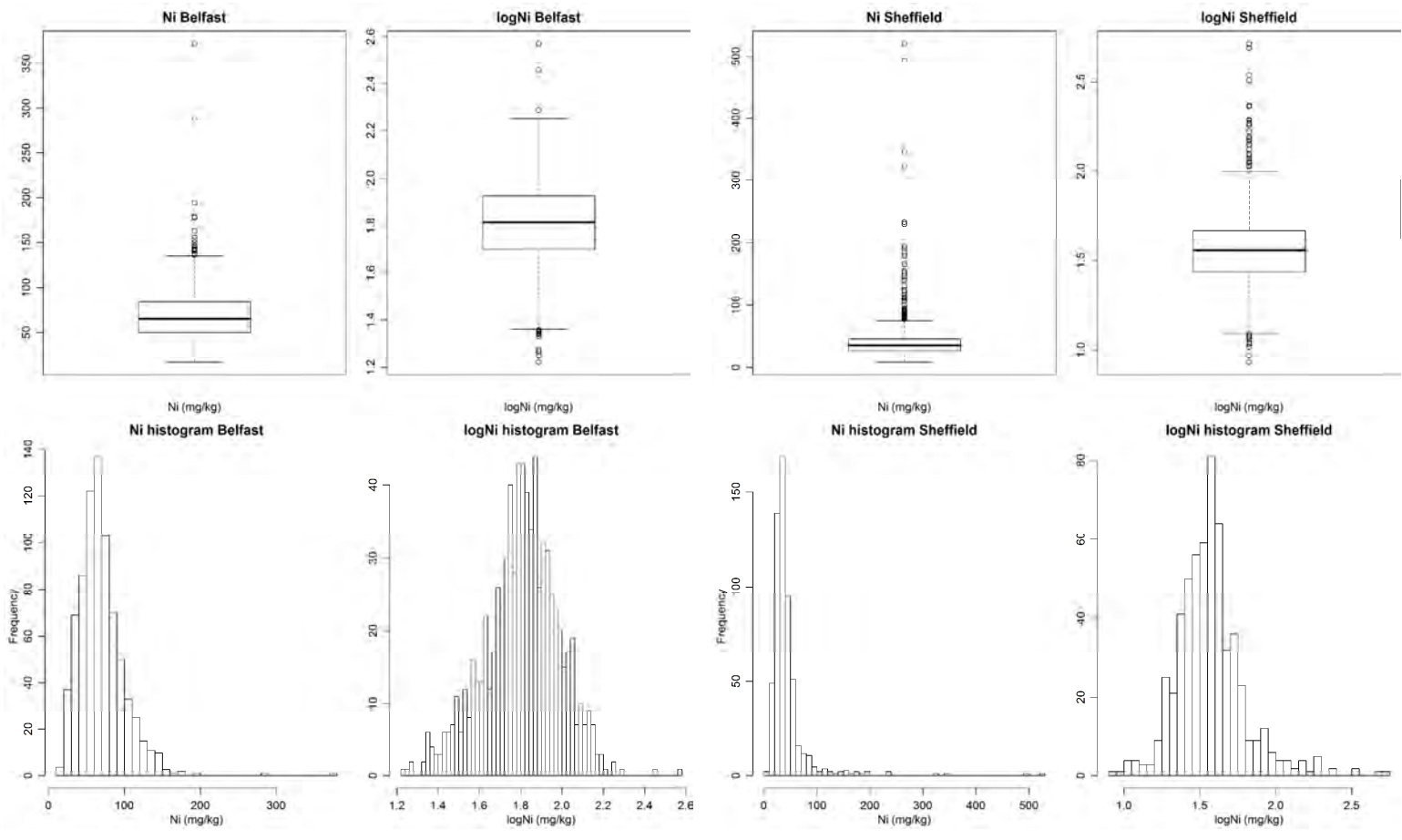
Appendix A – Boxplots and histograms



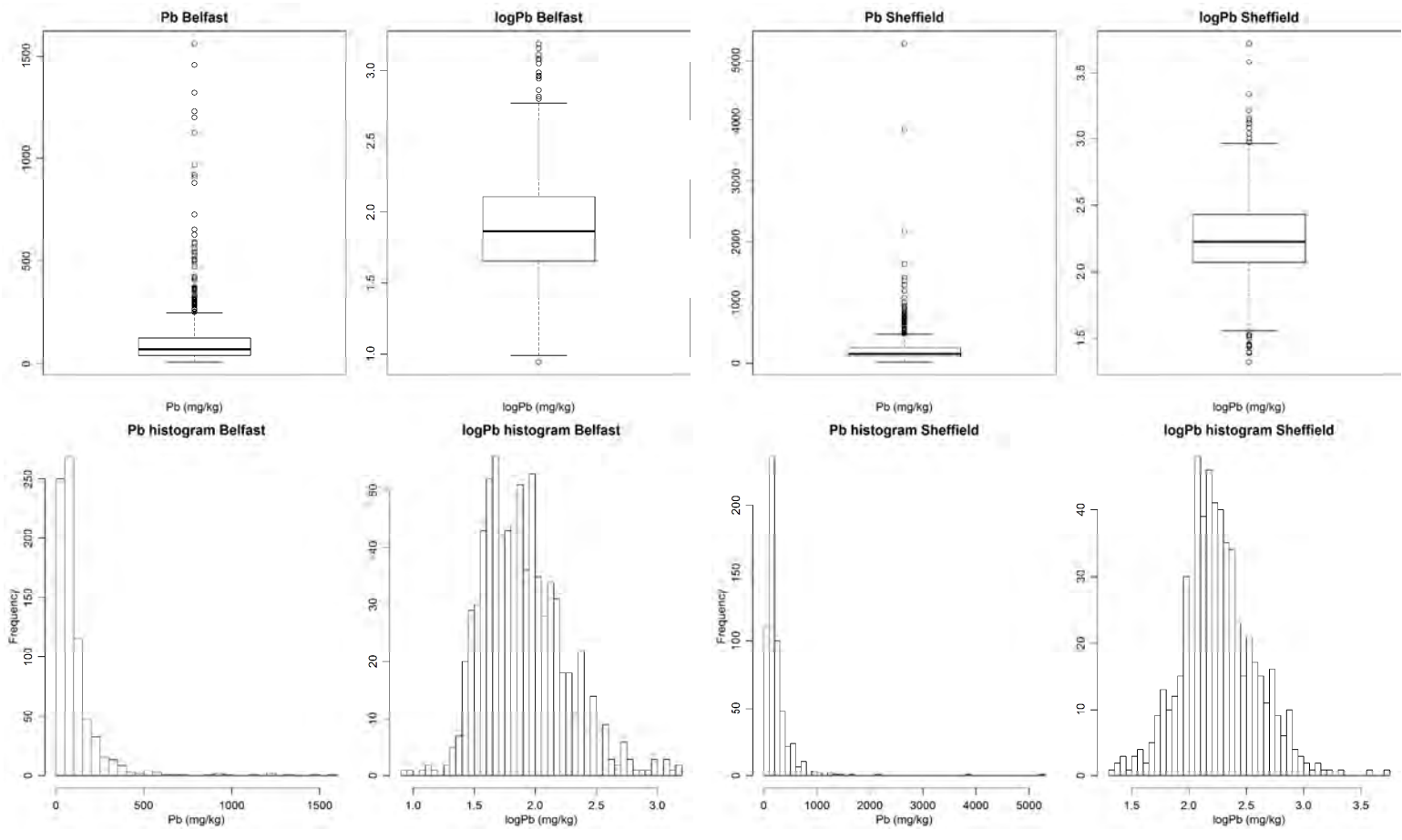
Appendix A – Boxplots and histograms



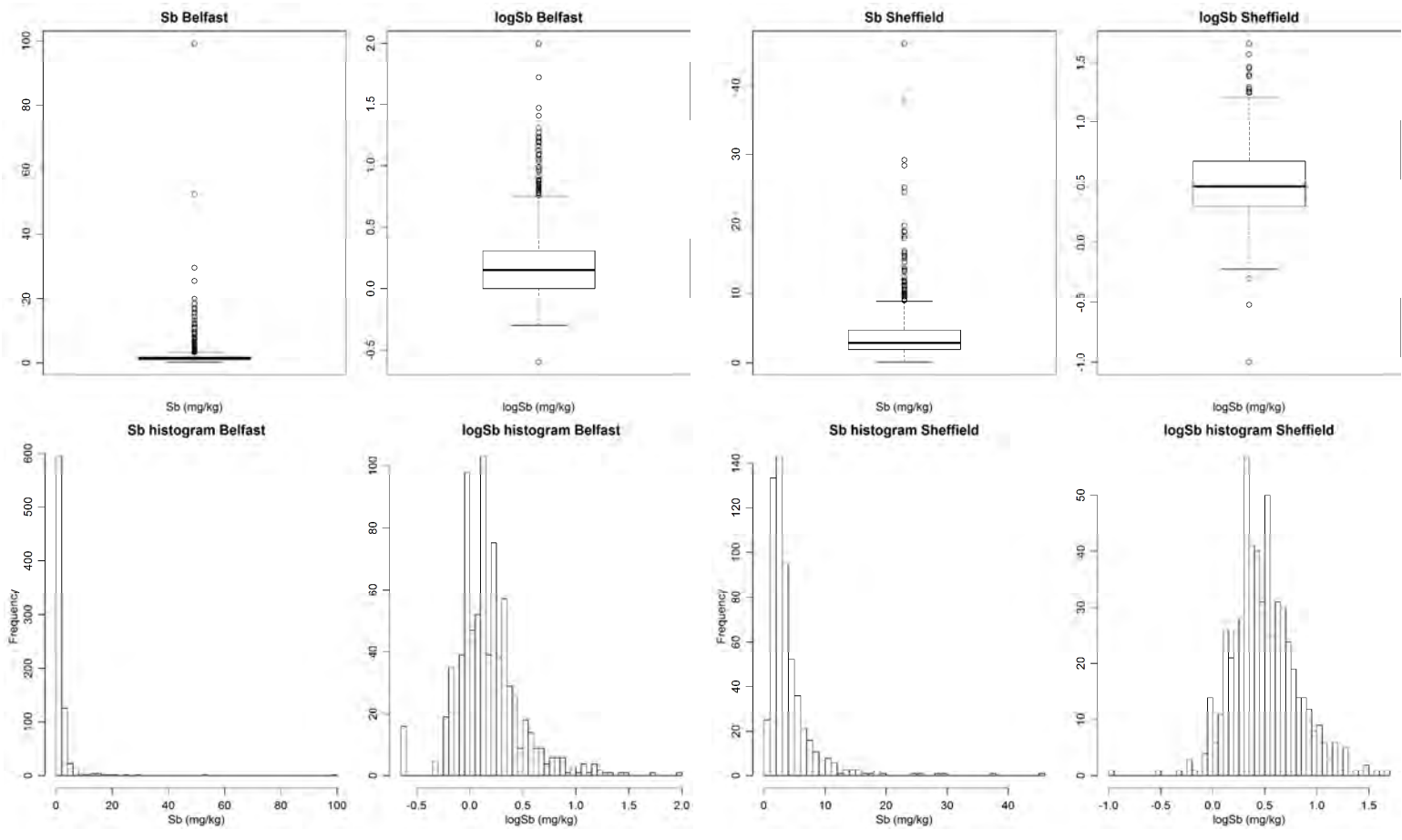
Appendix A – Boxplots and histograms



Appendix A – Boxplots and histograms

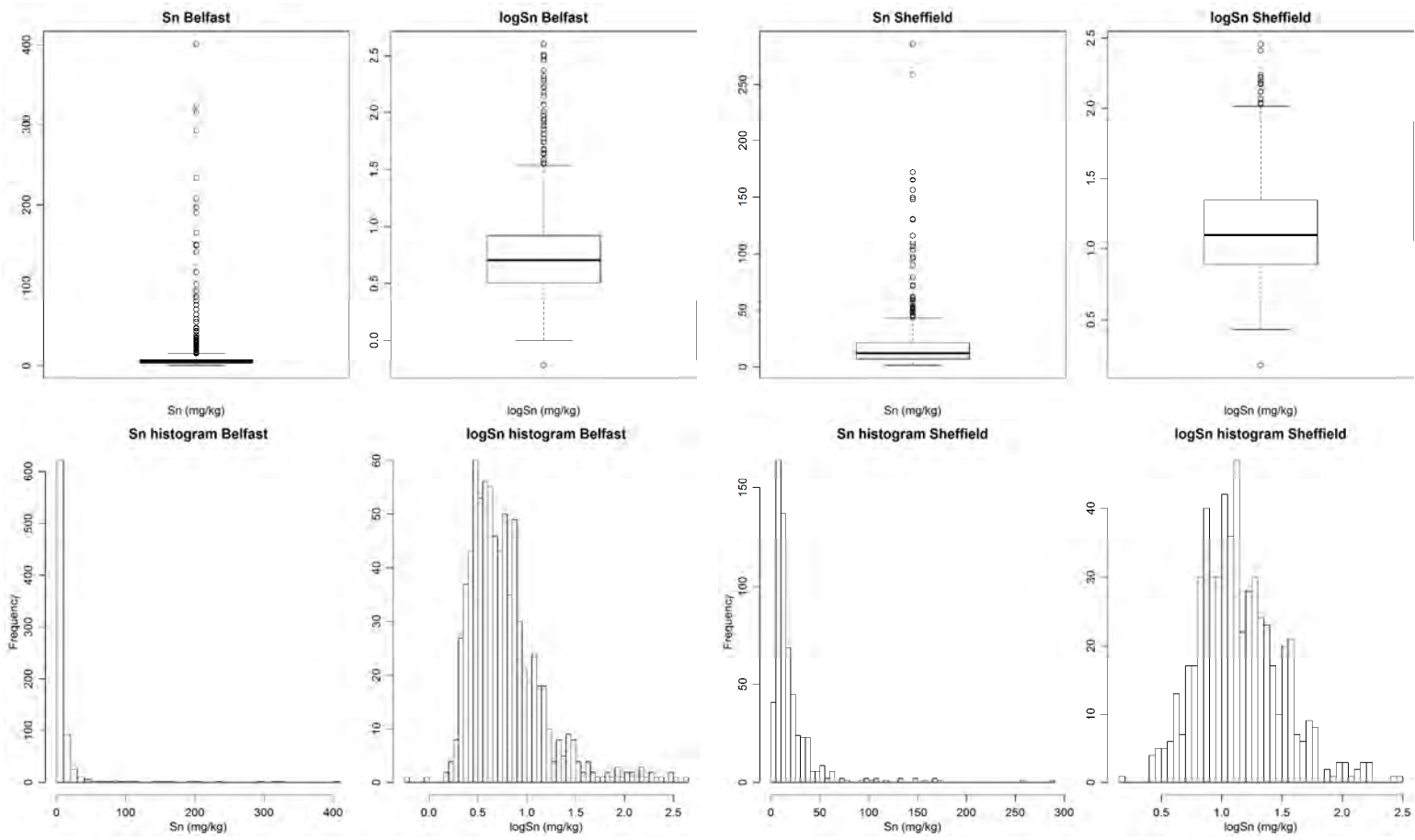


Appendix A – Boxplots and histograms

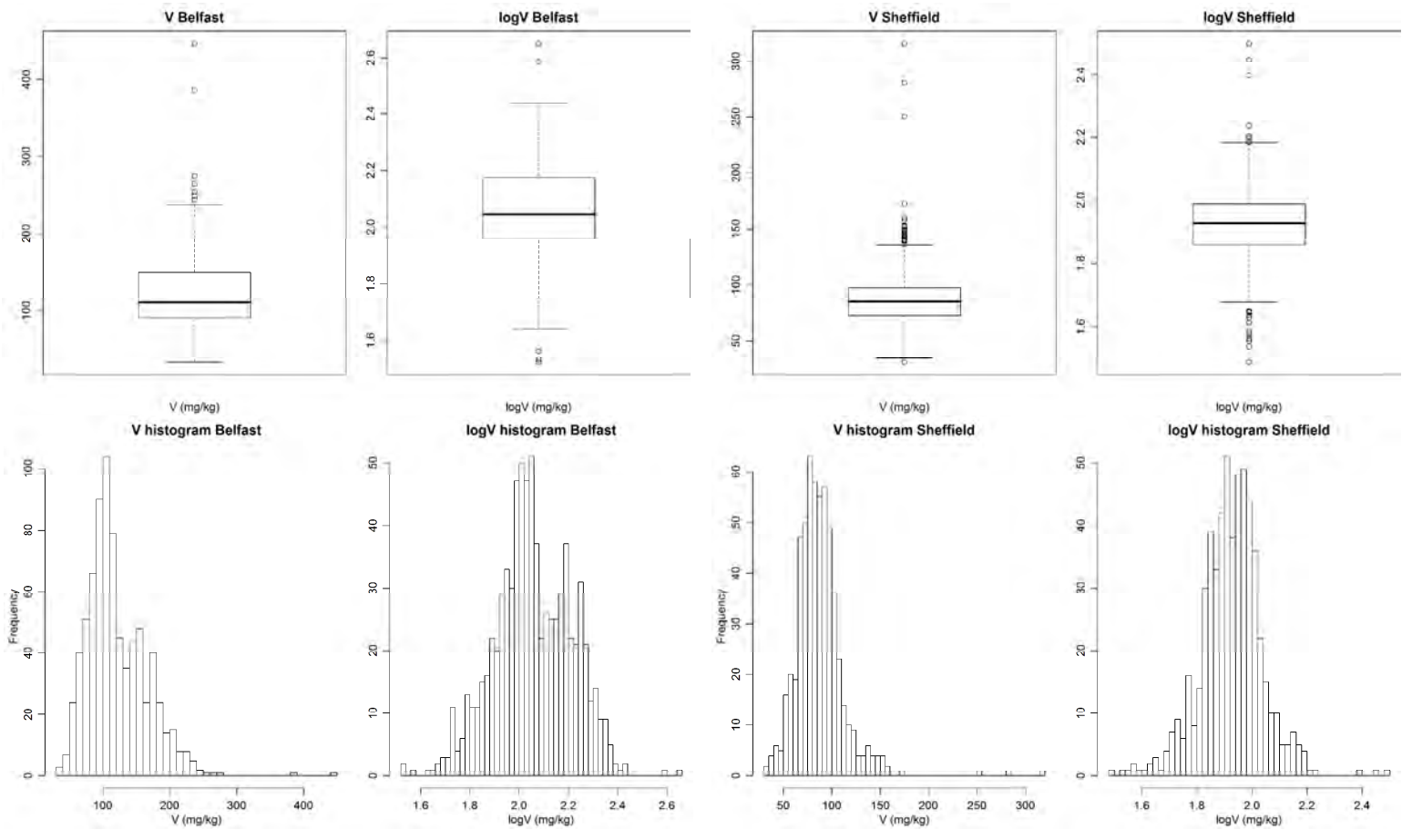


Historical development and its effect on soil contamination in urban areas

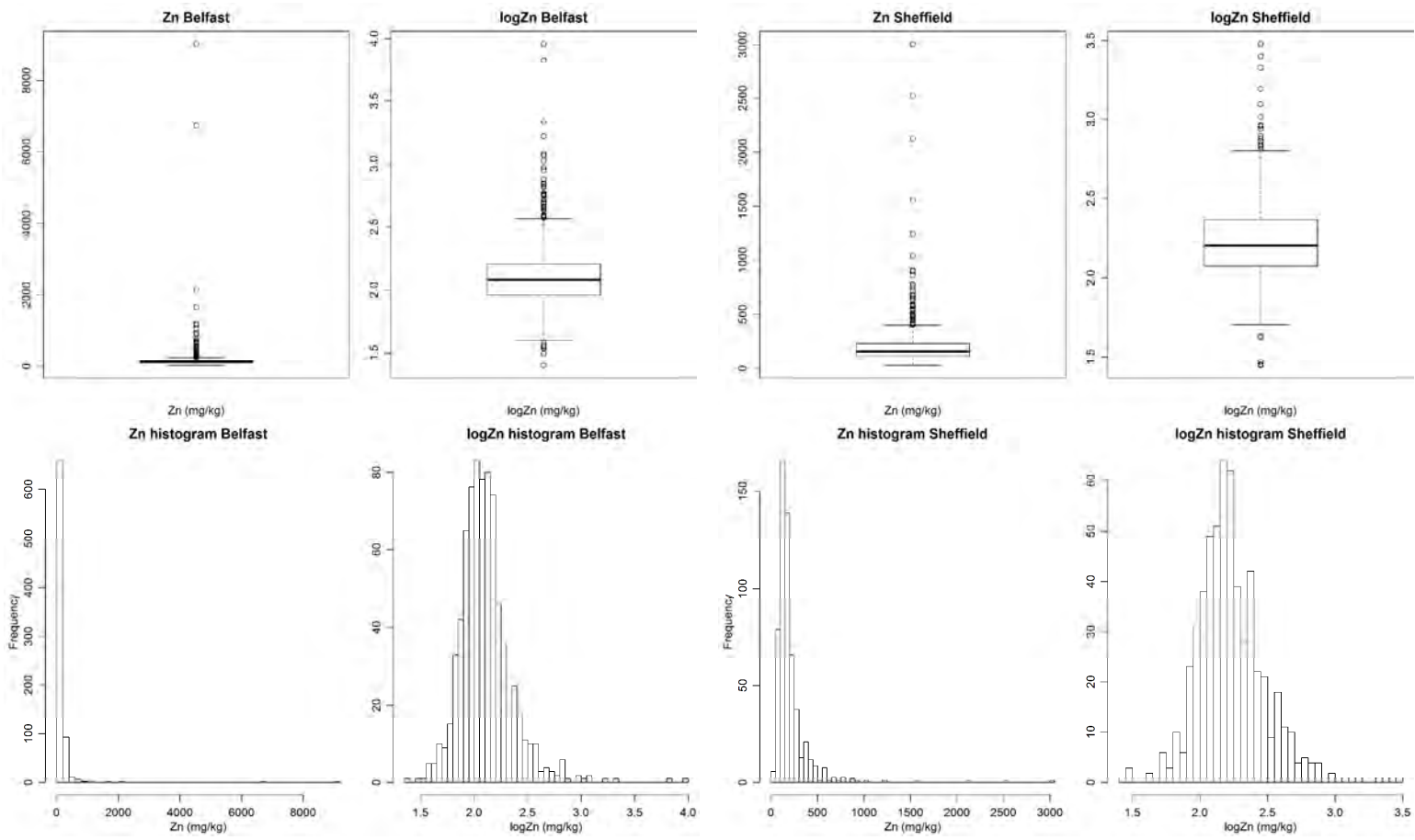
Appendix A – Boxplots and histograms



Appendix A – Boxplots and histograms

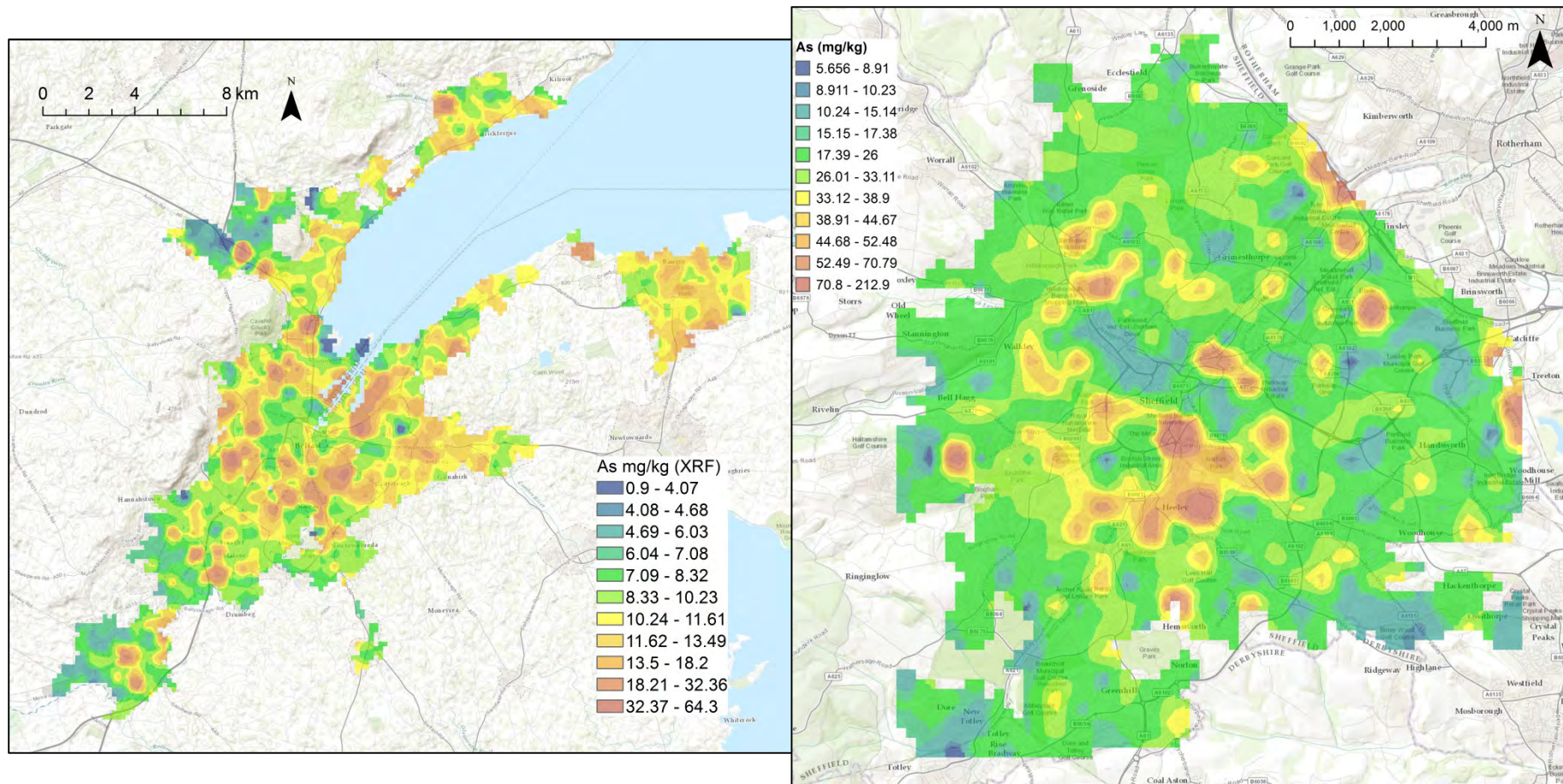


Appendix A – Boxplots and histograms



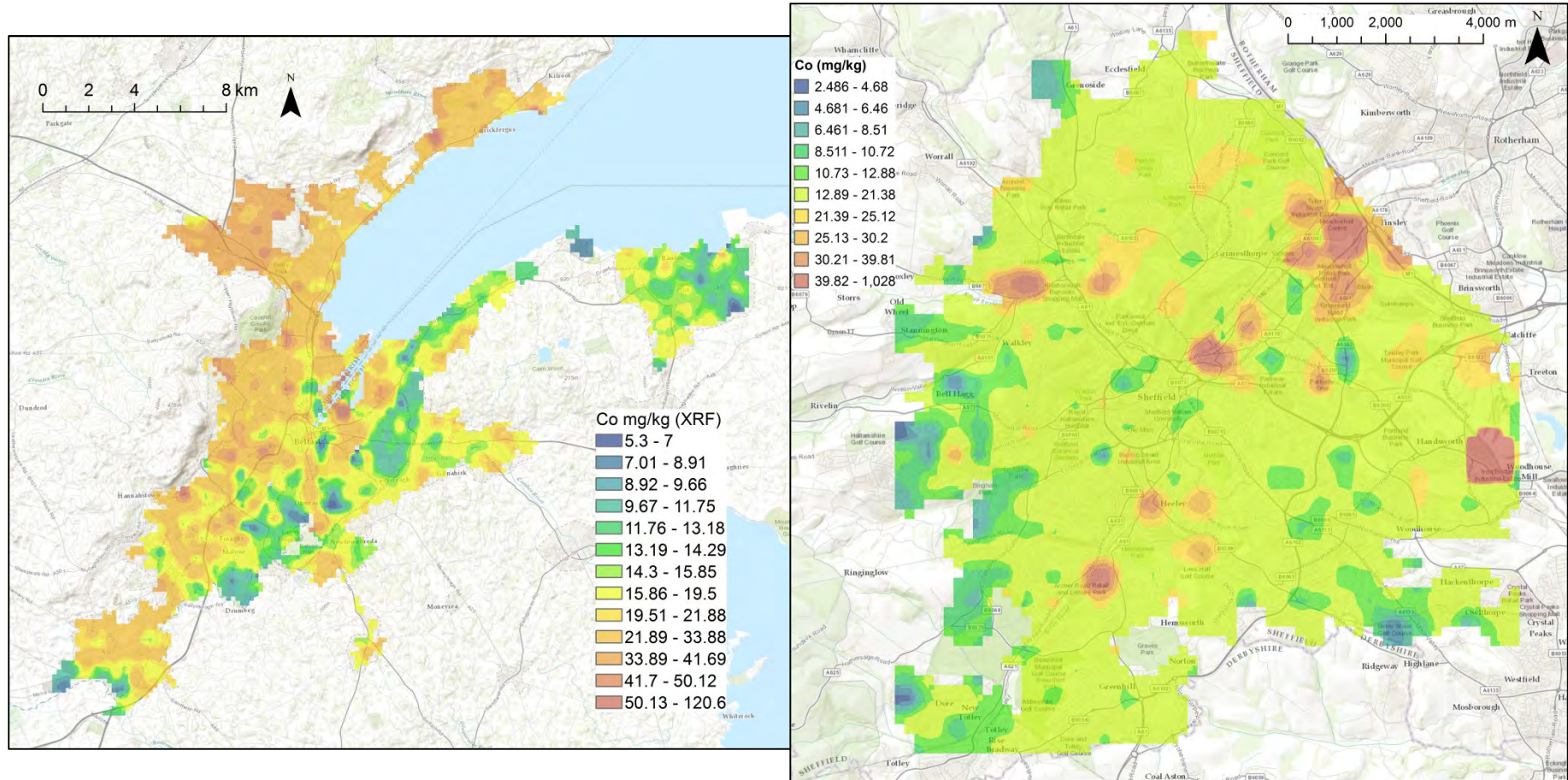
Historical development and its effect on soil contamination in urban areas

Appendix B – PTE concentration maps



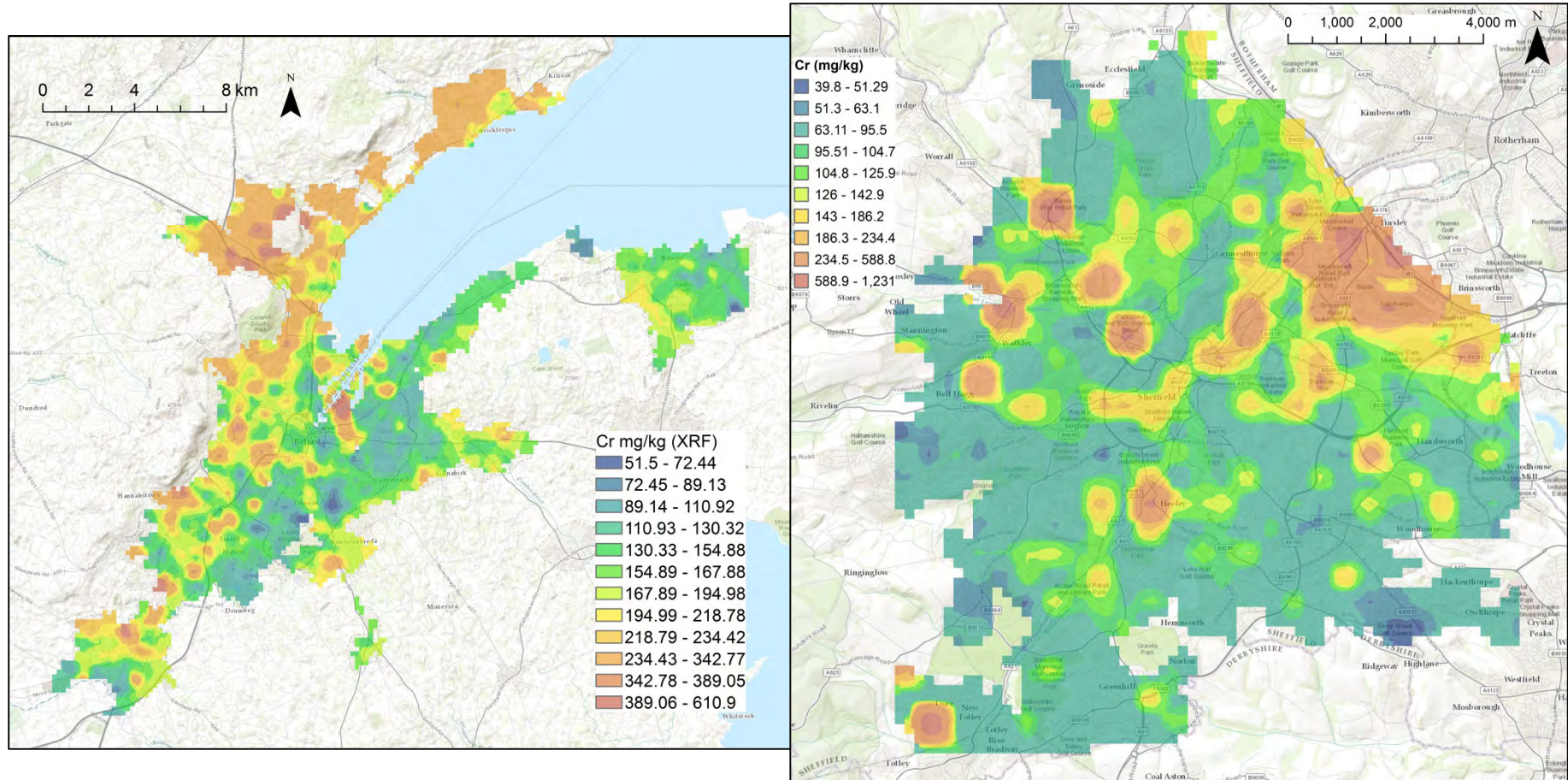
Historical development and its effect on soil contamination in urban areas

Appendix B – PTE concentration maps



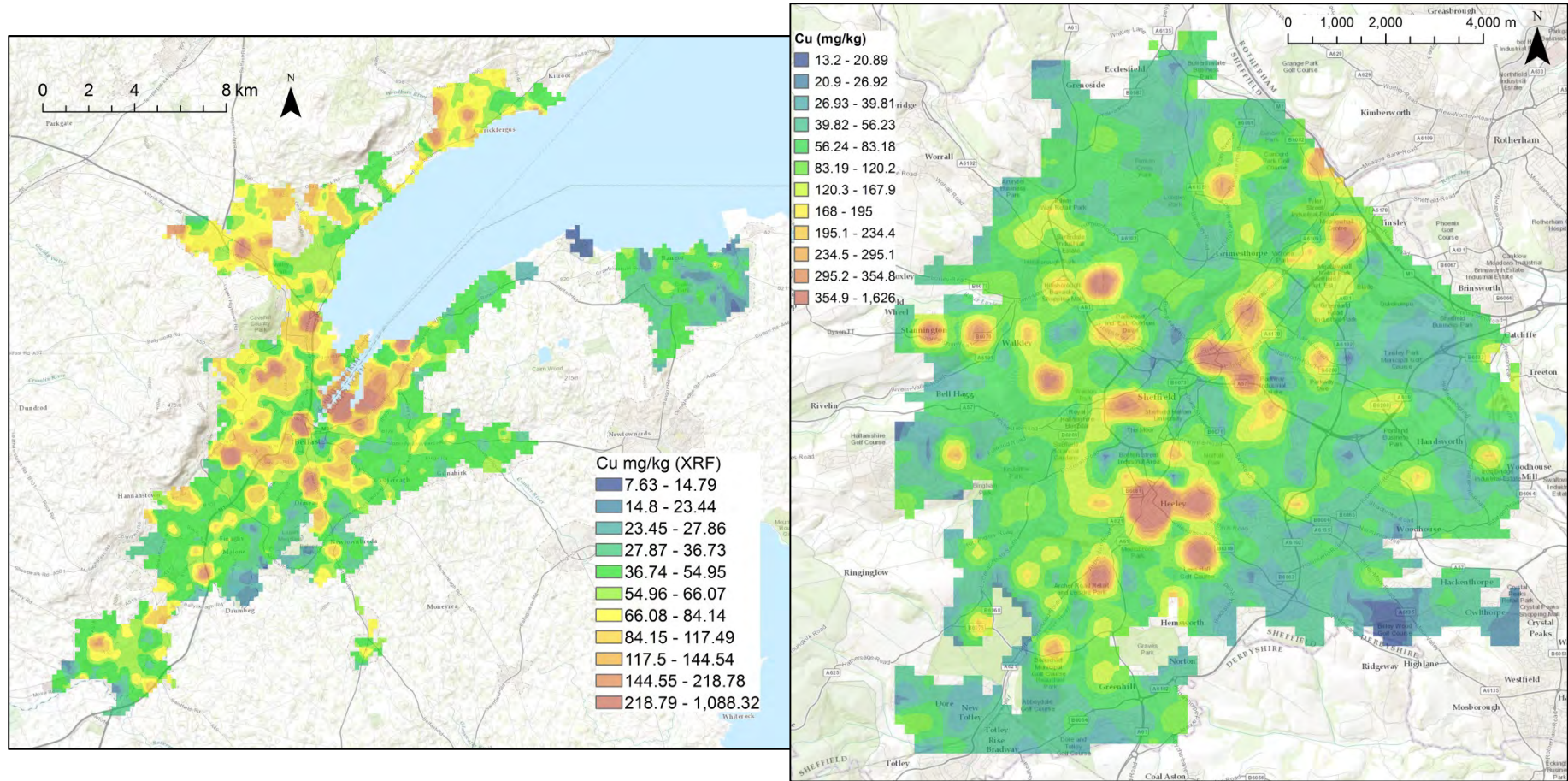
Historical development and its effect on soil contamination in urban areas

Appendix B – PTE concentration maps



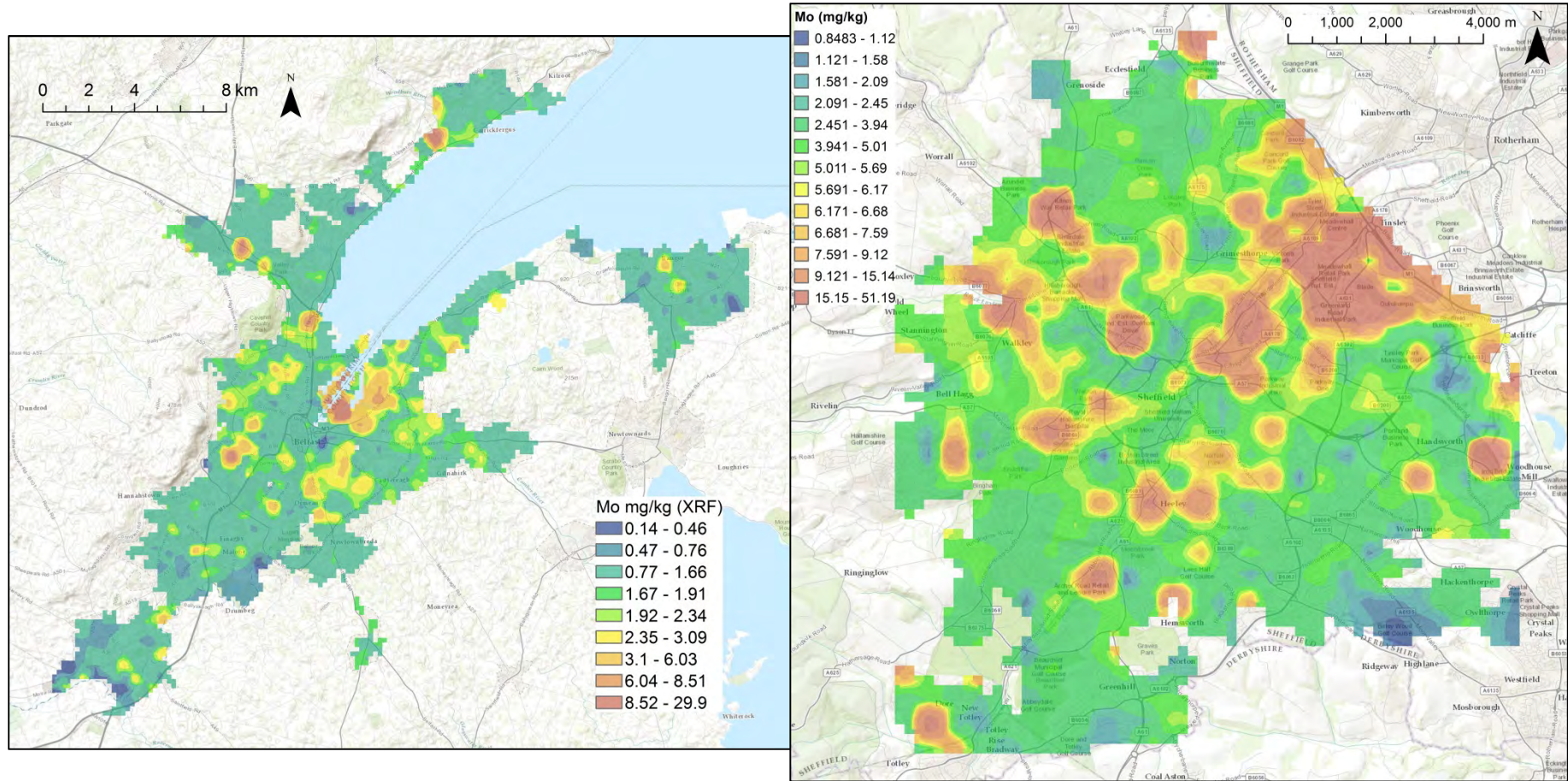
Historical development and its effect on soil contamination in urban areas

Appendix B – PTE concentration maps



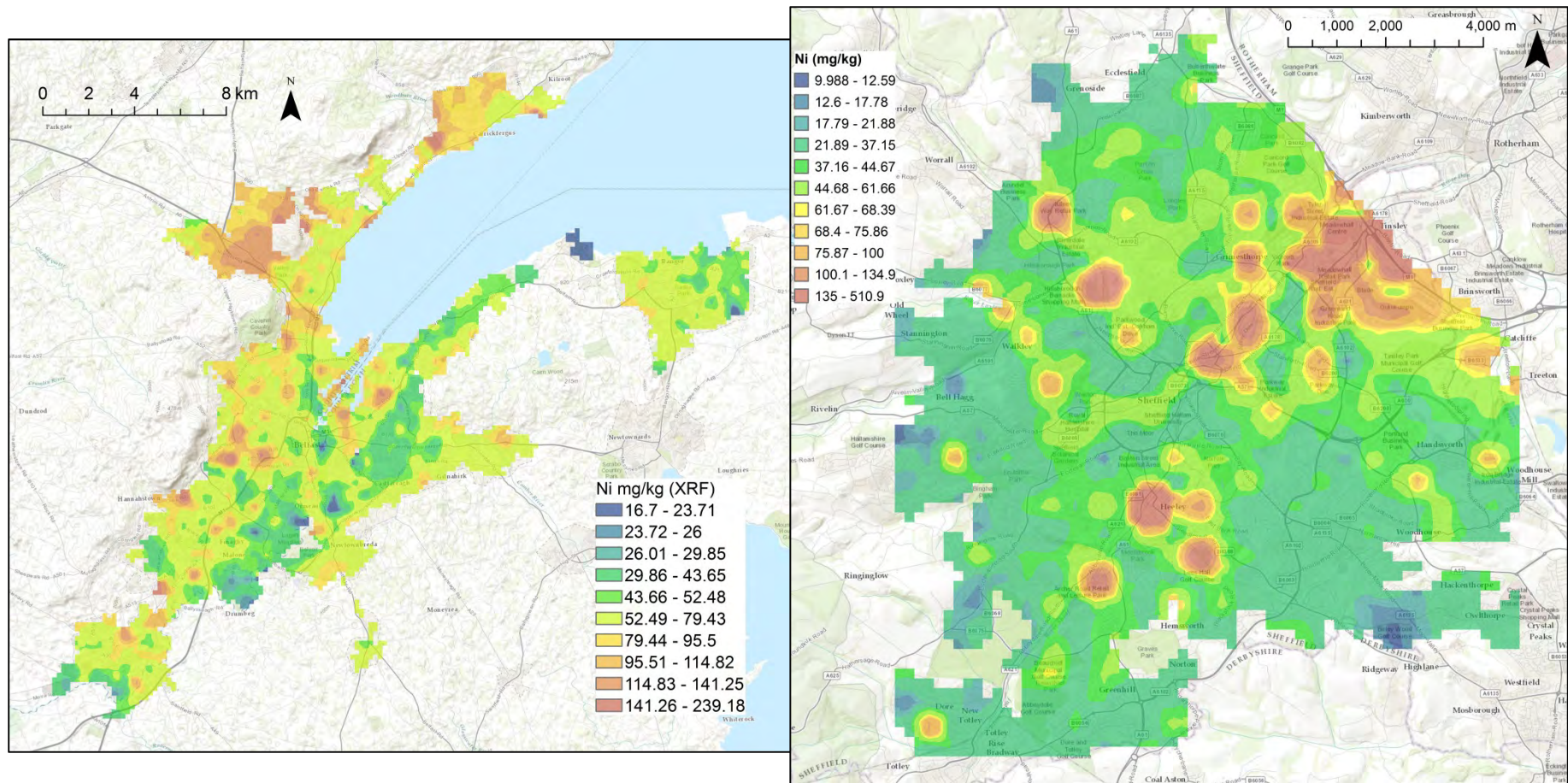
Historical development and its effect on soil contamination in urban areas

Appendix B – PTE concentration maps



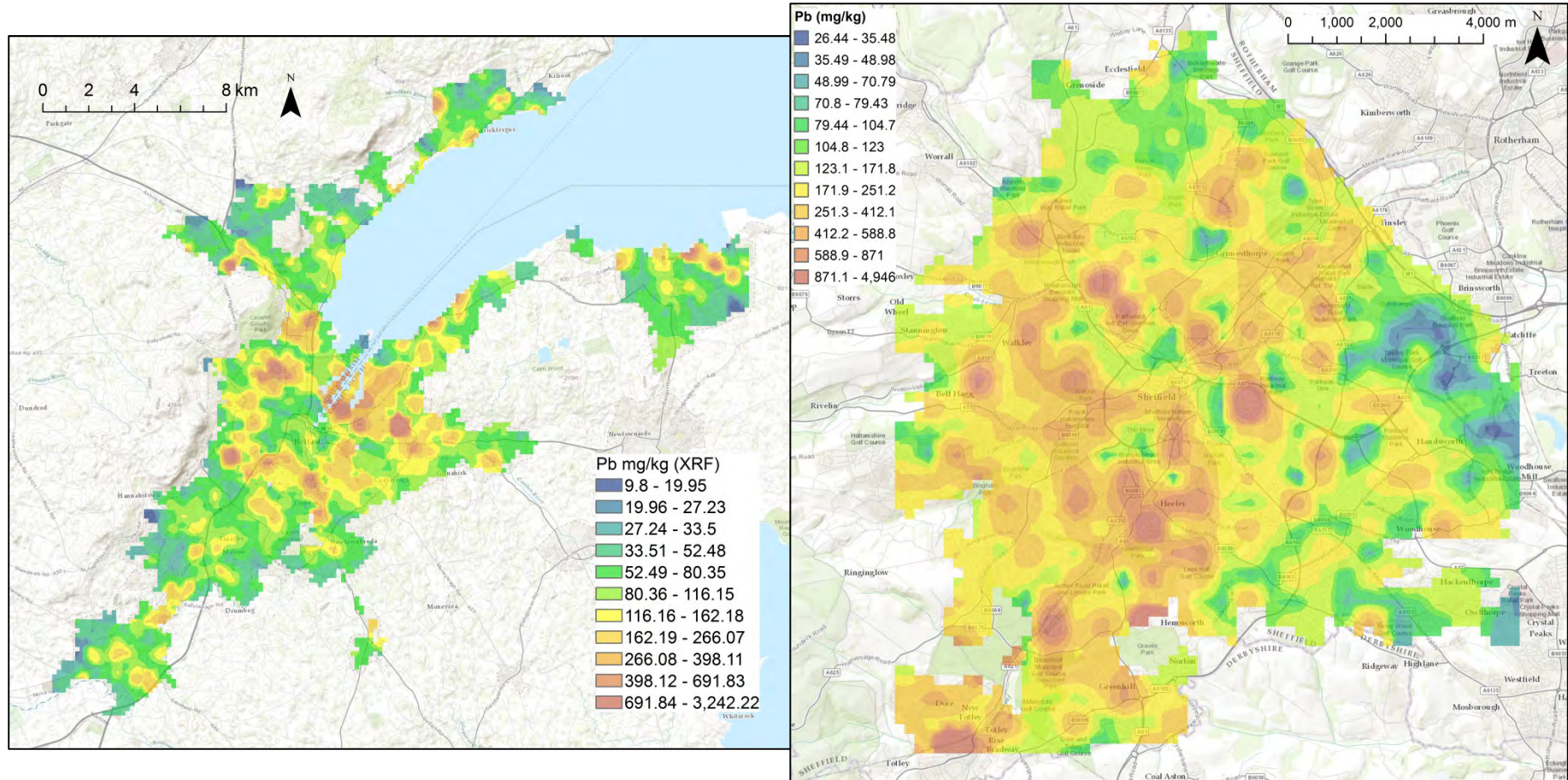
Historical development and its effect on soil contamination in urban areas

Appendix B – PTE concentration maps



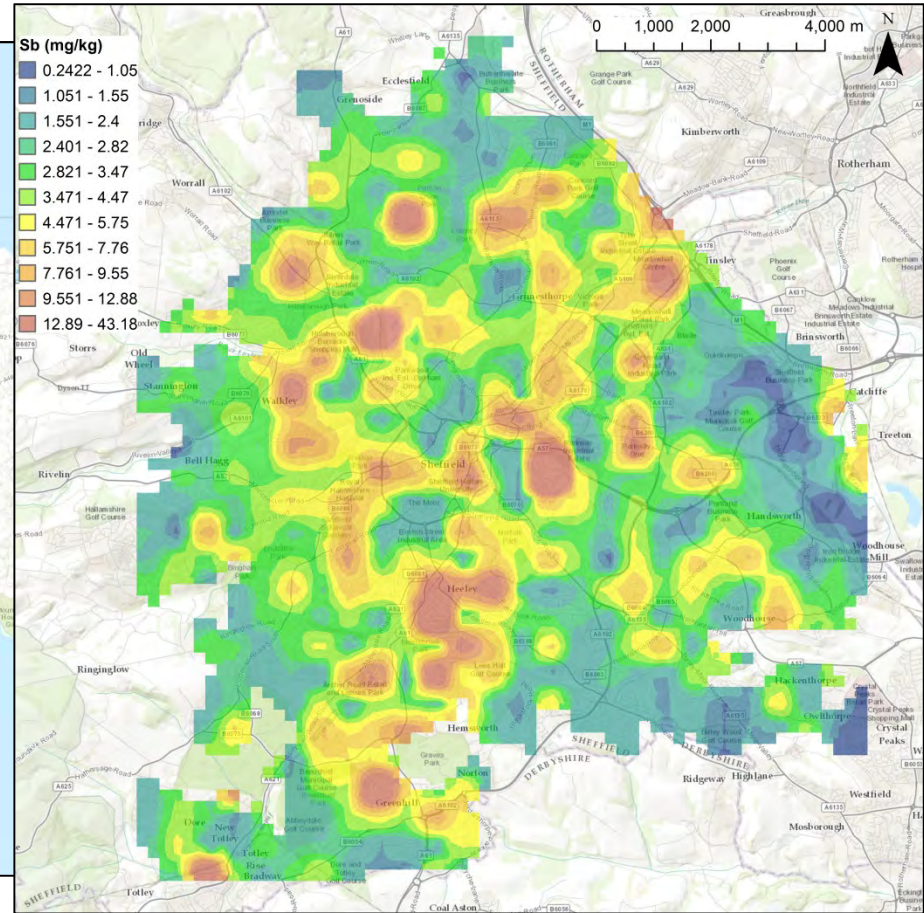
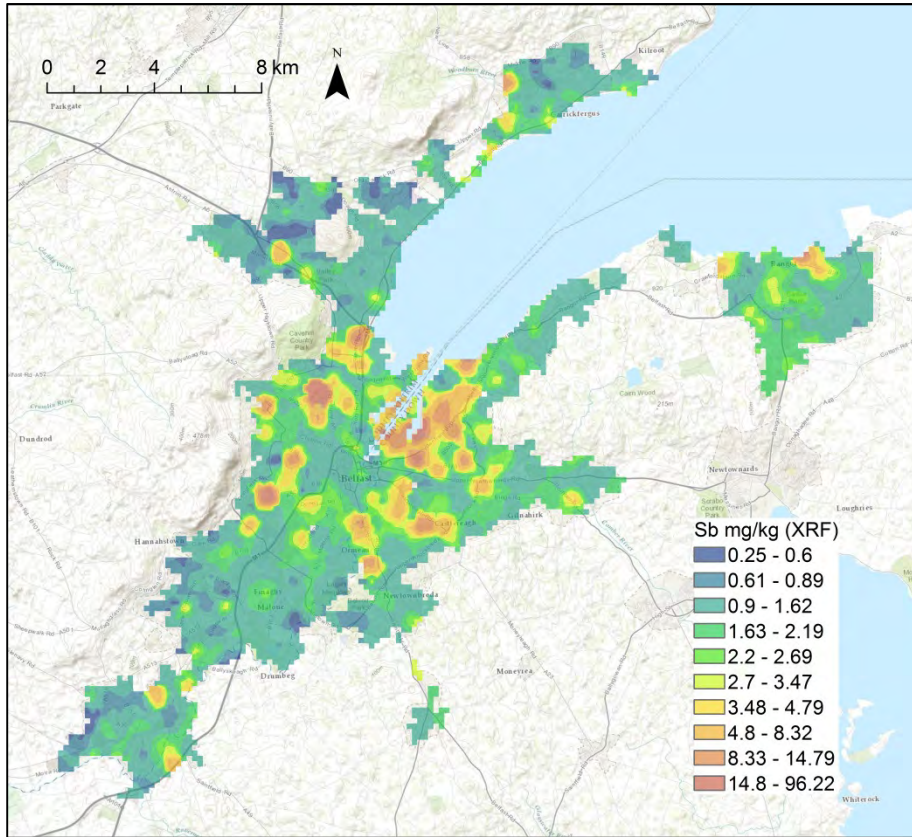
Historical development and its effect on soil contamination in urban areas

Appendix B – PTE concentration maps



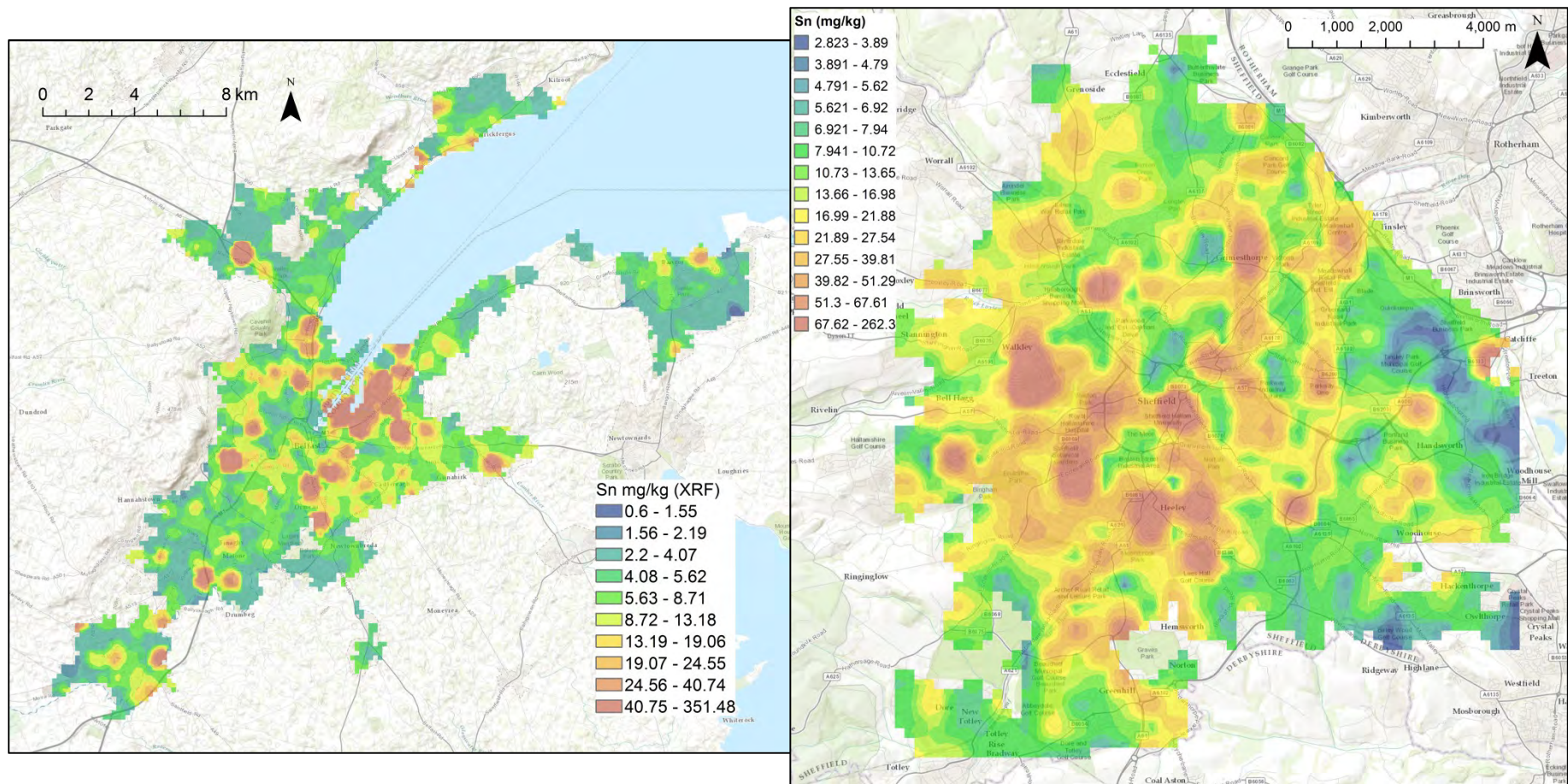
Historical development and its effect on soil contamination in urban areas

Appendix B – PTE concentration maps



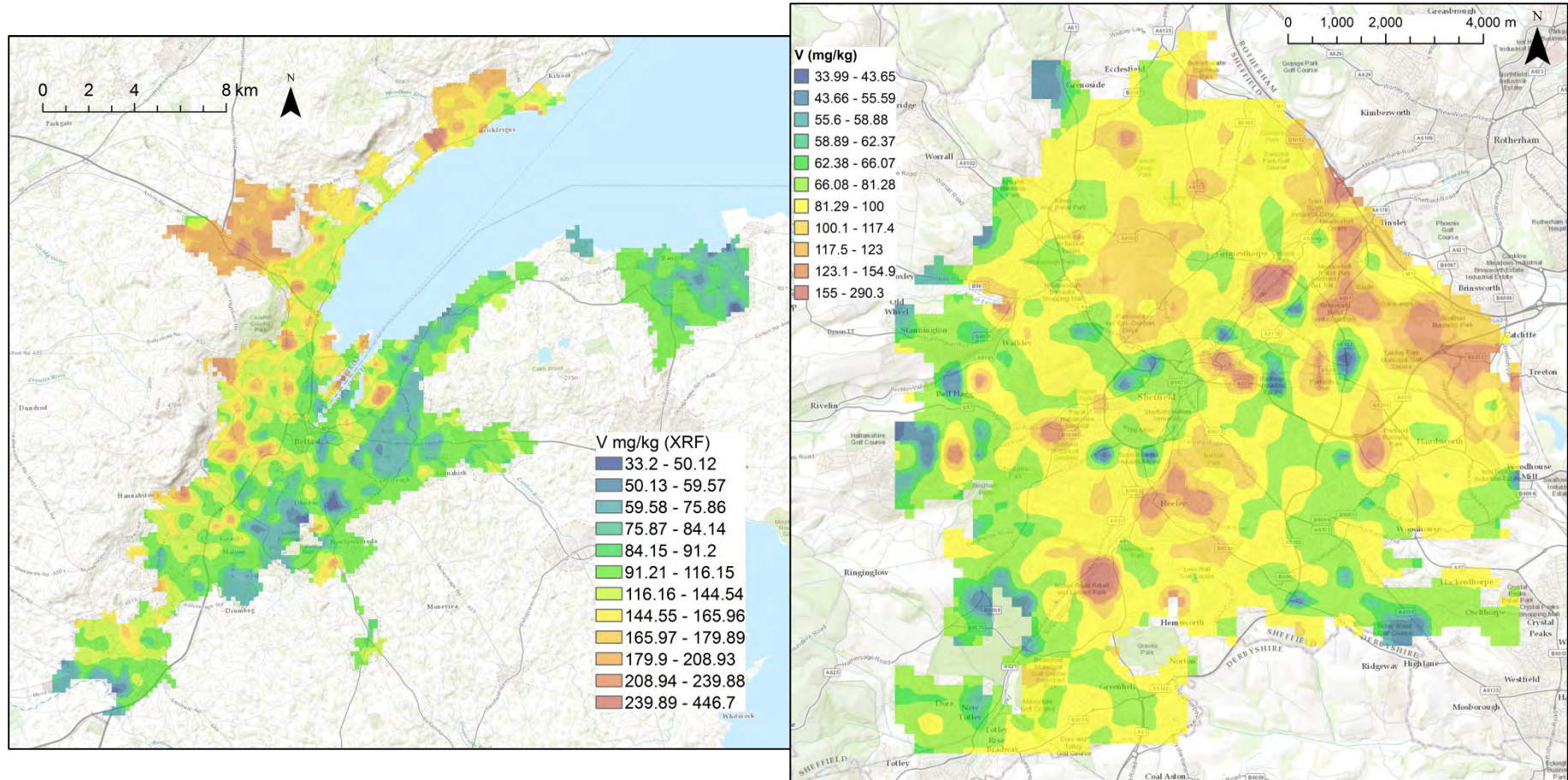
Historical development and its effect on soil contamination in urban areas

Appendix B – PTE concentration maps



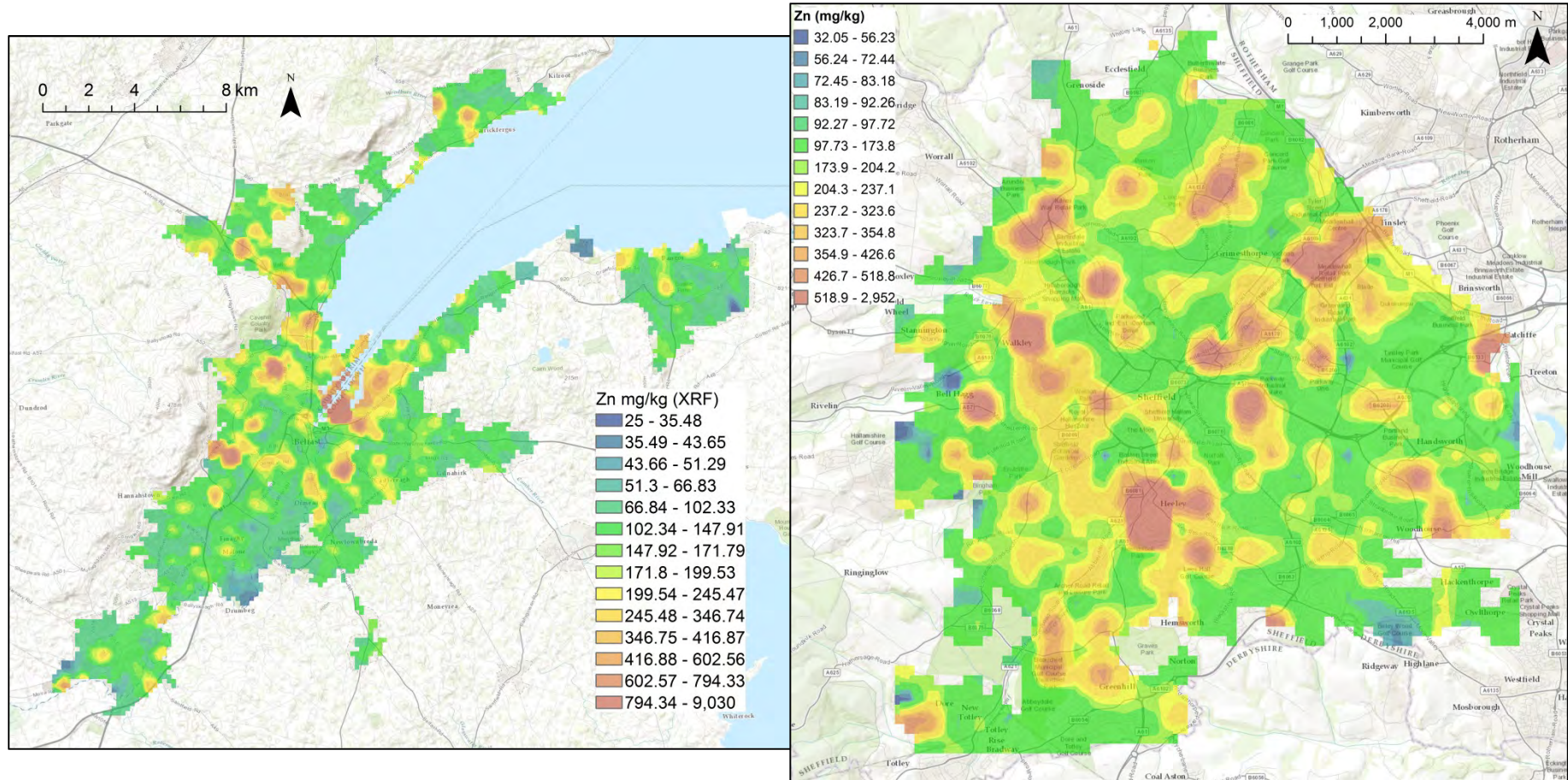
Historical development and its effect on soil contamination in urban areas

Appendix B – PTE concentration maps



Historical development and its effect on soil contamination in urban areas

Appendix B – PTE concentration maps



Historical development and its effect on soil contamination in urban areas

Depth Boxplot

```
rm(list=ls())
ppi = 300
library(psych)

## Read Belfast data
A <- read.csv("D:/Belfast.csv")

## Divide shallow concentration by deep concentration
Depth <- data.frame(A$Co_A_ICP/A$Co_S_ICP, A$V_A_ICP/A$V_S_ICP, A$Cr_A_ICP/A$Cr_S_ICP,
A$Ni_A_ICP/A$Ni_S_ICP, A$Cu_A_ICP/A$Cu_S_ICP, A$Zn_A_ICP/A$Zn_S_ICP,
A$Sn_A_ICP/A$Sn_S_ICP, A$Pb_A_ICP/A$Pb_S_ICP, A$Sb_A_ICP/A$Sb_S_ICP,
A$As_A_ICP/A$As_S_ICP, A$Mo_A_ICP/A$Mo_S_ICP)
colnames(Depth) <- c("Co", "V", "Cr", "Ni", "Cu", "Zn", "Sn", "Pb", "Sb", "As", "Mo")

## Generate boxplot image
png("Depth BPlot Belfast.png", width = 10*ppi, height = 8*ppi, res = ppi)
par(mfrow = c(1,1), mar = c(2.2, 3.2, 2, 0.5), mgp = c(2, 0.7, 0))
boxplot(Depth, ylab = "Depth Ratio (Shallow/Deep)", main = " ", log = "y", yaxt = "n", cex.lab = 1.5,
cex.axis = 1.5, cex.main = 1.5)
axis(2, at = c(0.01, 0.1, 1, 10, 100), label = c(0.01, 0.1, 1, 10, 100), cex.axis = 1.5)
t <- apply(FUN = median, X=Depth, MARGIN = 2)
t <- round(t, 2)
mtext(adj = c(0.05,0.14,0.23,0.32,0.41,0.50,0.59,0.68,0.77,0.86,0.94), side = 3, text = t, col = 1, cex = 1.5)
mtext(text = "M", adj = 0.01, side = 3, cex = 1.5)
abline(h=1, lty = 1, col = 2)
abline(h=0.2, lty = 5, col = 1, lwd = 1.5)
abline(h=1.5, lty = 3, col = 1, lwd = 1.5)
dev.off()
```

Cluster Analysis

```
rm(list=ls())
ppi = 300
library(vegan)
library(robCompositions)
library(dendextend)

## Read Sheffield data
A <- read.csv("D:/Sheffield.csv")

## Complete clr transformation
CLR_XRF = cenLR(A[,c("As", "Co", "Cr", "Cu", "Mo", "Ni", "Pb", "Sb", "Sn", "V", "Zn")])

k1 <- subset(A[,c("SAMPLE", "Easting", "Northing")])
k2 <- subset(CLR_XRF$x.clr[,c("As", "Co", "Cr", "Cu", "Mo", "Ni", "Pb", "Sb", "Sn", "V", "Zn")])
colnames(k2) <- c("As", "Co", "Cr", "Cu", "Mo", "Ni", "Pb", "Sb", "Sn", "V", "Zn")
k2 <- scale(k2)

png("Cluster Sheffield.png", width = 8*ppi, height = 4*ppi, res = ppi)
par(mfrow = c(1,1), mgp = c(1.7, 0.75, 0))
x=t(k2)
res=hclust(dist(x),method="ward")
dend <- as.dendrogram(res)
dend <- rotate(dend, order = c("Co", "V", "Cr", "Ni", "As", "Mo", "Sn", "Pb", "Sb", "Cu", "Zn"))
plot(dend, main = "Cluster analysis Sheffield", cex = 1, cex.axis = 1, cex.main = 1, cex.lab = 1, horiz = TRUE,
xlim = c(80,0))
dev.off()
```

Principal Component Analysis

```
rm(list=ls())
ppi = 300
library(vegan)
library(robCompositions)

## Read Sheffield surface data - clr transformed
A <- read.csv("D:/Sheffield.csv")

## Create subset As, Co, Cr, Cu, Mo, Ni, Pb, Sb, Sn, V, Zn
CLR_XRF = cenLR(A[,c("As", "Co", "Cr", "Cu", "Mo", "Ni", "Pb", "Sb", "Sn", "V", "Zn")])

k1 <- subset(A[,c("SAMPLE", "Easting", "Northing")])
k2 <- subset(CLR_XRF$x.clr[,c("As", "Co", "Cr", "Cu", "Mo", "Ni", "Pb", "Sb", "Sn", "V", "Zn")])
colnames(k2) <- c("As", "Co", "Cr", "Cu", "Mo", "Ni", "Pb", "Sb", "Sn", "V", "Zn")

k = cbind(k1, k2)
summary(k)

pca <- rda(na.omit(k[,-1:-3]), scale = TRUE)
PCA_S <- summary(pca)
PCA_S$cont

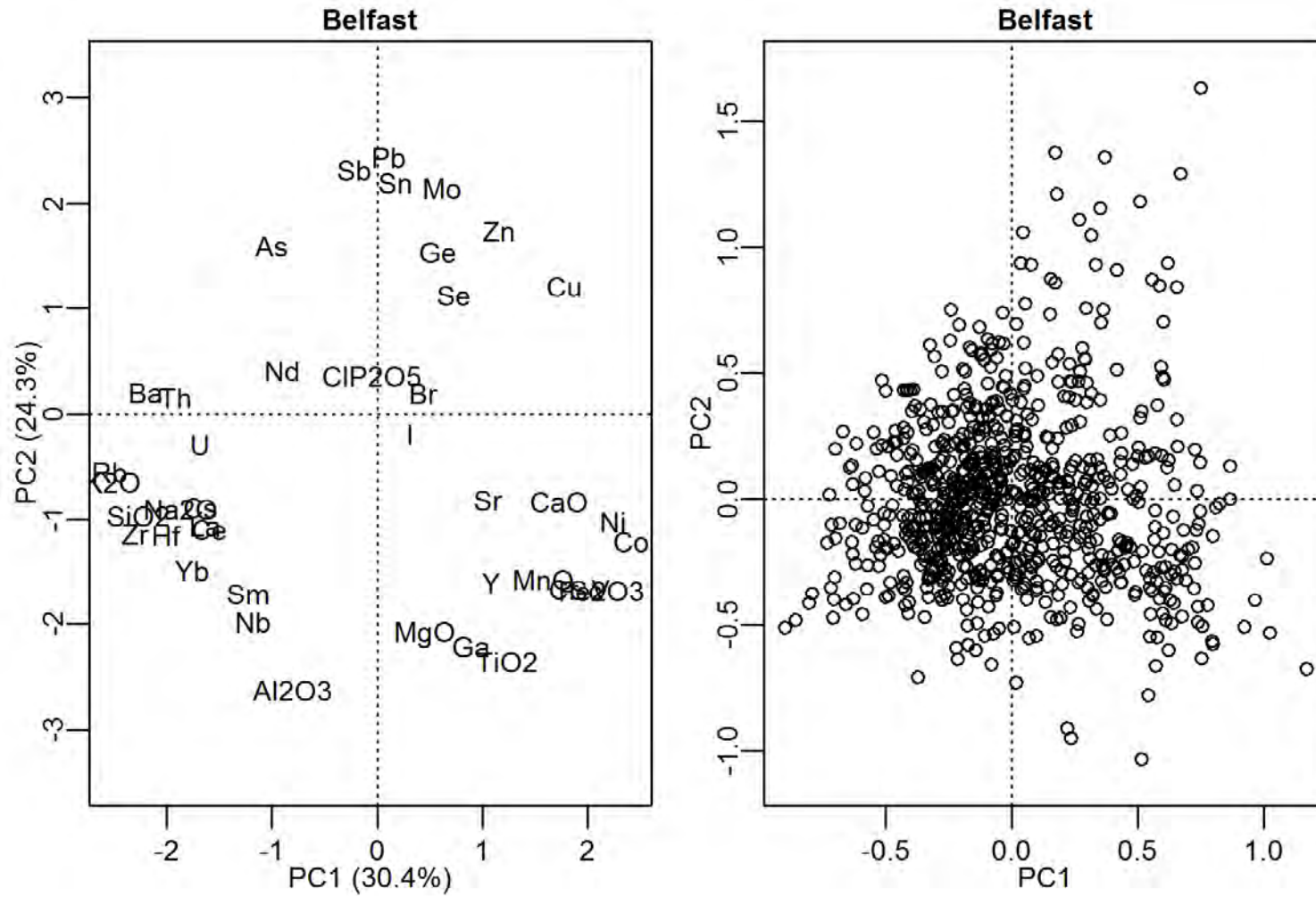
png("PCA 1&2.png", width = 6*ppi, height = 4*ppi, res = ppi)
par(mfrow = c(1,2), mar = c(2, 2, 0.5, 0.5), mgp = c(1, 0.4, 0))
plot(pca, scaling = 3, display = "species", type = "n", xlab = "PC1 (37.8%)", ylab = "PC2 (14.9%)", cex.axis = 0.75, cex.lab = 0.75)
text(pca, scaling = 3, display = "species", cex = 0.75, col = c(2, 3, 3, 4, 1, 3, 6, 6, 6, 3, 4))
plot(pca, scaling = 3, display = "sites", type = "n", cex.axis = 0.75, cex.lab = 0.75)
points(pca, scaling = 3, display = "sites", cex = 0.75)
dev.off()

png("PCA 2&3.png", width = 6*ppi, height = 4*ppi, res = ppi)
par(mfrow = c(1,2), mar = c(2, 2, 0.5, 0.5), mgp = c(1, 0.4, 0))
plot(pca, scaling = 3, display = "species", type = "n", choices = c(2:3), xlab = "PC2 (14.9%)", ylab = "PC3 (13.6%)", cex.axis = 0.75, cex.lab = 0.75)
text(pca, scaling = 3, display = "species", cex = 0.75, choices = c(2:3))
plot(pca, scaling = 3, display = "sites", type = "n", choices = c(2:3), cex.axis = 0.75, cex.lab = 0.75)
```

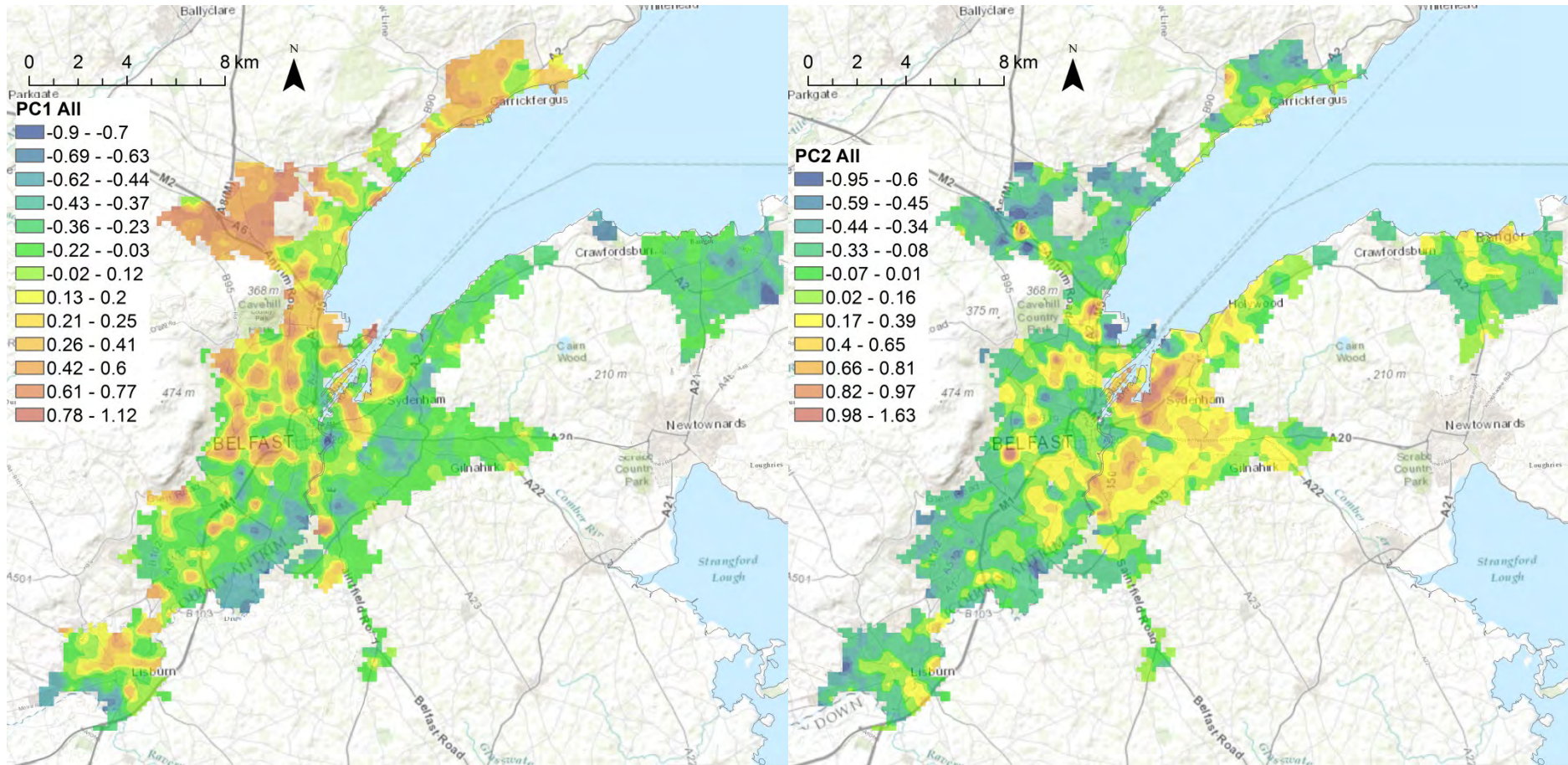
Appendix C – Sample R scripts

```
points(pca, scaling = 3, display = "sites", cex = 0.75, choices = c(2:3))  
dev.off()
```

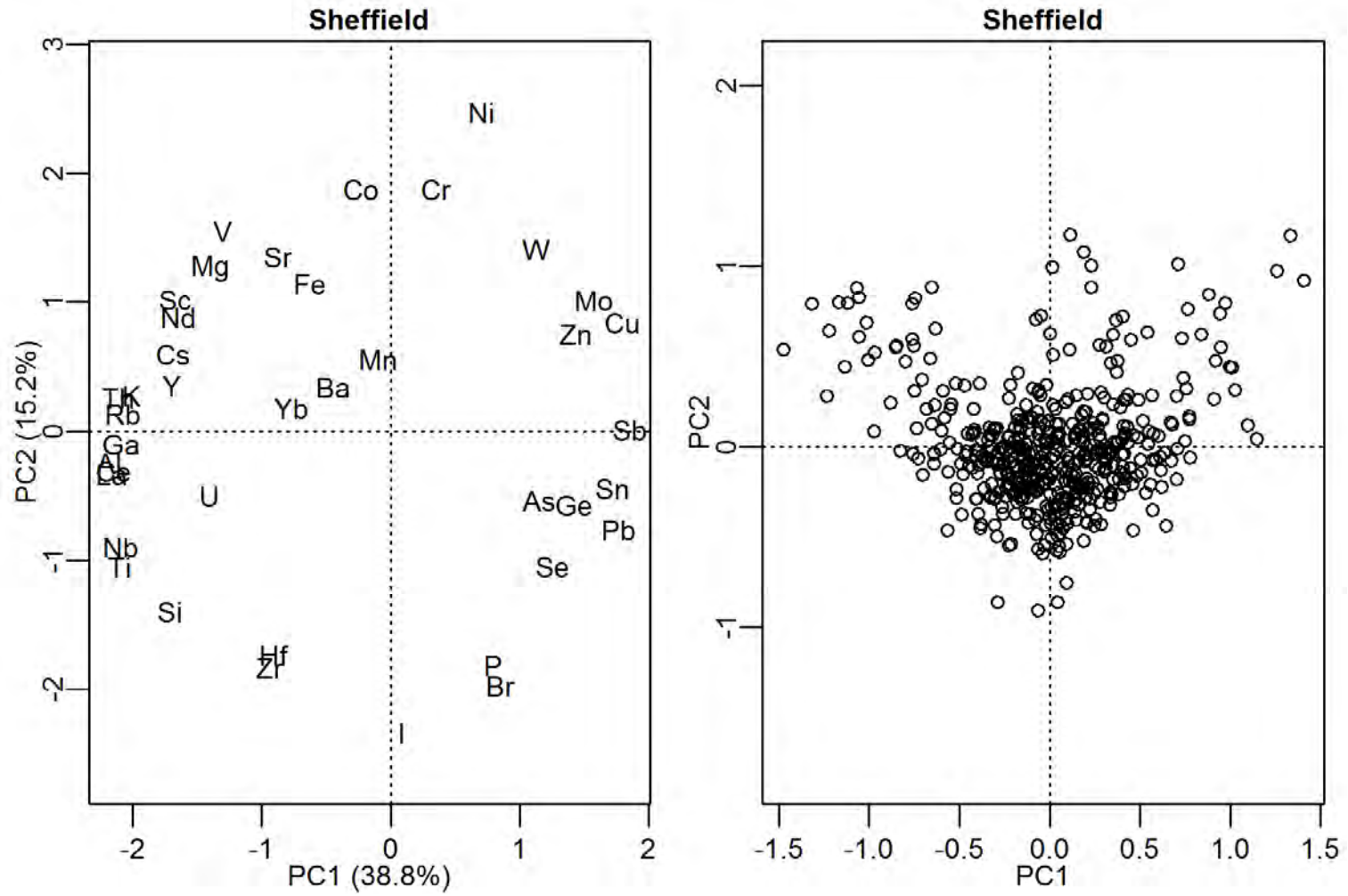
```
PCA = vegan::scores(pca, scaling = 3, display = "sites", choices = 1:3)  
write.csv((cbind(k, PCA)), file = "PCA.csv")
```



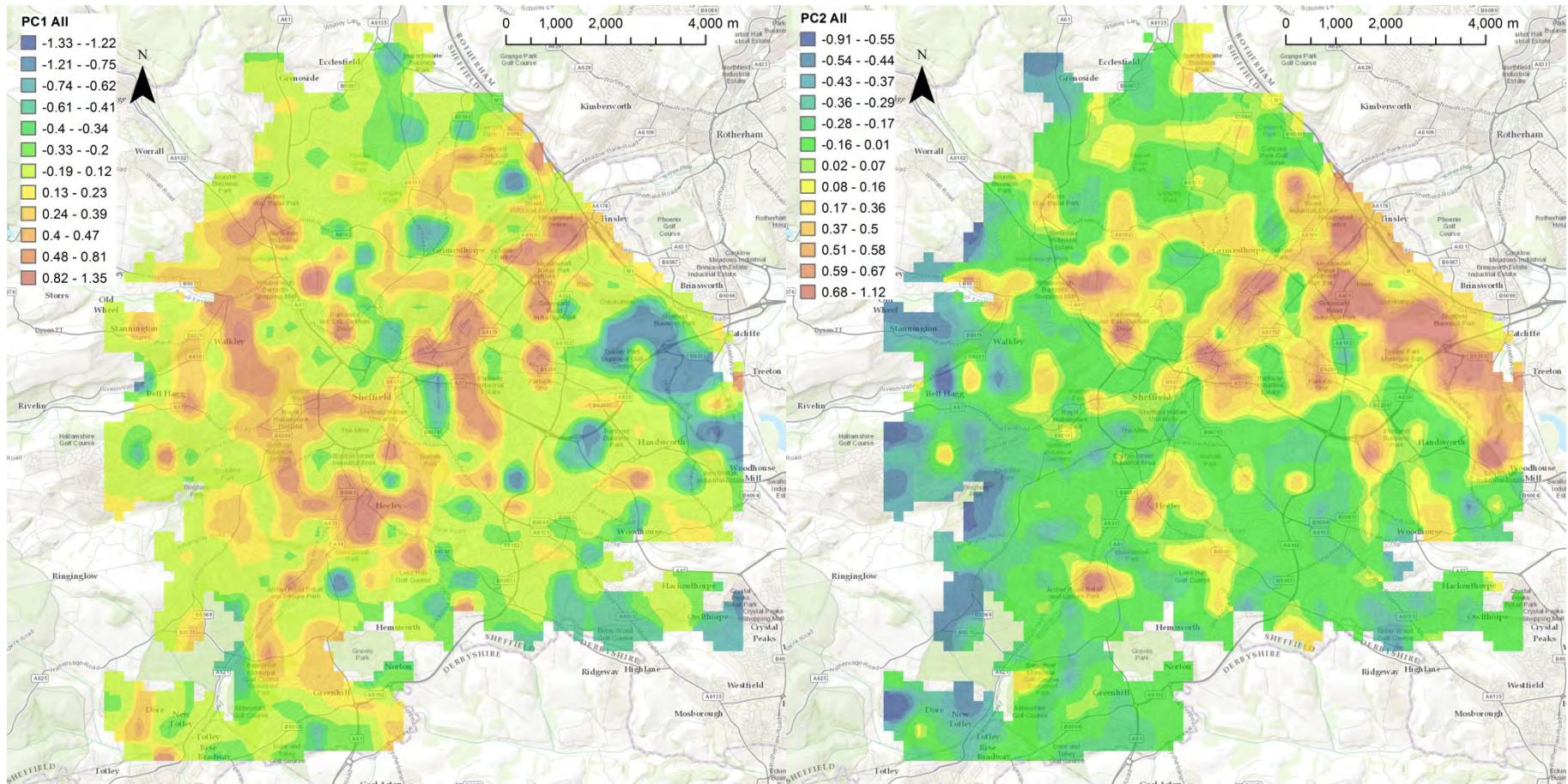
Appendix D – PCA with all available elements



Historical development and its effect on soil contamination in urban areas



Appendix D – PCA with all available elements



Historical development and its effect on soil contamination in urban areas