Morrison, Steven (2011) *Assessment of geo-environmental inequalities in the Glasgow conurbation.*
MSc(R) thesis.

Assessment of Geo-environmental Inequalities in the Glasgow Conurbation

Steven Morrison

A thesis submitted to the
University of Glasgow
for the degree of
Master of Science

Department of Statistics

Submitted July 2011

Abstract

Links between environmental quality and health have been of increased interest in recent years, with studies investigating the effects of air pollution, river-water quality and contaminated land. The current project investigated the spatial associations between potentially harmful metals in the soil; air pollution indicators and health indicators, in the context of Environmental Justice across Glasgow. The relationships between indicators of deprivation and the other variables were also assessed. The key aims of this research were to assess the spatial distributions of the environmental and health variables across Greater Glasgow and to carry out an overview investigation of the spatial associations between the variables on a city-wide basis.

Chapter 1 is a literature review of environmental and health issues including a discussion of Environmental Justice in Scotland, the effects of air pollution on health, contaminated land legislation and the potential health effects of soil contaminants.

Chapter 2 is an introduction to the datasets used in the study, presentation methods and geostatistical analysis methodology, including variograms and kriging. The modelling and inference methods used are also discussed.

Chapter 3 presents a preliminary analysis of the soil metals dataset along with an exploratory analysis of each of the soil contaminants. This analysis focuses on the geologically and anthropogenically controlled distributions of

each element, identifying and removing any trends which may be present and generating concentration predictions where necessary.

Chapter 4 includes the spatial presentation of each of the datasets, where each is presented at the Intermediate Geography (IG) level. In this chapter, the geometric mean of each soil contaminant is illustrated across Greater Glasgow along with an index of soil metal scores based on distribution percentiles.

Chapter 5 contains the results of the analysis whereby the relationships between the environmental and health indicators are assessed. This includes an exploratory analysis of the spatial associations between the variables before the relationships are then formally assessed with use of generalised linear models and relative risks. This Chapter also discusses issues regarding modelling health outcomes, modelling the soil contaminants and the ecological fallacy.

Chapter 6 summarises the findings of the research, some constraints to the analysis and recommendations for future work on environmental and health studies.

Acknowledgements

I would like to thank my project supervisor Professor Marian Scott for her help and guidance during the year of my research. I would like to thank my project leader, Fiona Fordyce from the British Geological Survey (BGS) in Edinburgh, for her continued assistance and support throughout the project and Jennifer Bow from BGS for providing a two day training course on how to use the ArcGIS software.

Thanks also go to the Statistics department of the University of Glasgow for acquiring funding for my research, which was granted by the Information Services Division of the NHS. I also thank the Clyde Urban Super Project (CUSP) for supporting the project and the BGS University Funding Initiative (BUFI) for providing additional financial support. I am privileged to have been given the opportunity to carry out such an interesting area of research and the experience it has provided me with.

Finally, I would like to thank my parents for supporting me throughout my education and taking the 300 mile trips down to Glasgow to visit me!

Contents

List of Figures

List of Tables

Chapter 1

Introduction and Literature Review on Environment and Health Issues

1.1 Introduction

Many studies have shown that populations exposed to high concentrations of potentially harmful metals in the environment can have their health affected (WHO, 1996), while others demonstrate no evidence of adverse health effects from contaminated land (RCEP, 1996). In Scotland, there has been an increased interest in the link between poor health outcomes and environmental inequalities such as poor housing, crime and industrial pollution. Previous work by Fairbairn et al. (2005) has shown an association between deprived areas in Scotland and air pollution, derelict land and river-water quality. The Environmental Justice Agenda has looked at the potential health impacts of environmental quality within neighbourhoods, although investigations to date into links with pollution have been preliminary. Land quality has not been assessed and may be of concern in terms of potential health effects. As a result, it is of interest to investigate whether inequalities in potential soil metal exposure have a spatial association with other indicators of environmental and social inequalities.

Glasgow was selected for this study as it has a long history of urbanisation and industrialisation resulting in elevated concentrations of potentially harmful metals in the soil. The distribution of these elements across the city has been mapped by the British Geological Survey (BGS) (Fordyce et al., 2005 and Fordyce et al., In Prep), forming the basis for this study.

The data provide an opportunity to test whether poor land quality is spatially coincident with indicators of poor health and deprivation in the largest city in Scotland. If a relationship is apparent from the study, then land quality may need to be taken into account in future Environmental and Social Justice Agendas. This is the first time that land quality, air pollution, deprivation and health datasets have been combined for a major UK city, and offers the challenge of creating a method of assessing the associations between land quality and health indicators. The project relates to the BGS Clyde Urban Super-Project (BGS, 2010a) and will aid further developments to the geo-environmental characterisation of the Clyde Basin.

1.2 Background

The main aim of the project is to examine the relationships between indicators of poor health, deprivation, land quality and air quality in an urban environment. However, it must be stressed that the purpose of this investigation is not to relate particular soil metals or pollutants in air with health problems, but rather to consider a spatial association between poor land, air quality, deprivation and poor health indicators in the context of the Environmental Justice Agenda. If a link is found, it may lead to further studies and remediation strategies to improve the environment and help with deprivation-related issues in the future. As part of the project, a literature review of environment and health issues; the Environmental Justice Agenda; air quality; the distribution of potentially harmful metals in the environment, urban environments in particular, and the possible health impacts of exposure to potentially harmful metals in soils and poor air quality was carried out.

1.3 Environment & Health

It is known that there is a close relationship between the health of the public and the quality of the environment. There are ongoing concerns about the risks to human health due to environmental issues such as pollution and flooding. It is important to inform the public of the risks associated with a poor environment, although the links between pollution, environment and health are fairly complicated. The Scottish Government has a new strategy to meet the economic, environmental and social needs of the people in Scotland. This is heavily based on Environmental Justice. In this context, 'justice' is about the distribution of positive and negative factors affecting environmental quality as well as providing opportunities for people to contribute in decision making about their environment (Scottish Executive Environment Group, 2005).

1.3.1 Environmental Justice Agenda

The issue of Environmental Justice in Scotland was launched in a major speech in 2002 by Jack McConnell, who was First Minister at the time. He recognised that "*people who suffer most from a poor environment are those least able to fight back"* and explained that the "*gap between the haves and have-nots is not just an economical issue."* In his speech, he insisted that we should provide fewer opportunity gaps between those with the most and those who have the least in terms of the quality of environment, while using examples of energy inefficiency in pensioner's homes and fuel poverty as cases of injustice. He highlighted the issues that needed to be dealt with, which included industry and pollution, education, transport and housing regeneration (McConnell, 2010).

Since this speech the Environmental Justice Agenda has come a long way and has addressed several issues across Scotland. A project published in 2005 aimed to consider the extent of people in Scotland living at different levels of deprivation and living in proximity to factors affecting environmental quality (Scottish Executive Environment Group, 2005). Although this was an extensive research project which aimed to develop evidence to plan future policies and address these issues, the analysis focused across a breadth of issues, rather than an in-depth analysis, which resulted in limitations within the study (Scottish Government, 2005). The investigation covered eight environmental topics: industrial emissions, derelict land, landfill, quarries, woodlands, green space, river water quality and air quality. These topics were analysed in conjunction with the Scottish Index of Multiple Deprivation (SIMD, 2010). The analysis showed that industrial emissions, derelict land and river water quality all had a strong relationship with deprivation. It also showed that people in deprived areas were less likely to live near areas of woodland. No association was found between deprived areas and greenspace, although there was evidence of a relationship between several air pollutants and deprived areas in Scotland (Scottish Government, 2005). Since this study demonstrated an association between deprived communities and poor environmental quality in Scotland, the key objective was to improve the quality of life and health of individuals as well as securing Environmental Justice for all of Scotland's communities (Scottish Government, 2005). Several other agencies and government departments in Scotland such as the Scottish Executive, Scottish Environment Protection Agency (SEPA) and Scottish Natural Heritage (SNH) have also taken up Environmental Justice themes.

The Environmental Justice Policy attempts to create a *"*healthy local environment, free from pollution, flooding and degraded streetscapes, and rich in attractive, safe public spaces*"* (Scottish Executive, 2005). To contribute information to the Environmental Justice debate, the present project will examine possible associations between soil quality, air pollution and indicators of deprivation and poor health in the Greater Glasgow region. Therefore, these issues are discussed in more detail in the following sections of this thesis.

1.3.2 Air Quality

One key concern under the Environmental Justice Agenda is air quality since it is known to have potential effects on public health. Air pollution is a complex mixture of chemicals in the atmosphere that causes damage to human health and the environment. Throughout the 19th century and most of the 20th century air quality in the UK was very poor in urban areas as a result of heavy industrialisation. The introduction of steam engines, coal use and motor vehicles all contributed to the increased levels of air pollution. The scale of air pollution problems became apparent following the smog of 1952 in London which resulted in over 4000 people losing their lives and many others suffering from illnesses (Wilkins, 2006). The British Government brought in new regulations to limit the use of fossil fuels and generation of black smoke. Clean Air Acts were introduced in 1956 and 1968 to reduce atmospheric pollution from coal combustion (Farmer and Jarvis, 2009).

In more recent times, the health effects of air pollution on public health have been thoroughly investigated and guidelines have been set in order to improve air quality, with the latest standards set in July 2007 (UK Air Quality Archive, 2007). The declines in industrial processing and coal combustion have resulted in considerable improvements in air quality over the last few decades. However, despite this, the increased use of vehicles is now the most prevalent source of emissions (Scottish Government, 2005). Table 1.1 provides the UK air quality objectives based on protecting the health of the public (UK Air Quality Archive, 2007).

Today, air pollution is unlikely to have any serious effects on healthy individuals; however, those with lung or heart problems are at greater risk. Exposure to high levels of air pollution can lead to irritation of lungs, can result in premature death for those who are seriously ill and is more likely to lead to an attack for asthmatics (DEFRA, 2010).

Table 1.1 – UK Air Quality Objectives for protection of human health, July 2007

Pollutant	**Air Quality Objective**		**To Be Achieved By**
	Concentration	**Measured as**	
Benzene			
All authorities	16.25 µg m^{-3}	Running annual mean	31 December 2003
England and Wales Only	5.00 µg m^{-3}	Annual mean	31 December 2010
Scotland and N. Ireland	3.25 µg m^{-3}	Running annual mean	31 December 2010
1,3-Butadiene	2.25 µg m^{-3}	Running annual mean	31 December 2003
Carbon Monoxide			
England, Wales and N. Ireland	10.0 mg m^{-3}	Maximum daily running 8-hour mean	31 December 2003
Scotland Only	10.0 mg m^{-3}	Running 8-hour mean	31 December 2003
Lead	0.5 µg m^{-3}	Annual mean	31 December 2004
	0.25 µg m^{-3}	Annual mean	31 December 2008
Nitrogen Dioxide	200 µg m^{-3} not to be exceeded more than 18 times a year	1-hour mean	31 December 2005
	40 µg m^{-3}	Annual mean	31 December 2005
Particles (PM10) (gravimetric)			
All authorities	50 µg m^{-3}, not to be exceeded more than 35 times a year	Daily mean	31 December 2004
	40 µg m^{-3}	Annual mean	31 December 2004
Scotland Only	50 µg m^{-3}, not to be exceeded more than 7 times a year	Daily mean	31 December 2010
	18 µg m^{-3}	Annual mean	31 December 2010
Particles (PM2.5) (gravimetric) *	25 µg m^{-3} (target)	Annual mean	2020
All authorities	15% cut in urban background exposure	Annual mean	2010 - 2020
Scotland Only	12 µg m^{-3} (limit)	Annual mean	2010
Sulphur Dioxide	350 µg m^{-3}, not to be exceeded more than 24 times a year	1-hour mean	31 December 2004
	125 µg m^{-3}, not to be exceeded more than 3 times a year	24-hour mean	31 December 2004
	266 µg m^{-3}, not to be exceeded more than 35 times a year	15-minute mean	31 December 2005
PAH *	0.25 µg m^{-3}	Annual mean	31 December 2010
Ozone *	100 µg m^{-3} not to be exceeded more than 10 times a year	8 hourly running or hourly mean*	31 December 2005

(new objectives are highlighted in shading)
*** not included in regulations at present**

The World Health Organisation (WHO) produces global standards in environmental health by setting air quality standards, using relevant scientific evidence. However, there are uncertainties in several aspects of air quality in relation to health. One is that strategies from other organisations differ in terms of the thresholds being set. Another difficulty is that these standards differ from the defined objectives, since the thresholds are also set with reference to economic efficiency, practicability etc. An additional problem is that meeting these air quality standards does not mean that a health effect is absent.

The main pollutants of concern are carbon monoxide, nitrogen dioxide, sulphur dioxide, ozone, particulate matter and benzene. Carbon Monoxide (CO) is a colourless and odourless poisonous gas, released into the atmosphere when incomplete combustion occurs, normally in vehicles. It disrupts oxygen flow in the blood, and is a major problem for those with heart disease (Scottish Air Quality, 2010). Nitrogen Dioxide (NO_2) is produced from a variety of combustion processes and leads to irritation of the lungs which can result in respiratory infections. Children are also at risk to respiratory diseases when frequently exposed. Sulphur Dioxide (SO_2) also causes damage to the lungs and is formed by the burning of coal and oil. It is believed that SO_2 causes further harm when the concentrations of other air pollutants are high. Another air pollutant which affects lung function is Ozone (O_3), as it is known to irritate the pathways to the lungs. Ozone is formed by chemical reactions generated by sunlight, through oxidation of compounds in the presence of nitrogen oxides (Scottish Air Quality, 2010). Particulate matter (PM) is a harmful pollutant made from a mixture of very small particles of solid or liquid, suspended in a gas. The focus is on PM_{10} as it represents particles of 10 µm or less and these are most likely to penetrate the lungs. Exposure to PM_{10} can cause inflammation and leads to deterioration in the condition of those with heart or lung diseases. Benzene is a compound which is present in petrol and is mainly released into the atmosphere by the burning of petrol in vehicles. Exposure to high concentrations of benzene can result in damage of the liver and kidney, cancer and birth defects (Scottish Air Quality, 2010).

As mentioned in Section 1.3.1, the 2005 project carried out under the Environment Justice Agenda found evidence of an association between air pollution and those living in deprived areas in Scotland. Listed below are some of the key findings from the investigation into the links between air quality and deprivation (Scottish Government, 2005).

- The most deprived decile has the poorest air quality, for 80% of the pollutants.
- 80% of pollutants meet Air Quality Strategy Objectives but still cannot assume an absence of health impact.
- For NO_2, those in the most deprived decile are 12 times more likely to be living in an area of higher pollution that those in the least deprived decile.
- Those in the least deprived decile are 3 times more likely to experience highest PM_{10} concentrations than those in the least deprived decile.

Therefore, previous studies had distinguished a potential relationship between air quality and deprivation. Part of the current project was to investigate this further; focussing on spatial associations that air pollutants may have with health indicators in Glasgow. Derelict land was another of the seven topics covered in the Environmental Justice Agenda project (Scottish Government, 2005) and it was important to look into the research that has already been carried out on land quality for the current project before going on to consider the possible health effects and associations with social deprivation.

1.3.3 - Contaminated Land

Potentially harmful elements (PHEs) such as arsenic (As), chromium (Cr), copper (Cu), nickel (Ni), lead (Pb), selenium (Se) and zinc (Zn) occur in soil naturally, while human activities also contribute to the concentration of these elements. Many are essential for health and are taken up by plants and animals into the human food chain. As such, there are links between soil and human health and insufficient concentrations of these key nutrients can result in dietary deficiencies. On the other hand, many of these essential elements are potentially toxic at high concentrations. Mainly as a result of man-made pollution from industrial and urban sources, soils can be contaminated by these harmful substances, which can cause health problems in some cases if high-level exposure occurs over long periods of time. Although they are regarded as harmful elements, the health effects of some are not clear and remain uncertain (Selinus and Frank, 2000).

Several studies have demonstrated elevated concentrations of PHEs in urban areas. Sampling and analysis in several Norwegian cities revealed that soils in older areas of the cities were highly contaminated with PHEs, Pb in particular (Ottesen et al., 2008). For example, in Oslo, 38% of day-care centres required remediation due to high concentrations of metals, while 45% of day care centres in Bergen needed remediation because of elevated Pb concentrations. These high levels of PHEs were due to the accumulation of metals over many years from the demolition and redevelopment of buildings, road traffic, fires, industry etc. Day-care centres were targeted for the study as children are more vulnerable to health effects from soil metals due to their small body size and developing nervous system. They are also more exposed to soils due to their regular contact with school playgrounds, and have typical hand to mouth behaviour (Mielke et al., 2005). A study, based in New Orleans, demonstrated that there was a strong inverse association between the quantity of multiple metals accumulated in the community of elementary schools and the learning achievement of students attending schools in the same community (Mielke et al., 2005). The findings also established that Pb was one of the most prominent urban contaminants and that increased blood Pb concentrations in children were highly related to learning and behaviour difficulties, along with other health effects. A further study on the relationship between children's blood Pb and learning achievement was carried out in New Orleans (Mielke et al., 2009). The performance was measured across English, Mathematics, Science and Social Studies and the results showed that blood Pb concentrations were negatively correlated with achievement in all these subjects. Another observation was that blood Pb appeared to be a more influential predictor of learning achievement than class numbers and poverty.

In the UK, Fordyce et al., (2005) have shown that metal concentrations in soils from 14 urban centres were elevated up to five times in comparison to those in rural soils. Of the 14 cities included in the study, Glasgow was recorded to have the highest median Cr concentration. This was partly related to the volcanic bedrock underlying parts of Glasgow, which is naturally high in the element, but mainly because of the history of metal processing in the city. The world's largest chromite processing works was located in southeast Glasgow from 1830 to 1968. In the past, waste from the plant was used as fill material around the city leading to concerns about potential health impacts on the local population (Farmer and Jarvis, 2009). An age-standardised study on lung cancer incidence found no evidence of adverse health effects and exposure levels did not create any significant risk to human health (Watt et al., 1991), while a subsequent study on the association between Cr contaminated land and leukaemia found no evidence of a relationship between soil-Cr and leukaemia (Ezzaguirre-Garcia et al., 1999). However, problems with both studies were that particle size effects on inhalation and transportation of Cr pollution by winds were not taken into account. In the following years, Cr contaminated land sites were capped to reduce the exposure to airborne dust and remediation strategies were put in place to control the retention of Cr on the sites (Farmer and Jarvis, 2009).

Despite the results from the above risk assessments, it is very difficult to examine the possible health effects related to contaminated land due to uncertainties about exposure and risk. Most of these studies show insufficient evidence of exposure and raise a doubt that other confounders could play a part. There are many complex issues surrounding contaminated land in relation to health problems and our understanding of these issues is fairly poor. The ecological fallacy is one such problem, where population-based results are not necessarily true at the individual level. This issue is discussed in further detail in Chapter 5 of this thesis. Another difficulty which Kibble and Saunders (2001) highlight is that, when exposure exists, it is often of a small degree, difficult to measure, or insignificant.

None the less, it is important to predict human exposure to PHEs in soil and there are three key routes for these to enter the body. Ingestion of dust and soil from home-grown vegetables through the mouth is one possible route. This is more common in children due to deliberate eating of soils and hand to mouth behaviour. On average, we each consume approximately 100mg of soil every day (EA, 2009a). Another exposure route is inhalation through the mouth and nose. Soil particles can be blown by the wind and inhaled, while contaminated dust and vapour can all be inhaled either outdoor or indoor. Exposure to soil contamination can also occur via absorption through the skin. The Environment Agency (EA) have developed a model that can be used to assess the health risks from long term exposure to soil contamination, which is discussed in the following section.

1.3.4 – UK Contaminated Land Legislation

The UK government's main approach to improving the environment is to identify areas of contaminated land, make appropriate decisions for sustainable development and therefore improve environmental quality. Land can be contaminated by metals, organic substances or radioactive and chemical materials, all of which could have harmful effects on four key receptors identified in the legislation; namely humans, water quality, buildings and ecosystems. There are estimated to be 300,000 hectares of contaminated land in the UK from industrial processes (Scottish Government, 2005), which equates to 3000 km^2. To tackle the problem of contaminated land, the Government introduced new legislation in 1990 (Environmental Protection Act, Part IIa, 1990) whereby Local Authorities are required to examine their area to identify contaminated land and make sure it is dealt with. If metal concentrations are found to exceed the soil guideline values (SGV), further inspection and testing of the land must be performed to assess whether the elements in the soil present a risk to human health (Fordyce et al., In Prep).

To help estimate and assess the risks from long term exposure to soil contamination, the UK government developed the Contaminated Land

Exposure Assessment (CLEA) model (EA, 2009a). Health Criteria Values (HCV) were obtained from expert evaluations and literature, and were used to describe the levels at which long term human exposure to chemicals in soil is tolerable or posed a minimal risk to health (EA, 2009a). The Soil Guideline Values (SGV) were calculated by estimating the concentration of a particular element in the soil entering the body, which is equivalent to the relevant HCV (EA, 2009b). These can be used at specific sites to assess the likely health risks from contaminated land. Different SGV were evaluated for separate land uses depending on the activity of a site. In terms of soil metal contamination, the main elements deemed to be of concern under the CLEA guidelines are As, Cd, Cr, Pb, Hg, Ni and Se, while the soil exposure route for Cu and Zn are not thought to be a concern for human health. The SGV for each element is given in Table 1.2 (EA, 2010a). With the exceptions of Hg and Cd, these chemical elements plus the potentially harmful metals Cu and Zn form the focus of the present study. Mercury (Hg) is not included as there are no existing data available for Glasgow in the geochemical dataset provided for the present project. Cadmium (Cd) is not included as very few of the measurements were above the limit of detection in the dataset provided for the project (Fordyce et al., In Prep). The potential health effects of these metals are described in the following section of this thesis.

Table 1.2 – CLEA soil guideline values for metals of concern for selected land uses

Elements	Units	CLEA Soil Guideline Values	Notes
Arsenic (As)	mg kg^{-1}	32	Residential Land
Cadmium (Cd)	mg kg^{-1}	1.8	Allotments
Chromium (Cr)	mg kg^{-1}	130	Old SGV, new SGV awaited
Lead (Pb)	mg kg^{-1}	450	Old SGV, new SGV awaited
Mercury (Hg)	mg kg^{-1}	80	Inorganic Hg, allotments
Nickel (Ni)	mg kg^{-1}	130	Residential Land
Selenium (Se)	mg kg^{-1}	120	Allotments

From: EA (2010a). Soil Guideline Value Reports published using the new approach.

1.4 Health Effects of Potentially Harmful Metals

As already discussed in the previous sections of this review, exposure to high concentrations of PHEs in the environment has the potential to adversely affect human health. A history of industrialisation in many cities can result in higher metal concentrations in soil, which contributes to the poor quality of land. Under current UK contaminated land legislation, contact with contaminated soil is considered as one of the main pathways in which harmful substances can enter the body. The sources of these metals in the environment, methods of exposure and potential health impacts are outlined as follows, taken from Fordyce et al. (In Prep) as well as the EA CLEA reports.

1. Arsenic (As)

Inorganic arsenic is a naturally occurring metal element found in many minerals and is widely distributed in rocks, soils and sediments. It is also released naturally into the environment from forest fires and volcanoes. The public is exposed to As from these natural sources as well as in drinking water from the leaching of rocks and soils (EA, 2009c). However, concentrations of As can be elevated in the environment as a result of anthropogenic activity. Arsenic is used in the manufacture of glass, alloys, pesticides and wood preserving products and is released into the environment via the burning of fossil fuels and waste disposal (EA, 2009c). Arsenic is one of the oldest known poisons to man and large intakes can lead to health problems. Both oral and inhalation exposures to inorganic As pose a carcinogenic hazard and exposure via both routes result in an increased overall risk of cancer (EA, 2009c).

2. Chromium (Cr)

Chromium is found in many minerals and is widely distributed in rocks, soils, water, and sediments in the natural environment, while the main natural source of Cr in the atmosphere is dust from the continents. However, far larger amounts of Cr are contributed to the atmosphere from human activities including metal industries, burning of coal and oil, cement works and road dusts, while uses of Cr include wood preservatives and pigments in the chemical industry, and the production of steel in the metallurgical industry (EA, 2002a). Chromium is present in meats, fruit and vegetables and is an essential dietary nutrient. Shortages can lead to heart conditions and diabetes (EA, 2002a). Conversely, Cr is toxic in high concentration. The toxicity of Cr depends upon its oxidation state. Hexavalent Cr (VI) is more toxic than the trivalent form Cr (III). Most naturally occurring Cr is in the Cr (III) state in soils and foodstuffs. Although Cr (VI) rarely occurs naturally, it is produced from industrial sources. Chromium (VI) is a known human carcinogen as inhalation has been linked to cancers of the respiratory tract. Exposure routes of concern for Cr include inhalation of air and dusts, ingestion of food, water and dusts and dermal contact with soils and dusts. Excess exposure to Cr can result in allergies, rashes, ulcers and respiratory problems (EA, 2002a).

3. Copper (Cu)

Copper is a naturally occurring trace element that is widely distributed throughout the environment in rocks, soils, sediments and natural waters. However, exposure to Cu-contaminated soil is not usually a concern since only a small amount of a person's daily Cu intake comes from soils (MOE, 2001). In the UK, the main anthropogenic sources of copper are coal-fired power stations, iron and steel industries, non-ferrous metal industries, waste incinerators, agricultural chemicals and the application of sewage sludge to land. It is widely used in conductors, wires, plumbing and piping, paints and coins (EA, 2010b). Copper is also present in water, fruit and vegetables and is an essential nutrient for humans and animals, helping the production of blood in the body. Although Cu is an essential dietary element, high concentrations can be harmful and may lead to Cu poisoning which can cause gastric pain, vomiting, diarrhoea, acute renal failure and liver damage (WHO, 1996). Deficiencies of Cu can result in anaemia, whitening of the skin, bone and vascular abnormalities and brittle hair.

4. Nickel (Ni)

Nickel is found in many minerals and is widely distributed in rocks, soils, water, and sediments in the natural environment. It is distributed in the soil from the weathering of rock but is also released into the environment through oil and coal combustion, and the burning of waste materials (EA, 2009d). Nickel is also used in producing alloys and platings, commonly found in vehicles, electrical equipment, jewellery and coins (EA, 2009d). Eating foods containing Ni is the main source of human exposure, although other routes include air, drinking water, tobacco and contact with soil (ATSDR, 2005). The most common health problems associated with Ni are allergic reactions, and high intakes of Ni can result in more serious effects such as lung cancer and bronchitis (ATSDR, 2005).

5. Lead (Pb)

Lead is a naturally occurring metal present, in trace amounts only, in most natural rocks, soils, sediments and waters. Therefore, anthropogenic sources of Pb are more important than natural sources in most circumstances. Lead is released into the atmosphere by industrial emissions and burning of fossil fuels as well as via leaded petrol, although the use of leaded petrol has declined in recent years (EA, 2002b). Due to atmospheric fallout, these all contribute to the Pb content of the soil. Lead is mainly used in lead-acid batteries although other uses include ammunition, sheet lead, cable sheathing and solder (EA, 2002b). Lead is one of the more damaging metals to the human body and can enter via the food chain, drinking water and air. Exposure to higher concentrations of Pb can result in problems such as nervous disorders, impaired foetal development, reduced learning ability in children, high blood pressure and kidney damage (EA, 2002b).

6. Selenium (Se)

Selenium is a naturally occurring metalloid that occurs, in trace concentrations only, in most natural rocks, soils, waters and sediments. In addition, it is released into the environment from coal combustion and the mining and refining of various metals (EA, 2009e). Selenium is used in the production of black and red glass and also as a catalyst in pharmaceutical products. It also has semiconductor and photoelectric properties making it useful in electronic and photocopying production (EA, 2009e). Exposure to Se is common as it occurs naturally in air, soil and water. Selenium is present in many foods including meat, grains and cereals and is an essential element for human health. Deficiency can lead to heart disease and immune system and reproductive disorders. Conversely, it is toxic at higher concentrations. People who live closer to waste sites and metal industries are more likely to be exposed to larger concentrations of Se, while some of the potential health effects from long term exposure include deformed nails, nausea and vomiting, hair loss and neurological effects (EA, 2009e).

7. Zinc (Zn)

Zinc is a naturally occurring element in the environment; mainly found in soil, rocks and waters. However, the majority of Zn in the environment is a result of human activities such as metal production, tyre wear on roads, and some emissions from the burning of coal and waste (EA, 2010c). Zinc is commonly used as a coating to other metals in order to prevent corrosion in batteries and several alloys, while Zn compounds are also used in the manufacture of plastics, rubber and cosmetics. Excessive exposure to zinc compounds can be harmful to human health and can affect the kidney, lungs, pancreas and reproductive system (EA, 2010c).

1.5 Project Rationale

This review has summarised the environment and health issues in relation to air pollution, contaminated land and Environmental Justice. The evidence illustrates the need to carry out the present project and look further into the associations between environmental and social inequalities in Glasgow. This study will use a Geographic Information System (GIS) in developing spatial-based techniques to make statistical inferences about the associations between land quality, air quality, deprivation and health indicators. If this project finds strong relationships further examinations and remediation strategies to reduce poor environmental quality may be required in the future.

The main aims of the research were as follows:

- Investigate the spatial distributions of environmental and health variables across Greater Glasgow.
- Carry out an overview assessment of the spatial associations between the environmental and health variables.
- Explore and discuss the key findings of the research and suggest what could be done to aid Environmental Justice studies in the future.

Chapter 2

Materials and Methods

2.1 Materials

This chapter includes a detailed description of the datasets used in the investigation, namely the soil metals, air quality, deprivation and health indicator datasets collated for the project.

2.1.1 Soil Metal Data

Land quality information for the project was provided by the British Geological Survey (BGS) Geochemical Baseline Survey of the Environment (G-BASE) project in the form of a soil geochemistry database for Glasgow. According to Fordyce et al. (In Prep) and Johnson et al. (2005), the BGS is responsible for the national geochemical survey of the UK, known as the G-BASE project. Starting in the late 1960s in the north of Scotland and working southwards, G-BASE is a systematic survey which aims to characterise the chemistry of the UK surface environment. One of the main aims of the G-BASE project is to generate information for sustainable development in the UK. The data have a wide range of environmental applications including prioritising the remediation of contaminated land; protection of ecosystems; sustainable use of mineral resources; water resource management and enhancing our understanding of the association between the quality of the environment and health. Sampling is primarily based on the collection of rural stream sediment, stream water and soil samples at a density of 1 per 2 km^2 across the country (Johnson et al., 2005). However, in response to demands for geochemical data motivated by legislation from governments and agencies to reduce the impacts of PHEs in the environment, urban soil sampling was initiated in 1992. To date, systematic urban soil sampling at a density of 4 per km^2 has been completed for 27 cities including Belfast, Cardiff, Glasgow and London (BGS, 2010b).

The G-BASE soil sampling of Glasgow was carried out between 2001 and 2002 on a systematic grid across the conurbation. The aim of the survey was to characterise the soil geochemistry of the Glasgow conurbation (Fordyce et al., In Prep).

Adhering to G-BASE procedures (Fordyce et al., In Prep), 1381 urban soil samples in Glasgow were collected at a density of 4 per km^2, while 241 rural samples were collected around the outskirts of Glasgow from every second km^2. In the urban areas, each kilometre national grid-square on 1:25,000 scale OS map was divided into four sub-squares of dimension 500 m x 500 m, with a sample collected as close as possible to the middle of each sub-square. Samples were collected from sites such as parks, gardens, road verges, open spaces, school yards, sports fields and waste ground. In British National Grid co-ordinates, the extent of the sampled area ranges from 233550 to 275570 west to east and from 649160 to 680600 south to north. The minimum distance between any two sampled locations was approximately 76 m and the furthest distance between any two points at the extremities of the area was 50,255 m. The short distance of 76 m is likely to be due to restrictions in sampling sites, since roads, housing etc. may be in the way of the desired sampling location.

At each site, two separate soil samples were collected, a top soil (5 – 20 cm) and a deeper soil (35-50 cm). A handheld Dutch auger was used to collect five sub-samples from the corners of a 20 m x 20 m square. These five sub-samples collected from each site were homogenised to create one top soil sample and one deeper soil sample (Fordyce et al., In Prep). However, for this study only the top soils were used as the public are more likely to come into contact with surface soils rather than deeper soils. Therefore, top soils are more important for human interaction.

After collecting the samples, the soils were air dried at < 30 °C to avoid volatilisation of the Se and then sieved to fractions of size < 2 mm. Before analysis of the samples, they were homogenised, coned and quartered. The next step was to grind 30 g of the samples in an agate planetary ball mill so that 95 % of it was < 53 µm, and then further divided the sample to 12 g. The samples were then analysed by X-ray Fluorescence Spectrometry (XRFS) for total concentrations of 46 elements (Fordyce et al., In Prep).

The XRFS lower limits of detection (LLD) for each of these elements are shown in Table 2.1. These are theoretical quantities for the element concentration, which are equivalent to three standard deviations above the background count rate for the substance being analysed. The data underwent thorough quality control procedures during the sampling and analysis process according to standard G-BASE policy. As part of this treatment the data below the LLD were assigned to a value of half the detection limit (Fordyce et al., 2005; Johnson et al., 2005; Fordyce et al., In Prep). There are other possible statistical approaches to deal with values reported below the detection limits. One common approach, adopted by Helsel, D. R. (2005), is to treat the data as censored observations. However, for the current project, half the detection limit was used as this is standard G-BASE practice.

A sub-set of the G-BASE Glasgow soil geochemistry database was provided for this project, containing spatially registered information on the total element concentrations of the PHEs As, Cr, Cu, Ni, Pb, Se and Zn. In addition to these elements, Potassium (K_2O) was also included in the study as a non-harmful control element in top soils. As well as the element concentrations at each sampling site, the dataset also comprised national grid coordinates (Eastings and Northings); the BGS sample numbers for each location, and whether the sample was collected from a rural or urban location.

Table 2.1 – G-BASE Glasgow soils determinands and limits of detection

Element	Name	Method	Units	Detection Limit
As	Arsenic	WD-XRFS	mg kg^{-1}	0.9
Cr	Chromium	WD-XRFS	mg kg^{-1}	1.3
Cu	Copper	WD-XRFS	mg kg^{-1}	0.8
Ni	Nickel	WD-XRFS	mg kg^{-1}	0.6
Pb	Lead	WD-XRFS	mg kg^{-1}	0.5
Se	Selenium	WD-XRFS	mg kg^{-1}	0.2
Zn	Zinc	WD-XRFS	mg kg^{-1}	0.5
K_2O	Potassium	WD-XRFS	wt %	0.05

WD-XRFS = Wavelength Dispersive X-ray Fluorescence Spectrometry

From: Fordyce et al. (In Prep)

2.1.2 Air Quality Data

The data on air quality for this study consisted of ambient air pollution concentrations of Nitrogen Dioxide (NO_2) and Particulate Matter (PM_{10}), which were regularly measured across Scotland. The focus is on these particular pollutants as they are routinely available. More information on the methods and monitoring locations of the air pollutants are available from Scottish Air Quality (2010). Ambient concentrations were recorded at locations at a specific time, instead of being averaged over a period of time. However, the data used for this study were modelled average pollution concentrations from these ambient concentrations at the Intermediate Geography (IG) level, and were available to download from the Scottish Neighbourhood Statistics website, SNS (2010). Further discussion on the Scottish Neighbourhood Statistics (SNS) and small area geographies such as Intermediate Geography and Datazone are provided in the following section of this thesis. Modelled estimates were used since data and monitoring sites were not concentrated on a small enough scale for this study. However, there was no uncertainty available on these measurements so they were assumed to be known mean concentrations within the corresponding regions. The NO_2 and PM_{10} concentrations were both measured in micrograms per cubic metre (ugm^{-3}) and the concentration recorded was a population weighted mean over the three year period between 2002 and 2004. As part of the process to determine the concentrations, a detailed procedure was carried out as outlined in the following manner. The maps of the modelled pollutants were computed using dispersion models in addition to the data measured from the monitoring stations, with concentrations estimated for every square kilometre on the national grid (SNS, 2010). Each address in Scotland was allocated a population with the addresses merged together with the air quality data. The population weighted average background concentration was then derived for each Datazone (DZ). For these data, the concentrations were estimated using the population census from 2004 (SNS, 2010). Background concentrations were selected for the present study as they were recorded distant of potential sources, so were more representative of the general air pollution levels in the city than kerbside or roadside locations; therefore being more appropriate for this study.

2.1.3 Health Data

The Scottish Neighbourhood Statistics (SNS) is a website run by the Scottish Government to enhance the availability and reliability of information within different geographical areas across Scotland (SNS, 2010). The statistics include data and reports on issues such as education, employment, health, poverty, population and crime. The creation of this long term programme offers support for assessing the government's plan for reducing the gap between deprived areas and the rest of the country (SNS, 2010). Data for up to 100 indicators are available to download from the website and are presented for a variety of different small area geographical levels including Health Boards, Electoral Wards, Intermediate Geography (IG) and Datazone (DZ). Datazone level is a very useful in providing many small area statistics across Scotland and the area typically covers populations of approximately 500 to 1000 household residents, while IG level covers an average of 4000 household residents and is used to produce small area statistics which are not suitable at the DZ level (Scottish Government, 2011). Since the current study was dealing with health data, it was very difficult to obtain data at such a small scale as Datazone due to confidentiality reasons. Therefore, the smallest geographic scale suitable for the purposes of this study and for which health data were routinely available was the IG level. Health data for this study were also available from the Information and Services Division (ISD) of the National Health Service (NHS), which is Scotland's national organisation for providing health information, statistics and information technology (IT) services (ISD, 2010).

Numerous studies have carried out research on the health effects of the physical environment, some of which have been discussed in the previous chapter. Two of the main health effects are problems with lung and respiratory systems, although others which are less common include allergies, abnormalities, liver damage, high blood pressure, kidney damage etc. Therefore, the health data selected for this study focussed on two particular outcomes:

- Lung Cancer
- Respiratory Disease

The reason for selecting these was partly that they were the most common diseases, had substantial counts and were part of the Scottish Government health strategy. More importantly, lung cancer and respiratory disease were known to have biologically plausible links with the soil contaminants and air pollutants selected for this study. Since likely exposure levels to the soil contaminants were not available for Glasgow, it is difficult to determine any direct association soil metals may have with health. However, one cannot overlook the possibility of exposure via blown soil dust contributing to these health outcomes, along with other factors such as air pollution, poor diet and smoking.

The numbers of lung cancer registrations during the five year period from 1998 to 2002, including both sexes, with population denominators by each IG level from the census in 2001, were selected. Similarly, data for lung cancer were available at the IG level. The corresponding rate per 100,000 for lung cancer registrations was also used for simplicity in the further analysis of health indicators. The data for respiratory disease consisted of the total number of hospital admissions in the year 2002, also recorded at the IG level. The health data used for this investigation were downloaded from the SNS website, with corresponding age and sex specific data available from ISD (ISD, 2010).

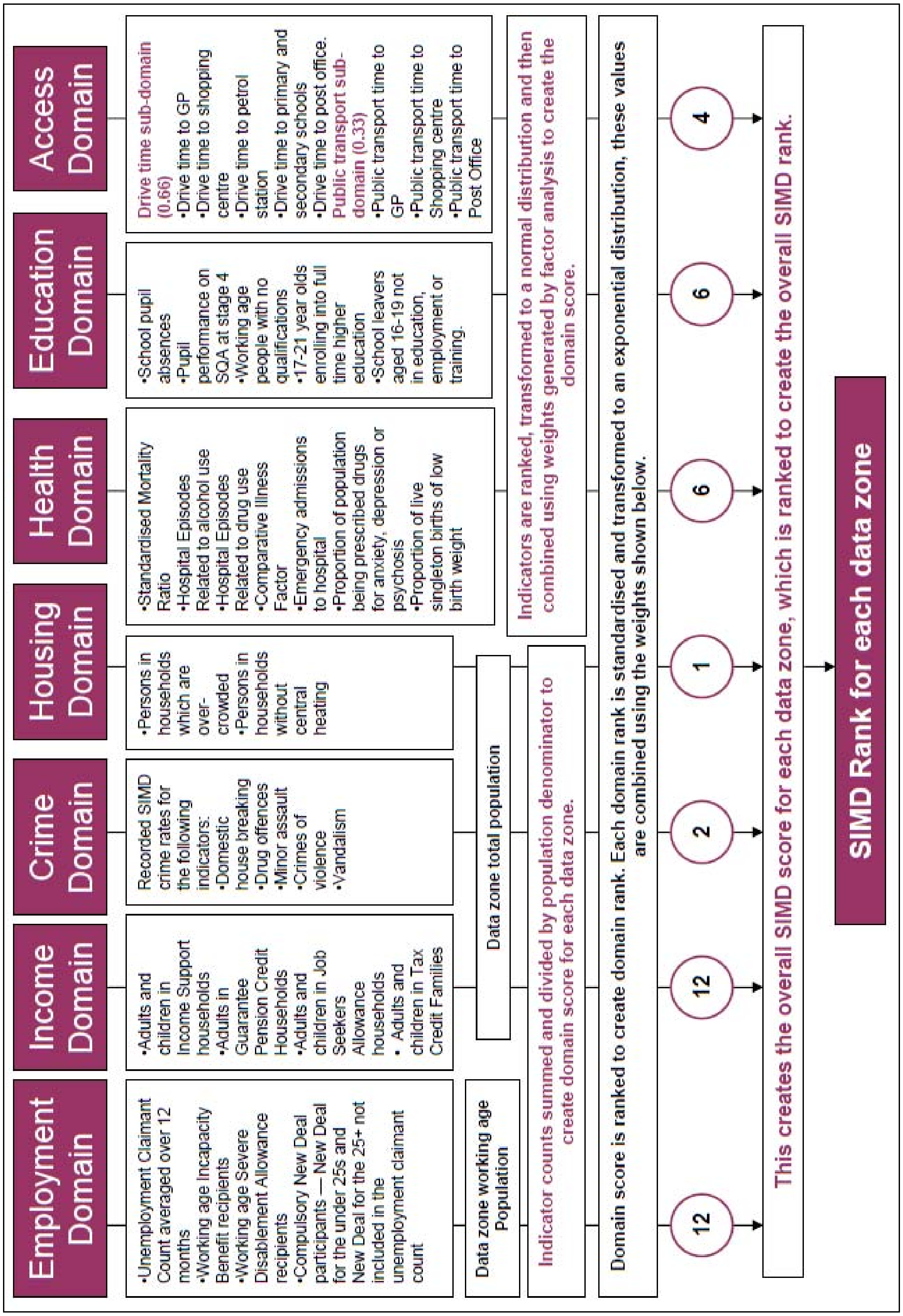
2.1.4 Deprivation Data

As discussed in Chapter 1, deprivation is an indicator that has been shown to have associations with health outcomes. The Scottish Index of Multiple Deprivation (SIMD) was developed by the Scottish Government and aims to recognise areas with higher deprivation across Scotland, providing information and statistics on the circumstances of those living in more deprived areas (SIMD, 2009a). It offers a relative measure of how deprived an area is, with Datazones in Scotland being ranked from 1 (most deprived) to 6505 (least deprived). However, this ranking system does not offer a way of determining how much more deprived one area is in comparison to another, e.g. rank two is not twice as deprived as rank four. The latest version of the SIMD was developed in 2009 and was constructed using several indicators across seven domains: health, education, employment, housing, income, access to services, and crime. The index was created by weighting each domain according to its relative importance, how robust the data were, and the time difference between collecting the data and creating the index (SIMD, 2009b). The diagram of SIMD 2009 Methodology in Figure 2.1 outlines the key steps taken in constructing the SIMD ranks for each DZ. From this deprivation index of ranks, a classification of deciles was also created where the ranks were divided into ten equally distributed percentiles grouped by areas with similar characteristics of deprivation, as shown in Table 2.2. In this index, decile 1 represents most deprived, while decile 10 corresponds to the least deprived DZ.

Table 2.2 – Allocation of deprivation deciles across Scotland

Decile (10%)	SIMD Rank From	To
1	1	651
2	652	1301
3	1302	1952
4	1953	2602
5	2603	3253
6	3254	3903
7	3904	4554
8	4555	5204
9	5205	5855
10	5856	6505

From: SNS (2010)



From: (SIMD, 2009c)

Figure 2.1 – SIMD 2009 methodology

2.2 Presentation Methods

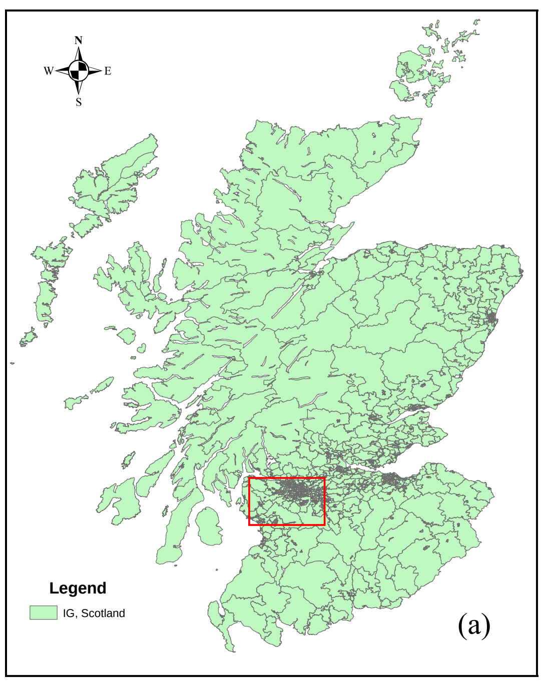
This section describes how the data were presented spatially, along with a description of the software used. An important aspect of the project was to select a geographical scale that was appropriate to the data being examined. Therefore, this section of the thesis discusses the options when deciding upon an appropriate geographical level.

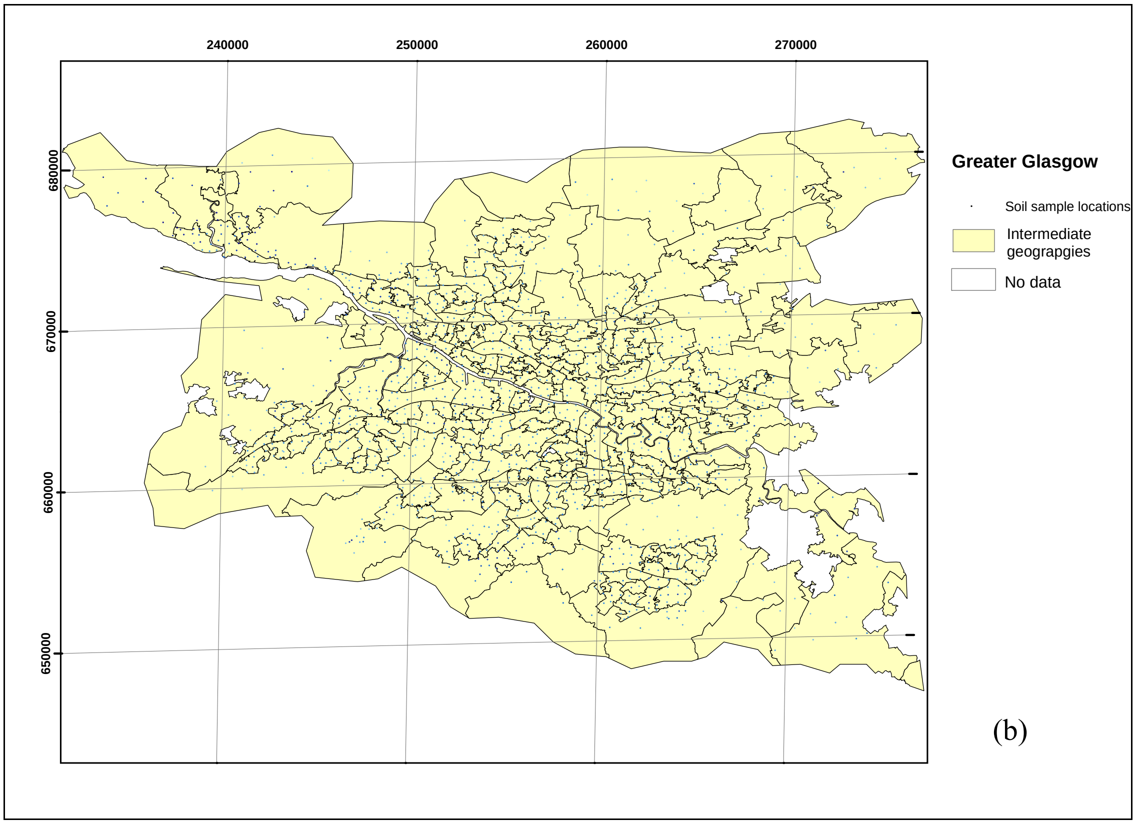
2.2.1 Spatial Presentation of Datasets

Presentation of all geographical-based maps including top soil metal distributions, air pollutants, deprivation index and health indicators for the present project were produced in a Geographical Information System (GIS) software package: ArcGIS version 9.2 (Environmental Systems Research Institute, ESRI®). A GIS is a computer-based tool for analysing and mapping spatial objects. It allows management and visualisation of geographically referenced datasets. Examples of uses of GIS software include the identification of crime hotspots, assessing the most suitable location of a wind farm, locating oil reserves etc.

2.2.2 Spatial Extent of Data

One problem in this investigation was the way in which the data for health indicators, air pollutants, land contaminants and deprivation indices were reported. The soil metal concentrations were measured at 1622 discrete point locations across the Greater Glasgow region and were spatially referenced data. In contrast, the health data were given as a count, or equivalently as a rate per 100,000, recorded for Intermediate Geography (IG) areas. Similarly, the modelled data for air NO_2 and PM_{10} concentrations were recorded at IG level. However, the SIMD (2009b) reports the deprivation indices at DZ level, which is at a finer scale than IG areas. In Scotland, there are 1235 IG areas and G-BASE soil metal concentration data for Greater Glasgow were available for 279 of them. Figure 2.2 (a) illustrates the IG areas across Scotland while Figure 2.2 (b) shows IG areas in Greater Glasgow, along with sampling locations of the G-BASE soil metals dataset. Details of the names and locations of the IG areas across Glasgow are provided in Figures A1 and A2 and Table A1 in the Appendix of this thesis. The map of Glasgow in Figure 2.2 (b) shows that in some of the rural IG areas there are very few G-BASE soil sampling points, while urban IG areas have a larger number of samples. This resulted in greater uncertainty of summary measures within the rural IG areas. The large rural IG areas around Glasgow were clipped to the spatial extent of the G-BASE soils dataset for the purposes of presentation. Intermediate Geography levels were preferred to Postcode and Datazone for the health datasets, since using these smaller geographical scales had fewer registrations per area, which could have lead to less stable incidence ratios. Intermediate Geography areas were also more suitable for air pollution data, due to the improved availability of data at this geographic scale. In terms of the soil contaminants, the most convenient approach in dealing with the difference in this data type to the health and air pollution datasets was to aggregate up the soil metal data. This was achieved by averaging the concentrations for each metal within each IG area (see Chapter 4).





Rural IG areas on the edge of Glasgow have been clipped to the extent of the soil metal dataset

Figure 2.2 – (a) Intermediate Geography areas across Scotland and (b) Map of G-BASE soil sampling locations within Intermediate Geography areas across Glasgow

2.2.3 Exploratory Data Analysis and Spatial Examination of the Soil Metal Dataset

Before looking at the spatial distribution of the metals, a preliminary analysis was necessary to consider the concentration ranges of the soil metal data. The results of this preliminary analysis are outlined in Chapter 3 and made use of summary statistics and graphical tools such as histograms and boxplots, which were produced in the statistical software package R. The process was particularly important for the exploratory geostatistical assessments that were carried out on the soil data as the geostatistical analysis of environmental features such as soil requires that the data be approximately normally distributed, since departures from normality can lead to unstable estimates and doubts about the inference and interpretation of results (Webster and Oliver, 2001). A suitable transformation was required to make the data conform to a normal distribution and stabilise the variability, before carrying out further tests.

Exploratory data analysis is also an integral part of a geostatistical analysis and is used in investigating the spatial aspects of the data as well as informally checking the assumptions of any pre-defined models. Typically this involves plotting the response variable (in this case soil metal concentration) in relation to the spatial coordinates (Easting, Northing) to check for any spatial trends and potential outliers. Such plots can also identify a possible spatially varying mean in the response. The exploratory analysis, described in Section 2.3, involved using the spatial methodology to assess the spatial variability in the data. In addition, maps of the distribution of metals in soils across Glasgow were already available from the G-BASE survey (Fordyce et al., In Prep) and examination of these distributions provided a means of identifying areas of high soil metal concentration and of potential greater exposure to metals in soil.

2.3 Geostatistical Analysis Methods

Natural geological and soil forming processes as well as man-made inputs control the distribution of metals in soil and these change over relatively short distances in the Glasgow area (Fordyce et al., In Prep). Therefore, a geostatistical analysis was useful in assessing the spatial variation in the soil. All the geostatistical analysis in this study was carried out in the software package R, using a library called geoR, which was downloaded from the Comprehensive R Archive Network (CRAN, 2009).

In the analysis of these types of data it is important to have a clear methodology to examine and treat the data before proceeding to more complex analysis. The data were summarised with the aid of tables and plots, with any features likely to cause problems in the further analysis, identified. After observing any spatial patterns in the soil metals data, the trend and distribution of the residuals was analysed and modelled using variograms to describe the spatial trend and correlation. The simplest spatial model was defined as

$$Z(s) = m(s) + \varepsilon(s), \tag{2.1}$$

where *Z(s)* was the response measurement at location *s*, *m(s)* represented the spatial trend over space *s* and *ε(s)* was the spatial error term, with possible spatially correlated errors. Spatial trend models were applied to the metals where appropriate and are outlined in Chapter 3 of this thesis.

2.3.1 Variograms

Variograms were used to assess the spatial continuity of the soil metals dataset in terms of the correlations between points at different distances apart. The results are outlined in Chapter 3 of this thesis. All variogram plots were generated in R. The underlying theory behind spatial variograms is described here. In order to make predictions in unsampled areas of a region, the spatial correlation must be estimated. In spatial problems of this type, it is likely that locations close together are more similar while those far apart are not so similar, in terms of soil metal concentration in this case. To describe the associations between the sampled points, an estimation of the covariance is required. However, the problem with the covariance is that the mean at a location cannot be estimated as only one measurement was taken at each location. To deal with this problem, the assumption of stationarity is made. A stationary process should be employed so that the distribution has attributes which do not change over space (i.e. the mean is constant or the trend is constant over space). To compute the semi-variances and estimate the variogram, a stationary stochastic process is used. This requires a spatial trend surface to be estimated, which is then removed so that the residuals are modelled to determine the spatial correlation. In general, deviation from stationarity usually indicates that the mean response will depend on geographical location, and if there is evidence that a directional trend is present for a particular parameter (soil metals in this case), the spatial trend should be eliminated by adding a suitable polynomial to the model and examine the variogram of the residuals. This detrending process allows assessment of the spatial correlation effects present in the data without the interference of a trend. A linear or quadratic trend should provide a satisfactory description of the spatial trends and one should avoid using higher degree surfaces since complex trends can be described by the stochastic component in a model (Diggle and Ribeiro, 2007). It can sometimes be very difficult to distinguish what type of trend is present in the data, or even whether an estimated trend ought to be removed at all, before making assessments on the spatial correlation. However, the variance between points does depend on the distance between them so that the covariance is assumed constant for a given lag *h*. This dependence between values with a difference in lag is known as the autocovariance function, which has a simple relationship with autocorrelation. The semi-variance, which is used to compute the variogram, can be related to the autocorrelation (Webster and Oliver, 2001).

The variogram is the expected squared difference (in the response values) between the coordinates at two locations, and is the variance per point when considered as pairs. The distance measured is typically the Euclidean distance. The variogram summarises the spatial relations in the data by obtaining an ordered set of values of semi-variances at different distances (Webster and Oliver, 2001). The semi-variance is denoted by γ and is calculated for each and every pair of points, say x_i and x_j, given by equation 2.2.

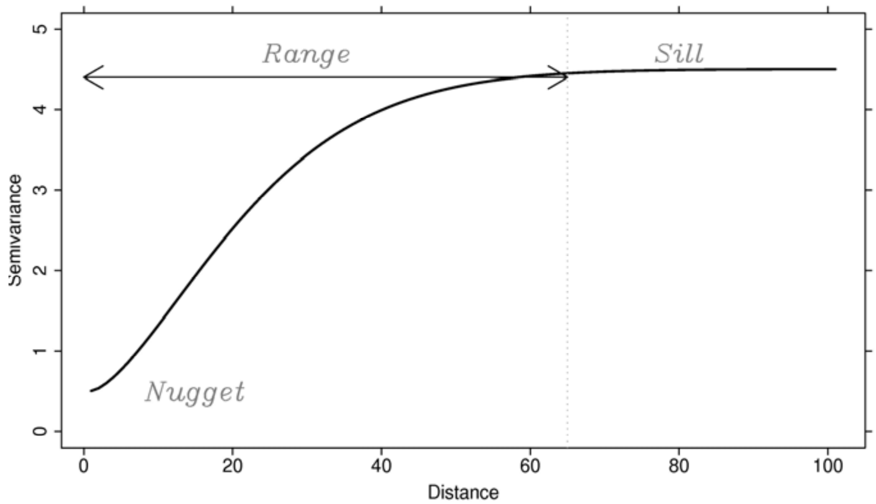
$$\hat{\gamma}(h) = \frac{1}{2m(h)} \sum_{i=1}^{m(h)} \left[z(x_i) - z(x_i + h) \right]^2 , \tag{2.2}$$

where h is the distance between x_i and x_j and $m(h)$ is the number of pairs of points separated by distance h. The semivariances are then plotted against the corresponding distance between the locations to form an empirical variogram plot (see Figure 2.3). When the semi-variance reaches the sill asymptotically, the autocorrelation between observations becomes zero at this distance. An empirical variogram is a useful tool in geostatistical analysis and can form the basis of what type of model would be most appropriate in describing the spatial relationships in the data.

2.3.2 Modelling the Variogram

In the present study, when detrending the data by applying a suitable polynomial to the coordinates, sometimes this was not enough to eliminate the spatial correlation present in the data, in which case any remaining spatial correlation was modelled. Therefore, the spatial continuity was only modelled for data where there was a significant amount of correlation in the soil metal dataset not captured by a trend surface.

However, some properties of the empirical variogram demonstrated that this method was not entirely satisfactory. The variogram is calculated in such a way that each semivariance is only an estimate of the average for that particular lag, and is therefore subject to some variability (Webster and Oliver, 2001). As such, variograms in the present study produced erratic behaviour, particularly at larger distances, as illustrated in the empirical variogram plots in Chapter 3. Where appropriate, the true variogram was estimated by a suitable model and this characterised the regional variation. This was achieved by imagining a general trend in the empirical variogram plots and estimating the model parameters by eye. The three parameters are illustrated in Figure 2.3 and include the signal variance σ^2 (partial sill), measurement error variability τ^2 (nugget), and correlation function ϕ (distance parameter, where the range is said to be a function of ϕ). The theory states that the sill parameter is where the variogram reaches an upper bound and the range is defined as the distance at which the variogram almost reaches the sill (Webster and Oliver, 2001). The partial sill is the difference between the sill and the nugget variance. The range distance represents the limit to where locations are spatially dependent, while the nugget effect signifies a discontinuity at the origin of the variogram and is due to either measurement error or spatial variability on a smaller scale than the distance between the two closest points in the sampling region (Diggle and Ribeiro, 2007). When creating models to describe the spatial continuity, an effective correlation range is produced to represent the distance at which 95% of the spatial covariance is accounted for and this estimated range from the model gives an indication of the scale within which concentrations are likely to be correlated (Shinn et al., 2000).



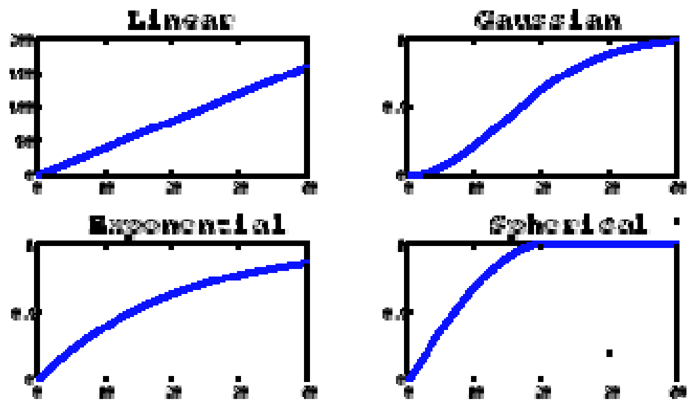
From: (School of Geosciences, 2009)

Figure 2.3 – Empirical variogram parameters

The theory behind variogram modelling is outlined here. To model the variogram one should fit an appropriate but simple function that represents the empirical variogram, which satisfies the functions of sill, nugget and range as estimated from the plots. The most common models used in geostatistics are the Linear, Spherical, Gaussian and Exponential, which are illustrated in Figure 2.4 below. Gaussian models are widely used for geostatistical data as they capture a range of spatial patterns according to the correlation structure specified and can also be applied to transformed data (Diggle and Ribeiro, 2007). The Gaussian model for the variogram is defined by equation 2.3.

$$\gamma(h) = \sigma\left\{1 - \exp\left(\frac{-h^2}{\phi^2}\right)\right\}, \tag{2.3}$$

where σ and ϕ represent the sill and distance parameter respectively, and h is the distance between the two point locations.



From: (Sierra, 2001)

Figure 2.4 – Variogram model examples

A similar model to the Gaussian which is also widely used in geostatistics is the negative exponential, defined by equation 2.4. In this model the variogram function tends towards the sill asymptotically and approaches the origin at more of a slope than the Gaussian model.

$$\gamma(h) = \sigma\left\{1 - \exp\left(\frac{-h}{\phi}\right)\right\}. \tag{2.4}$$

A variogram which increases and does not have a sill can be represented by the linear variogram model. Another feature of the linear model is that the range is random as it depends on the greatest distance in the variogram. This is the simplest of the variogram models and is given by equation 2.5.

$$\gamma(h) = h\left(\frac{\sigma}{\phi}\right). \tag{2.5}$$

The spherical model has a general shape which is similar to the Exponential and Gaussian, where the strong spatial correlation becomes less apparent as the distance between locations increases. However Spherical models do not approach the sill as asymptotically as these two models. This model is also illustrated in Figure 2.4 and is described by the following equation.

$$\gamma(h) = \sigma\left\{\frac{3}{2}\left(\frac{h}{\phi}\right) - \frac{1}{2}\left(\frac{h}{\phi}\right)^3\right\} \qquad \text{for } h \le \phi \qquad (2.6)$$

$$\gamma(h) = \sigma \qquad \text{for } h > \phi$$

Since distances are always positive, the covariance function for any chosen model should be positive definite. Furthermore, the correlation between locations should decrease as the distance between them increases. The Matérn family of correlation functions satisfies these properties and is defined by equation 2.7.

$$\rho(d) = \frac{(\mu / \phi)^{\kappa} k_{\kappa}(\mu / \phi)}{2^{\kappa-1}\Gamma(\kappa)} \qquad (2.7)$$

where μ is a vector containing the values of the distances between pairs of points, ϕ is the value of the range parameter, and κ is a smoothness parameter. This function is a special case where a sensible model can be selected to describe the spatial correlation. The constant chosen for kappa (κ) influences the relationship between the range and the sill. When $\kappa = 0.5$, the Matérn correlation function is represented by the Exponential function while the Gaussian correlation function is approached as κ tends to infinity. In general, the smoothness of the model increases with increasing κ. When using a large value for κ (>2), the Gaussian model is too smooth and is unrealistic in determining predictions. On the other hand, using a small value such as 0.5, the Exponential is too sensitive in modelling the correlation structure. In most cases, it is suitable to choose a value somewhere between these, where Diggle and Ribeiro, (2007) suggest approximately 1.5.

In the present study a variety of variogram models were used, with a Gaussian and Exponential always considered. The kappa value was also suitably adjusted to give a correlation structure somewhere between where it was not too smooth or too responsive. For all models, ordinary least squares was used to fit the variogram model as this was the default option of the parameter estimation in the variofit command of the library geoR.

The theory states that a variogram model is isotropic if the covariance between any two locations, x_i and x_j, depends only on the distance between them (Diggle and Ribeiro, 2007) and is anisotropic if there are directional differences present. When features at nearby locations appear to be correlated, this is referred to as spatial dependency, while points located further apart tend to be spatially independent. Monte Carlo (MC) confidence envelopes are used to assess the significance of spatial dependence in the variogram and are constructed by simulating independent realisations from the data values, based on their spatial location. This is carried out by randomly allocating data values to the spatial locations, where the minimum and maximum values from the simulations are taken as the lower and upper confidence bands of the MC envelopes. The appropriate function is R automatically computes the MC confidence bands. When the points in the variogram fall outside the confidence bands of the MC plot, one rejects the null hypotheses of spatial independence in favour of the alternative, that there is some spatial correlation present in the data. A personal judgement needs to be made on these MC plots as one would certainly expect there to be some spatial correlation present in the variogram, even when most of the spatial trend is removed. In cases where a more complex trend does not have a considerable effect on the data, it is more sensible to estimate the parameters from the simpler variogram. As mentioned previously, a linear or quadratic trend should be adequate in describing the trend component. However, when the removal of a trend effect still reveals a significant amount of spatial correlation, a suitable model is produced from the estimated parameters.

In this study, the observed parameters from the most suitable variogram were used to make predictions of soil metal concentrations using kriging. However, this interpolation method only applied when the variogram could be modelled adequately and the nugget was sufficiently small. When the MC plots showed that spatial independence could not be rejected and indicated no significant spatial correlation, interpolation methods were not appropriate. Where a constant, linear or quadratic trend explained most of the correlation in the data, predictions were not undertaken.

2.3.3 Kriging

Kriging is a useful tool in the field of geostatistics to predict parameter values at locations that have not been sampled and convert point source data to continuous surface data. This approach simply provides a visual representation of the predictions, which can be used to aid data interpretation. Kriging uses the available data and takes into account the variability in the spatial features of the variogram model (Webster and Oliver, 2001). In this study, ordinary kriging was used in which the estimates were linear combinations of the data and weighted according to values at nearby locations. This approach minimises the variability of the kriging results and yields unbiased predictions. Where spatial correlation was not present in the soil metals data, kriging was not appropriate. Although the kriging analysis was carried out in R, the maps illustrating predictions across Glasgow were produced in ArcGIS.

The theory behind ordinary kriging is that a constant but unknown mean is assumed and the covariance parameters, from the estimated variogram model as described in 2.3.2, are treated as known. Let x_0 be an unsampled location in the region, then the predicted concentration $\hat{Z}$ at this unsampled location is given by equation 2.8.

$$\hat{Z}(x_0) = \sum_{i=1}^{N} w_i z(x_i) \qquad (2.8)$$

The weights w_i must sum to one to ensure that the estimates are unbiased, and $\hat{Z}(x_i)$ are the measured concentrations at each point location. These estimates are also allocated a kriging variance, which is given by equation 2.9.

$$\mathrm{var}\left[\hat{Z}(x_0)\right] = 2\sum_{i=1}^{N} w_i \gamma(x_i, x_0) - \sum_{i=1}^{N}\sum_{j=1}^{N} w_i w_j \gamma(x_i, x_j) \tag{2.9}$$

where $\gamma(x_i, x_0)$ is the semivariance of the concentration $\hat{Z}$ between the point x_i and the unsampled location x_0, while $\gamma(x_i, x_j)$ is the semivariance between two locations x_i and x_j (Webster and Oliver, 2001).

2.4 Modelling and Inference Methods to Assess Relationships between Health, Deprivation and Environmental Parameters

This section discusses the methods used in the defining and selection of statistical tests and models along with the process of checking necessary assumptions to assess the relationships between the health, deprivation and environmental parameters. Before any modelling was carried out, some statistical methods to assess patterns within the variables were applied as well as an assessment of the correlation between the variables. All modelling methods and plots of assumption checking were carried out in R.

2.4.1 Pearson's Correlation Coefficient

As part of the exploratory analysis of the spatial relationships between parameters in this project, it was of interest to assess the strength of association between the different metals in the soil as well as between soil metals and the other environmental and health variables. This was done using Pearson correlation coefficients. The Pearson product moment correlation coefficient is a simple method to assess the degree of correlation between two variables and measures the linear dependence between two covariates. A value in the range [-1, 1] is generated where zero represents no correlation and -1 and 1 indicate perfect negative and positive correlation respectively. The correlation coefficient is given by formula 2.10.

$$\rho = \frac{E[(X - \mu_X)(Y - \mu_Y)]}{\sigma_X \sigma_Y} \tag{2.10}$$

2.4.2 Moran's I

Another aspect of the environment and health indicators which was assessed as part of the exploratory spatial analysis was any occurrence of spatially determined patterns. For example, one would like to know whether concentrations of air NO_2 occur randomly across Greater Glasgow, or if there is a behaviour which appears to show some form of clustering. Moran's I test was carried out on the datasets used in this study to assess the randomness of any spatially determined patterns.

Moran's I tests for spatial autocorrelation in the data, and is often used as an indicator of spatial dependence in geographically referenced data (Bivand et al., 2009). It tests the significance of a spatial pattern or structure with reference to a statistical distribution, obtained by randomly generating observed values (Mitchell et al., 2007). The test of Global Moran's I was carried out in ArcGIS and measures spatial autocorrelation in the form of overall clustering from characteristics at spatial locations (ArcGIS Resource Centre, 2010). Moran's I statistic is defined as a ratio of quadratic forms in the normally distributed residuals from a regression of *y* on *X* (Bivand et al., 2009), with the estimated residuals given by equation 2.11.

$$\begin{aligned} \hat{\varepsilon} &= (I - X(X^T X)^{-1} X^T) y \\ &= Ay \end{aligned} \tag{2.11}$$

The estimated residuals $\hat{\varepsilon}$ are then used to compute the Local Moran's I statistic, given by equation 2.12, where V_i is a local spatial link matrix and Global Moran's I is the sum of the Local Moran's I_i's.

$$I_i = \frac{\hat{\varepsilon}^T \cdot V_i \cdot \hat{\varepsilon}}{\hat{\varepsilon}^T \cdot \hat{\varepsilon}} \tag{2.12}$$

The null hypothesis of the Global Moran's test is an indication of spatial independence so that the spatial pattern may be due to random chance and could therefore be dispersed across the region. The alternative hypothesis indicates spatial dependence of observations which may imply a spatial structure that is not random. The distributions under both hypotheses are required for the test if size and power are to be assessed, and the normal approximation is the most commonly used distribution (Bivand et al., 2009). However, the test does not offer a solution to where particular clusters are located, or whether the correlation arises between high or low concentrations (Mitchell et al., 2007).

2.4.3 Generalised Linear Models

Regression modelling was used in formally analysing the relationships between the health indicators and the environmental variables in this study. To allow more flexibility in respect to the distribution of the health outcome of interest, Generalised Linear Models (GLMs) were used. Theory states that a GLM is a generalisation of ordinary least squares regression which allows the model to follow a distribution for the response, other than the normal approximation, and is achieved via a link function. The model consists of a linear predictor and also assumes that observations are independent and follow some exponential family distribution (Wood, 2006). A GLM is of the following structure

$$g(E(Y_i)) = X_i \beta \tag{2.13}$$

where g represents a link function, X_i is the i^{th} row of the model matrix and β is a vector of unknown parameters. The right hand side of the equation is known as the linear predictor. The link function indicates the relationship between the mean of the distribution and the linear predictor. Distributions which are useful for modelling include the Normal, Binomial, Poisson, Gamma and Inverse Gaussian, where each of these are allocated a suitable link function such as the log, logit, inverse or identity (Wood, 2006). When a model uses the normal distribution with the identity link function, it represents an ordinary linear model. The distribution allocated to the response variable must be a member of the exponential family, which is represented in the form of equation 2.14.

$$f_\theta(y) = \exp\left[\frac{y\theta - b(\theta)}{a(\phi) + c(y,\phi)}\right] \tag{2.14}$$

where *a*, *b* and *c* are arbitrary functions, ϕ is a scale parameter and θ is the link function of the parameter.

2.4.4 Standardised Residuals

An integral part of statistical modelling is the process of model checking and this was carried out in the present project using the following method. In generalised linear models, the assumptions of mutually independent responses and constant variance should be assessed, while normality should be checked if the model follows a Gaussian distribution. The standardised residuals are examined due to the difficulty in assessing the validity of the relationship between the mean and variance from the raw residuals (Wood, 2006). The most common approach to standardising the residuals is by the Pearson method, where the standardised residuals are given by equation 2.15.

$$\hat{\varepsilon}_i^p = \frac{y_i - \hat{\mu}_i}{\sqrt{Var(y_i - \hat{\mu}_i)}} \tag{2.15}$$

The residuals $\hat{\varepsilon}$ should have a mean of zero and a constant variance ϕ when plotted against the fitted values. Residual versus fitted values plots are assessed to check the assumptions of mean equal to zero and constant variance. The plots should not reveal any trends in mean or variance for these assumptions to hold. Normality plots are used to check the assumption that the errors are approximately normally distributed.

2.4.5 Akaike's Information Criterion (AIC)

A common technique in the selection of GLMs is the Akaike's Information Criterion (AIC). Since it can be difficult to decide which model best describes the associations in the data, this method was used in the selection of the most appropriate models for the present project. AIC measures the goodness of fit of a model based on the number of parameters and the maximum likelihood function of the model. It is computed by equation 2.16.

$$AIC = 2p - 2\ln(ML) \tag{2.16}$$

where p is the number of parameters in the model and ML is the maximum likelihood of the model (Wood, 2006). The AIC is appropriate to estimate the fit of the model since it penalises the addition of an unnecessary parameter, which prevents over-fitting. The most suitable model is chosen by the minimum value of AIC.

2.4.6 Relative Risk

In order to make inferences about the associations between soil metals and the distribution of lung cancer and respiratory disease across Greater Glasgow, relative risk (RR) estimates were computed to give an approximate increased risk of developing these health outcomes if exposure to soil metals was to increase in any given region. Relative risk methods are commonly used in studies of air pollution with health outcomes, such as in Mitchell et al., (2009) and this method was applied to the soil contaminants in the current study. However, these RR estimates should be treated with caution. Although they may imply a cause-effect relationship between soil metals and health outcomes, this is not necessarily true since the health effects could also be explained by other covariates such as smoking, for which data were not available in this study.

In the context of the current study, RR estimated the risk of developing the health outcome of interest, relative to exposure to the soil metal concentration. The RR estimate for a particular soil contaminant was computed by taking the exponential of the model coefficient of exposure due to a realistic increase in a particular soil metal concentration, as shown in equation 2.17, whilst the 95% confidence interval for the RR was given by equation 2.18. A table of the increases in soil metal concentrations assumed for the study is provided with the analysis in Chapter 5. A confidence interval which is entirely greater than one implied a statistically significant health association with the particular soil metal covariate included in the model.

$$\exp[increase \times coeff] \qquad (2.17)$$

$$\exp\left[increase \times \left(coeff \pm 1.96 \times Std.Error\right)\right] \qquad (2.18)$$

Chapter 3

Exploratory Spatial Analysis of Soil Metals Dataset

3.1 Preliminary Analysis

This chapter presents the results of the geostatistical analysis of the soil metals dataset. As mentioned in the previous chapter, there were 1622 soil concentrations recorded at point locations for each of the eight elements. All the elements were measured in milligrams per kilogram (mg kg^{-1}) with the exception of potassium (K_2O), which was recorded as a weighted percentage of the oxide (wt %). This element was only included in the study as a control since it is not a threat to human health in soils and its distribution in Glasgow soils is different to the other metals as it is primarily controlled by the underlying geology (Fordyce et al., In Prep).

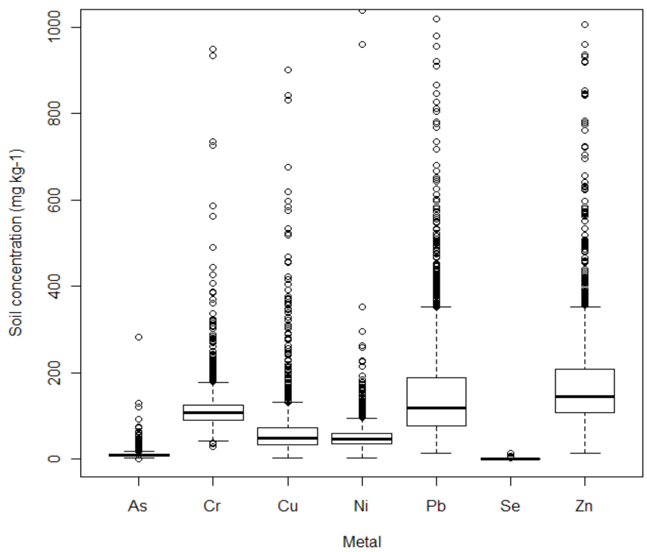
3.1.1 Descriptive Statistics

Table 3.1 contains the summary statistics for each of the elements, which show large differences in the concentration as expected, given differences in their typical geochemical abundances in the environment. For instance, Se ranges from 0.1 to 14.5 mg kg^{-1} while Cr ranges from 28 to 4286 mg kg^{-1}. Chromium, Pb, and Zn have the highest median soil concentrations of the metals in the Greater Glasgow region. The maximum observed concentration of the eight elements was a measurement of 5001 mg kg^{-1} for Pb. The standard deviations demonstrate a large amount of variability in the data and for many metals the standard deviation was at least as large as the mean value.

Table 3.1 – Summary statistics for soil metal concentrations in the G-BASE Glasgow dataset

Element	N	Mean	St.Dev	Min	Q1	Median	Q3	Max
As	1622	10.79	10.57	1.10	7.30	9.10	11.43	282.80
Cr	1622	121.57	130.31	28.47	91.00	107.00	126.00	4286.00
Cu	1622	70.53	120.02	2.90	33.58	47.85	72.90	3679.90
Ni	1622	52.56	43.75	2.30	35.00	45.70	59.10	1038.10
Pb	1622	167.91	210.49	13.40	77.58	118.30	187.65	5001.00
Se	1622	1.03	0.68	0.10	0.70	0.90	1.20	14.50
Zn	1622	189.35	175.37	13.66	107.47	144.41	207.28	1780.80
K_2O	1622	1.36	0.26	0.26	1.21	1.32	1.48	3.140

Figure 3.1 illustrates the distribution of the soil metal concentrations, with K_2O omitted as it was measured on a different scale (wt %) to the other elements. A limit of 1000 mg kg^{-1} was set for the y-axis of the plot, simply for observational purposes. Lower and upper quartiles of the boxplots represent the 25^{th} and 75^{th} percentile distribution respectively. The plot supports some of the features of the data distributions outlined in the previous paragraph. However, it also highlights the highly positive skewed distribution of the elements and the number of outliers present. This skewed distribution suggested that the data required a transformation to achieve approximate normality before the exploratory geostatistical analysis could be carried out.

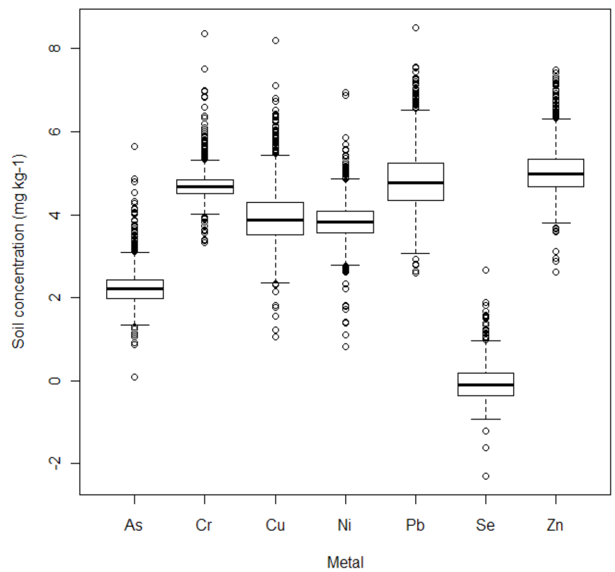


n = 1622. The box shows the 25th and 75th percentile and whiskers show the 10th and 90th percentiles.

Figure 3.1 – Boxplot of soil metal concentrations in the G-BASE Glasgow dataset

3.1.2 Transformation

The analysis of geostatistical data is more efficient when it is performed on data which are normally distributed. The boxplots in Figure 3.1 demonstrated a considerable departure from symmetry which can cause difficulties in the analysis. When the data are asymmetrical there are more likely to be dissimilar variances in different parts of the study region. A common transformation used in the analysis of this type of highly skewed data is taking the logarithm of the response. Applying a natural log transformation to the concentrations (see Figure 3.2) reduced the skewness in the data and yielded approximate normal distributions for each metal. Note again that K_2O is not included in Figure 3.2 as it was measured as weight percentage. The summary statistics for the log-transformed data are presented in Table 3.2.



n = 1622. The box shows the 25th and 75th percentile and whiskers show the 10th and 90th percentiles.

Figure 3.2 – Boxplot of log soil metal concentrations in the G-BASE Glasgow dataset

Table 3.2 – Summary statistics for log soil metal concentration in the G-BASE Glasgow dataset

Element	N	Mean	St.Dev	Min	Q1	Median	Q3	Max
As	1622	2.2477	0.4373	0.0953	1.9879	2.2083	2.4358	5.6447
Cr	1622	4.7024	0.3558	3.3322	4.5109	4.6728	4.8363	8.3631
Cu	1622	3.9477	0.6784	1.0647	3.5138	3.8681	4.2891	8.2106
Ni	1622	3.8304	0.4826	0.8329	3.5553	3.8221	4.0792	6.9451
Pb	1622	4.8171	0.7213	2.5953	4.3512	4.7732	5.2346	8.5174
Se	1622	-0.1002	0.5039	-2.3026	-0.3567	-0.1054	0.1823	2.6741
Zn	1622	5.0356	0.5871	2.6174	4.6773	4.9726	5.3340	7.4848
K2O	1622	0.2881	0.1899	-1.3471	0.1906	0.2776	0.3920	1.2296

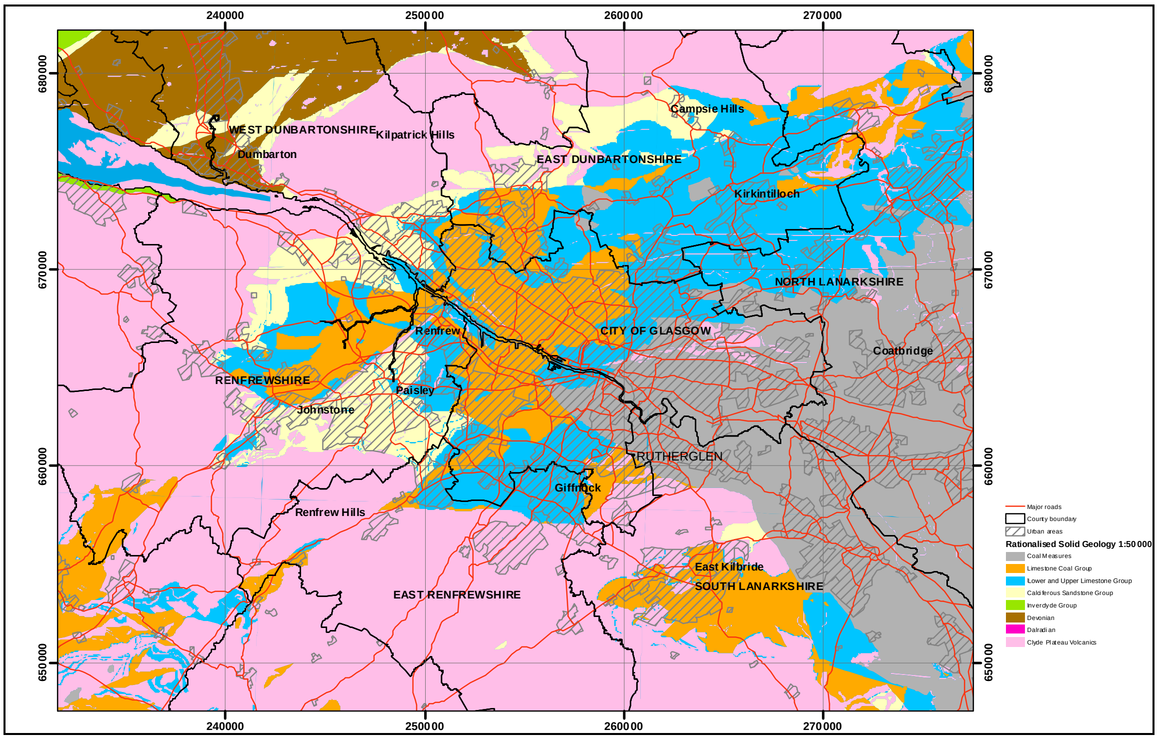
3.2 Soil Metals Exploratory Spatial Analysis

3.2.1 Overview

The remainder of this chapter presents the results of the exploratory spatial analysis of the eight soil metals, including examining possible spatial trends; assessing the spatial behaviour across the region with the use of variogram plots; modelling the variogram where there was spatial correlation present and generating predictions for unsampled locations based on the chosen correlation structure.

Maps showing the spatial distribution of the elements in top soils across Glasgow were already available as a result of the G-BASE survey of Glasgow (Fordyce et al., In Prep). These are based on the actual rather than the log transformed data with the individual sampling locations allocated different sizes and a colour scale, according to the soil metal concentration. These were included in the present study to show the distribution of soil metals across Glasgow. Brief descriptions were also included for each soil metal indicating the rock types and land use with which they are associated. These were taken from the work carried out by Fordyce et al. (In Prep). The bedrock geology of Greater Glasgow is also reproduced from the Fordyce et al. (In Prep) study in Figure 3.3 and was helpful in placing the soil geochemical distributions into context in terms of associations with particular rock types in the area.

For all further geostatistical analysis presented in this chapter, the log transformed data were used. It is generally very difficult to detect any directional trends by observing spatial maps, so a trend was fitted to the log transformed data allowing for a spatially varying mean for each of the elements. Careful consideration was taken of how to carry out the geostatistical analysis on the soil metals data in the present study. The resulting variograms in this section demonstrated erratic behaviour at larger distances, which presents a slight concern in interpretation. This is quite common in variogram plots and is likely to be caused by the sampling design as the sampling density is much lower in rural areas than in urban areas in the G-BASE dataset, so that there are fewer pair-wise points to be computed at greater distances. One possible approach to overcome this would have been to set a limit to the maximum range in the variogram. This 'binning' approach is illustrated in Diggle and Ribeiro, (2007) and can be defined using the *uvec* or *max.dist* command in R. They suggest that using the full extent of the variogram range is not entirely helpful since the estimates are unstable at larger distances. However, this approach overlooks data that may be of importance. In the context of this study, another plausible method could have been to include the urban data only, but this would discount 241 rural samples, which accounts for approximately 15% of the G-BASE soil dataset. Based on these grounds, both the rural and urban top soil data were included as part of the geostatistical analysis and it was accepted that some irregular behaviour would occur at larger distances.



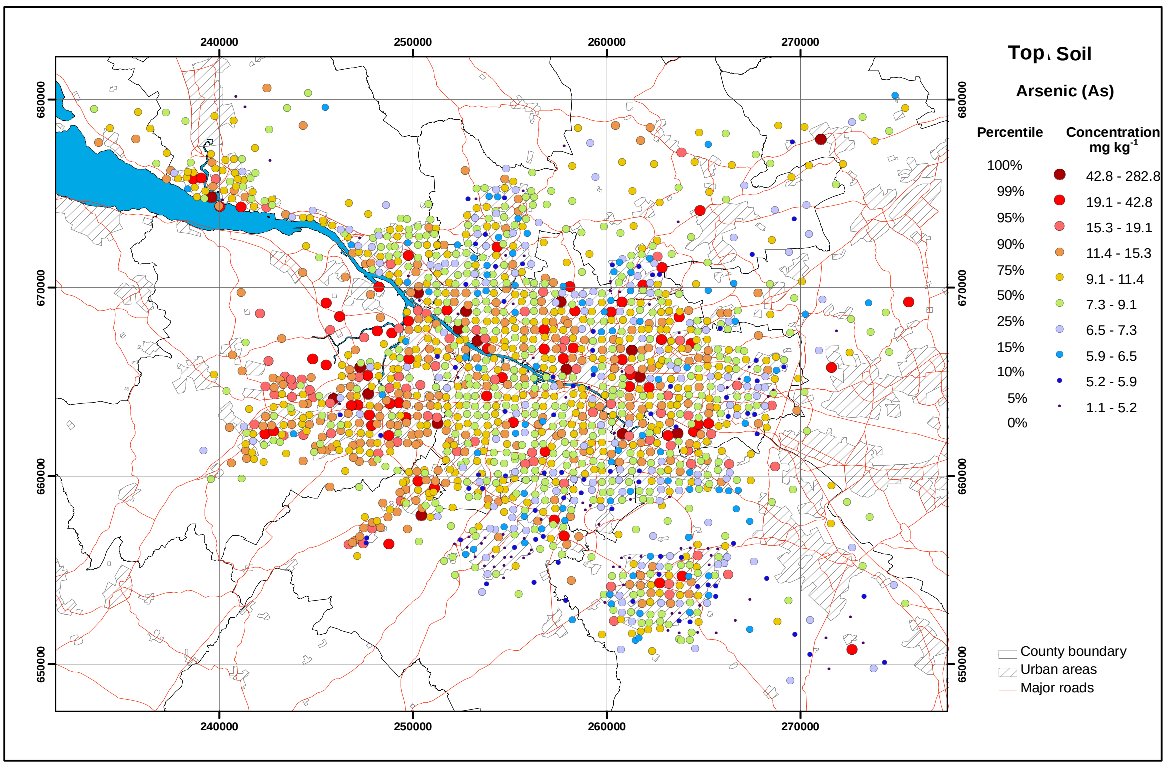
Grid squares represent 10 km
Linework derived from OS topography © Crown Copyright. All rights reserved. BGS 100037272/10
(From: Fordyce et al., In Prep) BGS © NERC

Figure 3.3 - Simplified bedrock geology map of Greater Glasgow from BGS DigiMap® data.

3.2.2 Geostatistical Assessments of Soil Metal Data

3.2.2.1 Arsenic (As)

Soil As concentrations across Greater Glasgow are indicative of the bedrock over which soils are developed and reflect man-made contamination. Arsenic is associated with peaty soils and with man-made ground in the south and south-west of Glasgow, and As values are also elevated in the East End where heavy industry was present in the past and with the former shipbuilding areas along the River Clyde (Fordyce et al., In Prep) (Figure 3.4). From Table 3.1, the mean soil As concentration is 10.8 mg kg^{-1} and the maximum recorded is 282.8 mg kg^{-1}.



Grid squares represent 10 km
Linework derived from OS topography © Crown Copyright. All rights reserved. BGS 100037272/10
(From: Fordyce et al., In Prep) BGS © NERC

Figure 3.4 – Map of As top soil concentrations in the G-BASE Glasgow dataset

Using geostatistical analysis and exploring a possible spatial trend more carefully, the log-transformed As soil concentrations were plotted against each of the spatial coordinates with a loess line added to help highlight any trend, as shown in Figures 3.5 (a) and (b). Overall there did not appear to be a strong difference in ln(As) soil concentration in different directions. Therefore, one would expect a constant mean concentration across the region.

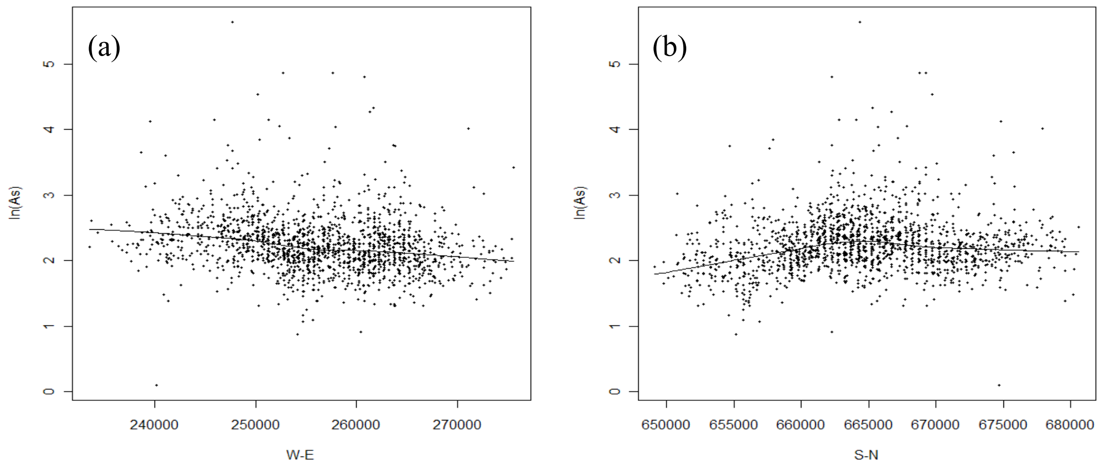


Figure 3.5 – ln(As) soil concentration plotted against each of the directional coordinates (a) west-east and (b) south-north, with a fitted loess line

Assessing the presence of a trend effect more formally, through fitting a linear regression model on the coordinates, the p-values in Table 3.3 for both easting and northing coordinates were statistically significant, indicating that a trend effect was present and should be removed. This did not agree with the initial impressions outlined in the previous paragraph of this thesis.

Table 3.3 – Regression on ln(As) soil concentration with directional coordinates

Coefficients	**Estimate**	**Std.Error**	**P-value**
Intercept	2.456	1.29	0.0571
Easting	-0.00001	0.0000014	<0.001
Northing	0.0000038	0.0000017	0.0256

Figure 3.6 (a) shows the experimental variograms for the original ln(As) soil concentrations and the observed residuals from a first order linear trend , with the latter fitted by ordinary least squares (OLS). The variogram containing the linear trend was fitted by ordinary least squares (OLS) as the specified default method. As discussed in Chapter 2, the variogram of a first

order polynomial allows us to investigate the spatial correlation in the region having removed the presence of a trend effect.

Erratic behaviour was seen in the variograms as the distance between points increases above 35 km. The variogram for the original data suggested that the variance between the locations only increased gradually as the distance between them increased. However, Table 3.3 showed that the p-values for the coordinates were statistically significant at the 5% level, so that a trend effect was indeed present for ln(As) and should be removed. For the detrended data where a linear trend effect was removed, the variogram appeared constant and all the points lay within the Monte Carlo (MC) confidence bands (Figure 3.6 (b)). Therefore, after the removal of a linear trend, ln(As) soil concentrations were spatially independent with no significant spatial correlation in the residuals. A linear trend was sufficient in describing the spatial correlation of ln(As) concentrations across Greater Glasgow and the residuals from a linear trend were log-normally distributed with mean zero and variance σ^2. Since no significant spatial correlation was present, there was no need to model the variogram and kriging was not carried out.

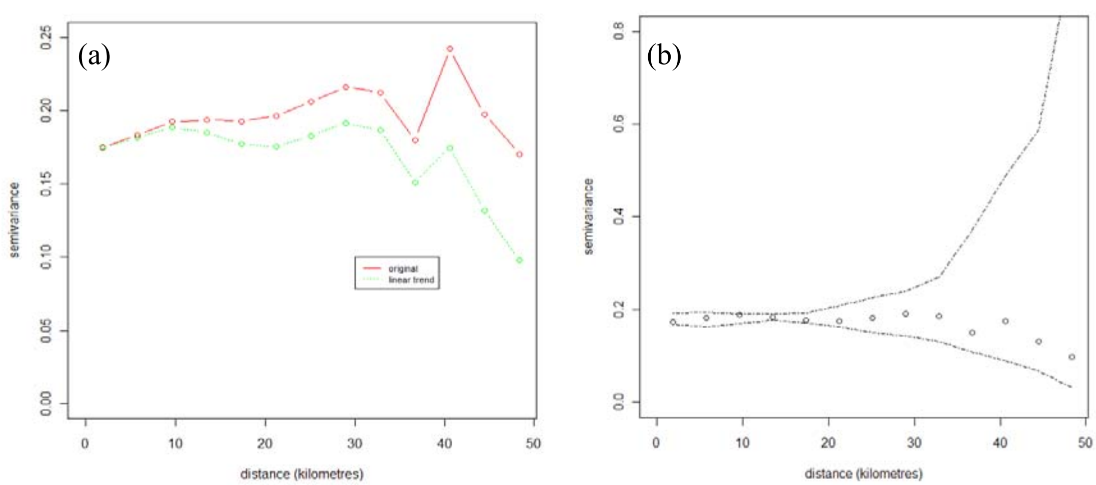
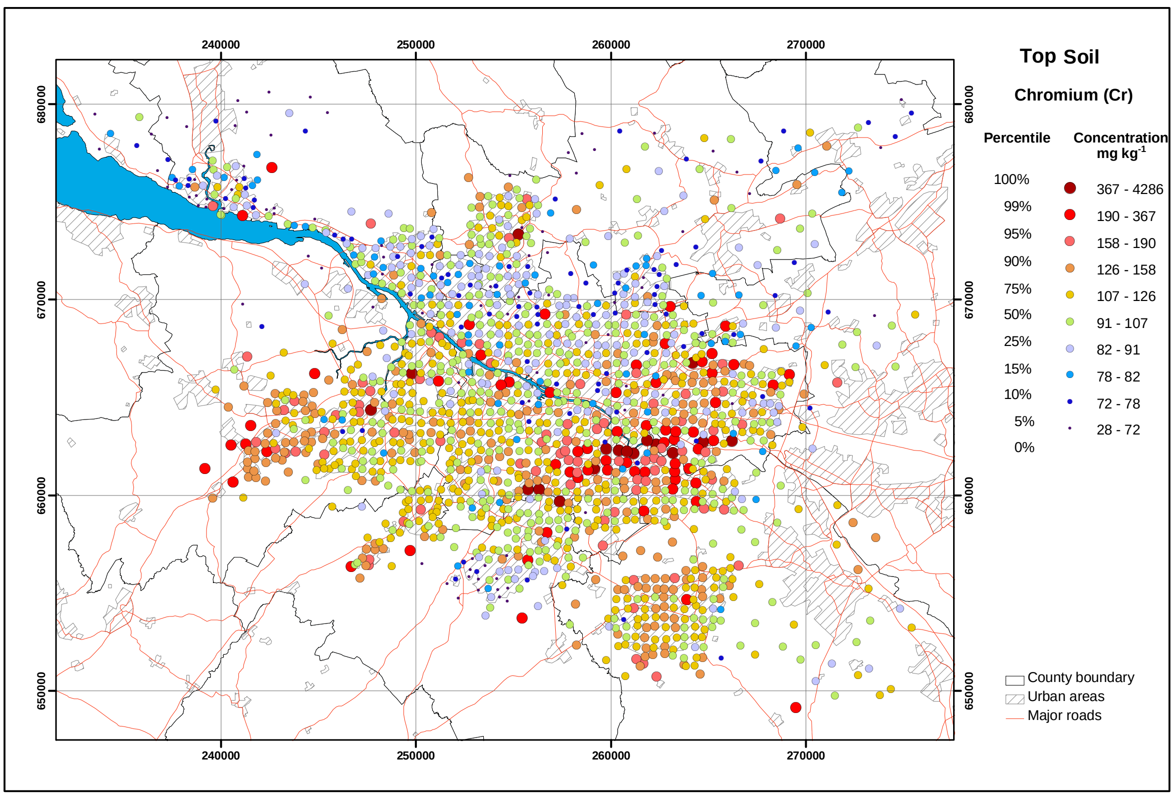


Figure 3.6 – (a) Empirical variograms for original data and removal of linear trend and (b) MC envelope for removal of a linear trend ln(As) soil variogram

3.2.2.2 Chromium (Cr)

Due to the former presence of heavy industries including foundries and a Cr-ore processing works, high concentrations of soil Cr are clustered in the east and south of Glasgow in areas such as Muirend, Rutherglen and Shettleston (Fordyce et al., In Prep) (Figure 3.7; see Figures 3.3 and A1 and A2 and Table A1 in the Appendix for locations.). Chromium content is lower in soil over Devonian Sandstones in Dumbarton (see Figure 3.3 for locations) and is elevated in the south of Glasgow due to the volcanic rock formations in that area (Fordyce et al., In Prep; Figure 3.7). From Table 3.1, the mean soil Cr concentration in Greater Glasgow is 121.6 mg kg^{-1} and the maximum is 4286 mg kg^{-1}.



Grid squares represent 10 km
Linework derived from OS topography © Crown Copyright. All rights reserved. BGS 100037272/10
(From: Fordyce et al., In Prep) BGS © NERC

Figure 3.7 – Map of Cr top soil concentrations in the G-BASE Glasgow dataset

Plotting the ln(Cr) soil values against each of the directional coordinates, Figures 3.8 (a) and (b) did not reveal any strong spatial trends across Glasgow.

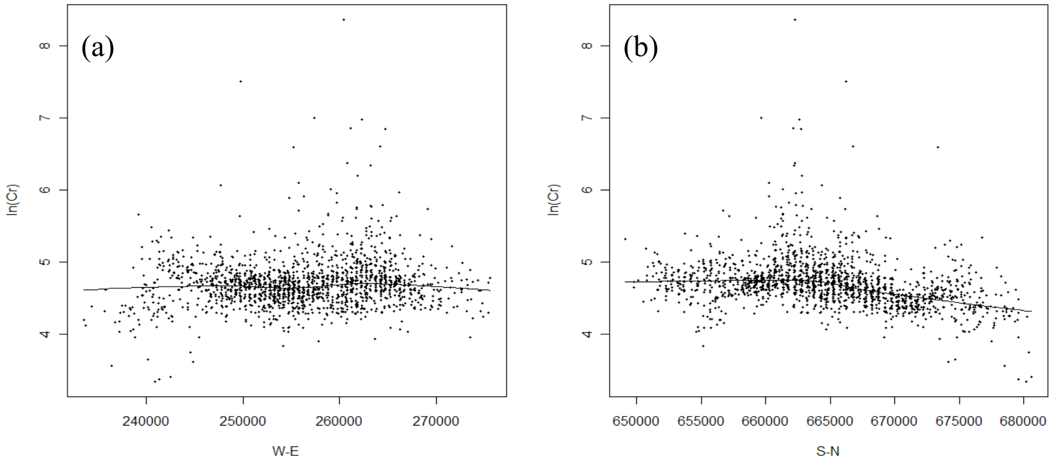


Figure 3.8 – ln(Cr) soil concentration plotted against each of the directional coordinates (a) west-east and (b) south-north, with a fitted loess line

A more formal assessment of the trend effects was performed, with the results shown in Table 3.4. Having fitted a linear regression model on the coordinates, the easting coordinate was not statistically significant indicating no significant spatial trend effect in this direction. However, the p-value for the northing coefficient was highly significant, suggesting that a trend effect was worth considering.

Table 3.4 – Regression on ln(Cr) soil concentration with directional coordinates

Coefficients	Estimate	Std.Error	P-value
Intercept	15.37	1.024	<0.0001
Easting	0.0000013	0.0000011	0.245
Northing	0.0000165	0.0000014	<0.0001

Figures 3.9 (a) and (b) display the experimental variograms for the original ln(Cr) soil data and also a variogram of 1st order linear trend which was fitted by ordinary least squares (OLS). The variogram for the original ln(Cr) soil data showed a fairly constant semi-variance at shorter distances and a rapid increase at distances greater than 25 km, while the linear trend variogram reduced the semi-variance at those larger distances. This erratic behaviour again arose due to the number of paired locations being smaller at

greater distances. Therefore, as the distance between points increased, a large amount of inconsistency in the variograms was to be expected.

The MC envelope for the original variogram (Figure 3.9 (b)) showed that the increase for the original variogram was not quite statistically significant as the points lay within the confidence bands. Therefore, the null hypotheses of spatial independence between observations could not be rejected and there was no spatial correlation present in the data, with ln(Cr) soil concentrations log-normally distributed with a constant mean zero and variance σ^2 across the region. Observing the original variogram, the increase in semi-variance suggested that some spatial correlation was present, although the MC plot demonstrated it was non-significant. Therefore, it was not necessary to model this non-statistically significant correlation and kriging was not used.

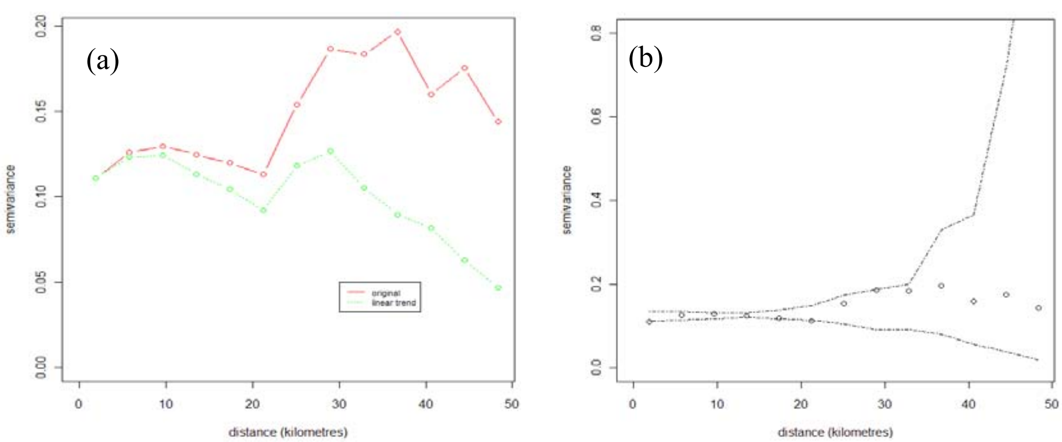
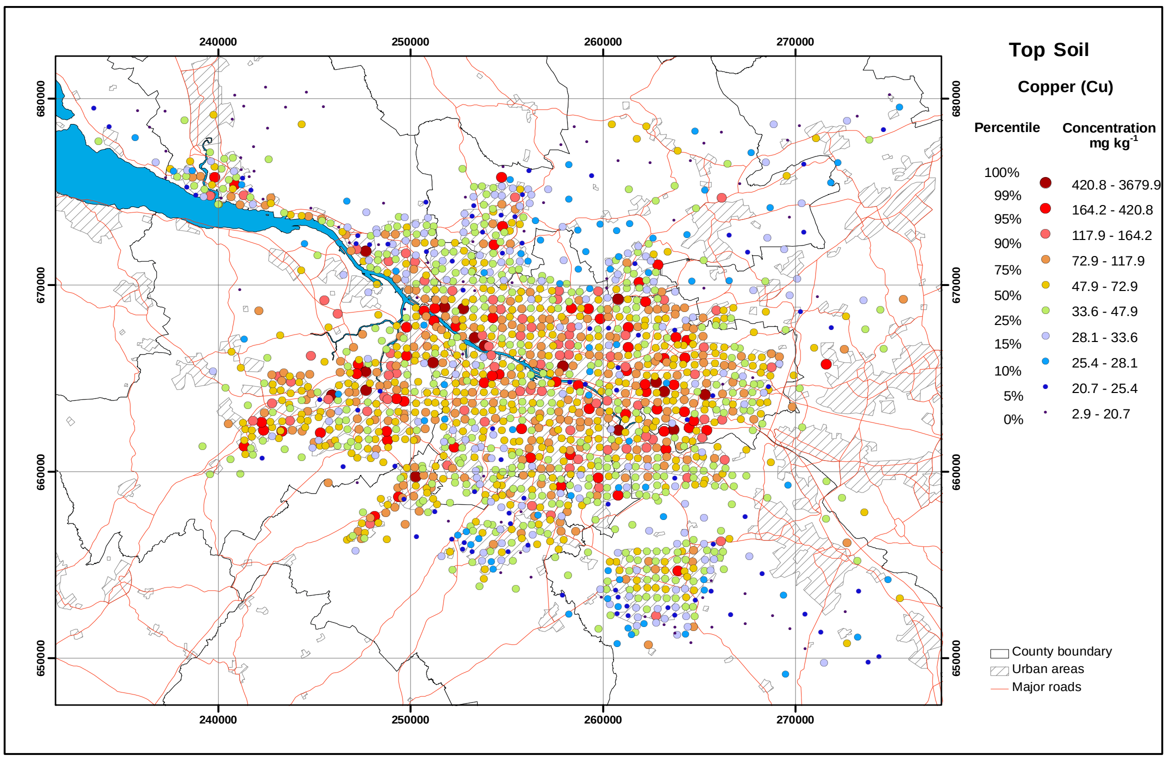


Figure 3.9 – (a) Empirical variograms for original data and removal of linear trend and (b) MC envelope for original ln(Cr) soil variogram

3.2.2.3 Copper (Cu)

Contamination of Cu from human activities is likely to contribute to the high concentration in urban soil compared to rural soils in the Glasgow area. Copper is generally high in concentration in the Rutherglen and Shettleston areas of the East End of Glasgow (see Figures 3.3 and A1 and Table A2 in the Appendix for locations) due to former industrial and metal working processes and in the former shipbuilding areas along the River Clyde. Copper is also associated with the organic rich soils and with man-made ground present in the south and south-west of Greater Glasgow (Fordyce et al., In Prep) (Figure 3.10). The average Cu concentration sampled in Greater Glasgow was 70.5 mg kg^{-1} with the maximum recorded concentration as 3679.9 mg kg^{-1} (Table 3.1).



Grid squares represent 10 km
Linework derived from OS topography © Crown Copyright. All rights reserved. BGS 100037272/10
(From: Fordyce et al., In Prep) BGS © NERC

Figure 3.10 – Map of Cu top soil concentrations in the G-BASE Glasgow dataset

The quadratic loess line fitted in Figure 3.11 (b) supported a possible spatial trend in that ln(Cu) soil concentrations were lower in the north and south in comparison to central areas. The potential trend effect was eliminated from the model before making any judgement on the spatial correlation. As a result of the quadratic relationship observed in Figure 3.11 (a) and (b), in particular for the north to south direction, a quadratic polynomial was fitted on the coordinates. Carrying out a regression on ln(Cu) soil concentrations, p-values for both the easting and northing coordinates in Table 3.5 were statistically significant, indeed indicating the presence of a trend effect.

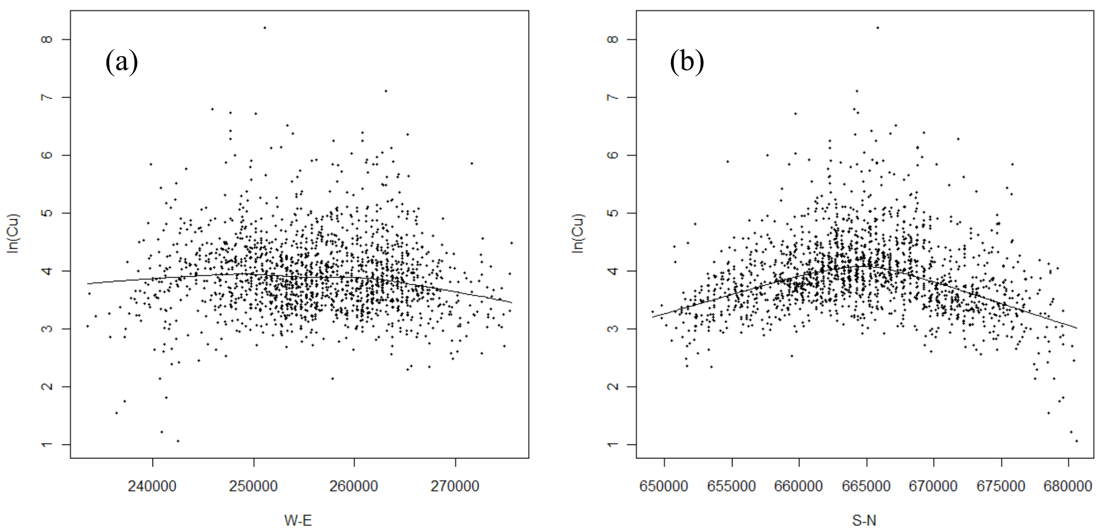


Figure 3.11 – ln(Cu) soil concentration plotted against each of the directional coordinates (a) west-east and (b) south-north, with a fitted loess line

Table 3.5 – Regression on ln(Cu) soil concentration with directional coordinates

Coefficients	Estimate	Std.Error	P-value
Intercept	10.64	2.044	<0.0001
Easting	-0.0000055	0.0000022	0.0138
Northing	-0.0000080	0.0000027	0.0036

Figure 3.12 (a) illustrates the experimental variograms for the original ln(Cu) data, the 1st order linear trend on the coordinates and a polynomial of 2nd order degree, with the latter two fitted by OLS. The plot suggested that the original and linear trend variograms were bounded by a semi-variance of

approximately 0.5 or 0.6, and had a nugget effect of around 0.35. The semi-variance only approached an asymptote at a distance of 20 km, or possibly at 30 km, while a large amount of fluctuation was observed at distances beyond this. This indicated that a stationary model was not appropriate. A linear trend was removed although this did not have much effect, as illustrated in Figure 3.12 (a). However, fitting a quadratic trend to the surface stabilised the semi-variance so that the variogram was constant throughout. The quadratic variogram had a nugget of approximately 0.35 and was relatively stable thereafter.

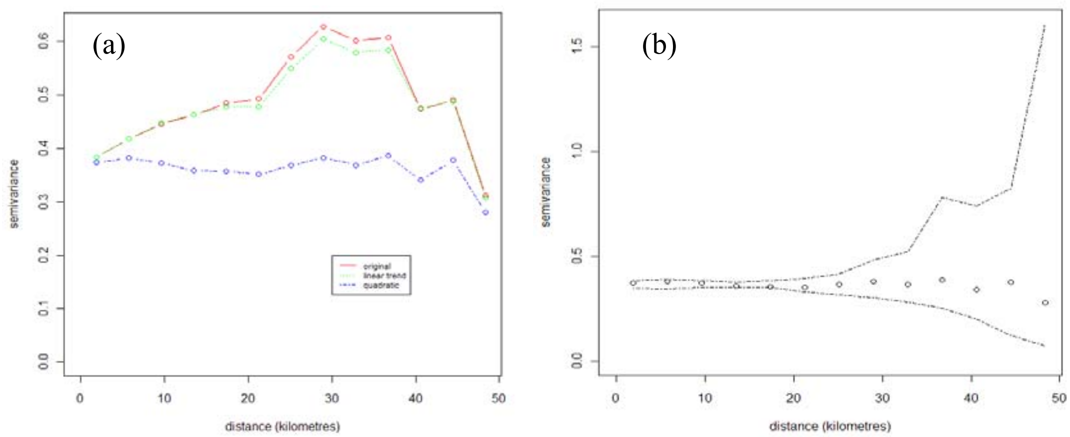
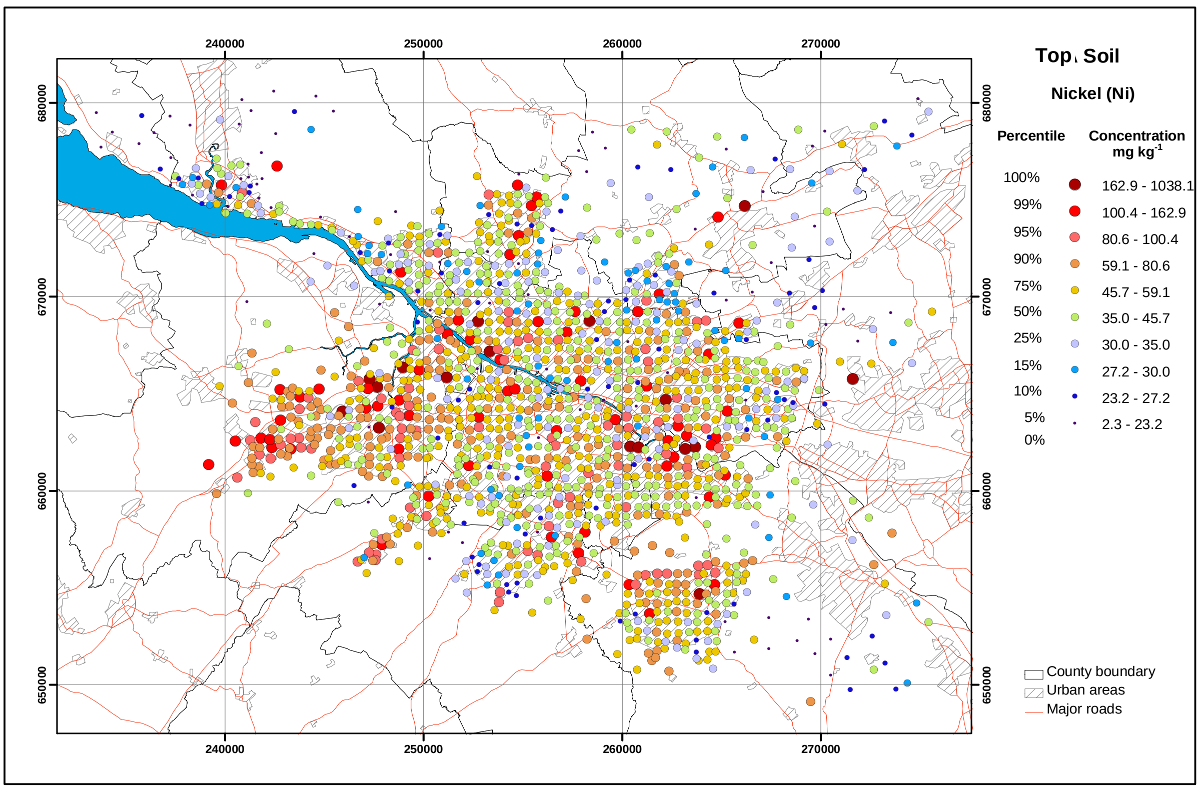


Figure 3.12 – (a) Empirical variograms for original data, removal of linear trend and fitted quadratic surface and (b) MC envelope for ln(Cu) soil variogram with a quadratic surface fitted

Figure 3.12 (b) shows MC envelopes from 99 random combinations of the residuals from a quadratic surface fitted to the ln(Cu) soil concentrations, using OLS. After the removal of a quadratic trend, the variogram appeared fairly constant and all the points lay within the confidence envelopes so that ln(Cu) soil concentrations were spatially independent. Therefore, a quadratic surface explained most of the spatial correlation in ln(Cu) soil concentrations and the residuals of ln(Cu) were log normally distributed with mean zero and variance σ^2 across the sampling region. There was no need to fit a model since the spatial correlation was not statistically significant and kriging was not carried out.

3.2.2.4 – Nickel (Ni)

Like Cr, Ni concentrations tend to be lower in soil developed over the Devonian Sandstone in the Dumbarton area (Figure 3.3), but values are elevated over the volcanic and limestone formation in south-west and southern Glasgow. Soil Ni content is also high in the Rutherglen-Shettleston area where heavy industrial processes took place in the past (Fordyce et al., In Prep) (Figure 3.13; see Figures 3.3 and A1 and Table A2 in the Appendix for locations). The average top soil Ni concentration recorded across the region was 52.6 mg kg^{-1}with a maximum measured concentration of 1038.1 mg kg^{-1} (Table 3.1).



Grid squares represent 10 km
Linework derived from OS topography © Crown Copyright. All rights reserved. BGS 100037272/10
(From: Fordyce et al., In Prep) BGS © NERC

Figure 3.13 – Map of Ni top soil concentrations in the G-BASE Glasgow dataset

The loess line fitted in Figure 3.14 (a) suggested no obvious directional trend effects of ln(Ni) soil concentrations from west to east and any possible trend, appeared to be relatively weak. Figure 3.14 (b) indicated a potential quadratic surface in explaining a spatial pattern from north to south.

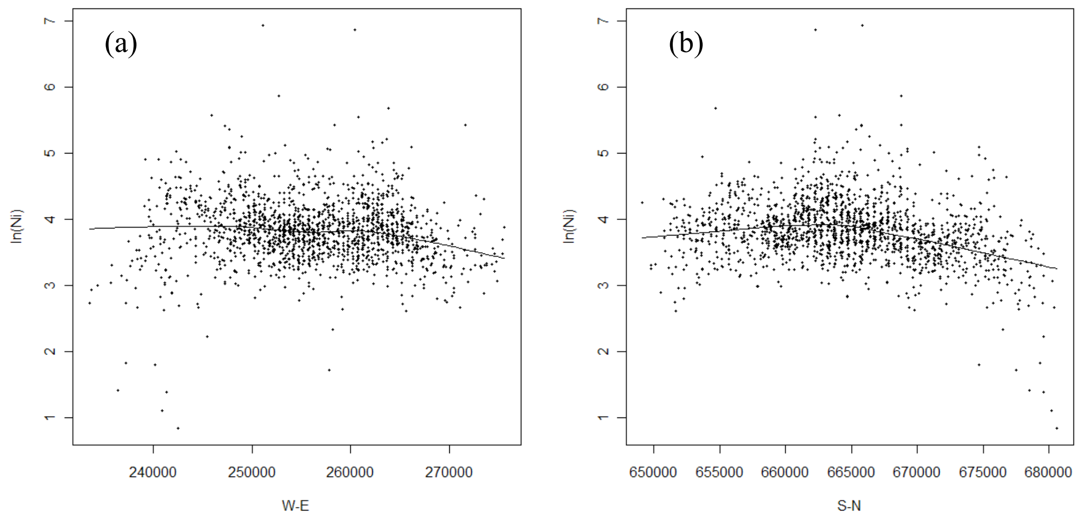


Figure 3.14 – ln(Ni) soil concentration plotted against each of the directional coordinates (a) west-east and (b) south-north, with a fitted loess line

The presence of a trend effect of ln(Ni) soil concentrations was assessed through fitting a linear regression model on the coordinates, with the results shown in Table 3.6. Since p-values for both easting and northing were statistically significant, a spatial trend existed and ought to be removed. Figure 3.15 (a) displays the empirical variograms for the original ln(Ni) soil data, the 1st order linear trend on the coordinates and the quadratic trend surface, with the latter two fitted using OLS. All the variograms have a nugget effect of approximately 0.2. The variograms may be unbounded due to the steep linear increase, particularly in the original and linear trend variograms. The autocorrelation does not approach zero asymptotically and this indicated spatial dependence between observations. However, fitting a quadratic surface reduced the semi-variance considerably and appeared to approach an upper bound asymptotically, at a large distance. From Figure 3.14 (b) it was expected that most of the trend would be removed by the quadratic surface and this variogram was modelled to assess the spatial correlation in the residuals.

Table 3.6 – Regression on ln(Ni) soil concentration with directional coordinates

Coefficients	Estimate	Std.Error	P-value
Intercept	19.64	1.405	<0.0001
Easting	-0.0000079	0.0000015	<0.0001
Northing	-0.0000207	0.0000019	<0.0001

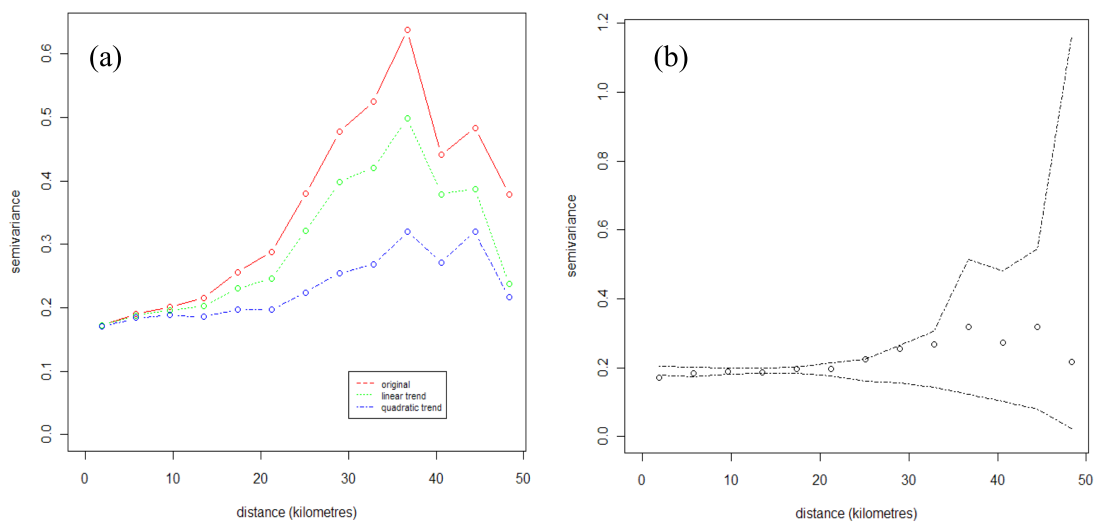
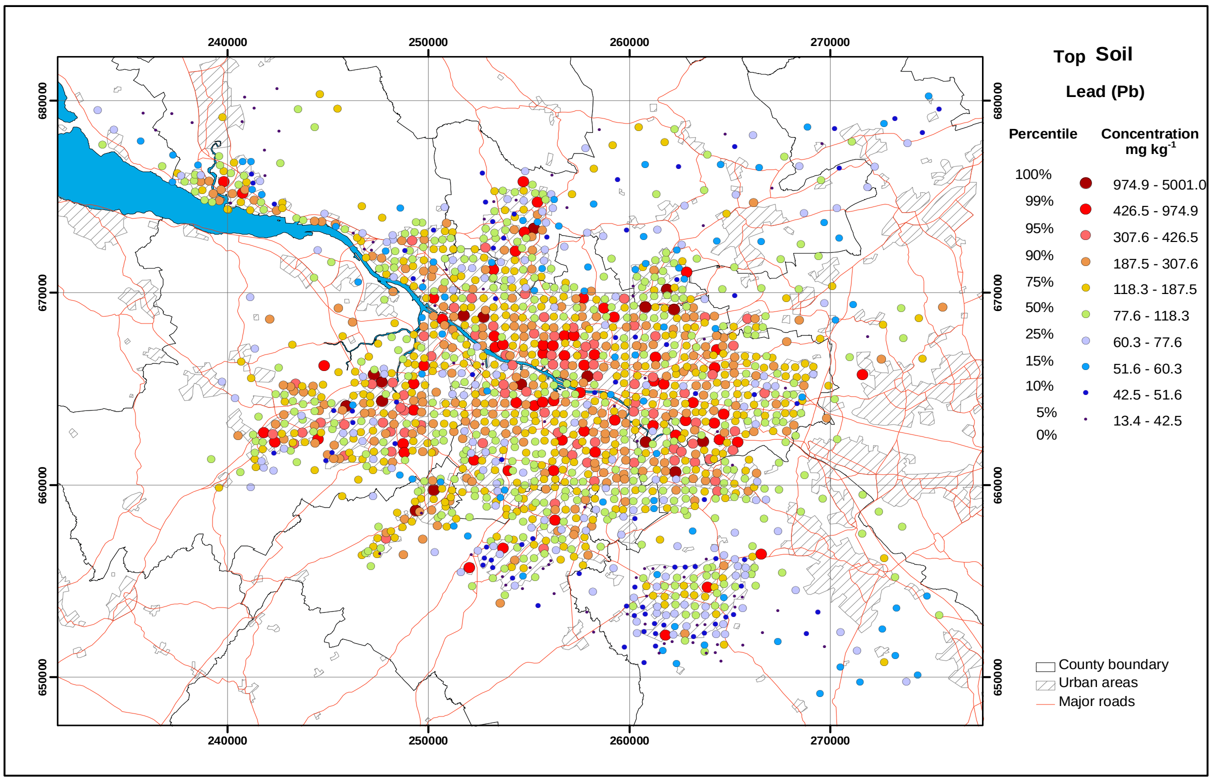


Figure 3.15 – (a) Empirical variograms for original data, removal of linear trend and fitted quadratic surface and (b) MC envelope for ln(Ni) soil variogram with a quadratic surface fitted

Figure 3.15 (b) shows the Monte Carlo (MC) confidence bands of the residuals from a fitted quadratic surface to the ln(Ni) soil concentrations. A highly statistically significant result was observed for the model including the linear trend so that spatial independence was rejected. However, when a quadratic surface was fitted the points of the variogram were all within the MC confidence bands so that the residuals of ln(Ni) soil concentrations were spatially independent and no spatial correlation was present after the removal of a trend effect. Therefore, a quadratic surface accounted for most of the spatial correlation and the residuals were log normally distributed with mean zero and variance σ^2 across Greater Glasgow. Again, there was no requirement to model the non-significant spatial correlation and kriging was not carried out.

3.2.2.5 Lead (Pb)

Lead concentrations in the natural environment are generally very low; therefore, Pb present in the urban environment is likely to have been produced by anthropogenic contamination. High concentrations of top soil Pb are present in the city centre as a result of traffic pollution and in the east and the south-east of Glasgow, as a result of metal working in these areas (Fordyce et al., In Prep). From Table 3.1, the mean top soil Pb concentration in Greater Glasgow was 167.9 mg kg^{-1} and the maximum was 5001 mg kg^{-1}. Figure 3.16 illustrates the spatial distribution of soil Pb concentrations across Greater Glasgow.



Grid squares represent 10 km
Linework derived from OS topography © Crown Copyright. All rights reserved. BGS 100037272/10
(From: Fordyce et al., In Prep) BGS © NERC

Figure 3.16 – Map of Pb top soil concentration in the G-BASE Glasgow dataset

Plotting the ln(Pb) soil concentration values against the directional coordinates, a possible spatial trend in the form of a quadratic surface was observed. Figure 3.17 (b) suggested higher concentrations in central Glasgow, with lower concentrations in the north and south. A similar but much weaker pattern was observed from west to east in Figure 3.17 (a).

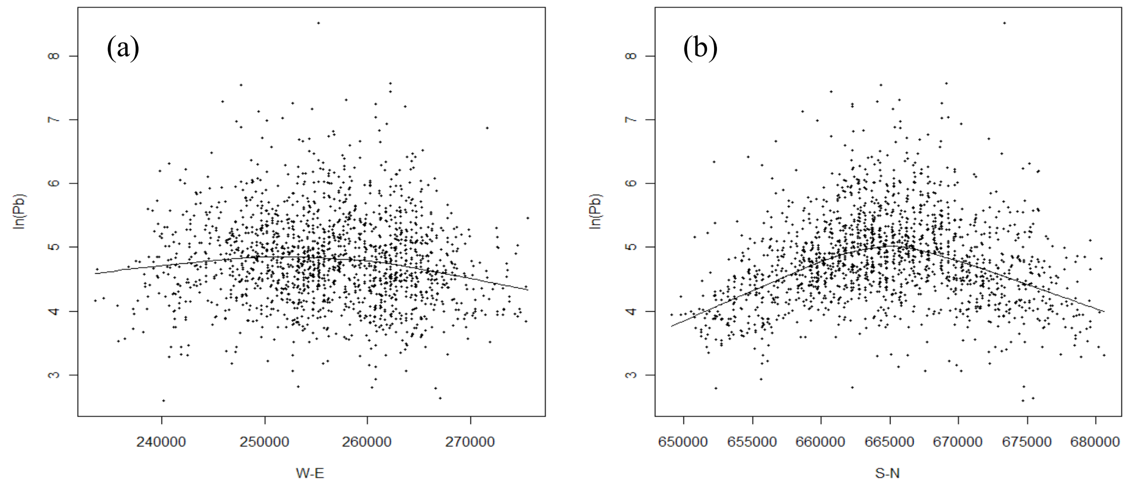


Figure 3.17 – ln(Pb) soil concentration plotted against each of the directional coordinates (a) west-east and (b) south-north, with a fitted loess line

To assess the presence of a spatial trend in ln(Pb) concentrations across Greater Glasgow, a simple linear regression on the coordinates was performed. The result in Table 3.7 shows that the northing coordinate was not statistically significant so that a trend was not deemed to be present. However, trend effects were assessed for this element due to the strange behaviour which occurred in the variograms.

Table 3.7 – Regression on ln(Pb) soil concentration with directional coordinates

Coefficients	Estimate	Std.Error	P-value
Intercept	3.873	2.174	0.0751
Easting	-0.0000055	0.0000024	0.0190
Northing	0.0000035	0.0000029	0.2207

Figure 3.18 (a) displays the empirical variograms for the original ln(Pb) soil data, a linear trend on the coordinates and a fitted quadratic trend. The variograms appear to behave in an unusual manner, where the original and linear trend variograms increase up to 20 km and a sharp decrease is observed where points are located further apart. Interpretation of any trend effect and spatial correlation in soil ln(Pb) is difficult since its behaviour suggests that the sampling density may not capture the variability in Pb concentration across the region. However, with the current data available one could argue that the points, with a quadratic surface fitted, all lie within the MC confidence bands in Figure 3.18 (c). Following this argument, the assumption of spatial independence was not rejected and a quadratic trend was removed so that a quadratic surface explained most of the spatial correlation across the sampling region. Again, the interpolation method of kriging was not required. However, the results of the geostatistical analysis carried out for soil ln(Pb) were dubious. The factors contributing to this are unknown but are likely to reflect the localised nature of Pb contamination within urban environments.

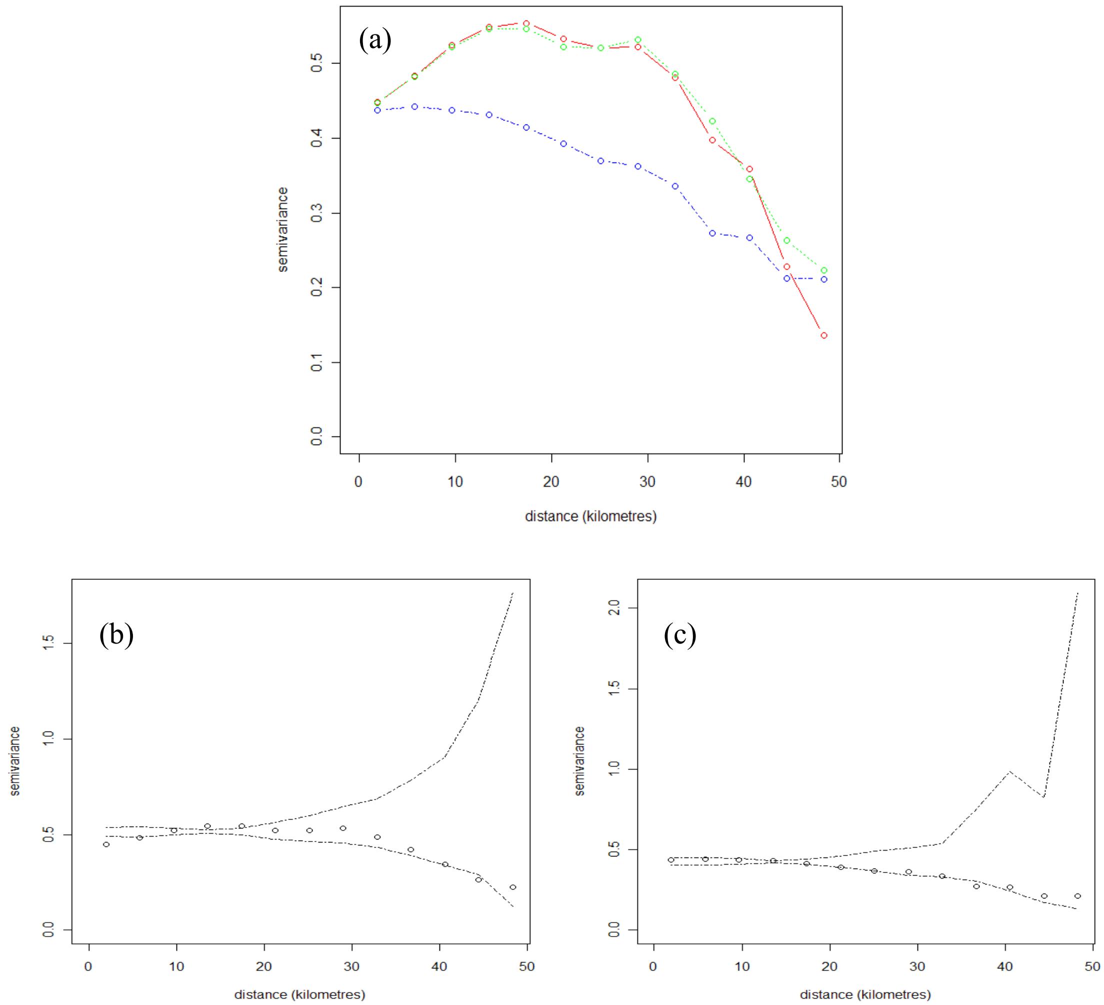
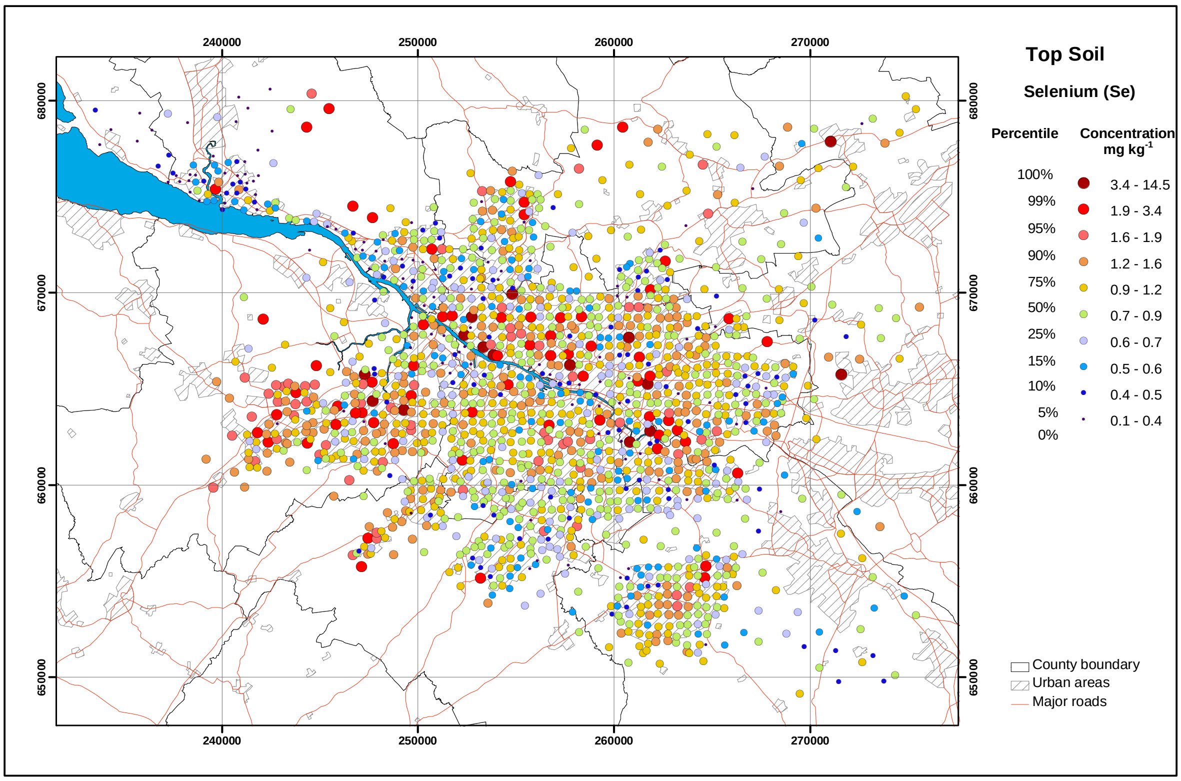


Figure 3.18 – (a) Empirical variograms for original data, removal of linear trend and fitted quadratic surface and (b) MC envelope for removal of linear trend ln(Pb) soil variogram and (c) MC envelope for ln(Pb) soil variogram with quadratic surface fitted

3.2.2.6 Selenium (Se)

Selenium is another element that is lower in concentration in soils developed over Devonian Sandstones, and is higher in peaty soils developed over the volcanic and limestone formations of south-west Glasgow (Figure 3.3). Higher concentrations in soils are also associated with the shipbuilding areas along the River Clyde and the East End sites of former heavy industry. (Fordyce et al., In Prep) (Figure 3.19). The mean Se concentration in Greater Glasgow was 1.03 mg kg^{-1} with a maximum of 14.5 mg kg^{-1} recorded (Table 3.1).



Grid squares represent 10 km
Linework derived from OS topography © Crown Copyright. All rights reserved. BGS 100037272/10
(From: Fordyce et al., In Prep) BGS © NERC

Figure 3.19 – Map of Se top soil concentrations in the G-BASE Glasgow dataset

The weak quadratic curve in Figure 3.20 (b) suggested that ln(Se) soil concentrations may have a quadratic relationship from north to south, with concentrations higher in central Glasgow in comparison to the north and south. In a similar manner to the other metals, any spatial trend was eliminated before assessing the spatial correlation in the data. Notice the pattern of the points in Figure 3.20, as lower measurements appear to be recorded at regular intervals, indicating problems caused by the limits of detection. From Table 2.1, the LLD for Se was 0.2, so that values close to this are likely to be less accurate. Therefore, the pattern of measurements observed in Figure 3.20 may arise due to the measurement accuracy and the range of the detector. In addition, Se only occurs in most soil environments in low concentrations, making it more difficult to record a precise value for this element.

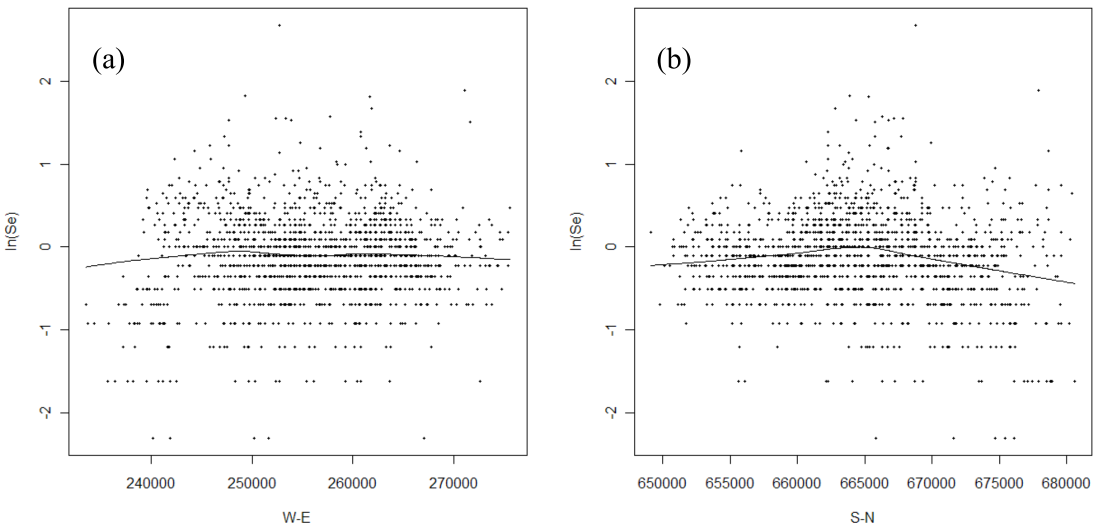


Figure 3.20 – ln(Se) soil concentration plotted against each of the directional coordinates (a) west-east and (b) south-north, with a fitted loess line

Examining the spatial trends more formally and fitting a linear regression line on the coordinates, Table 3.8 showed a statistically significant p-value was found for the northing coordinate while the easting coordinate was non-significant. Since the p-value for the northing coordinate was highly significant and taking into account the measurement problems discussed above, all variograms for this element were also assessed.

Table 3.8 – Regression on ln(Se) soil concentration with directional coordinates

Coefficients	Estimate	Std.Error	P-value
Intercept	4.613	1.514	0.0024
Easting	0.0000017	0.0000016	0.2925
Northing	-0.0000078	0.0000020	0.0001

The experimental variograms for the original ln(Se) soil concentrations, the observed residuals from a first order linear trend on the coordinates and a quadratic surface fitted are plotted in Figure 3.21 (a). For the original and linear trend variograms, the semi-variance between locations increased without decay as the distance between them increased. Again, erratic behaviour occurred at large distances. An unbounded variogram for Se was also observed, where the autocorrelation did not reduce to zero. This again indicated the presence of spatial dependence in the observations. However, the semi-variance was reduced considerably when a quadratic surface was fitted.

Using MC methods, Figure 3.21 (b) showed a statistically significant increase in the linear trend variogram as several of the points were outside the confidence envelopes so that the null hypothesis of spatial independence within the residuals was rejected. However once a quadratic surface was fitted, as shown in Figure 3.21 (c), the points were within the MC confidence bands. Therefore the variogram appeared fairly constant suggesting spatial independence and no spatial correlation in the residuals after the removal of a quadratic trend effect. A quadratic surface was suitable in describing the spatial correlation of soil ln(Se) across Greater Glasgow and the residuals of ln(Se) from a quadratic surface were log-normally distributed with a mean zero and variance σ^2 across the sampling region. Since there was no significant spatial correlation present, kriging was not used.

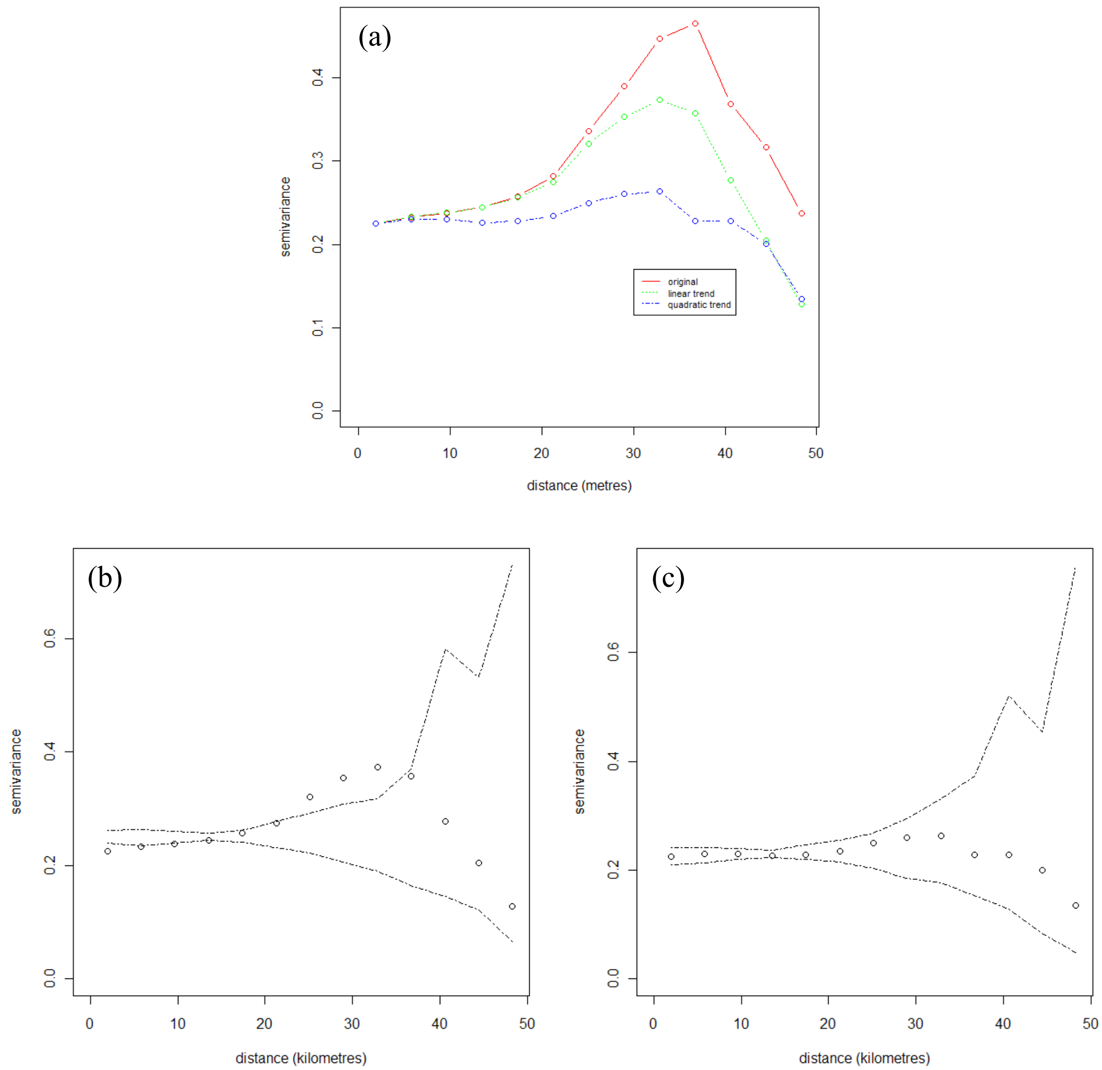
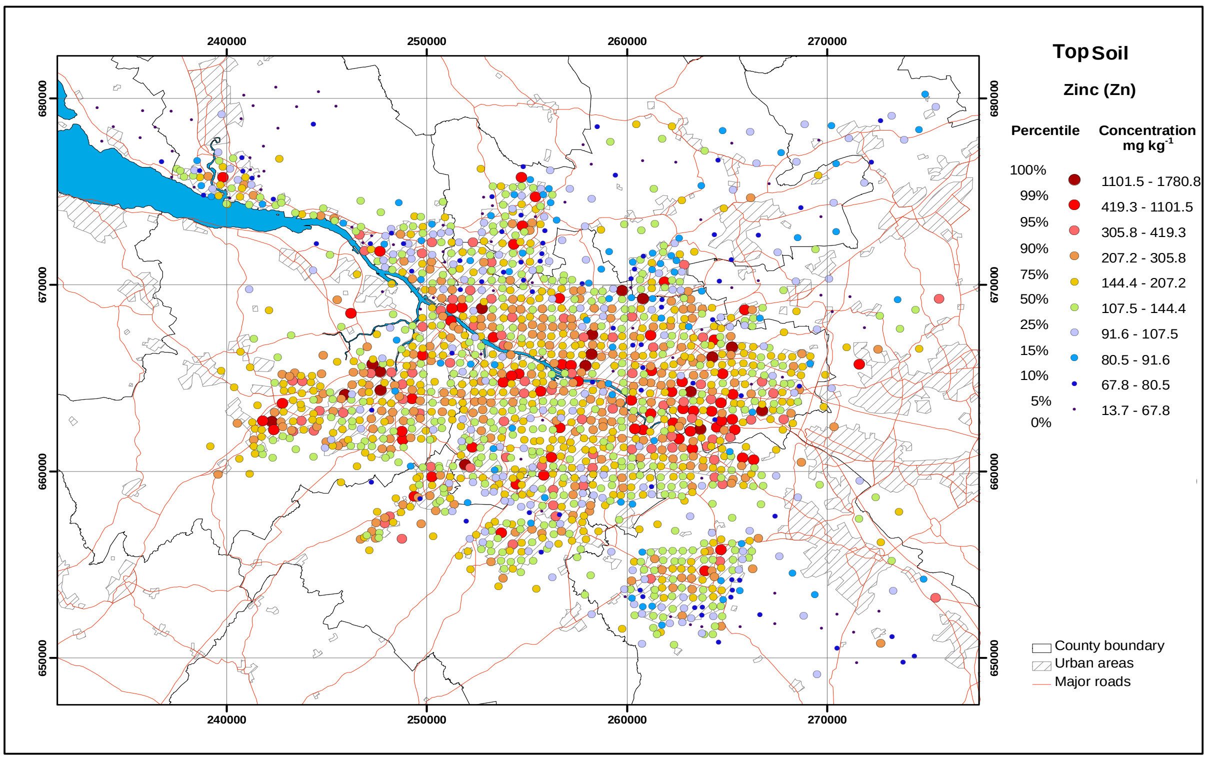


Figure 3.21 – (a) Empirical variograms for original data, removal of linear trend and fitted quadratic surface and (b) MC envelope for removal of linear trend ln(Se) soil variogram and (c) MC envelope for ln(Se) soil variogram with quadratic surface fitted

3.2.2.7 Zinc (Zn)

Zinc is present at a range of higher concentrations in soil over the Coal Measures, which is indicative of anthropogenic contamination. High concentrations are associated with the shipbuilding areas along the River Clyde and with the former heavy industry areas in the East End, while lower concentrations are found over Devonian Sandstones (Figure 3.3) (Fordyce et al., In Prep) (Figure 3.22). The average Zn concentration in Greater Glasgow was 189.4 mg kg^{-1} and the maximum recorded was 1780.8 mg kg^{-1} (Table 3.1).



Grid squares represent 10 km
Linework derived from OS topography © Crown Copyright. All rights reserved. BGS 100037272/10
(From: Fordyce et al., In Prep) BGS © NERC

Figure 3.22 – Map of Zn top soil concentrations in the G-BASE Glasgow dataset

Plotting the log-transformed Zn soil values against the coordinates, a possible trend in the form of a quadratic surface was observed. Figure 3.23 (b) suggested higher concentrations in central Glasgow with lower concentrations observed in the north and south. A similar but weaker pattern occurred from west to east as shown in Figure 3.23 (a).

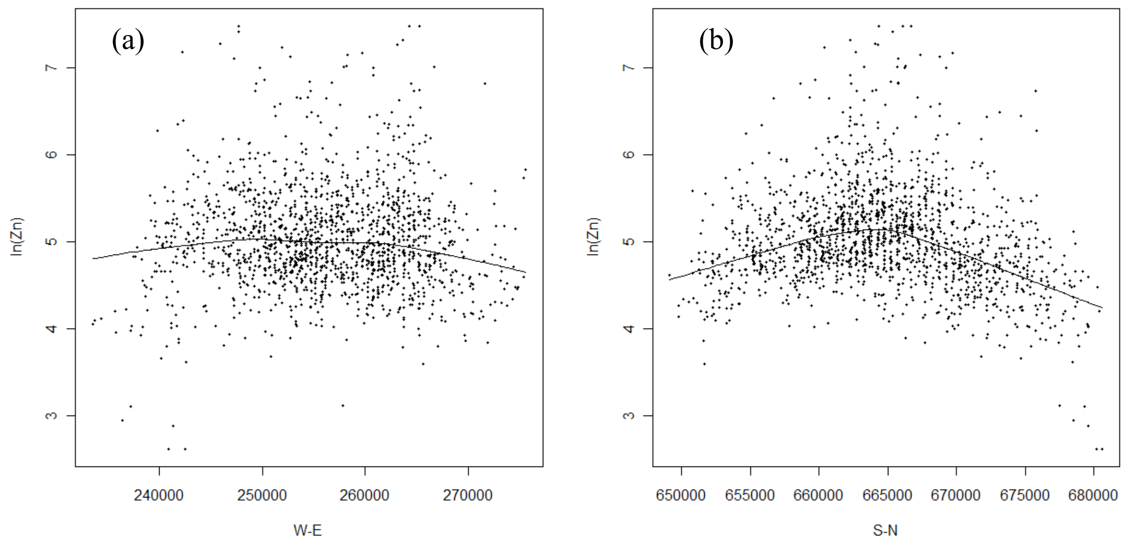


Figure 3.23 – ln(Zn) soil concentration plotted against each of the directional coordinates (a) west-east and (b) south-north, with a fitted loess line

The presence of a trend was more formally assessed by fitting a linear regression model on the spatial coordinates, with the results illustrated in Table 3.9. The table shows a highly significant p-value for the northing coordinate but a (just) non-significant value for the easting coordinate. Therefore, a spatial trend effect was present for ln(Zn) soil concentrations across Greater Glasgow and was assessed with the use of variograms.

Table 3.9 – Regression on ln(Zn) soil concentration with directional coordinates

Coefficients	Estimate	Std.Error	P-value
Intercept	16.12	1.752	<0.0001
Easting	-0.0000035	0.0000019	0.0696
Northing	-0.0000015	0.0000023	<0.0001

Figure 3.24 (a) displays the empirical variograms for the original ln(Zn) soil data, the 1st order linear trend on the coordinates and a quadratic surface. All variograms had a nugget effect of 0.3. Fitting a quadratic surface reduced the variogram to a constant decrease as distance increased, which should not occur. This implied that either the sampling density did not capture the variation of Zn across Glasgow or that the quadratic surface was too complex and over-fitted the variogram. In this case, the original and linear trend variograms behaved as one would expect so it was more likely that over-fitting had occurred. The original and linear trend variogram approached the sill at a range of around 30 km and seemed to have an upper bound of approximately 0.4 and 0.5, respectively.

Figure 3.24 (b) shows the MC envelope for the original data and several of the points lie outside the confidence bands. Therefore, this indicated some positive spatial correlation in the ln(Zn) soil concentrations across Greater Glasgow. Figure 3.24 (c) displays the corresponding MC envelopes for the variogram with a linear trend removed. Since only one or two points lay (just) outside the confidence envelopes, one could accept with a little uncertainty that the residuals were spatially independent after the removal of a linear trend. Therefore, a linear trend effect was removed and was sufficient in eliminating most of the spatial correlation in the data. The residuals of ln(Zn) soil concentrations from a linear trend were log-normally distributed with mean zero and variance σ^2 across the sampling region. There was no need to model the non-significant spatial correlation and kriging was not carried out.

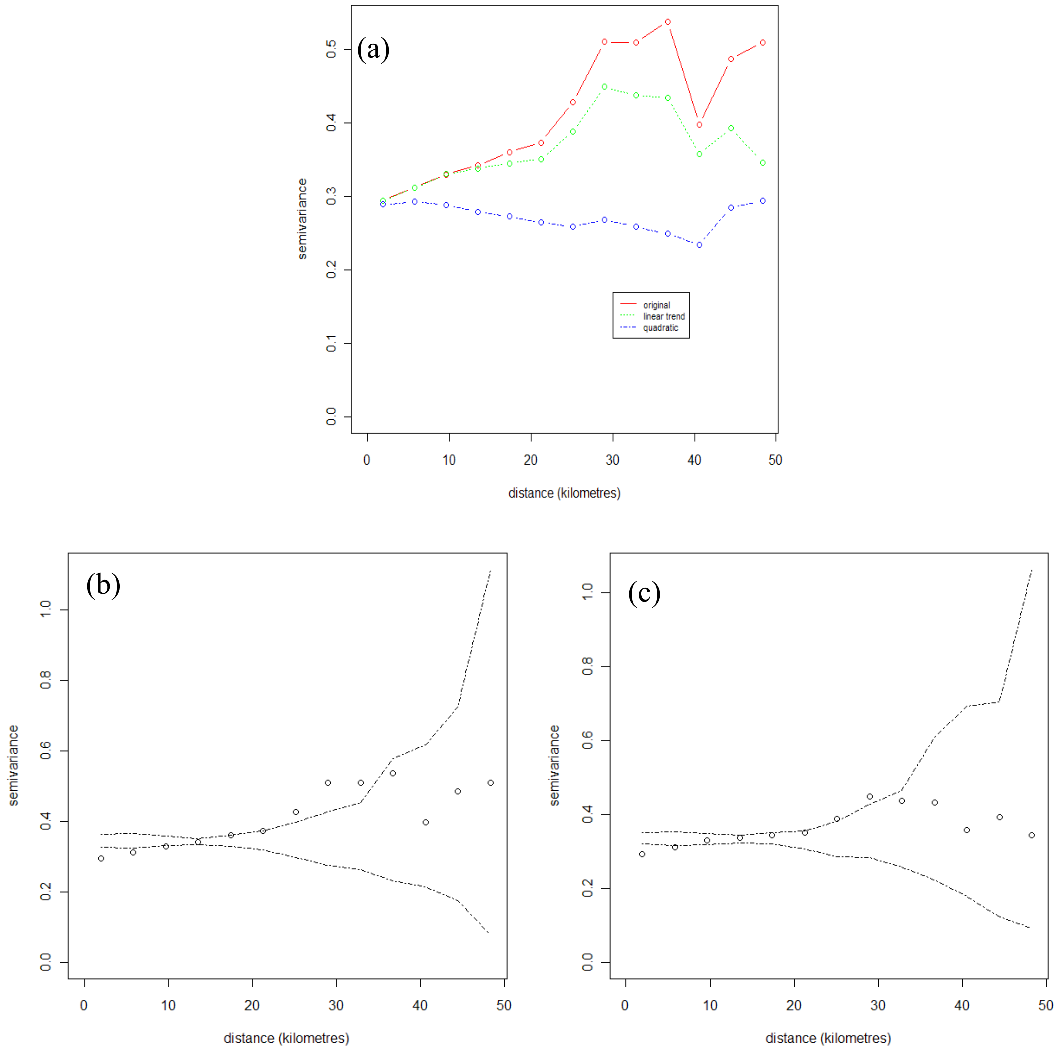
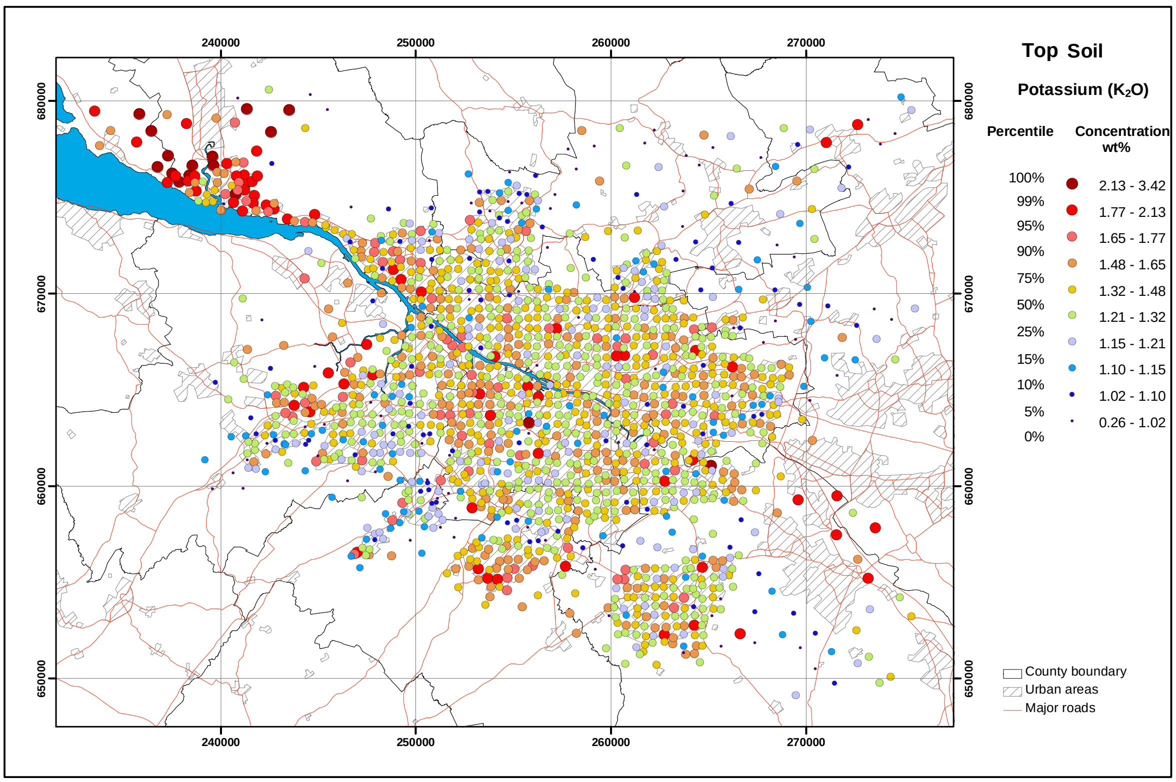


Figure 3.24 – (a) Empirical variograms for original data, removal of linear trend and fitted quadratic surface and (b) MC envelope for original ln(Zn) soil variogram and (c) MC envelope for removal of linear trend ln(Zn) soil variogram

3.2.2.8 Potassium

A similar analysis was carried out on K_2O, although dissimilar results were expected since its spatial distribution is different to the other elements. Potassium concentrations are higher over Devonian Sandstones present in the Dumbarton area and over the glaciofluvial sand and gravel deposits in Bothwell to the south-east of Greater Glasgow (Fordyce et al., In Prep) (Figure 3.25) (see Figures 3.3 and A1 and Table A1 in the Appendix for locations). The mean top soil K_2O concentration was 1.36 wt% with a maximum of 3.42 wt% (Table 3.1).



Grid squares represent 10 km
Linework derived from OS topography © Crown Copyright. All rights reserved. BGS 100037272/10
(From: Fordyce et al., In Prep) BGS © NERC

Figure 3.25 – Map of K_2O top soil concentrations in the G-BASE Glasgow dataset

Figure 3.25 indicates a cluster of elevated values in the far north-west area of Greater Glasgow, which is the area of Dumbarton. In general, it appears that lower concentrations are found in the north-east. The loess line fitted in the left hand plot in Figure 3.26 supports this possible west to east pattern. Empirical variograms were used to analyse the presence of spatial correlation in the data, firstly removing any significant trend effect.

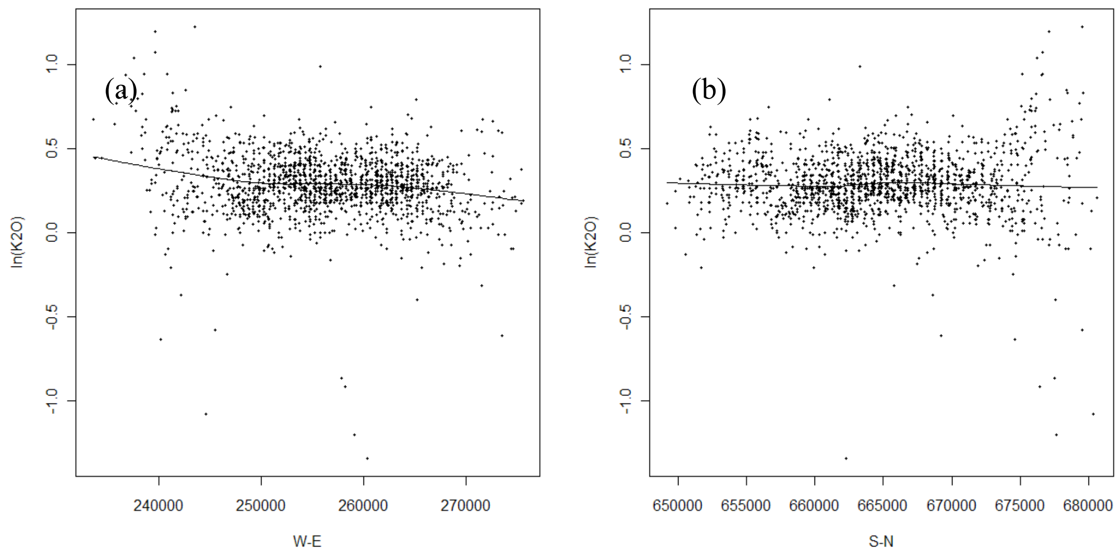


Figure 3.26 – ln(K_2O) soil concentration plotted against each of the directional coordinates (a) west-east and (b) south-north, with a fitted loess line

The presence of a trend effect was formally assessed by fitting a linear regression model, with the results shown in Table 3.10. For soil ln(K_2O), the p-value for the easting coefficient was highly significant while that of the northing was non-significant. This trend in the easting direction was likely to have been caused by the high concentrations in the north-west which were observed in Figure 3.25. Spatial trend effects were assessed with the use of variograms.

Table 3.10 – Regression on ln(K_2O) soil concentration with directional coordinates

Coefficients	Estimate	Std.Error	P-value
Intercept	1.759	0.5648	0.0019
Easting	-0.0000044	0.0000006	<0.0001
Northing	-0.0000051	0.0000008	0.4997

In a similar manner to the other elements, Figure 3.27 (a) displays the original $\ln(K_2O)$ variogram, one of a 1^{st} order linear trend on the coordinates and a quadratic trend, with the latter two fitted by ordinary least squares (OLS). The three variograms all showed a steep increasing semi-variance as the distance between points increased, suggesting a stationary model was not appropriate for these data, although one could argue that fitting a linear trend stabilised the variogram. However, eliminating further trends by fitting a quadratic did not reduce the semi-variance between points by much more. The variograms appeared to be unbounded, which was indicative of unpredictable behaviour at larger distances. Due to this steady linear increase, the autocorrelation did not approach zero and any possible quantity of spatial correlation would be expected at larger distances.

Figure 3.27 (b) and (c) display the MC envelopes, for the linear and quadratic variograms, respectively. Both produced a statistically significant result so that the null hypothesis was rejected and there was indeed a considerable amount of spatial correlation in the residuals between locations. The addition of a polynomial did not remove the trend effect and modelling the spatial continuity allowed concentrations to be estimated at any location, by kriging. ForK_2O, soil concentrations were log-normally distributed with mean zero depending on location and variance σ^2 according to a specified correlation structure $\sum$. The less complex variogram with a linear trend fitted was selected as an appropriate model as there was not much improvement in fitting a quadratic surface. To construct a model for the presence of spatial correlation, the parameters were estimated from Figure 3.27 (a). Using this plot, the residuals obtained after fitting a linear trend model were possibly bounded by a semi-variance of approximately 0.08 at a range of 30 km and a small nugget effect of around 0.02 was observed. The two input parameters for the kriging were 0.06 for the partial sill and 30 km for the range.

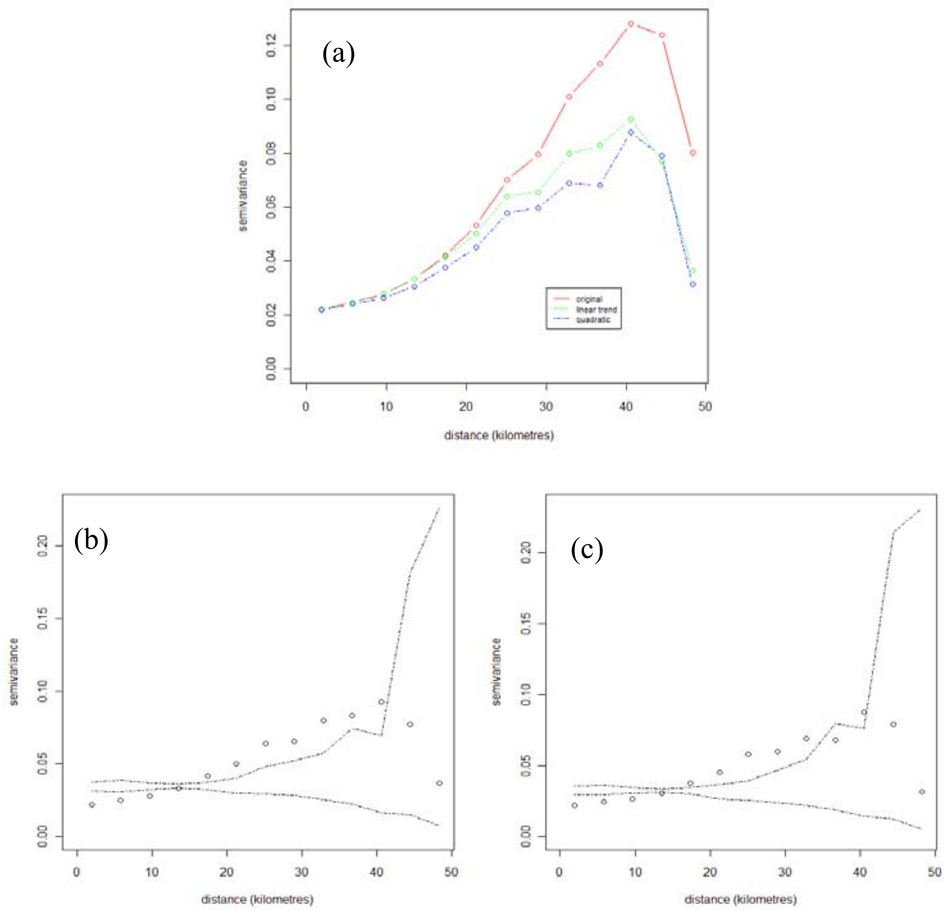


Figure 3.27 – (a) Empirical variograms for original data, removal of linear trend and fitted quadratic surface and (b) MC envelope for removal of linear trend ln(K_2O) soil variogram and (c) MC envelope for ln(K_2O) soil variogram with quadratic surface fitted

Using the estimated parameters and fitting a Gaussian and Exponential model to the empirical variogram, the plot in Figure 3.28 displays the models with an estimated nugget effect. A Matérn model with κ = 1.5 was assessed to model the spatial correlation. There appeared to be very little difference between the models although the Gaussian seemed to capture the behaviour in the variogram slightly better than the Exponential. The model which described the spatial continuity best was the model with κ = 1.5 and represented an unbounded variogram over the defined range. Table 3.11 contains the estimated parameters of the three models. The correlation range generated in R produced a distance of 142 km and this suggests the geographical scale at which K_2O concentrations are likely to be correlated. However, the scale of this study did not go beyond Greater Glasgow and the two furthest points are just over 50 km apart. Nonetheless, at a range of 142 km, the Matérn model with κ = 1.5 estimated the variogram to approach the sill asymptotically at a semi-variance of 0.139. The identified spatial correlation present was indicative of a K_2O distribution pattern caused by a non random process, most likely being controlled by geology.

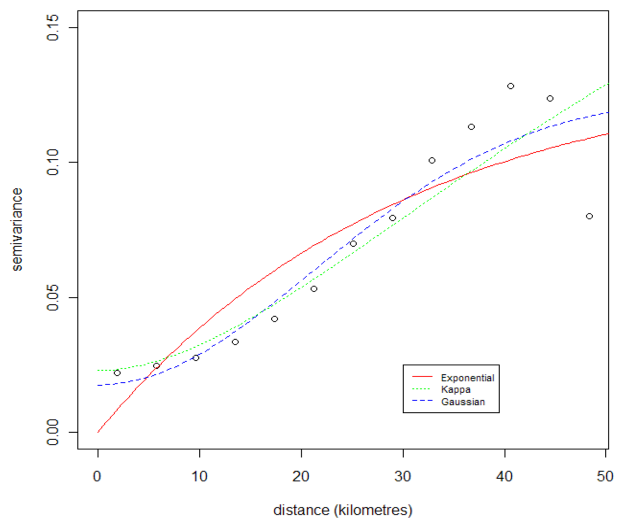


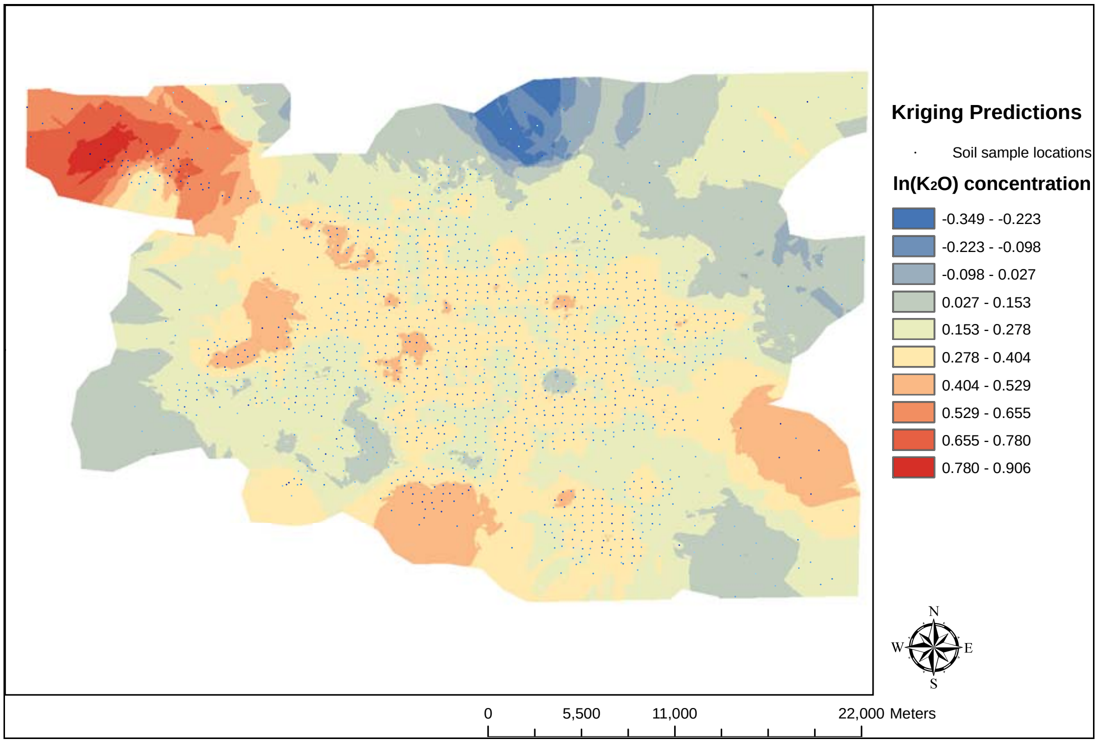
Figure 3.28 – Models to fit empirical variogram for ln(K_2O)

Table 3.11 – Estimated model parameters from ln(K_2O) soil variogram

	nugget	partial sill	range (km)
Exponential	0.0115	0.0819	89.9
Matérn (κ = 1.5)	0.0302	0.1086	142
Gaussian	0.0258	0.0581	51.9

Figure 3.29 displays the ordinary kriging estimates after fitting a linear trend model to the data and a Matérn correlation function to the residuals, where the estimated parameters of the Matérn model in Table 3.11 were used to predict ln(K_2O) soil concentrations. . To define the calculation of the soil metal predictions, a fixed search radius of 100 m with eight as the maximum number of real soil sample points to include in the calculation were selected. The map indicated elevated concentrations in the north-west of Greater Glasgow in particular, with high concentrations also present in the Bothwell area in the

south-east (see Figure A1 and Table A1 in the Appendix for locations), while lower ln(K_2O) soil concentrations were estimated in the north-east and south-west (Figure 3.29). The map concurs with the spatial distribution of K_2O in Glasgow top soils outlined in Figure 3.25.



Rural IG areas on the edge of Glasgow have been clipped to the extent of the soil metal dataset

Figure 3.29 – Map of predicted soil ln(K_2O) concentrations across Greater Glasgow

3.3 – Chapter 3 Summary

The exploratory analysis in this chapter was necessary to gain an understanding of the spatial distribution of the soil metals, while also highlighting any issues regarding correlations over space. The spatial distributions of most of the soil metals were explained by a simple linear trend or a quadratic function. A linear trend was sufficient in detrending the data for the analysis of As and Zn, while a quadratic surface was required to detrend data for Cu, Ni, Pb and Se. However, the analysis of Cr did not require a trend effect to be fitted. Erratic behaviour was observed at large distances in the variograms for all soil contaminants and this was largely due to the reduction in the number of pairwise comparisons at greater distances. Since the distribution of K_2O soil concentration was primarily controlled by geology, one would expect the results to show some spatial correlation. However, with the majority of the soil data being in urban areas, K_2O top soil concentrations did not vary too much which resulted in an unbounded variogram. In order to detrend the data, a Matérn correlation function with $\kappa = 1.5$ and the parameters estimated in Table 3.11 was fitted to the residuals, along with a linear trend model fitted to the data. This complex model (referred to as a 'Matérn' model) was used to predict kriging estimates of the $\ln(K_2O)$ soil concentrations. Ideally, a greater sampling density could improve the accuracy and robustness of the geostatistical results for the soil metals dataset and may lead to different conclusions. The relationships between the soil metal datasets and health factors are presented in Chapter 5 of this thesis.

Chapter 4

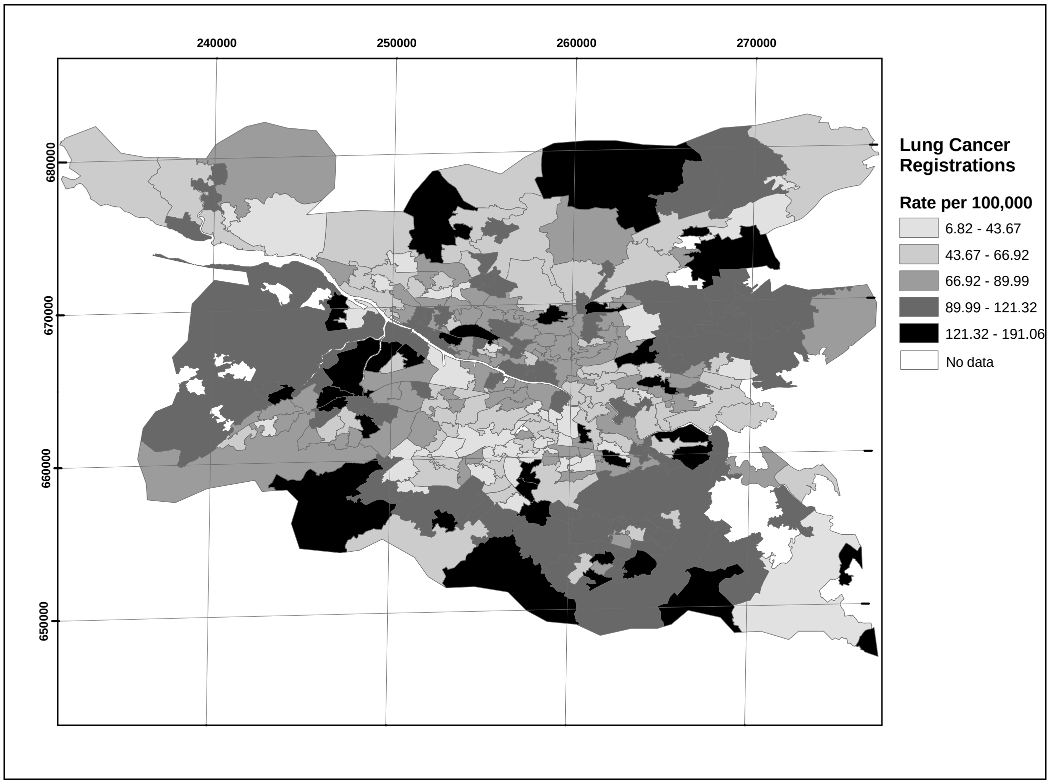
Spatial Presentation of the Soil Metal, Air Quality, Deprivation Index and Health Data

In order to carry out the statistical assessments of the associations between the health and environmental datasets outlined in Chapter 5, it was necessary in the first instance to present the soil metals, air pollution, deprivation and health datasets using spatial mapping at the Intermediate Geography (IG) level. These provided a subjective impression of the relationships between the variables of interest.

4.1 Spatial Presentation of Health Data

In this project, map representation tools in ArcGIS were used in the analysis of health outcomes since the development of these methods in public health has made substantial progress in recent years (Lawson, 2001). GIS-based maps of the distribution of health outcomes are also valuable in the risk assessment of disease patterns related to changes in environmental exposures (Jarup, 2004). For the data in this study, this involved mapping the crude rates of the health outcomes for each IG area across Greater Glasgow. This type of map gave an overall impression of the distribution of the particular health outcome over the region of interest. This was useful in being able to visually compare health outcomes to areas where higher exposure to air pollution and soil metal concentrations were more likely; although this approach does not offer an assessment of the statistical significance of areas with higher health incidence (Lawson, 2001). After performing an initial visual assessment on the distribution of the health outcomes, a model-based statistical approach was carried out in Chapter 5 to analyse the relationships between the environmental variables and health outcomes more fully.

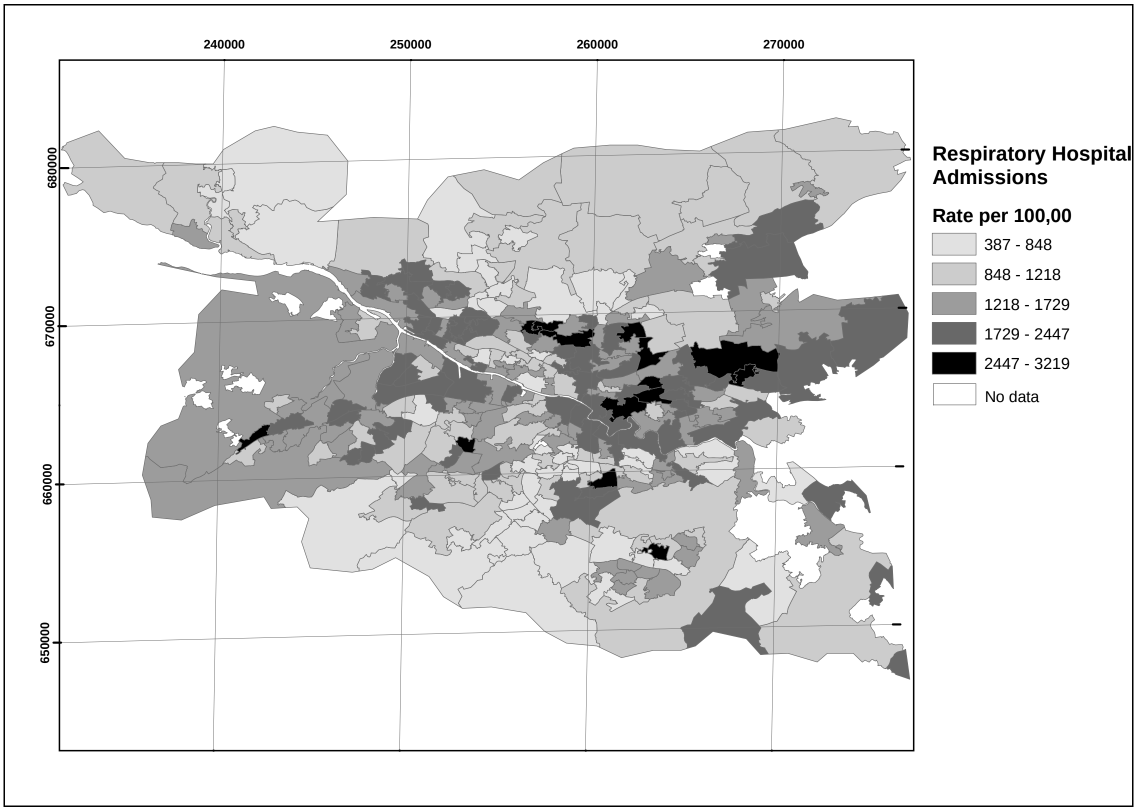
The indicators of health used in this study were lung cancer registrations and hospital admissions of respiratory disease. The rate per 100,000 persons for lung cancer registrations in Glasgow, for each IG area, was a five year average from the period 1998 to 2002. A typical summary measure used in mapping health outcomes is the raw counts or rates of the disease of interest. However, using counts would require a relatively large number of cases in each IG level and in some IG areas in the present project, very few lung cancer registrations were recorded. A map of raw rates of lung cancer incidence across the region is presented in Figure 4.1. From the map, areas of higher lung cancer registrations were identified in the north-east and south-east of Glasgow, although several other regions in urban and rural areas were also revealed to have higher incidence of lung cancer. Lower incidence rates were observed in several small-areas in the inner-city and north-west. However, in general there were no obvious patterns or regions with high rates of lung cancer registrations and the distribution was relatively dispersed.



Rural IG areas on the edge of Glasgow have been clipped to the extent of the soil metal dataset

Figure 4.1 – Map of raw rates of lung cancer registrations across Greater Glasgow for each IG area

Figure 4.2 illustrates the raw incidence rates of respiratory hospital admissions across the region, where a rate per 100,000 persons in 2002 was given for each IG area. Intermediate Geography (IG) areas with higher respiratory admission rates were observed in the East End of Glasgow including Parkhead West and Barrowfield; Old Shettleston and Parkhead North; Carntyne; Easterhouse Central and Garthmalock and Auchinlea and Gartmalock (See Figures A1 and A2 and Table A1 in the Appendix for locations). In general, lower incidence of respiratory admission rates were present in rural areas of Greater Glasgow.



Rural IG areas on the edge of Glasgow have been clipped to the extent of the soil metal dataset

Figure 4.2 – Map of raw rates of respiratory hospital admissions across Greater Glasgow for each IG area

However, these raw incidence rates do not take the age and sex distribution of each geographic area into account. It is accepted that older people are more likely to develop lung cancer and studies have demonstrated that smoking is also highly correlated, whereby trends in smoking preference are reflected in lung cancer incidence (ISD, 2009). There is also the possibility of a difference between the likelihood of males and females in developing these health outcomes. The same study (ISD, 2009) reported that there was a long term decline in male lung cancer incidence, but a 10% increase in the lung cancer incidence of females over the last 10 years. As a result of these issues, it was deemed appropriate to compute incidence ratios of the health outcomes according to the population structure across Greater Glasgow.

This was done using standardised incidence ratios (SIRs). These are a common summary measure used in mapping health outcomes, which compare the observed and expected number of cases in each region. The SIR is a form of indirect standardisation using age and sex specific rates from the standard population. Standard incidence ratios can be problematic in small areas or in cases of rare disease, where the estimates can be dominated by the variability across the sampling region. One other concern that can arise with SIRs is that they may not be directly comparable between regions if they are based on different populations, giving misleading results if the age and sex structures of the populations are dissimilar; although this rarely occurs in practice (Jarup, 2004). Therefore, it is only sensible to use SIRs when population structures are fairly similar between IG areas. The age and sex specific counts in each IG area are provided in Tables A2 and A3 in the Appendix of this thesis and showed a fairly similar population structure across Greater Glasgow and hence the use of SIRs was valid for the present study.

Therefore, to eliminate the effects of age and sex in the population, the health data were adjusted for age and sex using the total populations within each IG area. The age and sex specific populations used were from the 2001 census and were available to download from the SNS website (SNS, 2010).

In terms of lung cancer, the ratio between the observed and expected number of registrations defines the SIR for each IG area. This procedure was carried out using the following formulae.

$$SIR = \frac{Observed}{Expected} = \frac{\sum_{A}^{k=1} O_k}{\sum_{A}^{k=1} \left(\frac{y_k \lambda_k}{100{,}000} \right)}, \tag{4.1}$$

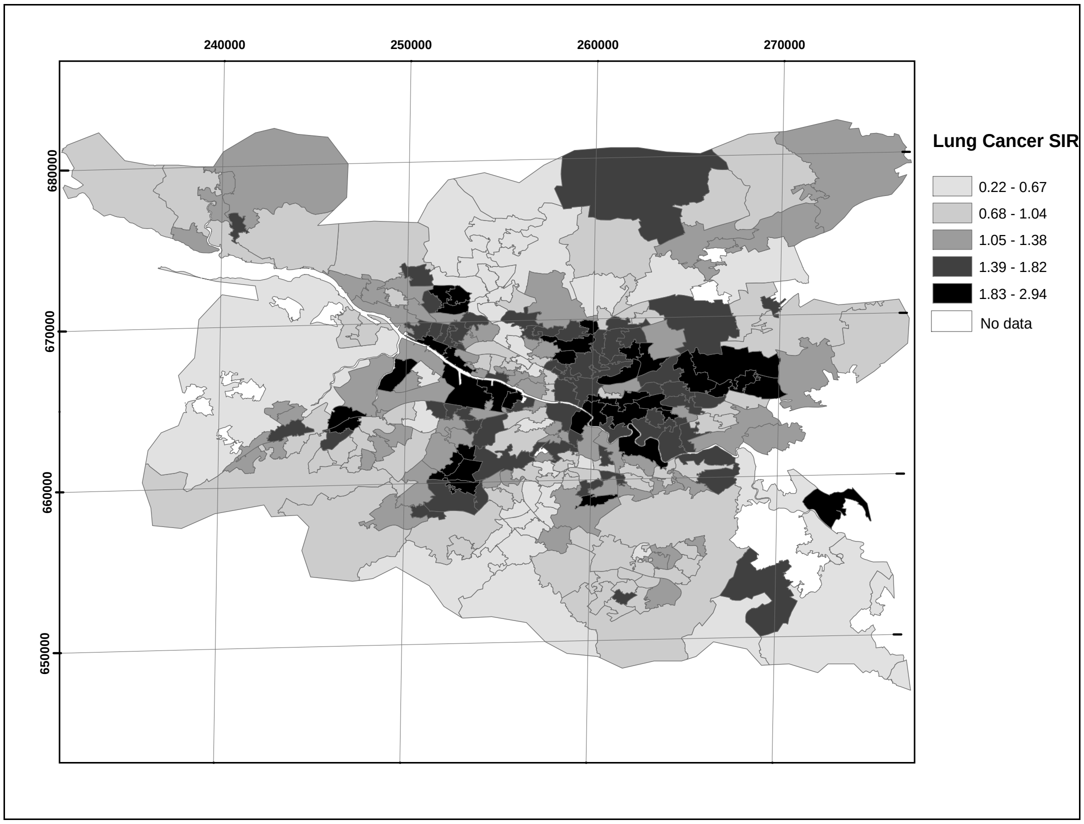
where k=1,….,A are the age groups 0-4,5-9,…,80-84,85+, for both male and female. The value O_k is the observed number of lung cancer registrations within each IG area, across the five year period from 1998 to 2002, as reported in the SNS data, y_k represents the expected number of lung cancer registrations in each age group provided by the ISD data and λ_k is the population of each age and sex group in each IG area. Since the age and sex specific lung cancer data were given as a count, the rate was computed by dividing the count by 100,000. The expected number of lung cancer registrations in each IG area was computed by summing the rates across each age group, for both sexes.

A slightly different approach was applied to the respiratory dataset since the age and sex specific hospital admissions from the ISD data were given as a rate per 100,000 for this outcome, rather than a count. Again, the age and sex specific populations from the 2001 census were used. In this case, the SIR for respiratory hospital admissions in each IG area was calculated by the following equation

$$SIR = \frac{Obs}{Exp} = \frac{\sum_{A}^{k=1} O_k}{\sum_{A}^{k=1} y_k \Big/ \lambda_k}, \tag{4.2}$$

O_k is the observed respiratory hospital admission rate within each IG area from 2002, y_k represents the expected respiratory admission rate in each age group provided by the ISD data and again λ_k is the population of each age and sex group. The expected rate of respiratory hospital admissions in each IG area was computed by summing the rates across each age group, for both sexes. Due to the advantages of the SIR approach over the raw counts and rate, SIRs were used in the remainder of the study.

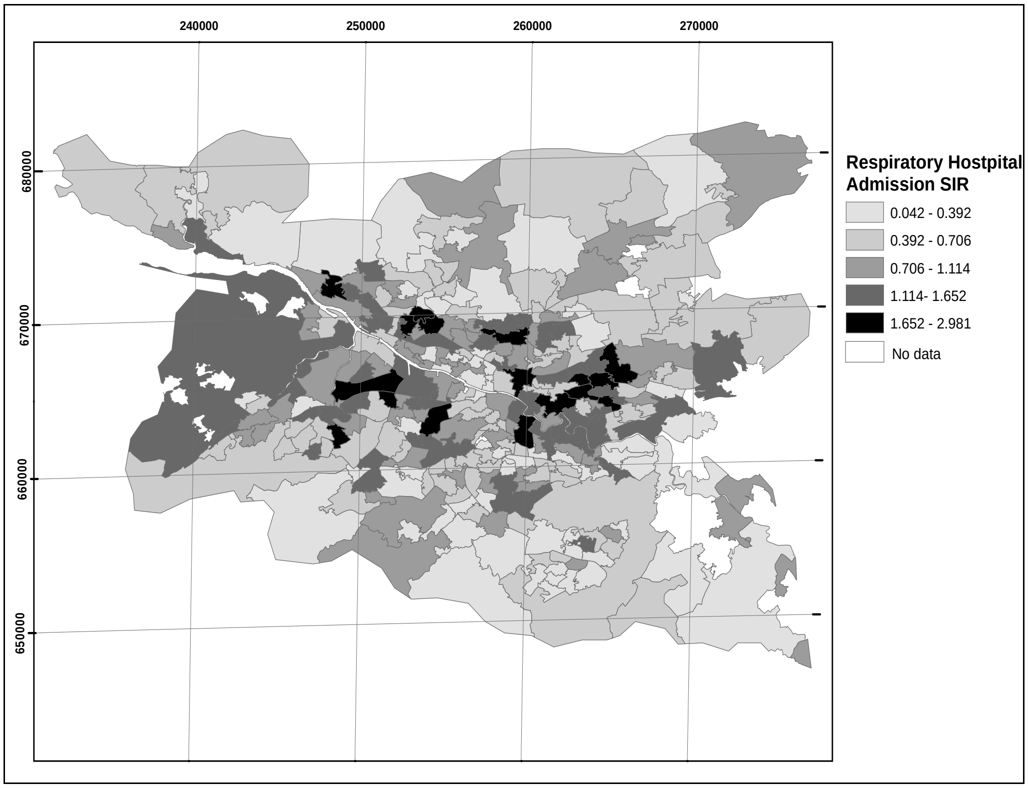
Lung cancer SIRs for each IG area are mapped in Figure 4.3 and show a very different distribution of lung cancer registrations to the raw rates observed in Figure 4.1. The map in Figure 4.3 indicates higher lung cancer incidence primarily in the East End of Glasgow, including areas such as Parkhead, Old Shettleston and Easterhouse, whilst areas of high lung cancer incidence south of the River Clyde include Govan, Ibrox, Gorbals and Crookstone (see Figures A1 and A2 and Table A1 in the Appendix for locations).



Rural IG areas on the edge of Glasgow have been clipped to the extent of the soil metal dataset

Figure 4.3 – Map of lung cancer SIRs across Greater Glasgow for each IG area

Standardised incidence ratios for respiratory hospital admissions are mapped in Figure 4.4. Areas with the highest incidence include parts of Paisley, Possil Park, Anniesland, Knightswood, Parkhead and Old Shettleston. Some urban areas with low incidence of respiratory admissions include Partick, Woodlands, Finnieston and Kelvinhaugh (see Figures A1 and A2 and Table A1 in the Appendix for locations).



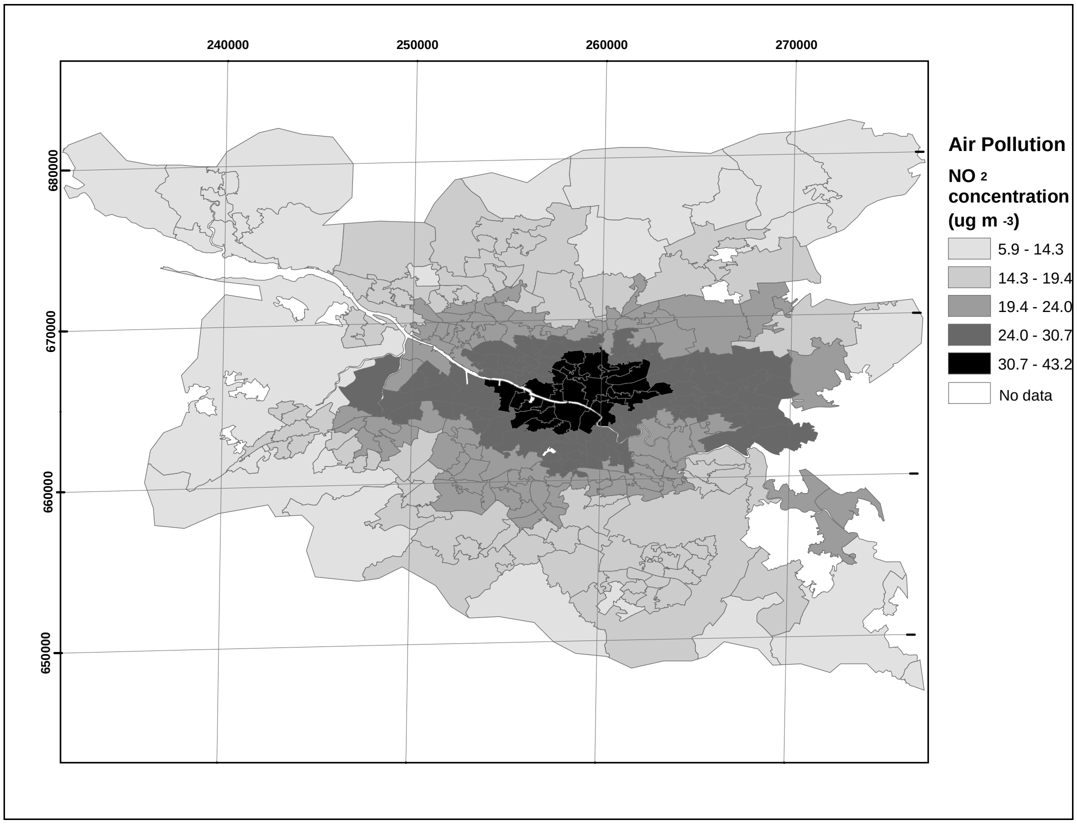
Rural IG areas on the edge of Glasgow have been clipped to the extent of the soil metal dataset

Figure 4.4 – Map of respiratory hospital admission SIR across Greater Glasgow for each IG area

4.2 Spatial Presentation of Air Pollution Data

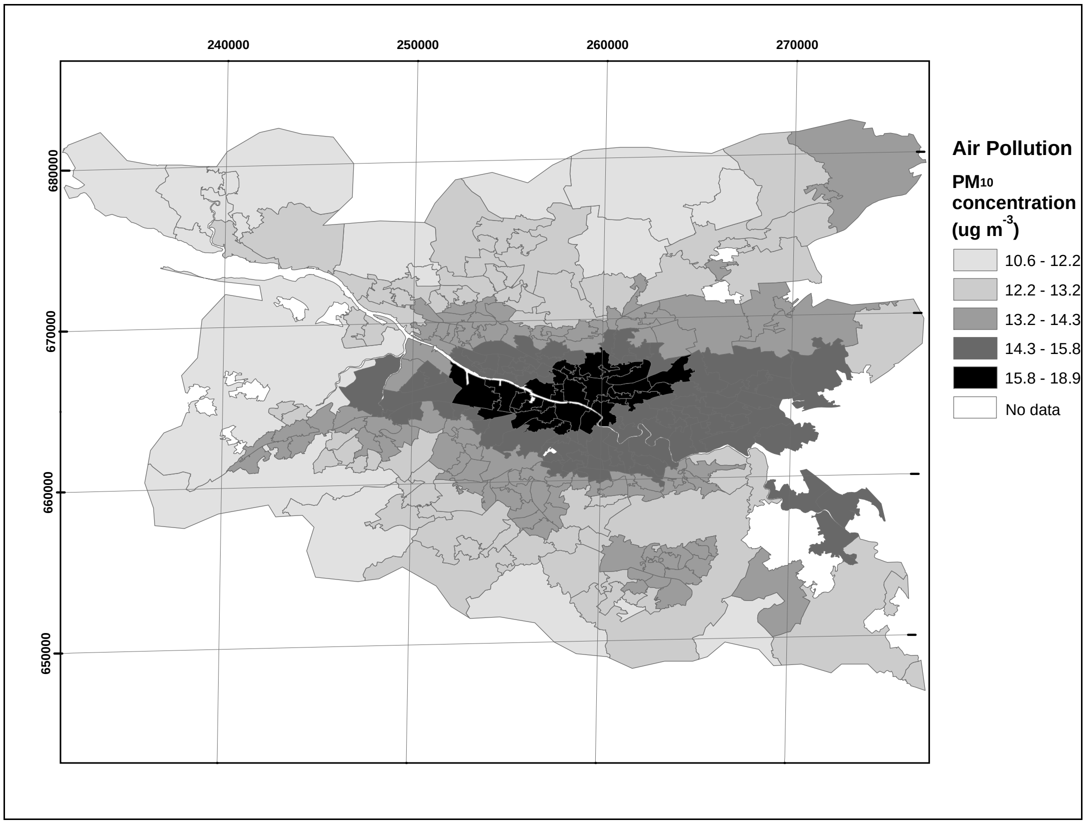
As mentioned in Chapter 2, the modelled data for air NO_2 and PM_{10} were treated as known average concentrations in each IG area. The UK Air Quality Objectives set by the WHO, outlined in the literature review (Chapter 1), stated that for air NO_2, the annual mean concentration to be achieved by December 2005 was 40 µg m^{-3}, while for PM_{10} the annual mean for Scotland was 12 µg m^{-3}, to be achieved by 2010.

When mapping at IG level, the modelled concentrations of air NO_2 and PM_{10} from 2001, a similar distribution for both pollutants was observed where elevated concentrations were found in the Inner City and East End of Greater Glasgow, while lower mean concentrations were recorded in rural areas (Figures 4.5 and 4.6). Annual mean air NO_2 levels were as high as 43 µg m^{-3}, while the highest recorded concentration for PM_{10} was 18.9 µg m^{-3}. In 2001, regions which exceeded the 2005 air quality objectives for NO_2 included the City Centre and Anderston (Figure 4.5). These same areas also exceeded the targets for air PM_{10} to be met by 2010 (Figure 4.6). Other regions which were close to surpassing the limits for both pollutants were Laurieston and Tradestone; Woodside; Finnieston and Kelvinhaugh, all of which are within urban areas of Greater Glasgow (see Figure A1 and Table A1 in the Appendix for locations).



Rural IG areas on the edge of Glasgow have been clipped to the extent of the soil metal dataset

Figure 4.5 – Map of mean air NO_2 concentrations across Greater Glasgow for each IG area



Rural IG areas on the edge of Glasgow have been clipped to the extent of the soil metal dataset

Figure 4.6 – Map of mean air PM_{10} concentrations across Greater Glasgow for each IG area.

4.3 Spatial Presentation of Deprivation Data

Since part of this project was to assess relationships between health outcomes and deprivation across Greater Glasgow, it was important to note that health was one of the domains included in the deprivation SIMD (SIMD, 2009a). Directly comparing these two variables would have produced erroneous conclusions, as health would have been accounted for in both datasets. Therefore, a new ranking system similar to SIMD 2009 was constructed for the current project that excluded the health domain. The same methods and weighting scheme adopted by the SIMD 2009 were applied, but computing the ranks across six domains, minus health. Ranks were weighted according to the values given in Figure 2.1. To construct the SIMD without health effects, ranks of the six domains were downloaded for each Datazone (DZ) from SNS (2010). The six domains were obviously measured on different scales and the data were already standardised so that the domains had the same distributions, ranging from 1 to 6505 (SIMD, 2009a). It is important to outline the problem that these rankings could not be used directly to create overall DZ ranks since the distributions were symmetrical and deprivation in one DZ and lack of deprivation in another, would cancel each other out; hence giving more weight to lower ranks. To account for this, an exponential transformation of the ranks was carried out using the following formula (SIMD, 2004).

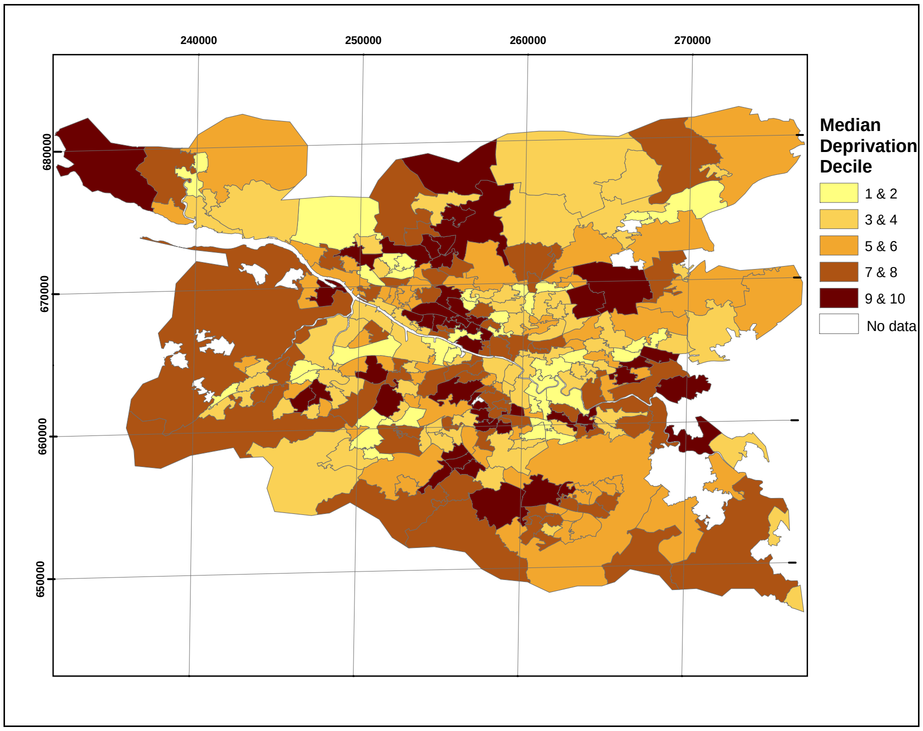
$$X = -23 \times \log\left\{1 - R \times \left[1 - \exp\left(\frac{-100}{23}\right)\right]\right\}, \tag{4.3}$$

where R is the rank transformed to lie in the range [0,1]. The constant -23 provides a 10% cancellation property to avoid equally weighted domains cancellation, when one area is deprived and the other not so deprived. After transforming to an exponential distribution, the ranks for each domain ranged between 0 and 100. These ranks were summed for each DZ, according to the weightings for each of the six domains given in the SIMD (2009a) methodology (Figure 2.1). The resulting scores were ordered to provide a ranking system from 1 to 6505, which excluded the health domain. Each ranked DZ was allocated a decile according to the classifications in Table 2.2.

In order to convert the deciles from DZ to IG level, a relevant summary measure was taken. The median value of the DZs within each IG area was selected as the most appropriate statistic to use. Using another measure such as the arithmetic mean would have generated an average decile with decimal places. However, the classification of deprivation needed to remain consistent as a categorical variable rather than a continuous variable. The median deprivation deciles of the original SIMD (2009a) dataset with health included are mapped in Figure 4.7 across Greater Glasgow. The median deprivations deciles for the re-constructed SIMD are shown in Figure 4.8. There are some differences observed between the two maps, with more deprived areas occurring in the East End of Glasgow in the re-constructed

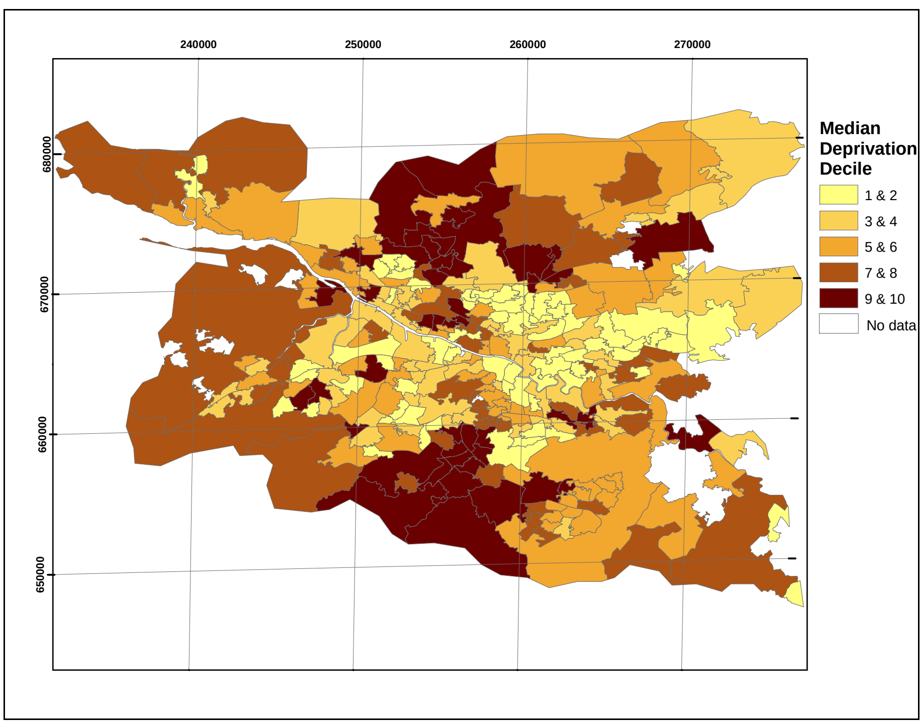
SIMD (Figure 4.8). Clusters of less deprived areas are also more apparent in Figure 4.8 in comparison to the original SIMD. These differences highlight the importance of removing the health domain from the SIMD, as this would have affected the results of this study.

The re-constructed SIMD map in Figure 4.8 indicated areas of higher deprivation in several areas of the East End including Easterhouse; Barlanark and Garthamlock and Auchinlea and Gartloch. Many areas within the Inner City were also highly deprived including Laurieston and Tradeston; Gorbals and Hutchesontown; Ibrox; Govan and Hillhead. Hillhead may seem somewhat surprising to be highly deprived, but this may be a result of the high population density of students living in the area, contributing to the lack of employment and low income. Some of the least deprived areas across Greater Glasgow included Dowanhill, Westerton, Kessington and Kilmardinny (see Figures A1 and A2 and Table A1 in the Appendix for locations).



Rural IG areas on the edge of Glasgow have been clipped to the extent of the soil metal dataset
Deprivation Deciles: 1 = least deprived, 10 = most deprived

Figure 4.7 – Map of median deprivation deciles across Greater Glasgow for each IG area, based on the original SIMD 2009 with all domains



Rural IG areas on the edge of Glasgow have been clipped to the extent of the soil metal dataset
Deprivation Deciles: 1 = least deprived, 10 = most deprived

Figure 4.8 – Map of median deprivation deciles across Greater Glasgow for each IG area, based on the re-constructed SIMD excluding the health domain

4.4 Spatial Presentation of Soil Metals Data

Since soil metal concentrations were recorded at point locations, a suitable summary measure was required so that this dataset could also be represented at the IG level, comparable to the other datasets in the study. As outlined in Chapter 3, the geochemistry datasets are highly skewed; therefore, the most appropriate summary measure was chosen to be the geometric mean. The advantage of the geometric mean is that it is robust. It is not dominated by high concentrations and outlying values, unlike other measures such as the arithmetic mean. The geometric mean (GM) in each IG area was calculated by multiplying the n raw values of soil metal concentrations within each of the 279 IG areas and then taking the n^{th} root of the product. Geometric means of each soil metal were calculated for all IG areas and are provided in Appendix A4. The GM for any element within a particular IG area is defined by equation 4.4, where c_i represents the soil metal concentration at point location i.

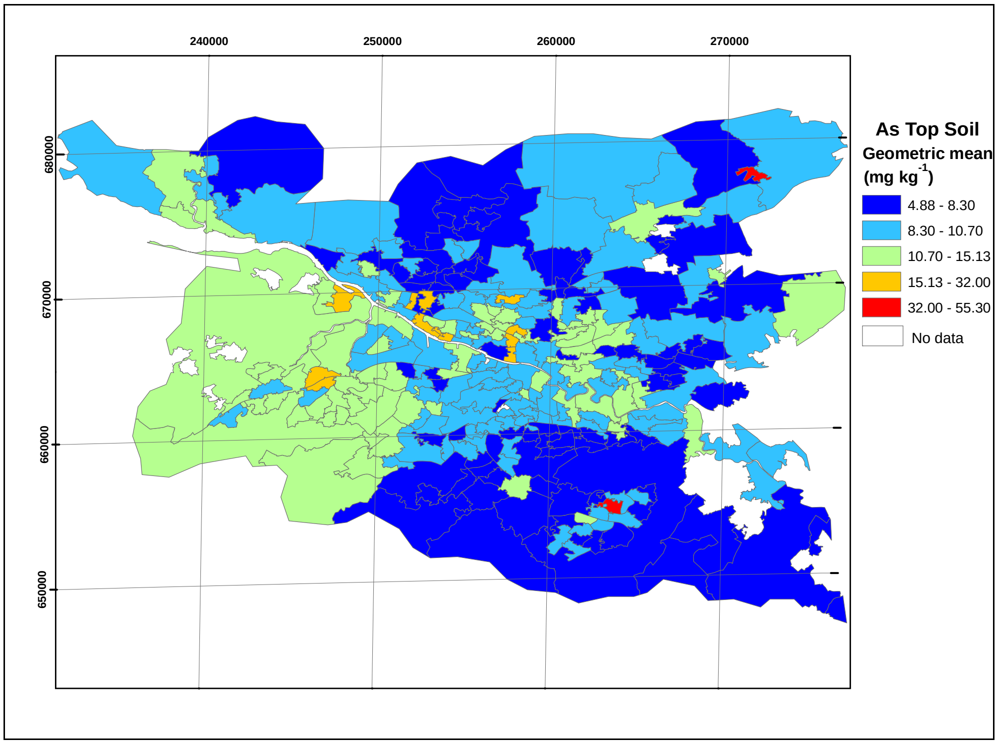
$$GM = \sqrt[n]{c_1 \times c_2 \times .... \times c_n}\ , \qquad (4.4)$$

For the purposes of observing patterns at IG level across the sampling region to compare with the other datasets, this was a more appropriate method of displaying the data than the spatially referenced concentrations.

Figures 4.9 to 4.16 illustrate maps of the geometric means for each IG area for each soil metal across Greater Glasgow. The highest bin-class in the soil metal concentration map legends for As, Cr, Pb and Ni were set to the CLEA soil guideline values (SVG) (Table 1.2) for these elements. The other bin-classes were selected by natural breaks as suggested by ArcGIS, with minor adjustments in rounding the values. The CLEA SGV above which soil-Se may be a potential health concern was 120 mg kg^{-1} (EA, 2009e) and none of the Se data in the Glasgow soils dataset exceeded this value (maximum 14.5 mg kg^{-1}). This is because Se is largely an essential element to health and is only toxic at very high concentrations in relation to its abundance in soil (Fordyce, et al, In Prep). Therefore, it was not possible to set the highest bin-class to the equivalent SGV for this soil contaminant in the map scale. For this element and for elements where no SGVs are defined (Cu, Zn, K_2O), natural breaks from ArcGIS were used, again with minor adjustment of the values.

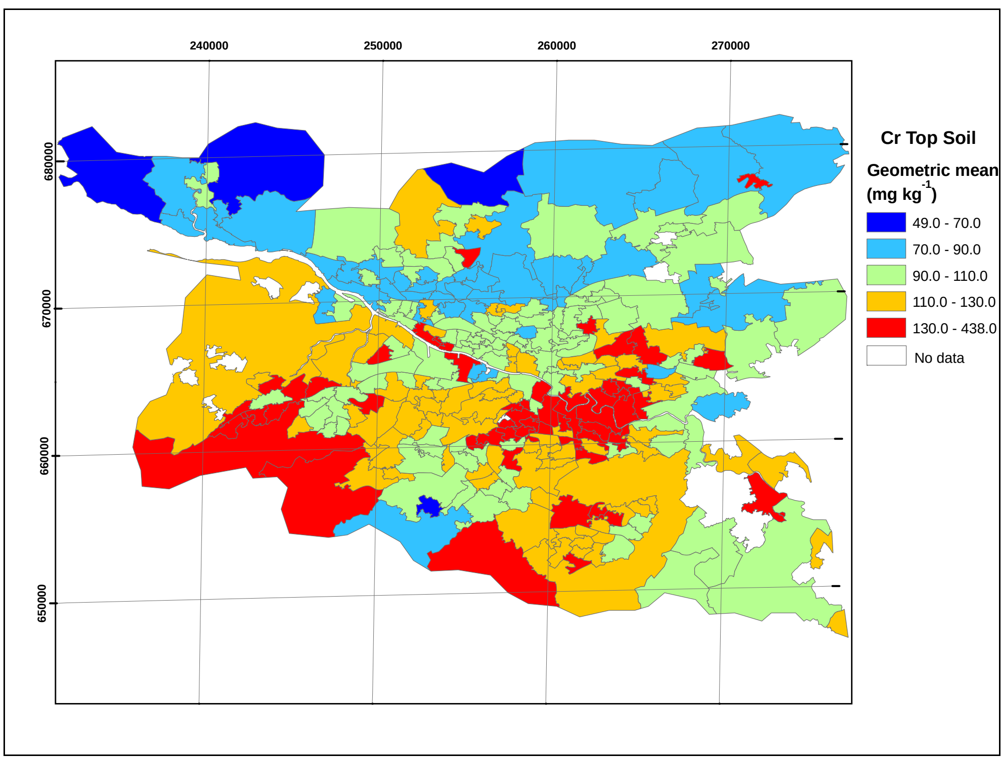
Figure 4.9 shows no obvious clusters of higher As geometric mean values across the region and only a couple of IG areas exceeded the CLEA SGV (32 mg kg^{-1}) for As soil concentration. However, lower concentrations are found in the rural areas of Glasgow, particularly in the south-east. Geometric mean Cr IG soil concentrations exceed the SGV (130 mg kg^{-1}) in many parts of Greater Glasgow, particularly the areas of Renfrewshire and the East End (Figure 4.10) (see Figures A1 and A2 and Table A1 in the Appendix for locations). This pattern is likely to be due to the industrial Cr-ore processing and steel making industries present historically in the East End of Glasgow and the volcanic rocks underlying large parts of Renfrewshire as highlighted in Chapter 3 of this thesis. Lower IG geometric mean soil Cr concentrations are mainly found in the northern rural areas of Greater Glasgow. Highest IG geometric means of soil Cu are located in the urban areas to the west of the city centre around the River Clyde, including Yoker, Scotstoun and Whiteinch (Figure 4.11). Other IGs with higher geometric mean concentrations are sporadically distributed and include Old Shettleston, Anderston, Cathcart and East Mains (see Figures A1 and A2 and Table A1 in the Appendix for locations). Geometric mean top soil Ni concentrations in only two of the IG areas (East Mains and Harestanes) exceed the SGV for Ni, although Whiteinch, and Scotstoun on the River Clyde as well as sporadic IG areas in the East End and south-west of Glasgow also have relatively high Ni concentrations (Figure 4.12) (see Figures A1 and A2 and Table A1 in the Appendix for locations).

For soil Pb, higher geometric mean concentrations are present in urban IG areas in particular, but only Anderston and East Mains exceeded the CLEA SGV of 450 $mgkg^{-1}$ (Figure 4.13). Highest IG soil geometric mean Se concentrations are located in upland areas at Kilsyth Bogside (see Figures A1 and A2 and Table A1 in the Appendix for locations), reflecting the affinity of this element for organic matter in soil (Fordyce et. al., In Prep). Moderately high geometric means are also clustered in the ship building areas on the River Clyde (Figure 4.14). Highest soil Zn IG geometric mean concentrations are associated with the urban centre of Glasgow including Finnieston, Anderston and Kinning Park. Geometric mean soil Zn concentrations are also high in Shettleston, Tollcross and East Mains as well as areas along the River Clyde including Yoker and Whiteinch (Figure 4.15) (see Figures A1 and A2 and Table A1 in the Appendix for locations). The map of soil K_2O geometric IG means across Greater Glasgow shows higher concentrations in the Dumbarton area in the north-west of Greater Glasgow, although Bothwell and Hamilton Centre in the south-west also have higher K_2O concentrations (Figure 4.16) (see Figures A1 and A2 and Table A1 in the Appendix for locations). As discussed in Chapter 3, the distribution of this element in soils is influenced by the Devonian sandstones present in Dumbarton and glaciofluvial deposits in south-west Glasgow.



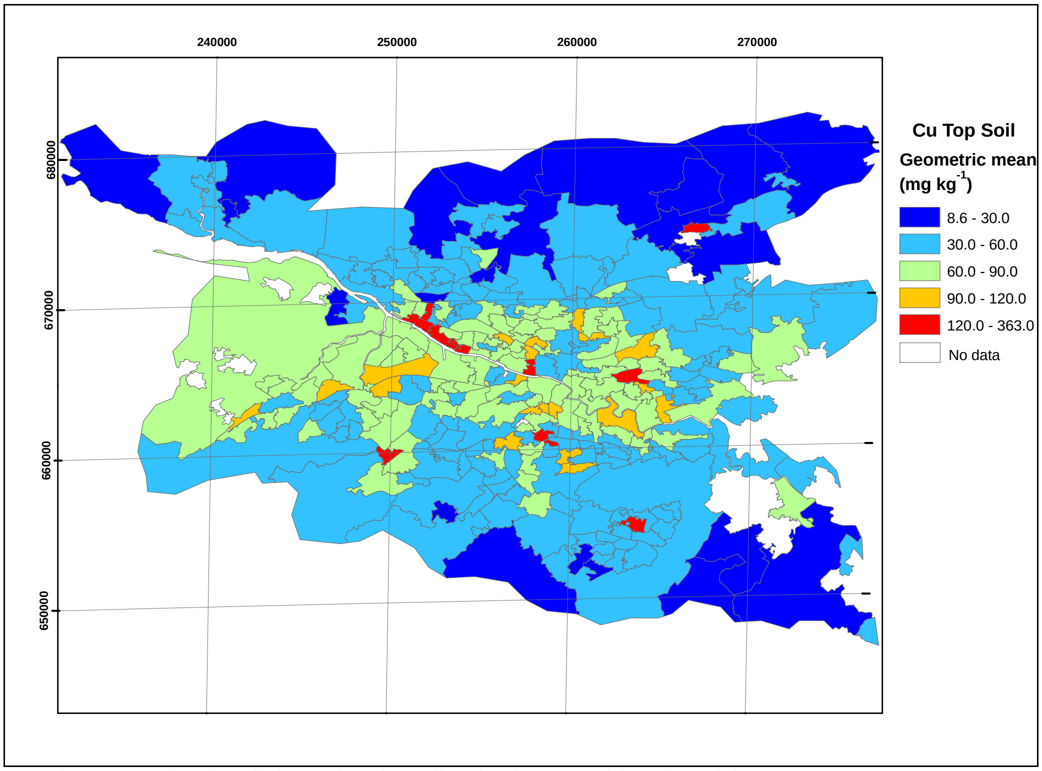
Rural IG areas on the edge of Glasgow have been clipped to the extent of the soil metal dataset

Figure 4.9 – Map of geometric mean soil As concentration in each IG across Greater Glasgow



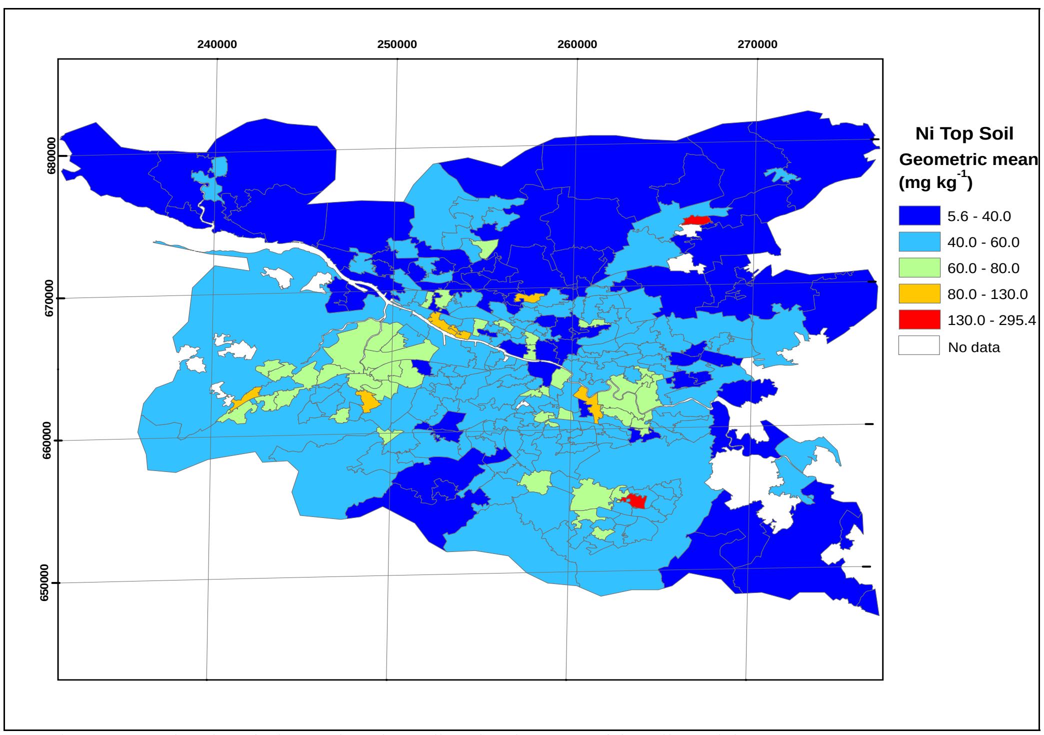
Rural IG areas on the edge of Glasgow have been clipped to the extent of the soil metal dataset

Figure 4.10 – Map of geometric mean soil Cr concentration in each IG across Greater Glasgow



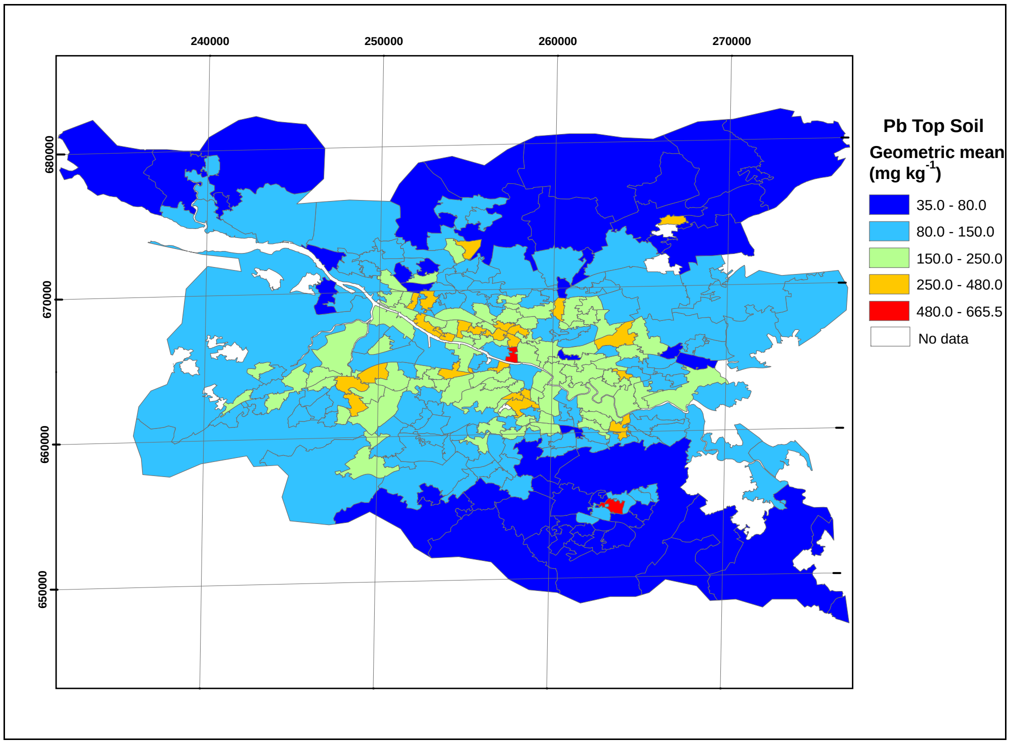
Rural IG areas on the edge of Glasgow have been clipped to the extent of the soil metal dataset

Figure 4.11 – Map of geometric mean soil Cu concentration in each IG across Greater Glasgow



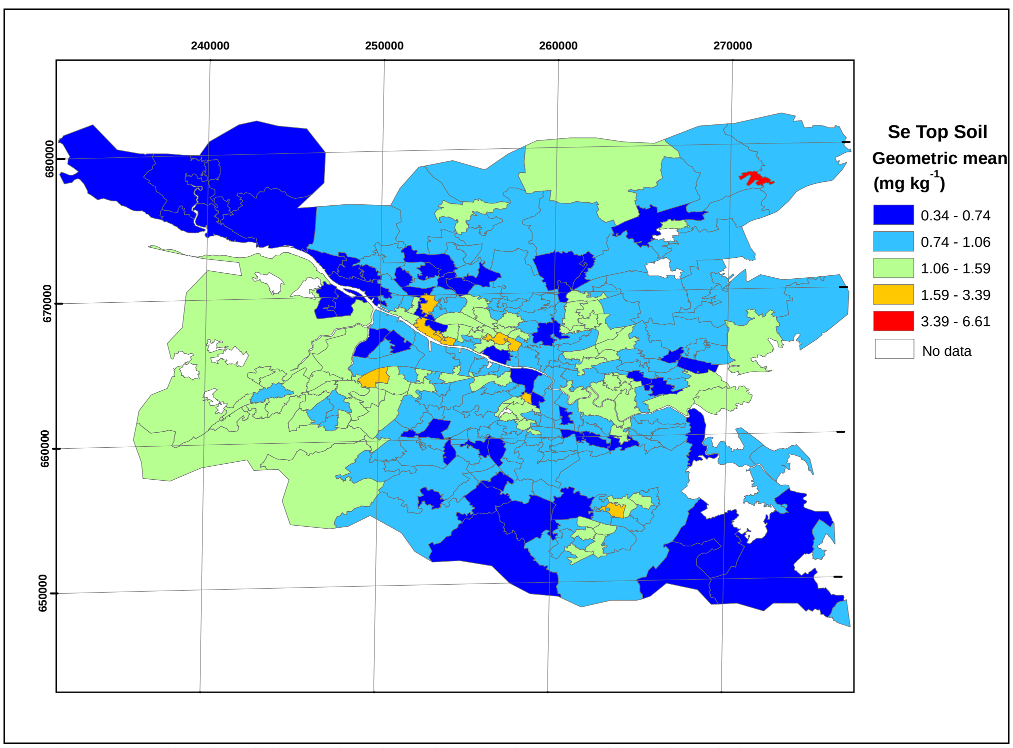
Rural IG areas on the edge of Glasgow have been clipped to the extent of the soil metal dataset

Figure 4.12 – Map of geometric mean soil Ni concentration in each IG across Greater Glasgow



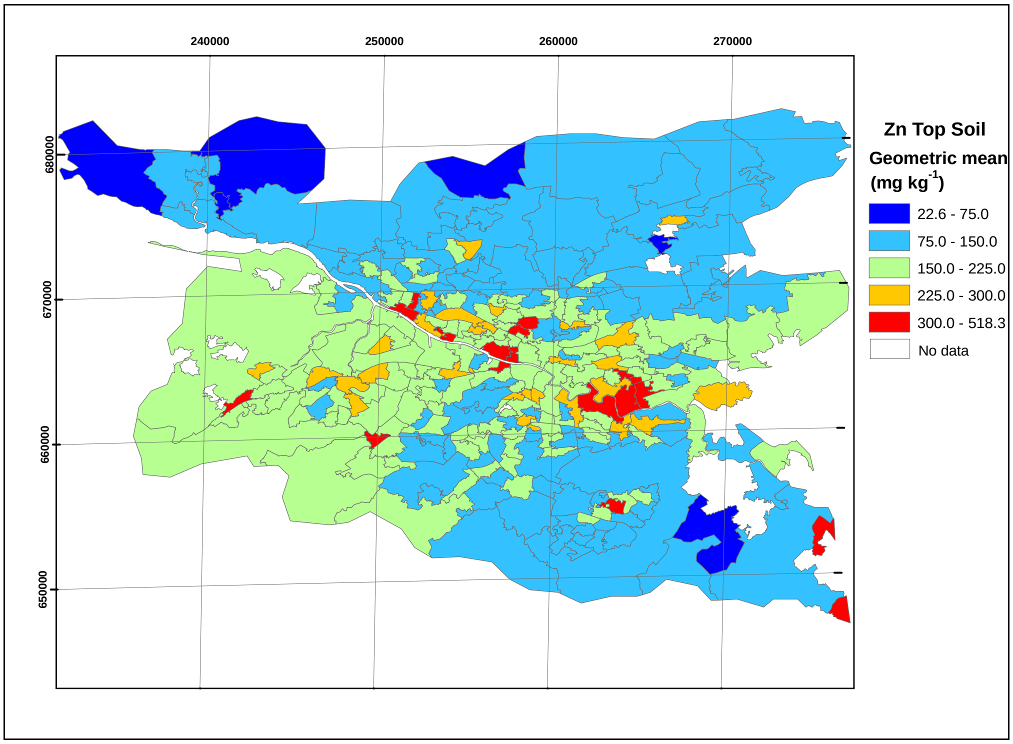
Rural IG areas on the edge of Glasgow have been clipped to the extent of the soil metal dataset

Figure 4.13 – Map of geometric mean soil Pb concentration in each IG across Greater Glasgow



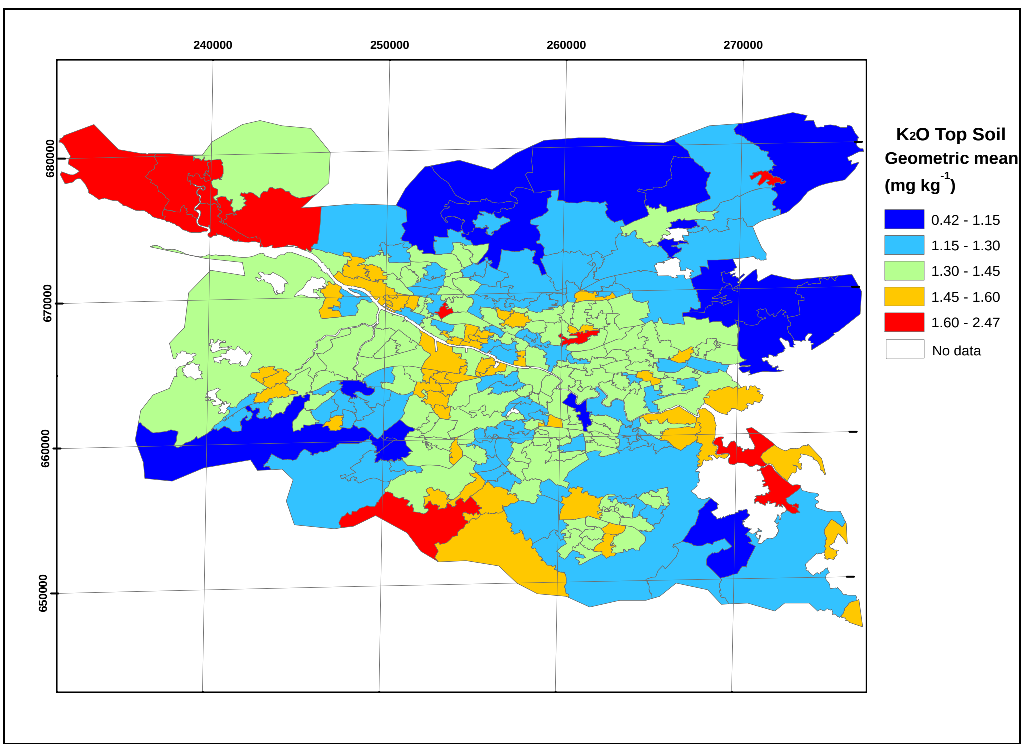
Rural IG areas on the edge of Glasgow have been clipped to the extent of the soil metal dataset

Figure 4.14– Map of geometric mean soil Se concentration in each IG across Greater Glasgow



Rural IG areas on the edge of Glasgow have been clipped to the extent of the soil metal dataset

Figure 4.15 – Map of geometric mean soil Zn concentration in each IG across Greater Glasgow



Rural IG areas on the edge of Glasgow have been clipped to the extent of the soil metal dataset

Figure 4.16 – Map of geometric mean soil K_2O concentration in each IG across Greater Glasgow

4.5 Creating an Overall Index of Soil Metal Quality

As stated in Chapter 1, the purpose of the project was to investigate the possibility of a spatial relationship between soil metal levels and health outcomes. Therefore, in addition to examining the relationships between each of the soil metals and health/deprivation/air quality, for the purposes of statistical analysis, an overall measure combining soil metal concentrations into one indicator of soil quality across the region was constructed at the IG level. Using appropriate statistical techniques, it was possible to characterise multiple metals into a single index and this was one of the key stages of this study.

In constructing an index of soil metal concentration, the elements with relevance to health were identified. Only As and Pb have no known biological function, while Cr, Cu, Ni, Se and Zn are all essential elements for human health (Fordyce et al., In Prep). They are all potentially toxic at high concentrations, although as mentioned in Chapter 1, under the current CLEA guidelines, Cu and Zn are not thought to pose a health risk via the soil exposure route (EA, 2010a). The metals deemed of potential concern to human health under CLEA guidelines and for which soil data were available for this project were As, Cr, Ni, Pb and Se and these were used to construct the index.

A variety of methods could be adopted to construct an overall measure of soil metal concentration. Richardson et al. (2010) developed an indicator of various environmental dimensions called MEDIx where exposure to detrimental dimensions such as high air pollution were allocated a score of +1, while beneficial indicators such as greenspace were given a score of -1. A similar strategy could have been considered to create a soil metal index in this project, where low concentrations of soil metals could be allocated a negative score since some elements are essential nutrients to the human body, as discussed in Chapter 1. This could have been done on the basis of assigning a value of -1 to soil metal concentrations below the SGV and + 1 to those above the SGV. However, SGVs indicate the concentration above which further investigations, into whether soils pose a risk to health, should be carried out, and not the concentration at which soils are known to cause health problems (EA, 2009b). Many factors, particularly the bioavailability of metals in soil influence their uptake and subsequent toxicity. Furthermore, exposure to multiple metals (more than one metal at a time so that effects can be cumulative) is very poorly understood and is not taken into account in the CLEA guidelines because it is so complex (EA, 2009a). Therefore, the concentration ranges at which combinations of soil metals are beneficial or detrimental to human health are not well understood and if this approach was taken it would have caused ambiguity in constructing the index. Furthermore, the influence on health for each combination of soil metals is likely to vary, meaning a different soil index would be required for the various health indicators. The process of creating an overall soil metal index to relate to health could be highly complex.

Therefore, for this overview project, a comparatively simple index based on the relative concentrations of the metals in Glasgow soils was constructed as a summary measure of land quality *per se,* rather than of land likely to pose a possible threat to human health. The log transformed data for soil As, Cr, Ni, Pb and Se were used in creating the index, which was based on percentiles of the data distributions for each of these elements.

The following steps were taken in order to construct the index:

- Each of the 1622 soil concentrations for the five metals was allocated a score from 1 to 10 according to its percentile Fn(x) from the empirical cumulative distribution function (cdf) so that higher concentrations were allocated a higher score.
- The scores for the five metals were summed together to generate a total soil metal index for each point location.
- The mean total metal score within each IG area was computed.

In the case of Glasgow, one could argue that for the purposes of relating the soil data to health, Se should have been excluded from the soil index making it a sum of four rather than five metals since all soil Se concentrations across Glasgow were shown to be considerably below the CLEA SGV. Although soil Se concentrations in themselves are well below the SGV, Se in combination with other metals may have a cumulative exposure effect. Since the aim of the soil metal index in this project was to provide an overall measure of relatively high and low potentially harmful metal content in the soil, Se was included.

A property of the index is that the summed scores at each point location can range from 5 to 50, since the lowest and highest score each of the five metals can have is 1 and 10 respectively. Ideally, one would like the soil metal index to be valid and reliable. Validity refers to whether the index actually measures what is desirable and reliability that it is consistent to the measured concentrations. Since the index is based on percentile distributions, it is certainly reliable as this method could be applied to any other dataset of this nature.

Therefore, although the five elements As, Cr, Ni, Pb and Se were included in the index for the purposes of this study, in other cities or areas, fewer metals or more metals may be of potential concern to health and could be included in an index of this type. However, since the metal scores depend on the number of metals included, to compare two areas the number of metals in the index would have to be the same in each area.

Since the computation of this soil metal index is relatively straightforward, there are imperfections. One flaw is that a particular soil metal score cannot be interpreted in terms of concentration in mg kg^{-1}, also making it difficult to determine whether concentrations should be classified as high or not. However, it is difficult to avoid this when combining soil metal concentrations.

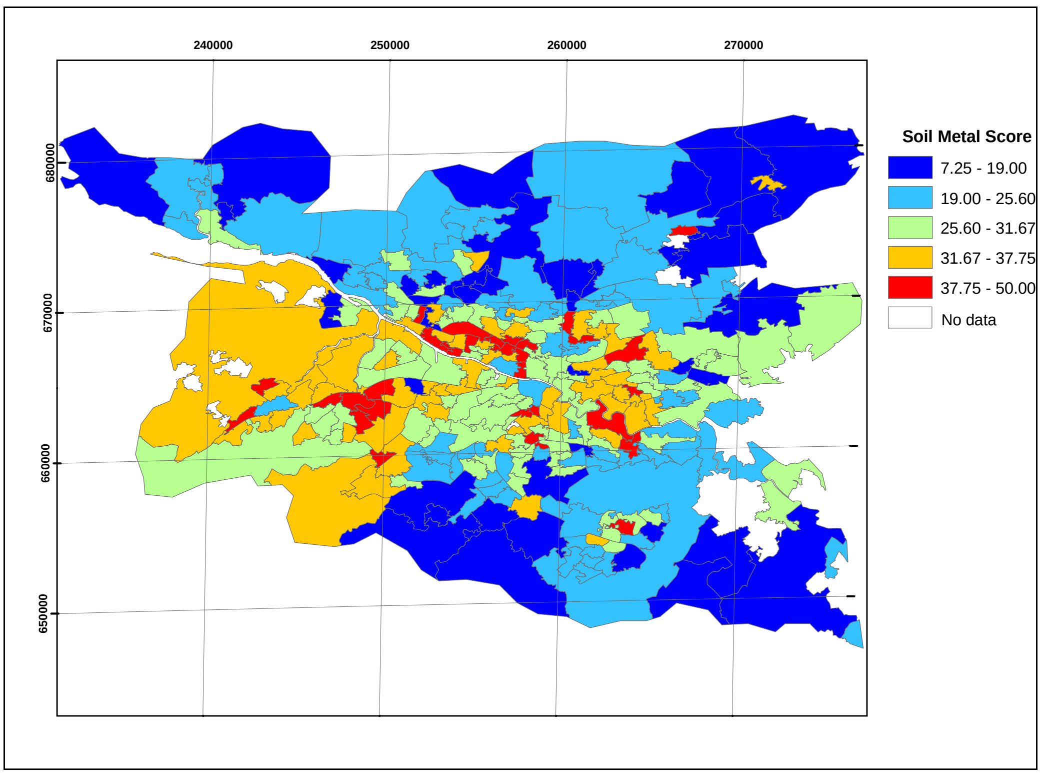
In the scoring system, each of the five soil contaminants effectively has an equal weighting on the total scores. One could argue that in the case of Glasgow, for example, the contribution of Se to the index should be negligible since all soil Se concentrations were shown to be considerably below the CLEA SGV, while geometric mean values for many IG areas exceeded the SGV for Cr. Therefore, the validity of the soil index in terms of links to health may have been improved by applying some form of weighting to the soil metals. However, the difficulties of trying to apply a weighting based on simply whether a soil metal concentration falls above or below the SGVs have been outlined above.

Similarly, devising a weighting system based on the relative concentrations of the SGVs for each of the metals would be problematic due to the differences in typical element abundances in soils. For example, the 1.8 mg kg^{-1} SGV for Cd (Table 1.2) is a reasonably high concentration of Cd relative to its typical abundance in soils, whereas it is a very low concentration for an element like Cr. Therefore, the SGVs could not be used as a weighting factor. For example, it is not true to say that since the SGV for As is 32 mg kg^{-1}, whereas the SGV for Cr is 130 mg kg^{-1}, that As is 4 times more toxic or problematic in soil than Cr.

However, as part of the CLEA guidelines, the EA reports (EA, 2010a) on the Tolerable Daily Intake (TDI) for a variety of soil contaminants, where the Index Dose (ID) of both inhalation and oral intake are outlined give an indication of the relative toxicity of the metals. These ID values could be used to create a weighting scheme for the soil contaminants of concern and to construct a soil metal index in the future. For example, the ID_{oral} intake for As is 0.3 $mgkg^{-1}$ while the ID_{oral} intake for Cr is 3 $mgkg^{-1}$, so that As could be deemed to be ten times as toxic as Cr and would therefore contribute more to the soil metal index. However, these weightings may have to be adjusted for different health outcomes and this would require further study and investigation beyond the scope of this thesis.

Although this investigation has applied a fairly simple approach to combining soil metal concentration data in order to make them comparable to other environmental and health datasets, this method of ranking according to percentiles of the data distributions has the advantage that is it readily applicable for a first pass assessment of relationships such as that being carried out in this study.

The soil metal index provides a summary measure of top soil metal concentrations across Greater Glasgow and is shown in Figure 4.17. The map clearly indicates higher scores to the south-west of Glasgow in Renfrew and Paisley. High combined soil metal scores are also present in urban West End areas of Glasgow, including Scotstoun, Whiteinch and Firhill. Relatively high soil metal scores are also prevalent in several areas in the East End, including Tollcross; Riddrie and Hoganfield; Dennistoun; Shettleston and Parkhead (see Figures A1 and A2 and Table A1 in the Appendix for locations).



Rural IG areas on the edge of Glasgow have been clipped to the extent of the soil metal dataset

Figure 4.17 – Map of average combined soil metal scores for each IG area across Greater Glasgow

4.6 Chapter 4 Summary

This chapter has presented the soil metal, air pollution, deprivation and health data used in this study to allow for further exploratory analysis and an assessment of the associations between the variables. Standardised incidence ratios were computed for the health outcomes, which accounted for the age and sex distributions in each IG area. After adjusting for population structure, higher incidence of both lung cancer registrations and respiratory hospital admission were present in urban areas and especially the east of Greater Glasgow. Elevated air pollution concentrations were also found in urban areas of Greater Glasgow, with comparatively lower concentrations in rural areas. A soil metal index was created to obtain a measure of general soil metal levels across the region that could be used to identify areas with high overall soil metal concentrations. The index was based on distribution percentiles of the five metals of concern under CLEA guidelines; As, Cr, Ni, Pb and Se. High metal scores were present in the rural areas of Renfrewshire and Paisley in the west of Glasgow and also several urban areas and the East End (see Figures A1 and A2 and Table A1 in the Appendix for locations).

Chapter 5

Relationships between Soil Metals, Deprivation, Air Pollution & Health Indicators

In addition to the mapped presentations outlined in Chapter 4, spatial associations between the environment and health/deprivation datasets were assessed by using exploratory statistics before the relationships between the variables were modelled more fully.

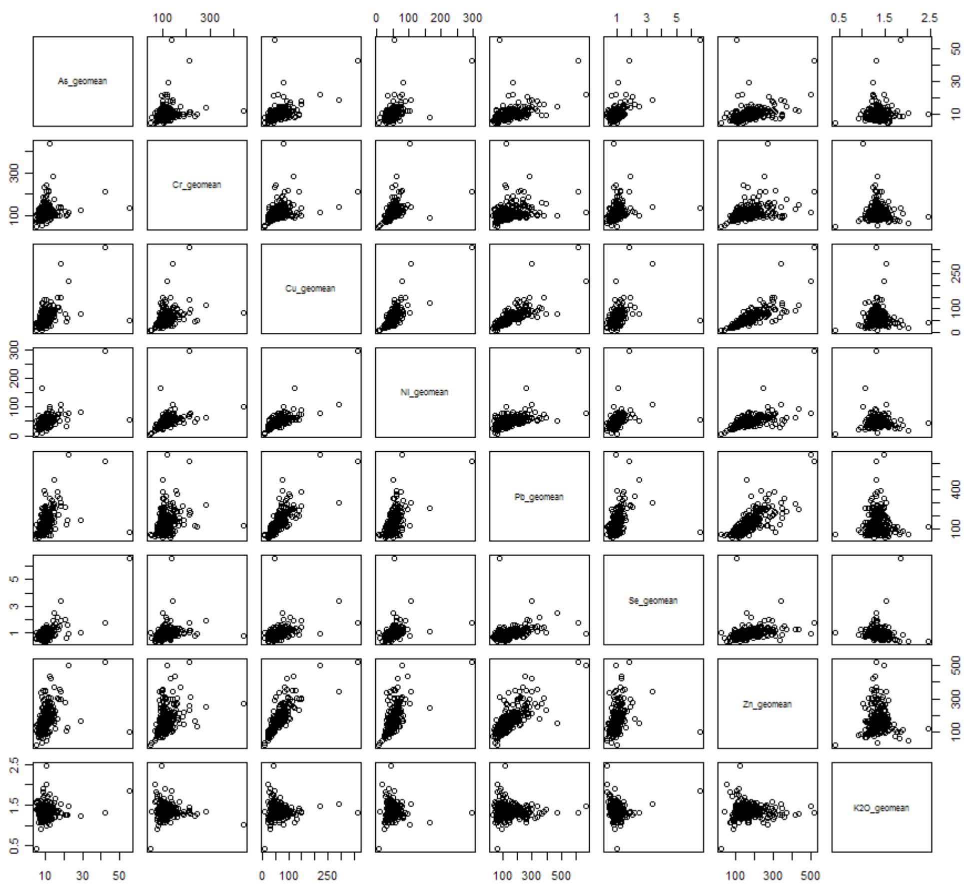
5.1 Initial Exploration of Spatial Associations between Environmental and Health Variables

Initial assessments of the spatial associations between environment and health/deprivation datasets were carried out using a series of scatterplots and statistical tests.

- Strength of association was assessed using the Pearson product moment correlation coefficient, where figures in italics represent those that are statistically significant.
- Moran's I test was used to investigate the relationships in further detail by testing the hypothesis of a possible spatial structure.

5.1.1 Comparisons between Metals In Soil

The matrix plot in Figure 5.1 illustrates the relationships between the geometric means of soil metals for the IG areas, and suggests a positive linear association between several of the elements, with the exception of K_2O. This particular element shows very little association with the other soil metals. However, this was expected since K_2O has a very different geological distribution to the other metals and was only included in the study as a control element. There are a few outliers present in all of the scatterplots, but this was expected since the geometric means in IG areas containing only two or three soil data points were dominated by any extreme outliers present. One could remove these outliers from the analysis for simplicity, but this would result in losing important information as these high soil metal concentrations are of interest.



n=279. There are 279 IG areas with a geometric mean of each soil metal taken for each area

Figure 5.1 – Matrix plot of geometric means of soil metals

Pearson correlation coefficients between each of the IG geometric means for soil metals are illustrated in Table 5.1 and reveal significant (95% confidence level) correlations between many of the soil metals. Statistically significant associations were observed between all soil contaminants, excluding geometric means of soil K_2O, as correlation coefficients exceeded the 0.118 threshold, based on 279 samples at the 95% confidence level. As expected, K_2O showed no significant relationship with the other metals in soils, with As being an exception. Both Cu and Ni have high (> 0.60) positive correlation coefficients with several of the other soil metals, while others demonstrate moderately positive correlations (0.200 – 0.600). Some of these high correlations may also be influenced by high geometric mean values in the dataset. This multi-collinear relationship between the soil metals created a problem when carrying out regression modelling and is discussed in Section 5.3.

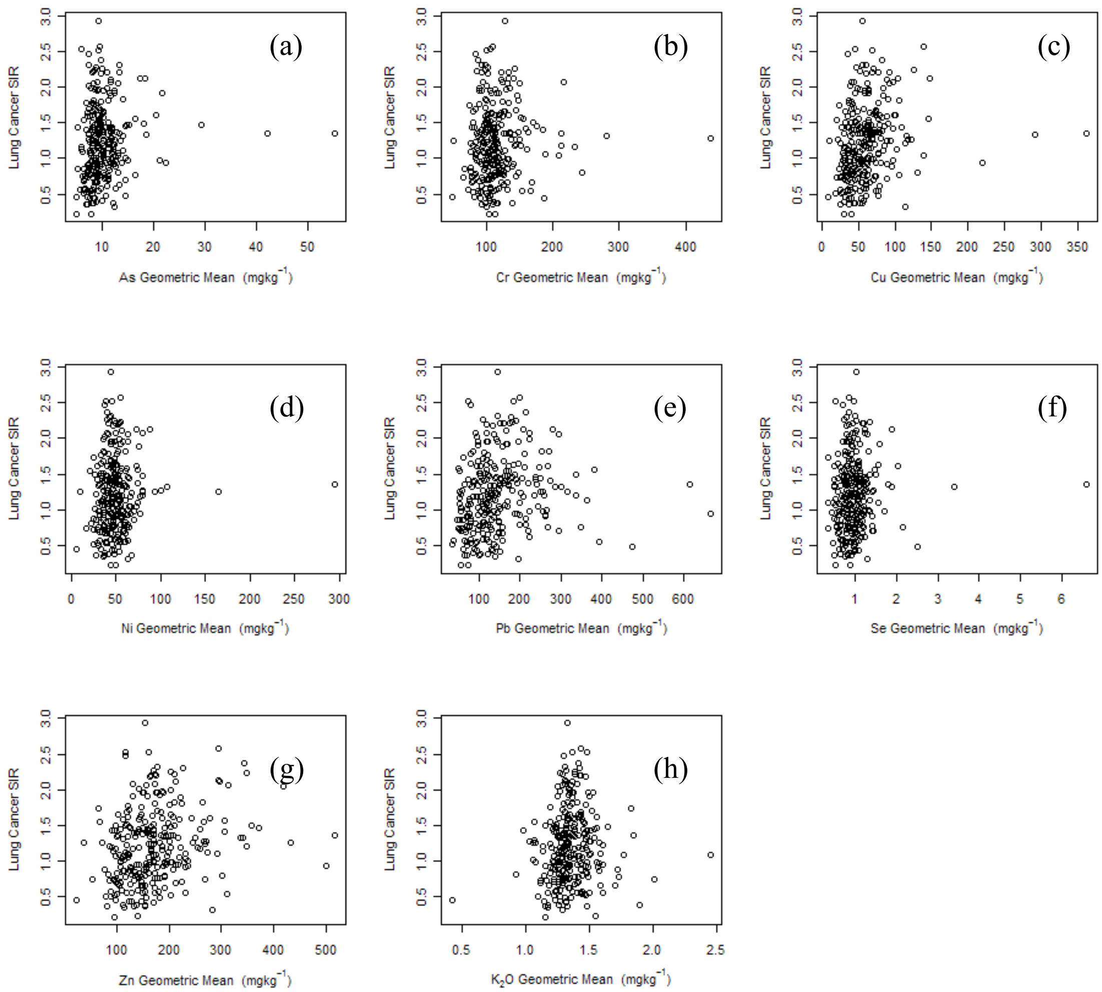
Table 5.1 - Pearson correlation coefficients between soil metal geometric means

	As	**Cr**	**Cu**	**Ni**	**Pb**	**Se**	**Zn**	**K_2O**
As	1							
Cr	*0.214*	1						
Cu	*0.520*	*0.356*	1					
Ni	*0.524*	*0.477*	*0.777*	1				
Pb	*0.469*	*0.260*	*0.794*	*0.581*	1			
Se	*0.706*	*0.214*	*0.390*	*0.343*	*0.365*	1		
Zn	*0.393*	*0.377*	*0.797*	*0.611*	*0.758*	*0.286*	1	
K_2O	*0.131*	-0.066	0.026	-0.010	-0.019	0.009	0.009	1

n = 279, r95% confidence level = 0.118. Figures shown in italics are statistically significant

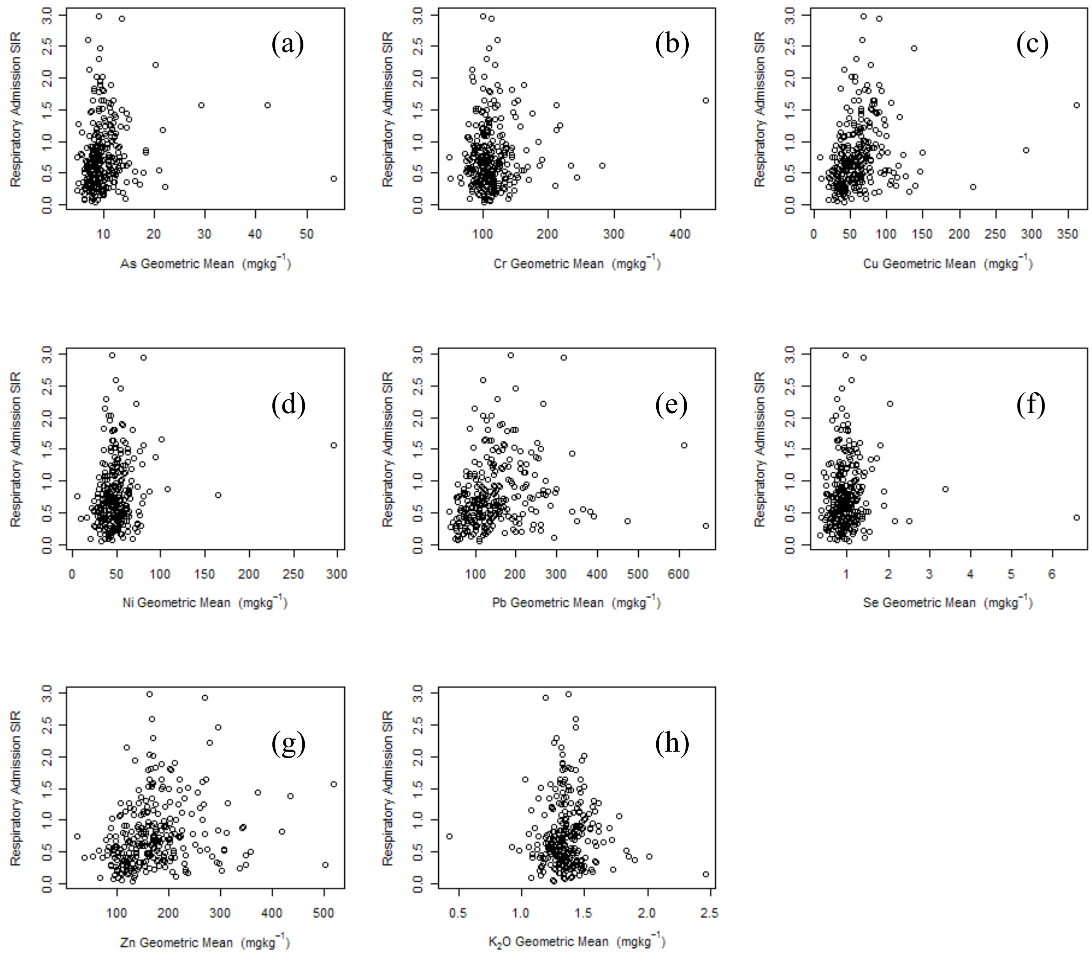
5.1.2 Relationships between Individual Soil Metals and Health Datasets

It is also of importance to assess the relationships between each of the soil metals and the health outcomes of interest. The plots in Figures 5.2 (a) to (h) demonstrate fairly weak relationships between the soil metals and lung cancer SIR. From observation of the plots, only Cu, Pb and Zn seem to have a moderately positive relationship with the lung cancer incidence. Figures 5.3 (a) to (h) illustrate similar weak relationships between most soil metals and respiratory hospital admission SIR and moderate associations for Cu, Pb and Zn.



n=279 for each scatterplot.

Figure 5.2 – Soil metal geometric means v lung cancer SIR



n = 279 for each scatterplot

Figure 5.3 – Soil metal geometric means v respiratory hospital admission SIR

The Pearson correlation coefficients between the soil metal geometric means and health outcomes are given in Table 5.2. Note that while some of the correlations are statistically significant, the sample correlation coefficients are relatively small. The results showed a statistically significant (95% confidence level) correlation between lung SIR and soil metals for Cu (0.215), Pb (0.141) and Zn (0.246), while none of the correlations for the other soil metals were significant. For respiratory hospital admissions, statistically significant correlations were found for As (0.125), Cu (0.239), Ni (0.191), Pb (0.199) and Zn (0.223). Although these might suggest a moderate correlation between areas of elevated soil metal concentration and health problems, this does not necessarily imply a direct causal relationship between these parameters. It may reflect an unrelated spatial coincidence. Therefore, relationships between soil metals and health outcomes were investigated further in Section 5.1.3 of this chapter.

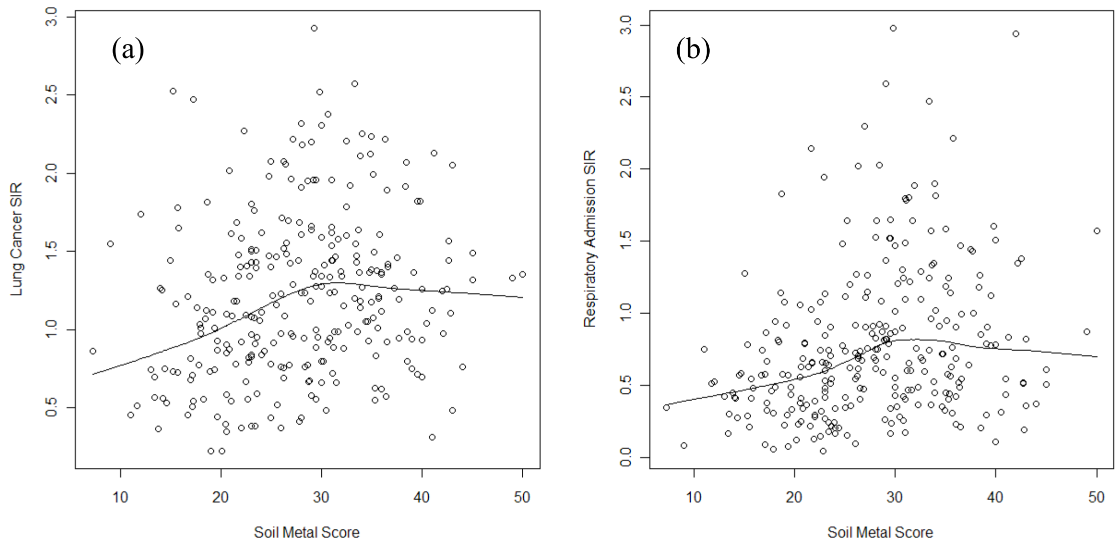
Table 5.2- Pearson correlation coefficients between soil metal geometric means and health outcomes in each IG

	As	**Cr**	**Cu**	**Ni**	**Pb**	**Se**	**Zn**	**K_2O**
Lung Cancer SIR	0.117	0.055	*0.215*	0.075	*0.141*	0.060	*0.246*	0.081
Respiratory Admission SIR	*0.125*	0.094	*0.239*	*0.191*	*0.199*	0.052	*0.223*	-0.004

n = 279, r95% confidence level = 0.118. Figures shown in italics are statistically significant

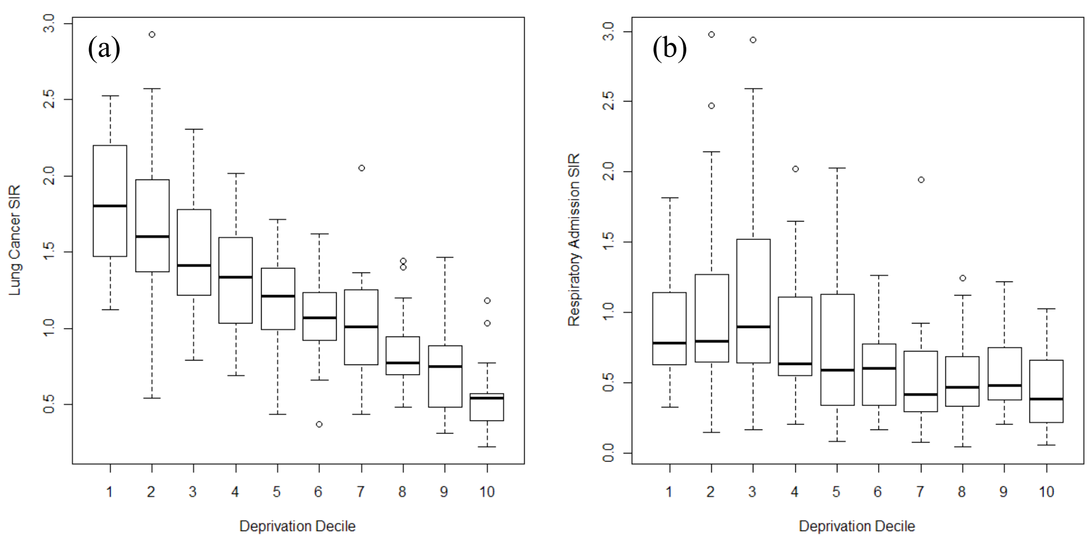
5.1.3 - Relationships between Soil Metal Score, Health and Deprivation

Having assessed the associations between the individual soil metals and health outcomes, it was necessary to look at the relationships the other environmental variables have with health as well as using the scoring index created in Section 4.5 to represent soil metal concentrations across Greater Glasgow. The associations between soil metal score and the SIR of both health outcomes are illustrated in Figure 5.4 (a) and (b), with a fitted loess line plotted. The pattern for both health outcomes suggests a weak to moderate positive linear relationship, with the variability in the both responses increasing as soil metal score increases. The relationship between deprivation deciles and lung and respiratory health indicators are illustrated in Figure 5.5. Figure 5.5 (a) clearly indicates a strong negative linear relationship between deprivation and SIR of lung cancer so that more deprived areas in Greater Glasgow tend to have higher incidence of lung cancer. Figure 5.5 (b) shows a similar but much weaker trend between deprivation and respiratory hospital admissions.



n = 279

Figure 5.4 - (a) Soil metal score v lung cancer SIR and (b) Soil metal score v respiratory hospital admission SIR, with fitted loess lines

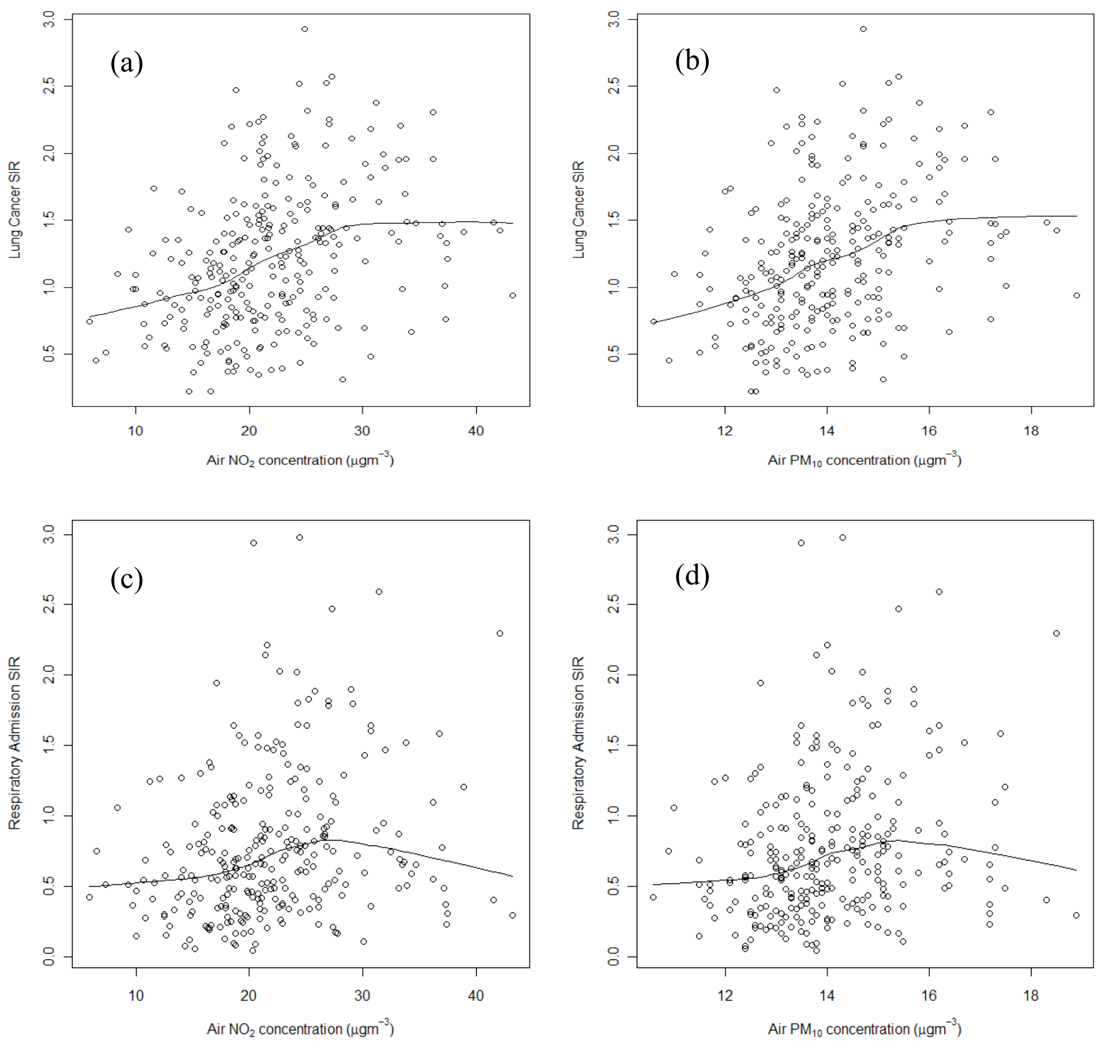


n = 279, Deprivation Decile scale: 1 = most deprived, 10 = least deprived

Figure 5.5 – (a) Median deprivation decile v lung cancer SIR and (b) Median deprivation decile v respiratory hospital admission SIR

Figures 5.6 (a) to (d) illustrate the relationships between the two air pollutants and both health outcomes. Figures 5.6 (a) and (b) reveal a moderate linear relationship between lung cancer incidence and both air NO_2

and PM_{10} across Greater Glasgow. The association between air pollution and incidence of respiratory hospital admissions is not so obvious. The fitted loess lines in Figures 5.6 (c) and (d) suggest a gradual increase in respiratory admissions with increasing air pollution, although there is a lack of data at high pollution concentrations to observe whether this trend continues.



n = 279

Figure 5.6 - (a) Air NO_2 concentration v lung cancer SIR, (b) Air PM_{10} concentration v lung cancer SIR, (c) Air NO_2 concentration v respiratory hospital admission SIR and (d) Air PM_{10} concentration v respiratory hospital admission SIR pollution, with fitted loess lines

The corresponding correlation coefficients for the plots (Figures 5.2 – 5.6) are shown in Table 5.3 and demonstrate the associations between the environmental and health variables. All variables showed a statistically significantly correlation with each other. Soil metal score was weakly to moderately correlated with both lung cancer incidence (0.208) and respiratory hospital admissions (0.262), while both air NO_2 and PM_{10} were also positive and moderately correlated with both health outcomes. The associations between the health indicators and deprivation deciles are perhaps even more interesting. A highly significant negative correlation coefficient (-0.729) was found between lung cancer incidence and deprivation across Greater Glasgow, with a similar but moderate correlation found for respiratory hospital admissions (-0.397). These correlations were negative since higher deprivation scores represent areas that are least deprived. Therefore, the correlation coefficient suggests that higher incidence of lung cancer is associated with deprived areas across Greater Glasgow.

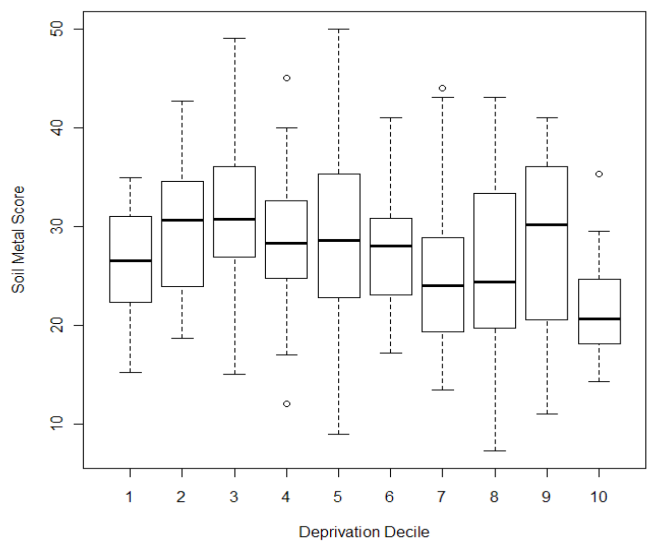
One other interesting result from the assessment of correlations between the variables was that soil metal score across Greater Glasgow also appeared to have a statistically significant (95% confidence level) negative association with deprivation (-0.213) and a significantly positive correlation with both air NO_2 (0.392) and PM_{10} (0.389). Figure 5.7 is indicative of a negative pattern between deprivation and soil metal score for most deciles, with the exceptions of deciles 1 (lower soil metal scores than trend) and 9 (higher soil metal scores than trend). This suggests that in general the more deprived areas across the region are more likely to contain elevated soil metal concentrations, in comparison to areas which are less deprived. Figures 5.8 (a) and (b) also reveal a moderate positive linear relationship between soil metal scores and the air pollutants, where areas with greater soil metal scores appear to have higher air pollution levels also. These results are interesting because one of the main aims of this study was to assess spatial associations and links between poor land quality and deprivation with a view to the Environmental Agenda. This study proves that there is a link between deprivation and poor land quality.

These links between deprivation, air pollution and soil metal are analysed and discussed in further detail in Sections 5.3 and 5.4 of this chapter.

Table 5.3- Pearson correlation coefficients between environmental variables and health outcomes

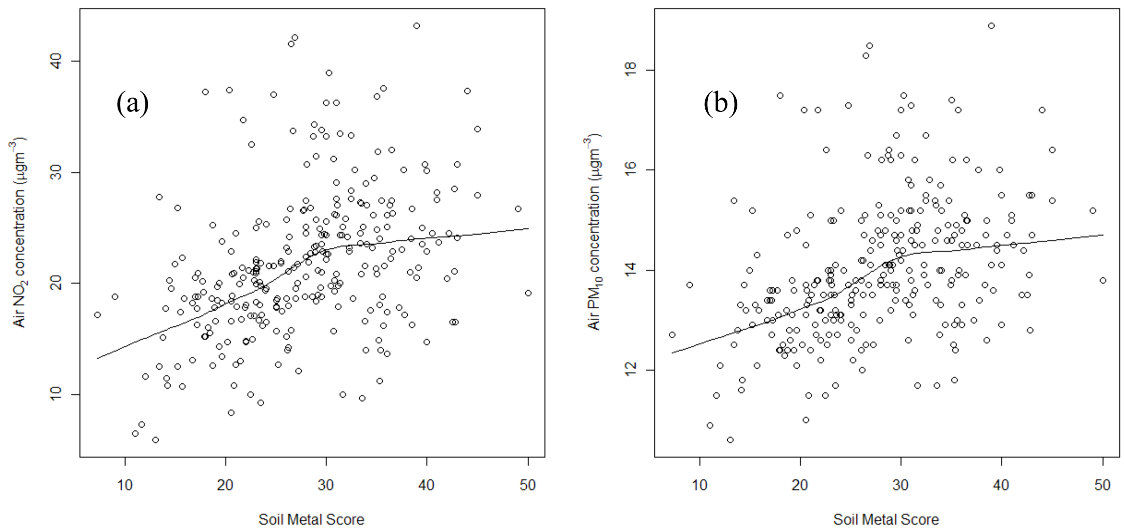
	Lung Cancer SIR	Respiratory Admission SIR	Metal Score	Deprivation	NO_2	PM_{10}
Lung Cancer SIR	1					
Respiratory SIR	*0.390*	1				
Soil Metal Score	*0.208*	*0.262*	1			
Deprivation	*-0.729*	*-0.397*	*-0.213*	1		
Air NO_2	*0.345*	*0.222*	*0.392*	*-0.286*	1	
Air PM_{10}	*0.341*	*0.230*	*0.389*	*-0.307*	*0.978*	1

n = 279, r95% confidence level = 0.118. Figures shown in italics are statistically significant



n = 279

Figure 5.7 – Soil metal score v median deprivation decile



n = 279

Figure 5.8 - (a) Soil metal score v air NO_2 concentration and (b) Soil metal score v air PM_{10} concentration, with a fitted loess line

Having presented maps of the environmental variables and health outcomes to gain an impression of their distribution across Greater Glasgow (Chapter 4), Moran's I test was carried out to investigate whether the observed spatial patterns occurred randomly or whether there was an underlying structure present in the form of clustering. The results from Moran's I in ArcGIS, using the normal approximation and performed at IG level, are illustrated in Tables 5.4, 5.5 and 5.6 for the health outcomes, the two air pollutants and geometric means of eight soil metals respectively. The p-values for all environmental and health variables were less than 0.1%, which indicated that there was less than a 0.1% likelihood that the clustered pattern of each could be due to random chance.

Bivand et al. (2009) investigated the functionality of Moran's I statistic and demonstrated the problems of using the Normal approximation in power calculations, where small neighbourhood areas can potentially lead to inference errors. In terms of an approximation to power, they suggest using the Saddlepoint approximation which uses a weighting structure and eigenvalues to evaluate the exact distribution. Another drawback of Moran's I statistic is that its limitations are not well recognised. Li et al. (2007)

developed an unconventional closed-form assessment of the strength of spatial autocorrelation, called the approximate profile-likelihood estimator (APLE) statistic. It was demonstrated that the APLE statistic is a better measurement of spatial dependence than Moran's I, provided that the correlation is not close to zero, and also that the APLE scatterplots are more informative than the equivalent Moran's scatterplot (Li et al., 2007). However in the current study, since the sampling area of Greater Glasgow was fairly large and measurements were recorded at only four per km^2, this inevitably resulted in correlations which approached zero.

Table 5.4 - Moran's I results for health outcomes

	Lung Cancer SIR	**Respiratory Admissions SIR**
Index	0.09	0.06
Z score	13.81	8.89
P-value	*<0.01*	*<0.01*

Figures shown in italics are statistically significant

Table 5.5 - Moran's I results for air pollution

	NO_2	**PM_{10}**
Index	0.61	0.35
Z score	29.81	50.63
P-value	*<0.01*	*<0.01*

Figures shown in italics are statistically significant

Table 5.6 – Moran's I results for soil metal geometric means

	As	**Cr**	**Cu**	**Ni**	**Pb**	**Se**	**Zn**	**K_2O**
Index	0.02	0.14	0.05	0.02	0.12	0.02	0.08	0.03
Z score	3.71	21.83	8.59	4.48	18.54	3.15	11.91	4.6
P-value	*<0.01*	*<0.01*	*<0.01*	*<0.01*	*<0.01*	*<0.01*	*<0.01*	*<0.01*

Figures shown in italics are statistically significant

Since it was shown that there was only a small probability that the structures within the datasets were due to random chance, a full analysis was carried out to assess the spatial relationships between health indicators, soil metal concentration, air pollution and deprivation as follows.

5.2 Relationships between Environment, Health and Deprivation Datasets

The final phase of this research was to investigate the relationships between the environmental data and health outcomes. The link between health and environmental variables is dependent on the accuracy of the exposure assessment and also the time lag between exposure and disease, where the associations between exposure and disease become more difficult to determine over time because of exposure changes and potential population migration (Jarup, 2004). However, the current project was not a time series study and instead focused on the geographical associations between the variables, where it was assumed that the data are representative of spatial relationships at a common point in time. Before examining the relationships between the datasets, it was important to consider the Ecological Fallacy.

5.2.1 Ecological Fallacy

The Ecological Fallacy often occurs in ecological studies and arises when inferences are made about a larger population group but may not hold for individuals (Lawson, 2001a). In this study the health data used were lung cancer registrations and respiratory hospital admissions, with the data originally obtained from specific patients. However, the health data was only available and presented at IG level for the purposes of confidentiality, altering the spatial resolution. Lawson, (2001) discusses how grouping influences the associations between exposure factors and health outcomes, and suggests using random effects to allow for changes to the resolution of the data. Other literature attempts to address the problem in other ways, including linking resolution levels in models (Plummer and Clayton, 1996) and a parametric approach of adjusting the model mean (Richardson et al., 1987).

Another problem was that as discussed in Section 2.1.3, individual level exposure to the environmental variables of soil contaminants and air pollutants were not available for Glasgow. Although geographically referenced soil metal concentrations and modelled air pollution concentrations were provided, these do not necessarily reflect individual exposure in that particular area since individuals may be exposed to different concentration levels at work, home etc. For example, populations working away from home all day may be less exposed to air pollutants than those staying at home. Similarly, populations living at ground level with bare soil in their gardens may be more exposed to soil than populations living in high rise flats. This uncertainty makes it difficult to determine any direct relationships which may exist between contaminated land and health. However, it is likely that exposure takes place in some way or form, either through direct contact with soil, eating vegetables or blown soil dust. This lack of exposure data is a well-known drawback in epidemiological studies of this type based on large summary datasets.

Although there are suggested strategies to account for ecological bias, it is very difficult to account for it in statistical modelling. Various types of ecological bias can exist in studies of this nature such as specification bias, within-group confounding and between-group confounding (Lee et al., 2009). Due to these complications and the difficulty of adjusting for ecological bias in the models, the current study has avoided this approach and retains simplicity of the statistical modelling process.

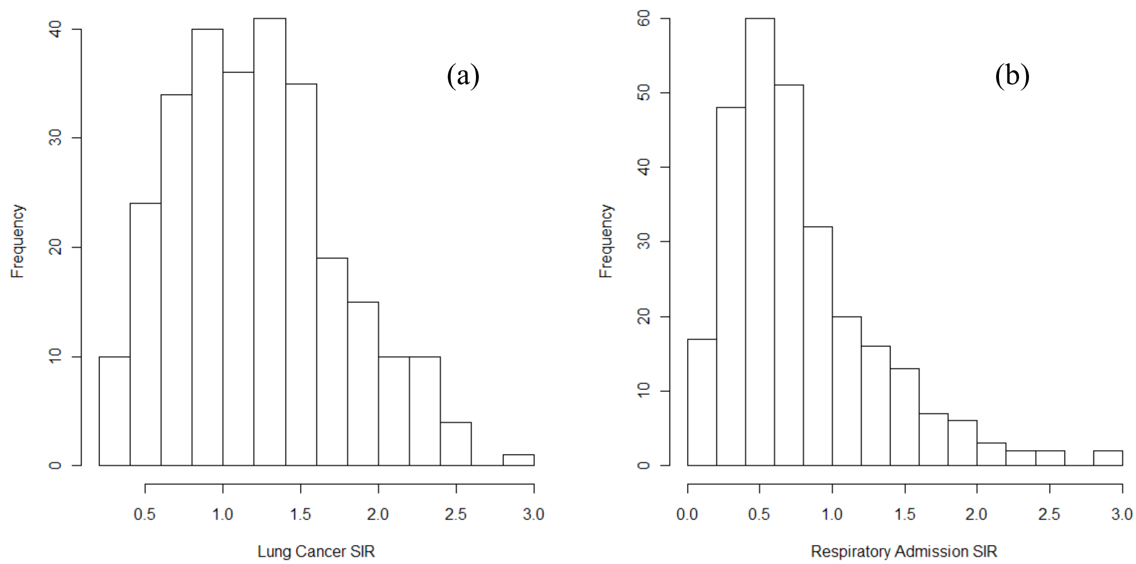
5.2.2 General Linear Modelling (GLM) and Relative Risk (RR)

Relationships between the variables were assessed using GLM and RR. However, as outlined in the previous section and Chapter 2, it is difficult to establish plausible causal links between soils and health and the RR results should be treated with caution. They are not used in this study to imply any direct links between soils and health. Rather they are used to assess the impacts on health outcomes of theoretical increases in soil metal concentration. Care was taken prior to the model building process in determining which environmental variables were to be included in the model, and in which order. As discussed in Chapter 2, air pollution is known to have health effects and was therefore included in the modelling procedure, in addition to the soil contaminants. Deprivation was also included, since similar studies on the health effects from environmental variables have shown that it is a significant confounder (Lee et al., 2009). There are a variety of ways in which one could approach the process of statistical modelling illustrated in this chapter. A systematic method has been adopted to demonstrate the possibilities of how to undertake the analysis. The rest of this chapter presents the modelling results of the effects of increases in soil metal geometric mean concentrations on health outcomes as well as making use of the index of soil metals created in Chapter 4. For both the geometric means and the index scores, a stepwise approach was taken of starting with a simple model including soil contaminants only and adding the deprivation and air pollution covariates. Due to the very high correlation between air NO_2 and PM_{10}, it was sensible to only include one of the pollutants in the model at a time. This same methodology was applied to both lung cancer registrations and respiratory hospital admissions.

5.2.3 Approach to Modelling Health Outcomes

A common approach to modelling health data in epidemiological studies of this kind is to use counts to represent the health outcome of interest. An example of such a study is (Lee et al., 2009). They used a Bayesian modelling framework, where the spatial regression modelled respiratory admission counts against other environmental indicators. In the study, the response was assumed to follow a Poisson distribution and carried out separate analyses for two spatial resolutions, IG and DZ levels. The study also used counts rather than standardised rates, which can also lead to greater variability in estimates from those IG areas with fewer counts, while this effect would be more apparent at DZ level due to the smaller population sizes. The current study has attempted to overcome these issues by taking the population structure into account and using standardised incidence ratios (SIRs) of the health outcomes (Chapter 4).

Observing the distribution of the health indicators, the SIR of lung cancer and respiratory hospital admissions appear to follow a log-normal distribution (Figure 5.9 (a) and (b)). In models where count data for health outcomes is preferred, a Poisson distribution with a log link is often specified (Lee et al., 2009). However, a different model was used in this study, where a log link function was identified along with a Gaussian family distribution, to represent the SIR of both health outcomes. Since the model followed a Gaussian distribution, normality of the residuals was assessed. The required model assumptions of independence and constant variance were also checked.



n = 279

Figure 5.9 – (a) Histogram of lung cancer SIR and (b) Histogram of respiratory hospital admission SIR

5.3 Approach to Modelling Soil Metals

When incorporating soil contaminants into the regression analysis, various approaches were explored. One possible method was to include all individual soil metal geometric means in the model, the second was to only include one of the soil metals in the model and the third was to use the soil metal index. There are drawbacks to each approach. Including all soil metals in the regression leads to the problem of multicollinearity. This occurs when two or more variables are highly correlated with one another and neither have a significant impact on the model when the other is accounted for. Since comparison between some of the soil contaminants demonstrated a high degree of correlation in Section 5.1.1, this issue became apparent in the regression analysis. The results for some soil contaminants showed a positive significant relationship between some soil metals and health outcomes, while other contaminants showed a negative relationship (see Section 5.1.2).

To avoid multicollinearity between soil contaminants, the two alternative approaches indicated above were carried out. Including one of the soil metals in the model made it considerably easier to interpret the results, since the contaminant was expressed in terms of concentration. However, the problem with including only one soil metal in the model was that a considerable amount of information was overlooked, in terms of the remaining metals which were not added. However, it was of interest to assess the associations between health and each soil metal individually. Therefore, the simplest model began with a soil metal, with the deprivation and air pollutant covariates added one at a time. Relative risks (RR) of a theoretical increase in soil concentration (see Section 5.3.1) were produced for each model, where soil contaminants were modelled separately.

The other alternative was to use the soil metal index to represent the land quality covariate in the model. In this case, the simplest model consisted of soil metal score only, with the deprivation and air pollutant covariates again added one at a time. Relative risks were also computed for a theoretical increase of soil metal score for these models. Although this method does not allow the possibility of interpretation in terms of soil concentration, it is more representative of general soil content across Glasgow rather than assessing the associations with one soil metal only.

5.3.1 Analysis of Lung Cancer Registrations SIR

Standardised Incidence Ratios of lung cancer registrations across Greater Glasgow were modelled against soil metals, deprivation deciles and air pollution. Due to the collinearity between soil metals, geometric means of the soil contaminants were assessed individually. To investigate general soil metal effects on health across Greater Glasgow, the soil metal index was also modelled. Relative risk (RR) estimates and corresponding 95% confidence intervals were provided for exposure to a theoretical but realistic increase in soil metal concentrations. However, these risk assessments imply no causal relationship and should be treated with caution. These increases are shown in Table 5.7, with each value being chosen to represent less than one standard deviation of the soil metal geometric mean concentration across the sampling region. These values also apply to the RR in the analysis of respiratory hospital admissions.

Table 5.7 – Assumed increases in soil metal geometric mean concentrations used in theoretical exposure risk assessments

Soil Metal	Exposure Increase
As	2 mg kg^{-1}
Cr	20 mg kg^{-1}
Cu	20 mg kg^{-1}
Ni	15 mg kg^{-1}
Pb	50 mg kg^{-1}
Se	0.3 mg kg^{-1}
Zn	50 mg kg^{-1}
K_2O	0.1 % wt
Soil Metal Score	5 units

5.3.2 Modelling Soil Metal Geometric Means against Lung Cancer SIR

With each soil metal modelled separately, only Cu, Pb and Zn were shown to have a significant association with lung cancer incidence across the region (Table 5.8). However, this model does not take the other covariates into account. When accounting for deprivation, only Zn had a statistically significant association with lung cancer incidence, while associations observed for Cu and Pb were non-significant. The p-values of all other soil contaminants were also considerably reduced, suggesting that deprivation is an important confounder. Therefore, deprivation explains most of the variability in lung cancer incidence across Greater Glasgow, when accounted for in the model along with soil metals.

At this stage, the model suggests that when taking deprivation into account, a theoretical increase of 50 mg kg^{-1} Zn soil metal concentration would increase the risk of being diagnosed with lung cancer by somewhere between 0.6 % and 5.4 % (Table 5.8). However, no causal link between soil Zn and lung cancer is actually implied here. This is because, once air NO_2 or PM_{10} were included into the models (Table 5.9), a non-significant association was found for each soil contaminant, including Zn. This implies that air pollution is also an important confounder in the study. Since the soil metals displayed non-significant relationships with lung cancer incidence across Greater Glasgow, when deprivation deciles and air pollution were accounted for, soil contaminants were removed from the models. This resulted in a final model of deprivation and air pollution, as illustrated in Table 5.10

The addition of deprivation to the models demonstrated that each decile had a significant relationship with lung cancer incidence, indicating an increasingly negative association between deprivation and lung cancer (Table 5.10). When air NO_2 or PM_{10} were also added to the models, a highly statistically significant effect of air pollution on lung cancer SIR across Glasgow was found.

Table 5.8 – RR models of soil metals against lung cancer SIR including deprivation

	Soil Metal		**+ Deprivation**	
Soil Metal Geometric Mean	**RR**	**95% CI**	**RR**	**95% CI**
As	1.017	(0.999, 1.035)	1.002	(0.989, 1.016)
Cr	1.012	(0.986, 1.039)	0.994	(0.974, 1.014)
Cu	1.036	*(1.015, 1.058)*	1.016	(0.999, 1.034)
Ni	1.018	(0.989, 1.049)	1.001	(0.977, 1.025)
Pb	1.029	*(1.001, 1.057)*	1.016	(0.995, 1.038)
Se	1.014	(0.986, 1.042)	0.997	(0.977, 1.018)
Zn	1.064	*(1.033, 1.096)*	1.030	*(1.006, 1.054)*
K_2O	1.017	(0.988, 1.046)	1.006	(0.985, 1.026)

Figures shown in italics are statistically significant

Table 5.9 – RR models of soil metals against lung cancer SIR, including deprivation deciles, air NO_2 and PM_{10}

Soil Metal Geometric Mean	+ Deprivation + Air NO_2		+ Deprivation + Air PM_{10}	
	RR	**95% CI**	**RR**	**95% CI**
As	1.002	(0.988, 1.016)	1.002	(0.988, 1.016)
Cr	0.992	(0.973, 1.012)	0.991	(0.971, 1.010)
Cu	1.008	(0.989, 1.027)	1.008	(0.990, 1.027)
Ni	1.002	(0.978, 1.026)	1.001	(0.977, 1.025)
Pb	0.998	(0.975, 1.021)	1.001	(0.978, 1.024)
Se	0.997	(0.976, 1.018)	0.996	(0.975, 1.017)
Zn	1.018	(0.994, 1.042)	1.019	(0.995, 1.043)
K_2O	1.007	(0.986, 1.028)	1.007	(0.986, 1.028)

Figures shown in italics are statistically significant

Table 5.10 – Summary of model including deprivation deciles and air PM_{10}

Coefficient	Estimate	Std. Error	P-value
Intercept	-0.03688	0.17931	0.8372
Decile 2	-0.11392	0.05148	*0.0277*
Decile 3	-0.21878	0.05311	*<0.0001*
Decile 4	-0.31679	0.07066	*<0.0001*
Decile 5	-0.44944	0.06347	*<0.0001*
Decile 6	-0.50975	0.07493	*<0.0001*
Decile 7	-0.58725	0.08447	*<0.0001*
Decile 8	-0.74625	0.09332	*<0.0001*
Decile 9	-0.90895	0.10536	*<0.0001*
Decile 10	-1.14901	0.15027	*<0.0001*
Air PM_{10}	0.04469	0.01220	*0.0003*
AIC			**220.21**

Deprivation Deciles: 1 = most deprived, 10 = least deprived. Figures shown in italics are statistically significant

5.3.3 Modelling Overall Soil Metal Score against Lung Cancer SIR

In addition to assessing the individual soil metal associations with lung cancer incidence, the index of soil metals was used to represent the soil metals covariate in the model. Again models assumed an approximately log normal distribution of the response, where a Gaussian family with a log link was specified.

When including soil metal index as the only covariate in the model, it was shown to have a statistically significant association with lung cancer incidence in Greater Glasgow, producing a very small p-value (Table 5.11). However, the inclusion of deprivation in the model removed the significant effect of the soil metal index, where all deprivation deciles were shown to have a statistically significant effect on lung cancer incidence across Greater Glasgow (Table 5.12). Again, the increasingly negative coefficients demonstrate the negative relationship between lung cancer and deprivation, where deprived areas tend to have higher incidence of lung cancer. A cross-sectional individual and area-based study in Norfolk, England carried out on the associations between socioeconomic status and lung function also found that living in deprived areas was predictive of reduced lung function, even when controlling for smoking (Shohaimi et al, 2004).

Air pollution also demonstrated a significant association with lung cancer incidence. Nitrogen Dioxide was shown to have a statistically significant effect on lung cancer incidence when accounting for the effects of deprivation and soil metal scores (Table 5.13), with a similar result produced for the addition of air PM_{10}. Even with the high degree of correlation between the air pollutants and deprivation illustrated in Section 5.1.3, a statistically significant effect was observed for both covariates. Since soil metal score had no significant association with lung cancer incidence across Glasgow when accounting for the effects of deprivation and air pollution, the soil metals covariate was removed from the final model, as illustrated in Table 5.14.

Table 5.11 – Summary of model including soil metal score

Coefficient	**Estimate**	**Std. Error**	**P-value**
Intercept	-0.09241	0.09414	0.32719
Soil Metal Score	0.01028	0.00312	*0.00111*
AIC			**418.56**

Figures shown in italics are statistically significant

Table 5.12 – Summary of model including soil metal score and deprivation deciles

Coefficient	Estimate	Std. Error	P-value
Intercept	0.51040	0.07348	<0.0001
Soil Metal Score	0.00337	0.00233	0.1483
Decile 2	-0.11174	0.05318	*0.0365*
Decile 3	-0.22368	0.05533	*<0.0001*
Decile 4	-0.32816	0.07244	*<0.0001*
Decile 5	-0.44289	0.06496	*<0.0001*
Decile 6	-0.54153	0.07590	*<0.0001*
Decile 7	-0.57612	0.08606	*<0.0001*
Decile 8	-0.77943	0.09451	*<0.0001*
Decile 9	-0.93427	0.10721	*<0.0001*
Decile 10	-1.19144	0.15262	*<0.0001*
AIC			**231.80**

Deprivation Deciles: 1 = most deprived, 10 = least deprived. Figures shown in italics are statistically significant

Table 5.13 – Summary of model including soil metal score, deprivation deciles and air NO_2

Coefficient	Estimate	Std. Error	P-value
Intercept	0.34191	0.08608	*<0.001*
Soil Metal Score	0.00075	0.00238	0.7519
Decile 2	-0.10839	0.05169	*0.0369*
Decile 3	-0.21010	0.05370	*<0.0001*
Decile 4	-0.31372	0.07073	*<0.0001*
Decile 5	-0.44119	0.06344	*<0.0001*
Decile 6	-0.49993	0.07495	*<0.0001*
Decile 7	-0.57889	0.08416	*<0.0001*
Decile 8	-0.73854	0.09315	*<0.0001*
Decile 9	-0.90403	0.10514	*<0.0001*
Decile 10	-1.14356	0.14974	*<0.0001*
Air NO_2	0.01012	0.00268	*0.0002*
AIC			**219.09**

Deprivation Deciles: 1 = most deprived, 10 = least deprived. Figures shown in italics are statistically significant

Table 5.14 – Summary of model including deprivation deciles and air NO_2

Coefficient	Estimate	Std. Error	P-value
Intercept	0.3566	0.0725	*<0.0001*
Decile 2	-0.1063	0.0510	*0.0382*
Decile 3	-0.2072	0.0527	*<0.0001*
Decile 4	-0.3113	0.0702	*<0.0001*
Decile 5	-0.4393	0.0630	*<0.0001*
Decile 6	-0.4987	0.0747	*<0.0001*
Decile 7	-0.5799	0.0840	*<0.0001*
Decile 8	-0.7380	0.0930	*<0.0001*
Decile 9	-0.9030	0.1049	*<0.0001*
Decile 10	-1.1457	0.1493	*<0.0001*
Air NO_2	0.0104	0.0026	*<0.0001*
AIC			**217.20**

Deprivation Deciles: 1 = most deprived, 10 = least deprived. Figures shown in italics are statistically significant

The final model (Table 5.14) consisted of only deprivation and air pollution, with NO_2 selected as this parameter satisfied the minimum AIC argument, as opposed to PM_{10}. All model coefficients were highly significant at the 5% level. Figure 5.10 (a) shows that the standardised residuals are evenly spread although the pattern suggests some ‘fanning out’ as the fitted values increases. However, this pattern is only influenced by a few individual points so that the assumptions of mean equal to zero and constant variance can be accepted. The plot in Figure 5.10 (b) displays a relatively straight line with the exception of a slight deviation at either end of the tails. Therefore, the assumption of normality holds for this model.

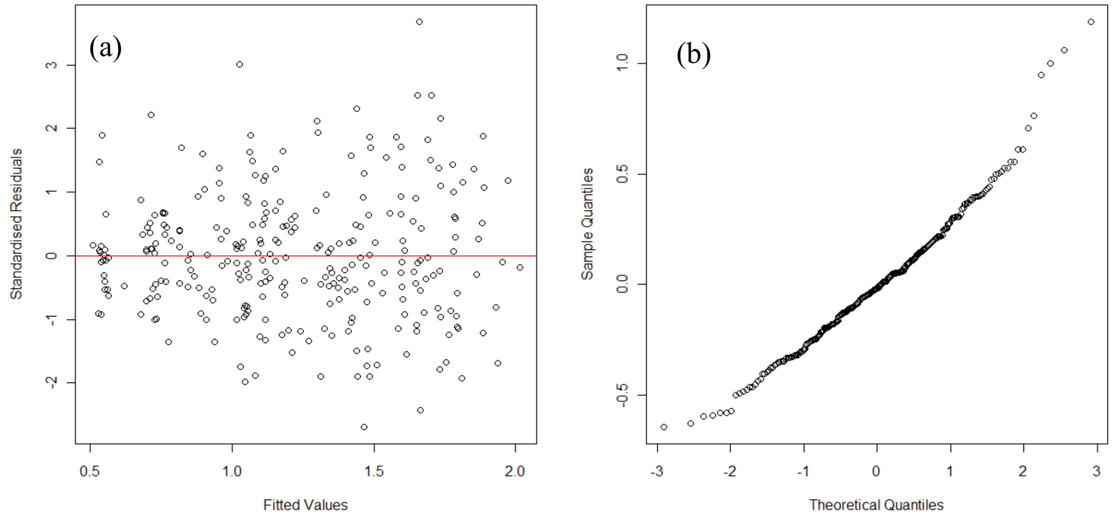


Figure 5.10 – (a) Residual v fitted values plot and (b) Normality plot for the final model in Table 5.7

For each model, Table 5.15 summarises the RR estimates and 95% confidence intervals of developing lung cancer assuming a theoretical exposure to a five unit increase in overall soil metal content. However, these results imply no causal relationship between lung cancer and soil metal concentration and no statistical relationship between soil metals and lung cancer was found in the present study.

Table 5.15 – RR models for lung cancer SIR

Model	**AIC**	**RR**	**95% CI**
Soil Metal Score	418.56	1.053	*(1.021, 1.085)*
Soil Metal Score + Deprivation	231.80	1.017	(0.994, 1.040)
Soil Metal Score + Deprivation + Air NO_2	219.09	1.004	(0.981, 1.027)
Soil Metal Score + Deprivation + Air PM_{10}	221.95	1.006	(0.983, 1.030)
Deprivation + Air NO_2	217.20		
Deprivation + Air PM_{10}	220.21		

Figures shown in italics are statistically significant

5.3.4 Analysis of Respiratory Hospital Admissions SIR

As with the analysis on lung cancer incidence, it was appropriate to model respiratory admissions with each soil metal separately, due to the problem of multicollinearity between metals. The index of overall metal scores was also modelled to assess the health associations with land quality across Greater Glasgow.

5.3.5 Modelling Soil Metal Geometric Means against Respiratory Hospital Admission SIR

When assessing soil contaminants individually, the addition of deprivation showed that only some deciles had a statistically significant association with respiratory hospital admissions, where increasingly negative coefficients were observed. A statistically significant relationship was also found for both air NO_2 and PM_{10}, when added to the model.

Soil Cr, Ni, Pb and Zn were shown to have a significant association with respiratory hospital admissions across Greater Glasgow (Table 5.16) at this stage. Table 5.16 provides RR estimates with corresponding 95% confidence intervals for the risk of being admitted to hospital due to respiratory causes for each soil contaminant, based on a theoretical increase in soil metal exposure.

However, no cause-effect link is implied by these relationships as other potential confounders need to be taken into account. Although deprivation was shown to be a statistically significant confounder, when added to each of the models, Cu, Ni, Pb and Zn still demonstrated a statistically significant relationship with respiratory hospital admissions. Since some deprivation deciles were shown to be statistically significant, deprivation was retained in the models.

However, with both air pollution and deprivation included in the models, only soil Ni demonstrated a statistically significant association with respiratory hospital admissions across Greater Glasgow, as illustrated in Table 5.17. Therefore, by minimum AIC, the final model in Table 5.18 shows that soil Ni, deprivation and air NO_2 are significant factors in explaining the respiratory hospital admissions across Greater Glasgow.

The RR confidence intervals in Table 5.17 for the soil Ni model suggests that, once accounting for deprivation and air NO_2, theoretical exposure to an increase of 15 mg kg^{-1} soil Ni concentration would increase the risk of being admitted to hospital due to respiratory causes by somewhere between 1.4% and 8.3%. Although direct causal relationships between soil and health are difficult to prove, these results are interesting because, as outlined in Chapter 2, Ni is a known respiratory irritant. This result may warrant further investigation.

Table 5.16 – RR models of soil metals against respiratory admission SIR, including deprivation

Geometric Mean	Soil Metal		+ Deprivation	
	RR	95% CI	RR	95% CI
As	1.024	(0.999, 1.050)	1.004	(0.979, 1.031)
Cr	1.035	(0.999, 1.070)	1.021	(0.986, 1.057)
Cu	1.056	*(1.027, 1.086)*	1.034	*(1.004, 1.066)*
Ni	1.055	*(1.024, 1.087)*	1.043	*(1.010, 1.078)*
Pb	1.056	*(1.017, 1.097)*	1.050	*(1.011, 1.090)*
Se	1.014	(0.972, 1.059)	0.992	(0.952, 1.033)
Zn	1.084	*(1.036, 1.133)*	1.050	*(1.050, 1.004)*
K_2O	0.999	(0.953, 1.047)	0.981	(0.939, 1.025)

Figures shown in italics are statistically significant

Table 5.17 – RR models of soil metals against respiratory admission SIR and air pollutants

Soil Metal Geometric Mean	+ Deprivation + Air NO_2		+ Deprivation + Air PM_{10}	
	RR	95% CI	RR	95% CI
As	1.004	(0.977, 1.032)	1.002	(0.976, 1.030)
Cr	1.019	(0.984, 1.055)	1.017	(0.983, 1.054)
Cu	1.025	(0.992, 1.060)	1.027	(0.994, 1.060)
Ni	1.048	*(1.014, 1.083)*	1.046	*(1.012, 1.081)*
Pb	1.034	(0.991, 1.078)	1.037	(0.995, 1.081)
Se	0.993	(0.952, 1.036)	0.992	(0.951, 1.034)
Zn	1.034	(0.986, 1.084)	1.037	(0.989, 1.087)
K_2O	0.977	(0.932, 1.024)	0.977	(0.933, 1.024)

Figures shown in italics are statistically significant

Table 5.18 – Summary of model including soil Ni, deprivation deciles and air NO_2

Coefficient	Estimate	Std. Error	P-value
Intercept	-0.5814	0.1785	*0.0013*
Soil Ni geometric mean	0.0031	0.0011	*0.0056*
Decile 2	0.1351	0.1264	0.2863
Decile 3	0.1841	0.1234	0.1368
Decile 4	-0.0517	0.1610	0.7484
Decile 5	-0.2037	0.1496	0.1744
Decile 6	-0.3451	0.1838	0.0616
Decile 7	-0.4926	0.2148	*0.0226*
Decile 8	-0.4959	0.2113	*0.0197*
Decile 9	-0.4637	0.2010	*0.0218*
Decile 10	-0.6089	0.2584	*0.0192*
Air NO_2	0.0138	0.0053	*0.0101*
AIC			**380.96**

Deprivation Deciles: 1 = most deprived, 10 = least deprived. Figures shown in italics are statistically significant

5.3.6 - Modelling Overall Soil Metal Score against Respiratory Hospital Admission SIR

The same model building process for the soil metal index used in the analysis of lung cancer SIR was also applied to the respiratory hospital admissions data. Beginning with a simple model of overall soil metal content only, a statistically significant relationship with respiratory hospital admissions was observed, as shown in Table 5.19. As illustrated previously, both deprivation and air pollution may act as confounders and were also added to the model (Table 5.20). The addition of deprivation had a different effect on the model (Table 5.20) to that of lung cancer registrations, since only the least deprived deciles had a significant (but not highly significant) negative relationship with respiratory hospital admissions. The most deprived deciles showed no significant association with respiratory admissions. This pattern is difficult to explain, although this study has questioned the domains which contribute to the SIMD, which could possibly be improved to give a more appropriate reflection of deprivation. The distribution of deprivation across Greater Glasgow may also result in these non-significant deciles, since several rural areas are classed as the most deprived areas and are unlikely to contain the highest incidence of respiratory hospital admissions.

Table 5.19 – Summary of model including soil metal score

Coefficient	Estimate	Std. Error	P-value
Intercept	-0.7992	0.1509	*<0.0001*
Soil Metal Score	0.0193	0.0048	*<0.0001*
AIC			**416.31**

Figures shown in italics are statistically significant

Table 5.20 – Summary of model including soil metal score and deprivation deciles

Coefficient	Estimate	Std. Error	P-value
Intercept	-0.4427	0.1655	*0.0079*
Soil Metal Score	0.0125	0.0047	*0.0081*
Decile 2	0.0999	0.1282	0.4365
Decile 3	0.1523	0.1254	0.2255
Decile 4	-0.0981	0.1632	0.5483
Decile 5	-0.1758	0.1465	0.2312
Decile 6	-0.4045	0.1823	*0.0274*
Decile 7	-0.4858	0.2133	*0.0236*
Decile 8	-0.5368	0.2072	*0.0101*
Decile 9	-0.5159	0.2010	*0.0108*
Decile 10	-0.6381	0.2579	*0.0134*
AIC			**383.14**

Deprivation Deciles: 1 = most deprived, 10 = least deprived. Figures shown in italics are statistically significant

When accounting for deprivation and soil metal score, neither air NO_2 nor PM_{10} concentration showed a significant association with respiratory admissions at the 5 % significance level, and is illustrated for air NO_2 in Table 5.21. The significant association between soil metal score and respiratory admission still prevailed, despite the addition of deprivation and air pollution. However, by minimum AIC arguments, air pollution was kept in the model and is justified since the p-value is fairly close to 5 % significance level.

Therefore, the final model in describing the relationships with respiratory admission SIR contained soil metal score, deprivation and air pollution based on NO_2, due to minimum AIC arguments. The model assumptions were assessed, with the appropriate plots illustrated in Figure 5.11 (a) and (b). The assumptions of constant variance and mean of the errors equal to zero were valid, despite the slight indication of 'fanning out' in the plot (Figure 5.11 (a)). However, the normality assumption is dubious since the points in Figure 5.11 (b) do not follow closely to a straight line, showing considerable curvature towards both tails.

Table 5.21 – Summary of model including soil metal score, deprivation decile and air NO_2

Coefficient	**Estimate**	**Std. Error**	**P-value**
Intercept	-0.6147	0.1956	*0.0019*
Soil Metal Score	0.0102	0.0048	*0.0367*
Decile 2	0.1069	0.1274	0.4022
Decile 3	0.1783	0.1242	0.1525
Decile 4	-0.0861	0.1628	0.5975
Decile 5	-0.1789	0.1465	0.2231
Decile 6	-0.3713	0.1841	*0.0448*
Decile 7	-0.4928	0.2140	*0.0221*
Decile 8	-0.5078	0.2100	*0.0163*
Decile 9	-0.4859	0.2014	*0.0165*
Decile 10	-0.5831	0.2580	*0.0246*
Air NO_2	0.0099	0.0055	0.0718
AIC			**381.62**

Deprivation Deciles: 1 = most deprived, 10 = least deprived. Figures shown in italics are statistically significant

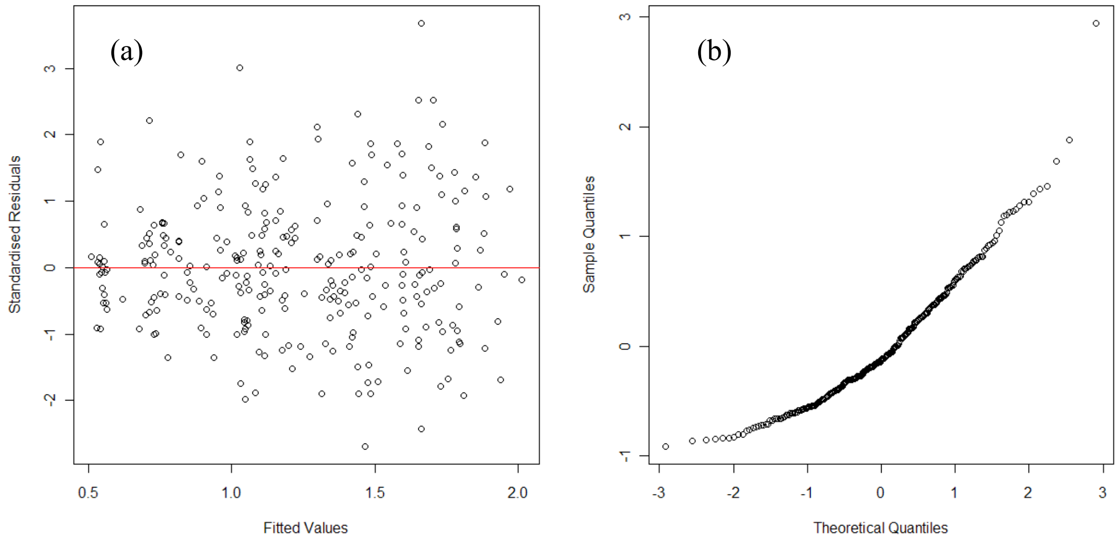


Figure 5.11 – (a) Residual v fitted values plot and (b) Normality plot for the final model in Table 5.14

Despite the model assumption doubts, the RR results for the final model including soil metal score, deprivation and air NO_2 are shown in Table 5.22. The results suggest that, for a theoretical increase of five units of soil metal score, the risk of being admitted to hospital due to respiratory disease is likely to increase by somewhere between 0.3% and 10.3%.

Rather than an actual health link, the apparent association between soil metal score and respiratory admissions is likely to represent the concurrent spatial association between lower soil metal concentrations and lower respiratory disease in rural areas compared to urban areas. However, it cannot be discounted that exposure to windblown metal-rich soil dust in the urban environment may contribute to respiratory disease, especially since recent studies carried out in Northampton suggest that 45% of metal-containing material in air PM_{10} particulates is derived from soil-dust (Cave and Cherney, 2010). However, further investigations would be required to assess this relationship more closely.

Table 5.22 – RR models for respiratory admission SIR

Model	AIC	RR	95% CI
Soil Metal Score	416.31	1.101	*(1.051, 1.155)*
Soil Metal Score + Deprivation	383.14	1.065	*(1.017, 1.115)*
Soil Metal Score + Deprivation + Air NO_2	381.62	1.052	*(1.003, 1.103)*
Soil Metal Score + Deprivation + Air PM_{10}	382.62	1.054	*(1.006, 1.105)*

Figures shown in italics are statistically significant

5.4 Chapter 5 Summary and Discussion

5.4.1 Spatial Associations between the Datasets

Initial spatial associations between the datasets were assessed. Pearson correlation coefficient tests between the soil contaminants found relatively high correlation coefficients between several of the soil metals, most notably between Cu, Ni and Pb, which flagged up the problem of multicollinearity within the dataset. Significant correlations were observed between some of the individual soil contaminants (Cu, Ni, Pb and Zn) and the health outcomes. One interesting result was that a strong negative association was found between deprivation and lung cancer incidence (-0.729), with a moderate negative relationship also observed for respiratory hospital admissions (-0.397).

One other interesting result was that soil metal score across Greater Glasgow appeared to have a statistically significant negative association with deprivation (-0.213) and a moderately positive correlation with both air NO_2 (0.392) and PM_{10} (0.389), indicating that land and air quality are poorest in the most deprived areas. This study has proved that there is a spatial association between deprivation and poor land quality in Glasgow.

This can be explained by the fact that historically in Glasgow large populations lived in proximity to heavy industry in the ship building, metal working and railway engineering centres of the city such as the East End. Many people still live in these areas but the industry has now gone, leading to high unemployment and low incomes contributing to deprivation. These are also the areas with a legacy of poor land quality and high soil metal concentrations.

Previous studies carried out by the Environmental Justice Agenda identified associations between poor air quality, water quality and proportion of derelict land with deprivation (Scottish Government, 2005). This project adds key information to these studies, namely that poor land quality is also associated with deprivation.

In addition to the increased risks of exposure to pollution, poor environmental quality has an impact on people's neighbourhood perception and quality of life (Scottish Government, 2005). However, none of the measures of deprivation currently include an environmental factor. For example, the SIMD is an index based on a number of social and health indicators, but does not have an environmental component. Environmental indicators such as land, air or water quality do not play a part in constructing the index.

Therefore, when policy makers are considering making improvements to the environment in terms of the Environmental Justice Agenda, a recommendation from this study is that land quality as well as environmental components such as air and water quality should be taken into account.

5.4.2 Relationships between Environmental and Health Indicators

Regression modelling was then used to investigate the relationships between the soil metals, air pollution, deprivation and health datasets. Models were produced separately to assess the effects of individual soil contaminants as well as using the index of soil metals as a land quality indicator. Relative risks were computed to approximate the theoretical risk of developing the health outcome of interest when exposed to an increase in soil metal concentration. However, it was noted that no cause-effect relationships were implied by these RR assessments which should be treated with caution as plausible exposure routes and links between soil health and Glasgow are difficult to prove.

The regression analysis showed that lung cancer was significantly associated with air pollution and deprivation across Glasgow as expected. However, none of the soil contaminants had a statistically significant association with lung cancer, when accounting for deprivation and air pollution. This was indeed the case for each of the soil metals as well as the soil metal index, indicating that deprivation and air pollution act as confounders in this study.

However, the equivalent analysis on respiratory hospital admission demonstrated that soil Ni had a significant association with respiratory admission, even when accounting for the effects of deprivation and air pollution. The models estimated that theoretical exposure to an increased soil concentration of 15 mg kg^{-1} Ni would increase the risk of being admitted to hospital due to respiratory problems by somewhere between 1.4% and 8.3%. Although direct causal relationships between soil and health are difficult to prove, this result is interesting because, as outlined in Chapter 2, Ni is a known respiratory irritant. This result may warrant further investigation.

Similarly, soil metal score had a statistically significant association with respiratory hospital admission at the 5% level, even when including deprivation and air pollution in the model.

The apparent association between soil metal score and respiratory admissions is likely to represent the concurrent spatial association between lower soil metal concentrations and lower respiratory disease in rural areas compared to urban areas, rather than a causal relationship. However, one cannot overlook the possibility of exposure to soil metal dust contributing to respiratory disease, particularly when recent studies demonstrated that 45% of metallic material in air PM_{10} comes from soil dust. However, this possible relationship should be investigated further.

A statistically significant relationship with respiratory disease was found for both air NO_2 and PM_{10}, when added to the individual soil metal models.

Interestingly only air NO_2 showed a statistical association with respiratory disease in the soil metal index model.

In both the individual soil metal model and the soil index model, only some of the deprivation deciles showed a significant relationship with respiratory disease and this was often a negative relationship between the less deprived areas and the disease. Some of the lower deciles showed no relationship with respiratory disease but this may reflect the presence of some deprived rural areas in the Glasgow area where respiratory disease is less prevalent.

Chapter 6

Conclusions

6.1 Summary and Conclusions

The main objective of this research was to investigate the spatial distributions of soil contaminants, air pollution, deprivation and health indicators across Greater Glasgow, while also carrying out an assessment of the spatial associations between the variables. The study has presented an initial examination of the distribution of these environmental and health variables and explored the links between them.

Soil metal concentrations (As, Cr, Cu, Ni, Pb, Se, Zn and K_2O); air NO_2 and PM_{10} and SNS health statistics for lung cancer incidence and respiratory hospital admissions, as well as the Scottish Index of Multiple Deprivation (SIMD), were compared across Glasgow. The analysis of relationships between these variables was carried out in a GIS at the Intermediate Geography (IG) level. This was considered to be more appropriate than Postcode, Datazone and Electoral Wards due to the sampling density of the soils dataset and greater availability of the other data at this spatial resolution. The SIMD index was recalculated to remove the health component and avoid accounting for the health indicators in this study twice.

Initial geostatistical analysis of the soil metal concentration dataset demonstrated that a linear or quadratic trend was sufficient in explaining any trends present for several of the soil contaminants; namely Cu, Pb, Ni, Se and Zn. However, the analysis for As and Cr showed that concentrations were spatially independent and there was no need to fit a trend effect. The results for K_2O revealed a considerable amount of spatial correlation in the concentrations across Greater Glasgow, even with the addition of a polynomial to describe the trend. For this particular element, the spatial continuity in the variogram was modelled, where the parameters were estimated from the variogram. Kriging was carried out to generate K_2O predictions across the sampling region and showed elevated concentrations in the north-west area in particular, with high concentrations also generated for the Bothwell area in the south-east of Greater Glasgow.

In order to convert the point source geochemistry dataset to IG areas compatible with the other datasets; the geometric mean soil metal concentration for each IG area was selected as a summary measure. This approach was deemed to be sensible as the soil metal for each IG was preserved in terms of soil concentration (mg kg^{-1}) and was also fairly robust to some outliers. In addition, an indicator of overall land quality was generated at the IG level for comparison with the other datasets. This comprised of a soil metal score for each IG area, based on the percentiles of the five elements (As, Cr, Ni, Pb and Se) of potential human health concern under current CLEA guidelines. However, the soil metal index was a relatively straightforward summary measure of land quality and should not be expected to give an accurate reflection of soils which pose a risk to human health.

The maps of the environmental and health variables provided an overview of their distribution across the sampling region for each IG area. The SIRs for the health outcomes revealed higher incidence, mainly in urban areas and parts of the East End of Glasgow, with modelled air pollution data for NO_2 and PM_{10} also showing higher concentrations in urban areas.

The SIMD, re-constructed with the health domain removed, indicated that most deprived areas were located in the east of Greater Glasgow, with only some rural areas being highly deprived. Mapped geometric means of the soil contaminants showed that some IG areas exceeded the CLEA soil guideline values (SGVs) for As, Cr, Cu, Ni and Pb. However, average Se concentrations across Greater Glasgow were very low relative to the SGV. The soil metal index showed higher soil metal scores were present in several areas in the west of Glasgow and were clustered in East End areas.

Initial assessments of the spatial relationships between the datasets showed significant correlations between several of the soil contaminants, highlighting that multicollinearity may be a potential problem in the analysis. Both air pollutants were also very highly correlated. Moran's I results for each of the environmental and health variables suggested that the patterns occurring were very unlikely to be due to random chance. Although there are some concerns about the functionality of Moran's I test and the problems of using small geographical areas, the large sample sizes within IG areas, along with the highly significant Moran's I results, supports the conclusions of a structure which is unlikely to be due to random chance.

The associations between the variables were explored further using regression modelling. The models demonstrated that deprivation had a significant relationship with both lung cancer and respiratory disease, with higher incidence occurring in more deprived areas. Air pollution also demonstrated a significant association with both lung cancer and respiratory disease as expected, in addition to deprivation and soil metal concentrations.

Soil metal concentrations showed no association with lung cancer, but soil metal score demonstrated a significant relationship with respiratory disease. However, this is most likely to reflect the spatial occurrence of low soil metals and low respiratory disease in the rural areas around Glasgow, as opposed to urban IG areas, rather than a plausible health link between these two datasets. None-the-less, it cannot be discounted entirely that exposure to windblown metal-rich soil dust in the urban environment may contribute to respiratory disease, with recent studies illustrating that 45% of metal material from air PM_{10} is derived from soil-dust. This relationship would require further investigation.

Interestingly, soil Ni concentrations showed a significant association with respiratory disease. Nickel is a known respiratory irritant. Whilst no causal link is implied in this study, these results also warrant further investigation.

One of the key findings from the study was the association between soil metal score and deprivation across Greater Glasgow. Although a relatively weak correlation of -0.213 was observed; the association was found to be statistically significant and warrants further investigation. In particular, this study has proved that there is a spatial association between deprivation and poor land quality in Glasgow. In other words, more deprived areas are likely to contain elevated soil metal concentrations. This relationship may be explained by the societal nature of large numbers of people still living in historic industrial areas in Glasgow, which tend to have higher soil metal concentrations. With industry now gone, these areas are more likely to have higher unemployment and low income, which is contributing to deprivation.

Previous investigations under the Environmental Justice Agenda demonstrated links between poor air and water quality and deprivation. However, this is the first time that relationships between poor land quality and deprivation have been assessed and identified. Therefore, this study contributes important information to the Environmental Justice debate.

This investigation has been constrained by the datasets available for the study. The project used health data which were routinely available, deprivation data, modelled air quality data and spatially referenced top soil metal concentrations. None of the above datasets were specifically collected and designed for the purposes of a study to assess relationships between environmental parameters and health. Therefore, it was only possible to carry out an overview assessment of these relationships that is subject to much uncertainty. This could be overcome in the future by specifically designing a study to investigate environmental and health relationships.

As outlined in this thesis, population exposures to contaminated land and air pollution are difficult to define and record. In particular, the soil metal data used in this study do not necessarily reflect direct exposure to the population within Glasgow. Lack of data on exposure is a common problem in epidemiological studies of this nature that are based on summary measure datasets. Therefore, interpretation of the health effects of soil metals was treated cautiously with concern to the Ecological Fallacy. Namely, that spatial coincidence does not imply a causal relationship. Furthermore, applying group level relationships to individuals should be avoided, although it is difficult to account for ecological bias in studies of this nature.

However, it cannot be ruled out that exposure to metals, via soil ingestion and inhalation of wind blown dust, adds to the metal loading of populations in Glasgow, which is why the associations between deprivation and poor land quality, and soil Ni and metal score with respiratory disease, this study has identified, are worthy of follow-up.

In addition to possible concerns about increased exposure to pollution, poor environmental quality has a negative impact on people's neighbourhood perception and quality of life and should be included in the Environmental Justice Agenda.

6.2 Recommendations

It is recommended that the associations between soil metal score and soil Ni and respiratory disease identified in the present study be investigated further to assess the links between environment and health in further detail.

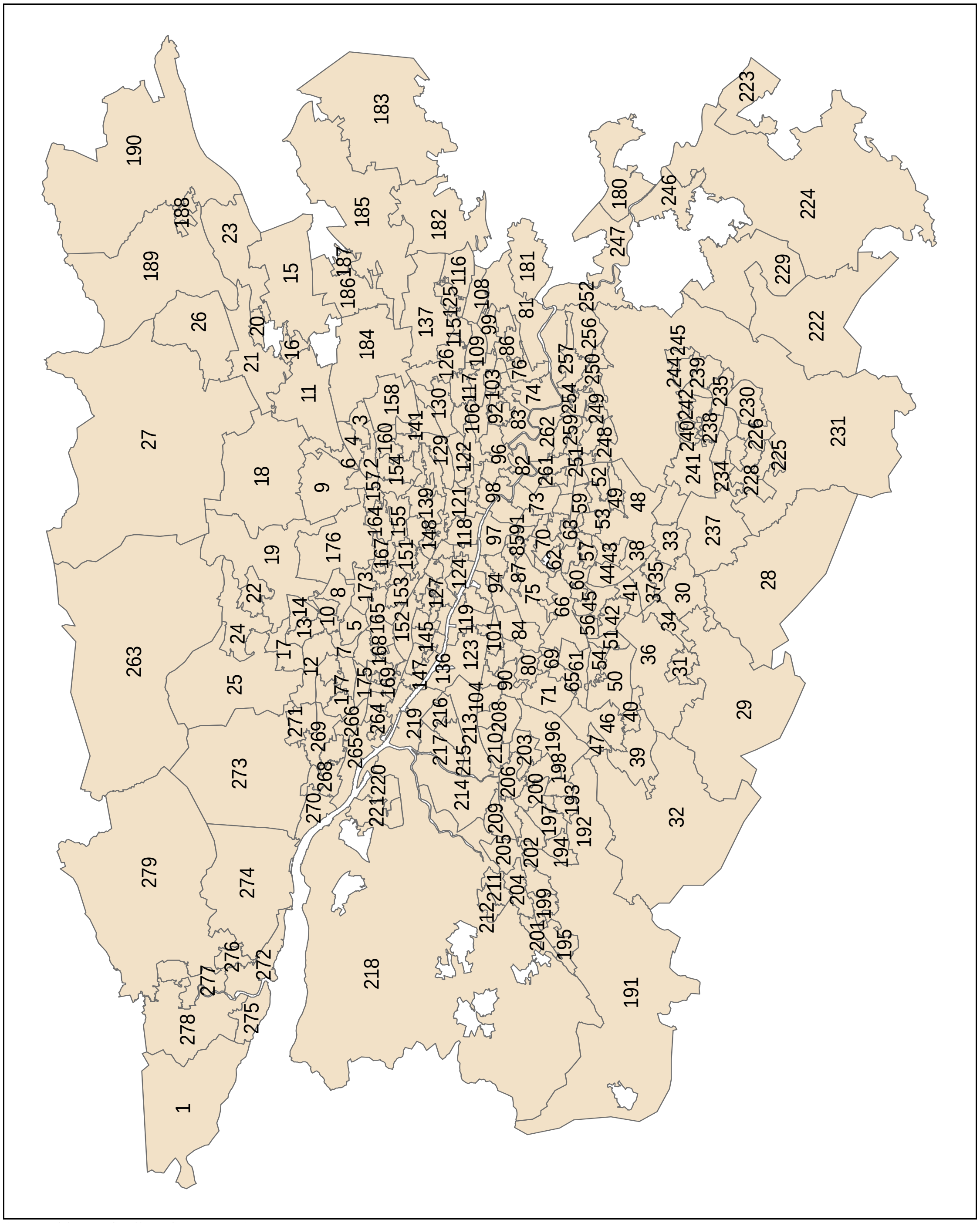
However, one aspect which this investigation has highlighted is the lack of information on soil exposure in the ambient environment. Since there were no data on population exposure to soil metals in this study, it was very difficult to determine the influence exposure to soil metals would have on the lung cancer and respiratory admission outcomes. To overcome these issues, it is recommended that further studies into soil metal impacts on health be carried out, based on specifically designed epidemiological investigations. This would involve individual level studies measuring metal concentrations in the soil and house dust as well as collecting lifestyle, smoking habit and dietary information. This could also eliminate some of the ecological bias which often arises in epidemiological studies.

In this study, only air pollution data for NO_2 and PM_{10} were available at this geographical scale, with both showing significant associations with the health outcomes. It is recommended that it may be beneficial to consider other air pollutants which may be of concern to human health in future investigations.

The current study has also highlighted the need for developing summary measures of contaminated land for use in environment and health assessments. The soil metal index created in this study was relatively straightforward and could be applied to many other similar datasets. However, there are several other approaches which may result in a better measure of land contamination. When constructing a summary measure of contaminated land, this study would recommend using a weighting scheme based on the Tolerable Daily Intake (TDI) values of the soil contaminants.

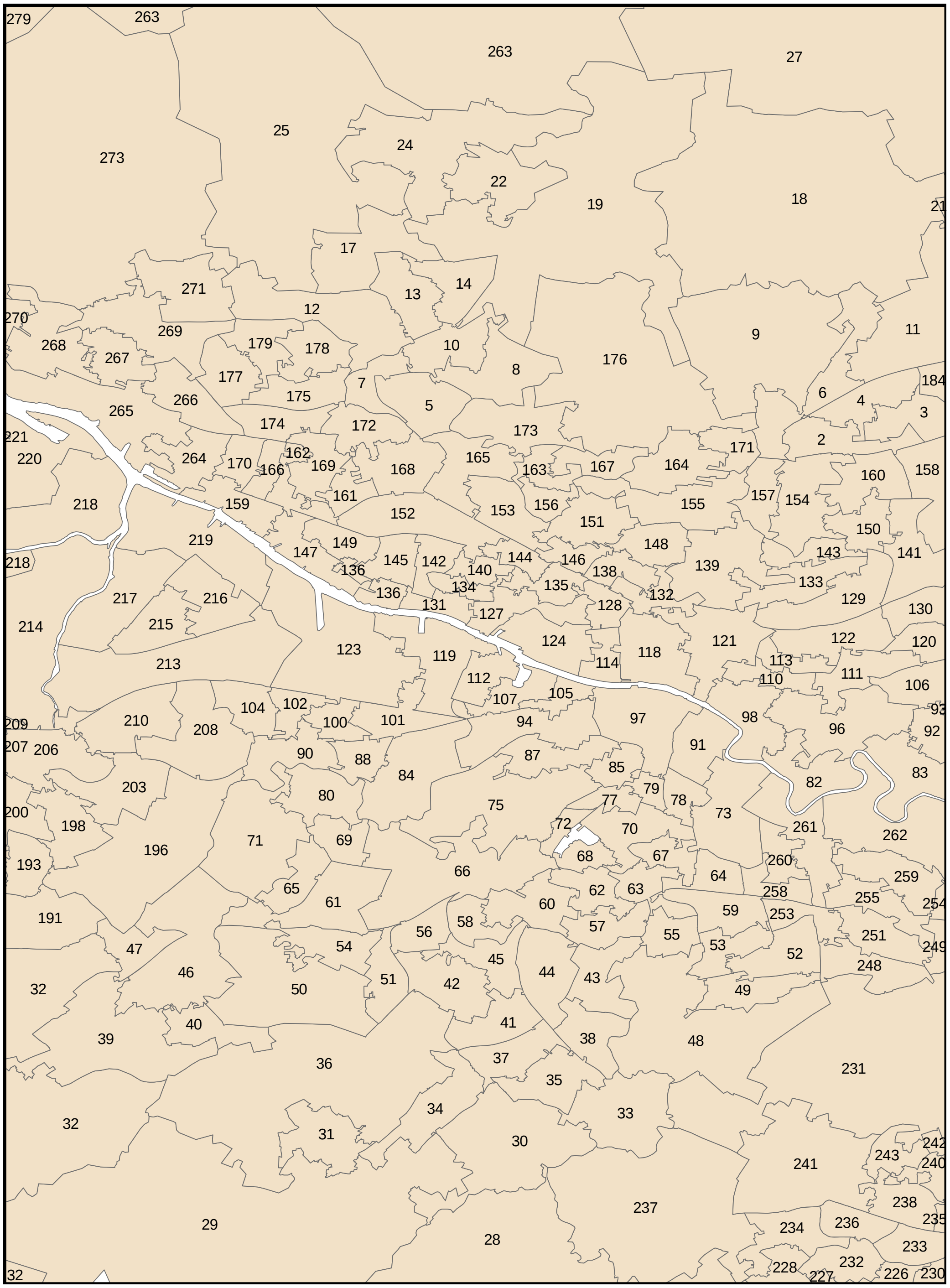
One of the key findings of this research was the apparent link between contaminated land and deprivation across Greater Glasgow. This link between soil metals and deprivation adds valuable information to the Environmental Justice Agenda. None of the measures of deprivation currently include an environmental component. For example, environmental indicators such as land quality do not play a part in constructing the SIMD index. In order to improve the quality of life and of the environment in deprived areas in the future, this study recommends that land quality should be taken into account.

Appendix



(See Table A1 for list of IG names)

Figure A.1 – Extent of Intermediate Geography areas used as part of this study across Greater Glasgow



(See Table A1 for list of IG area names)

Figure A.2 – Detailed map of Intermediate Geography areas in central Glasgow

Table A.1 – Intermediate Geography area names and number of soil samples in each

Intermediate Geography Code	Intermediate Geography Name	Figure Number	No. of Soil Samples
S02000144	Lomond Shore	1	9
S02000260	Auchinairn	2	2
S02000261	Woodhill East	3	5
S02000262	Woodhill West	4	4
S02000263	Westerton East	5	4
S02000264	Bishopbriggs West and Cadder	6	7
S02000265	Westerton West	7	2
S02000266	Kessington East	8	6
S02000268	Bishopbriggs North and Kenmure	9	10
S02000269	Kessington West	10	5
S02000270	Lenzie North	11	10
S02000271	South Castlehill and Thorn	12	5
S02000272	Kilmardinny West	13	6
S02000273	Kilmardinny East	14	6
S02000274	Rosebank and Waterside	15	5
S02000275	Kirkintilloch South	16	1
S02000276	North Castlehill and Thorn	17	8
S02000278	Torrance and Balmore	18	6
S02000279	Keystone and Dougalston	19	13
S02000280	Harestanes	20	1
S02000281	Kirkintilloch West	21	3
S02000282	Barloch	22	8
S02000283	Twechar and Harestanes East	23	4
S02000284	East Clober and Mains Estate	24	10
S02000285	West Clober and Mains Estate	25	7
S02000286	Milton of Campsie	26	3
S02000287	Lennoxtown	27	12
S02000310	Eaglesham and Waterfoot	28	4
S02000311	Mearnskirk and South Kirkhill	29	14
S02000312	North Kirkhill	30	7
S02000313	Mearns Village, Westacres and Greenfarm	31	4
S02000314	West Neilston and Uplawmoor	32	18
S02000315	Busby	33	5
S02000316	Whitecraigs and Broom	34	5
S02000317	Clarkston and Sheddens	35	5
S02000318	Crookfur and Fruin	36	23
S02000319	Williamwood	37	5
S02000320	Stamperland	38	7
S02000321	West Arthurlie and North Neilston	39	12
S02000322	Auchenback	40	4
S02000323	Lower Whitecraigs and South Giffnock	41	5
S02000324	South Thornliebank and Woodfarm	42	5
S02000325	Netherlee	43	6
S02000326	Merrylee and Braidbar	44	3
S02000327	North Giffnock and North Thornliebank	45	4
S02000328	Dunterlie, East Arthurlie and Dovecothall	46	12
S02000329	Cross Stobbs	47	4
S02000584	Carmunnock South	48	12
S02000585	Glenwood South	49	3

Intermediate Geography Code	Intermediate Geography Name	Figure Number	No. of Soil Samples
S02000586	Darnley West	50	9
S02000587	Darnley East	51	1
S02000588	Glenwood North	52	4
S02000589	Castlemilk	53	4
S02000590	Darnley North	54	6
S02000591	Carmunnock North	55	3
S02000592	Carnwadric West	56	4
S02000593	Muirend and Old Cathcart	57	4
S02000594	Carnwadric East	58	4
S02000595	Kingspark South	59	5
S02000596	Newlands	60	4
S02000597	Nitshill	61	7
S02000598	Merrylee and Millbrae	62	1
S02000599	Cathcart	63	2
S02000600	Kingspark North	64	2
S02000601	Crookston South	65	4
S02000602	Pollokshaws	66	14
S02000603	Mount Florida	67	3
S02000604	Langside	68	2
S02000606	Crookston North	69	4
S02000607	Battlefield	70	7
S02000608	Pollok South and West	71	11
S02000609	Shawlands West	72	1
S02000610	Toryglen and Oatlands	73	9
S02000611	Shettleston South	74	12
S02000612	Maxwell Park	75	13
S02000613	Carmyle and Mount Vernon South	76	5
S02000614	Strathbungo	77	2
S02000615	Govanhill East and Aikenhead	78	2
S02000616	Govanhill West	79	1
S02000617	Pollok North and East	80	4
S02000618	Baillieston East	81	10
S02000619	Dalmarnock	82	5
S02000620	Braidfauld	83	8
S02000621	Mosspark	84	10
S02000622	Pollokshields East	85	2
S02000623	Mount Vernon North and Sandyhills	86	5
S02000624	Pollokshields West	87	4
S02000625	Cardonald South and East	88	3
S02000626	Baillieston West	89	4
S02000627	Cardonald West and Central	90	5
S02000628	Gorbals and Hutchesontown	91	4
S02000629	Parkhead East and Braidfauld North	92	4
S02000630	Tollcross	93	1
S02000631	Kingston West and Dumbreck	94	6
S02000632	Shettleston North	95	3
S02000633	Parkhead West and Barrowfield	96	9
S02000634	Laurieston and Tradeston	97	8
S02000635	Calton, Galllowgate and Bridgeton	98	6
S02000636	Garrowhill West	99	4
S02000637	Cardonald North	100	2
S02000638	Craigton	101	3
S02000639	Hillington	102	2
S02000640	Greenfield	103	6

Intermediate Geography Code	Intermediate Geography Name	Figure Number	No. of Soil Samples
S02000641	Penilee	104	3
S02000642	Kinning Park and Festival Park	105	3
S02000643	Old Shettleston and Parkhead North	106	6
S02000644	Ibrox East and Cessnock	107	3
S02000645	Garrowhill East and Swinton	108	8
S02000646	Barlanark	109	3
S02000647	Gallowgate North and Bellgrove	110	2
S02000648	Carntyne West and Haghill	111	4
S02000649	Ibrox	112	3
S02000650	Dennistoun	113	1
S02000651	Anderston	114	2
S02000652	North Barlanark and Easterhouse South	115	4
S02000653	Easterhouse East	116	7
S02000654	Cranhill, Lightburn and Queenslie South	117	5
S02000655	City Centre West	118	4
S02000656	Govan and Linthouse	119	7
S02000657	Carntyne	120	2
S02000658	City Centre East	121	9
S02000659	Dennistoun North and Alexandra Parade	122	6
S02000660	Drumoyne and Shieldhall	123	21
S02000661	Finnieston and Kelvinhaugh	124	6
S02000662	Central Easterhouse	125	2
S02000663	Craigend and Ruchazie	126	12
S02000664	Hillhead	127	4
S02000665	Woodlands	128	3
S02000666	Roystonhill, Blochairn, and Provanmill	129	7
S02000667	Riddrie and Hogganfield	130	9
S02000668	Glasgow Harbour and Partick South	131	1
S02000669	Woodside	132	4
S02000670	Sighthill	133	4
S02000671	Partick	134	3
S02000672	Kelvingrove and University	135	2
S02000673	Whiteinch	136	2
S02000674	Garthamlock, Auchinlea and Gartloch	137	10
S02000675	Firhill	138	1
S02000676	Cowlairs and Port Dundas	139	7
S02000677	Partickhill and Hyndland	140	2
S02000678	Blackhill and Barmulloch East	141	7
S02000679	Broomhill	142	2
S02000680	Petershill	143	3
S02000681	Dowanhill	144	1
S02000682	Victoria Park	145	3
S02000683	North Kelvin	146	3
S02000684	Scotstoun South and West	147	5
S02000685	Keppochhill	148	3
S02000686	Scotstoun North and East	149	3
S02000687	Barmulloch	150	3
S02000688	Ruchill	151	6
S02000689	Kelvinside and Jordanhill	152	8
S02000690	Kelvindale	153	5
S02000691	Springburn	154	9
S02000692	Possil Park	155	5
S02000693	Wyndford	156	2
S02000694	Springburn East and Cowlairs	157	2

Intermediate Geography Code	Intermediate Geography Name	Figure Number	No. of Soil Samples
S02000695	Robroyston and Millerston	158	13
S02000696	Yoker South	159	2
S02000697	Balornock	160	5
S02000698	Knightswood Park East	161	1
S02000699	Knightswood East	162	1
S02000700	Maryhill West	163	2
S02000701	Milton West	164	6
S02000702	Anniesland East	165	6
S02000703	Knightswood West	166	3
S02000704	Maryhill East	167	2
S02000705	Anniesland West	168	5
S02000706	Knightswood Park West	169	4
S02000707	Yoker North	170	3
S02000708	Milton East	171	4
S02000709	Blairdardie East	172	5
S02000710	Summerston North	173	6
S02000711	Blairdardie West	174	3
S02000712	Drumchapel South	175	6
S02000713	Summerston Central and West	176	10
S02000714	Drumry West	177	6
S02000715	Drumchapel North	178	4
S02000716	Drumry East	179	3
S02000886	Forgewood	180	2
S02000901	Birkenshaw	181	1
S02000919	Townhead	182	6
S02000922	Glenmavis and Greengairs	183	4
S02000923	Stepps	184	12
S02000924	Gartcosh and Marnock	185	6
S02000925	Chryston and Muirhead	186	3
S02000926	Moodiesburn West	187	1
S02000940	Kilsyth Bogside	188	1
S02000941	Balmalloch	189	4
S02000942	Kilsyth East and Croy	190	7
S02000984	Renfrewshire Rural South & Howwood	191	17
S02000985	Paisley Glenburn West	192	3
S02000986	Paisley Glenburn East	193	5
S02000987	Paisley Foxbar	194	7
S02000988	Johnstone South West	195	8
S02000989	Paisley Dykebar	196	14
S02000990	Paisley South West	197	8
S02000991	Paisley South East	198	4
S02000992	Johnstone South East	199	6
S02000993	Paisley South	200	5
S02000994	Johnstone North West	201	4
S02000995	Paisley West	202	5
S02000997	Paisley East	203	6
S02000998	Johnstone North East	204	4
S02000999	Elderslie and Phoenix	205	12
S02001000	Paisley Central	206	5
S02001001	Paisley North West	207	3
S02001002	Paisley Ralston	208	5
S02001003	Paisley Ferguslie	209	9
S02001004	Paisley North East	210	7
S02001005	Linwood South	211	4

Intermediate Geography Code	Intermediate Geography Name	Figure Number	No. of Soil Samples
S02001006	Linwood North	212	5
S02001007	Paisley Gallowhill and Hillington	213	13
S02001010	Paisley North	214	18
S02001011	Renfrew South	215	5
S02001013	Renfrew East	216	7
S02001014	Renfrew West	217	7
S02001015	Renfrewshire Rural North and Langbank	218	35
S02001016	Renfrew North	219	14
S02001017	Erskine East and Inchinnan	220	2
S02001018	Erskine Central	221	1
S02001083	Chapelton, Glengavel and Sandford	222	3
S02001095	Ashgill and Netherburn	223	1
S02001098	Glassford, Quarter and Allanton	224	12
S02001105	Crosshouse and Lindsayfield	225	5
S02001108	Whitehills West	226	5
S02001109	Greenhills	227	2
S02001110	Mossneuk and Newlandsmuir	228	6
S02001111	Little Earnock	229	4
S02001112	Birniehill, Kelvin and Whitehills East	230	10
S02001114	Nerston and EK Landward Area	231	40
S02001116	Westwood South	232	3
S02001117	The Murray	233	5
S02001120	Hairmyres and Westwood West	234	5
S02001121	St Leonards South	235	5
S02001123	Westwood East	236	3
S02001124	Thorntonhall, Jackton and Gardenhall	237	8
S02001125	West Mains	238	5
S02001126	St Leonards North	239	5
S02001130	East Mains	240	1
S02001131	Stewartfield West	241	12
S02001132	Calderwood West and Nerston	242	6
S02001135	Stewartfield East	243	4
S02001136	Calderwood Central	244	3
S02001138	Calderwood East	245	7
S02001139	Hamilton Centre and Low Parks	246	2
S02001145	Bothwell South	247	2
S02001146	Fernhill and Cathkin	248	3
S02001148	Whitlawburn and Greenlees	249	5
S02001149	Vicarland and Cairns	250	6
S02001150	Burnside and Springhall	251	4
S02001151	Low Blantyre and Bardykes	252	2
S02001152	Spittal	253	3
S02001153	Cambuslang Central	254	4
S02001154	High Crosshill	255	6
S02001156	Halfway, Hallside and Drumsagard	256	9
S02001157	Westburn and Newton	257	6
S02001158	Bankhead South	258	2
S02001159	Burgh, Eastfield and Silverbank	259	5
S02001160	Burnhill and Bankhead North	260	2
S02001161	Shawfield and Clincarthill	261	6
S02001162	Farme Cross and Gallowflat North	262	10
S02001163	Blane Valley	263	1
S02001184	IZ One	264	5
S02001185	IZ Two	265	12

Intermediate Geography Code	Intermediate Geography Name	Figure Number	No. of Soil Samples
S02001186	IZ Three	266	5
S02001187	IZ Four	267	3
S02001188	IZ Five	268	7
S02001189	IZ Six	269	7
S02001190	IZ Seven	270	11
S02001191	IZ Eight	271	10
S02001192	IZ Nine	272	15
S02001193	IZ Ten	273	11
S02001194	IZ Eleven	274	20
S02001195	IZ Twelve	275	13
S02001196	IZ Thirteen	276	4
S02001197	IZ Fourteen	277	2
S02001198	IZ Fifteen	278	5
S02001199	IZ Sixteen	279	10

Note: The IG Codes from S02001184 to S02001199 represent IG areas in the West Dunbartonshire region. Since West Dunbartonshire did not provide IG names, they are just labelled as 'IZ One'.... 'IZ Sixteen' for the area covered in this study

Table A.2 – Population distribution of males in Intermediate Geography areas across Greater Glasgow

Intermediate Geography Name	Age																	
	0-4	5-9	10-14	15-19	20-24	25-29	30-34	35-39	40-44	45-49	50-54	55-59	60-64	65-69	70-74	75-79	80-84	85+
Lomond Shore	74	87	87	72	65	69	80	113	109	118	125	116	93	73	61	58	22	13
Auchinairn	123	143	163	126	1[illegible]	120	152	183	168	109	125	110	105	98	68	43	40	25
Woodhill East	84	93	130	125	80	66	87	135	125	124	86	70	60	35	14	11	4	1
Woodhill West	114	163	200	180	161	89	122	175	194	179	170	123	145	148	78	37	26	10
Westerton East	88	120	130	97	87	53	84	108	145	128	116	113	55	71	87	48	31	21
Bishopbriggs West and Cadder	161	164	202	174	143	128	155	213	244	215	175	175	166	159	134	95	38	17
Westerton West	77	93	101	94	80	53	91	110	106	116	112	105	79	78	47	36	11	2
Kessington East	75	92	96	112	87	45	69	88	106	109	117	106	97	73	64	36	18	9
Bishopbriggs North and Kenmure	139	186	215	205	131	88	137	217	241	203	167	153	187	197	159	93	37	23
Kessington West	72	92	97	127	113	56	70	86	89	120	155	127	107	96	76	62	38	27
Lenzie North	149	173	218	216	241	178	179	218	210	206	239	198	168	134	112	100	42	26
South Castlehill and Thorn	136	192	221	175	127	83	93	165	165	144	176	134	116	108	68	59	32	15
Kilmardinny West	84	125	114	114	76	38	63	105	119	120	129	115	94	104	86	53	43	42
Kilmardinny East	89	81	112	79	74	51	94	106	111	104	123	93	95	90	60	47	21	13
Rosebank and Waterside	104	143	135	124	105	74	120	148	147	124	137	93	91	86	46	23	9	6
Kirkintilloch South	66	92	124	121	95	89	110	141	123	116	97	83	81	83	45	46	30	12
North Castlehill and Thorn	146	165	184	150	129	113	169	187	174	169	181	125	96	73	51	28	10	10
Torrance and Balmore	93	120	114	111	86	40	67	127	129	141	131	90	58	47	38	34	13	8
Keystone and Dougalston	63	103	120	140	105	66	88	92	131	135	186	150	129	104	114	87	69	38
Harestanes	101	102	90	109	91	103	146	143	104	86	95	80	109	100	88	54	25	14
Kirkintilloch West	99	109	103	93	111	147	185	168	140	116	112	104	102	97	70	44	23	24
Barloch	80	109	80	75	72	46	72	102	109	111	108	97	85	99	78	59	39	17
Twechar and Harestanes East	99	104	124	117	143	82	121	111	127	111	126	105	79	63	42	24	10	6
East Clober and Mains Estate	95	157	133	151	136	74	91	140	146	126	108	94	77	90	71	44	27	10
West Clober and Mains Estate	138	123	152	105	67	55	86	144	162	126	92	78	53	46	30	14	5	6
Milton of Campsie	109	142	167	140	129	85	118	159	191	167	170	111	106	62	55	40	21	17

Intermediate Geography Name	Age																	
	0-4	5-9	10-14	15-19	20-24	25-29	30-34	35-39	40-44	45-49	50-54	55-59	60-64	65-69	70-74	75-79	80-84	85+
Lennoxtown	132	158	171	126	121	114	181	212	178	160	157	102	108	95	60	37	19	19
Eaglesham and Waterfoot	132	166	157	137	107	114	134	205	183	181	161	142	126	130	88	56	25	26
Mearnskirk and South Kirkhill	146	180	188	170	128	82	119	185	219	187	205	157	173	152	151	93	66	26
North Kirkhill	107	95	103	96	86	78	118	108	112	103	109	84	70	58	49	42	17	6
Mearns Village, Westacres and Greenfarm	247	263	231	193	119	108	217	282	280	174	138	126	93	76	44	50	24	12
West Neilston and Uplawmoor	184	229	242	224	187	127	161	226	262	254	238	171	127	119	84	51	27	10
Busby	104	120	92	83	92	127	118	134	131	113	109	70	75	71	69	60	38	27
Whitecraigs and Broom	88	110	136	135	99	46	62	90	126	137	182	133	112	97	78	57	28	14
Clarkston and Sheddens	161	243	247	208	185	120	169	215	270	258	199	169	149	119	103	66	43	22
Crookfur and Fruin	175	206	218	205	156	108	158	204	258	218	186	154	139	89	74	67	39	41
Williamwood	111	115	131	101	79	51	87	113	152	107	104	91	85	58	58	34	17	14
Stamperland	153	130	158	124	103	84	145	156	176	135	115	93	75	78	53	37	18	12
West Arthurlie and North Neilston	140	162	188	145	121	117	169	218	204	164	187	146	139	125	95	68	28	11
Auchenback	115	137	169	149	122	107	121	159	145	119	134	94	105	98	61	43	15	8
Lower Whitecraigs and South Giffnock	117	115	148	133	106	37	64	93	138	136	136	130	84	73	61	50	34	28
South Thornliebank and Woodfarm	117	156	152	144	105	91	117	142	172	162	129	91	98	80	69	56	25	9
Netherlee	155	177	164	124	103	64	108	162	165	172	164	125	97	108	85	94	35	37
Merrylee and Braidbar	145	151	183	177	136	69	99	181	192	172	194	151	121	100	99	84	45	28
North Giffnock and North Thornliebank	98	115	105	97	75	61	118	129	135	120	103	68	62	77	72	57	40	37
Dunterlie, East Arthurlie and Dovecothall	202	208	189	174	149	180	233	236	197	167	161	124	136	110	89	67	31	26
Cross Stobbs	97	110	112	109	79	91	129	148	137	118	140	108	89	59	51	41	18	12
Carmunnock South	91	88	99	117	74	79	63	103	99	124	77	79	89	70	66	44	29	23
Glenwood South	135	169	198	209	171	146	171	213	181	188	169	124	109	97	67	52	17	11
Darnley West	126	131	127	99	89	105	155	165	155	120	67	62	50	33	24	8	7	5

Intermediate Geography Name	Age																	
	0-4	5-9	10-14	15-19	20-24	25-29	30-34	35-39	40-44	45-49	50-54	55-59	60-64	65-69	70-74	75-79	80-84	85+
Darnley East	211	206	177	189	150	145	197	188	151	163	132	102	88	54	77	31	19	10
Glenwood North	193	173	220	148	85	91	129	118	79	78	99	74	55	55	46	27	19	5
Castlemilk	150	191	207	203	172	165	170	189	186	165	162	114	133	101	91	81	34	26
Darnley North	106	104	111	121	81	98	142	130	116	101	57	55	38	31	34	31	24	13
Carmunnock North	82	98	113	126	115	79	84	115	122	132	157	102	66	65	48	38	16	14
Carnwadric West	203	144	164	156	111	150	194	200	185	156	121	121	125	92	77	38	16	6
Muirend and Old Cathcart	134	137	154	178	141	138	170	232	183	191	176	121	115	95	87	87	39	28
Carnwadric East	74	79	101	123	95	85	101	75	81	100	100	102	81	80	63	46	33	16
Kingspark South	156	152	165	151	142	140	195	182	184	146	120	85	66	61	45	32	11	12
Newlands	175	177	235	217	177	110	142	200	224	239	193	154	130	117	105	50	32	26
Nitshill	155	227	229	231	155	102	143	165	175	134	131	120	135	107	95	39	20	17
Merrylee and Millbrae	92	93	117	132	102	88	134	174	133	127	124	83	88	75	67	55	28	29
Cathcart	111	101	98	100	150	264	228	180	169	143	148	94	77	66	94	64	40	18
Kingspark North	153	195	177	180	142	137	138	187	199	190	156	111	79	78	62	29	26	18
Crookston South	168	169	217	206	130	60	118	127	137	108	120	103	113	98	71	28	9	10
Pollokshaws	154	123	125	119	124	178	194	183	166	166	141	129	140	118	119	77	43	41
Mount Florida	107	89	92	102	138	202	226	172	150	139	108	107	96	72	57	40	32	18
Langside	84	62	56	66	104	300	265	196	163	111	98	69	56	46	52	44	30	22
Crookston North	70	110	116	91	80	58	71	108	97	97	79	80	101	104	90	57	26	17
Battlefield	143	93	119	124	193	304	340	238	209	143	142	99	76	72	67	80	40	25
Pollok South and West	165	163	148	147	119	139	219	234	238	193	145	108	116	90	60	43	19	11
Shawlands West	59	52	46	49	137	272	242	178	114	109	84	58	57	56	31	39	16	25
Toryglen and Oatlands	151	191	224	179	146	161	163	207	193	145	180	162	172	177	141	101	52	21
Shettleston South	110	142	152	152	94	96	124	154	187	156	128	91	91	94	62	60	36	23
Maxwell Park	143	174	187	208	182	110	199	224	222	237	231	122	131	110	99	75	39	43
Carmyle and Mount Vernon South	76	101	89	97	77	72	73	105	150	97	101	90	90	70	77	43	15	14
Strathbungo	184	131	116	114	241	366	428	315	221	191	157	98	70	74	46	50	21	17
Govanhill East and Aikenhead	107	110	128	106	131	133	150	158	151	136	107	91	92	90	68	46	32	29

Intermediate Geography Name	Age																	
	0-4	5-9	10-14	15-19	20-24	25-29	30-34	35-39	40-44	45-49	50-54	55-59	60-64	65-69	70-74	75-79	80-84	85+
Govanhill West	181	154	118	140	242	259	333	294	200	136	139	103	92	57	69	43	24	10
Pollok North and East	146	189	247	193	143	122	150	170	151	117	111	86	114	105	67	28	6	11
Baillieston East	131	151	138	114	121	129	163	196	174	143	134	99	100	93	72	61	32	17
Dalmarnock	102	109	115	134	101	104	101	121	119	95	128	119	133	111	92	45	28	9
Braidfauld	148	181	198	156	113	126	154	148	151	127	108	128	114	112	106	56	44	34
Mosspark	122	138	157	137	103	93	139	187	169	134	142	129	108	129	147	123	65	39
Pollokshields East	202	176	162	189	184	207	263	189	157	104	72	44	60	36	31	33	9	19
Mount Vernon North and Sandyhills	96	138	161	141	128	61	111	137	170	168	144	130	116	89	72	48	26	6
Pollokshields West	187	184	174	147	148	164	170	170	175	136	131	94	94	63	66	46	26	11
Cardonald South and East	120	112	126	119	111	88	117	159	142	119	95	82	57	72	72	40	29	21
Baillieston West	104	125	141	167	149	139	126	151	130	138	144	150	132	141	106	64	28	9
Cardonald West and Central	125	175	199	184	148	130	140	205	207	172	174	163	162	153	132	90	51	29
Gorbals and Hutchesontown	129	124	148	113	192	170	197	193	159	157	135	131	145	128	95	65	37	20
Parkhead East and Braidfauld North	116	128	133	105	97	120	125	140	116	111	91	85	65	68	40	31	16	11
Tollcross	112	111	141	126	111	131	159	158	143	137	99	94	80	78	65	43	23	15
Kingston West and Dumbreck	118	103	104	89	105	139	163	168	156	107	113	106	94	99	86	81	32	24
Shettleston North	112	130	113	107	113	161	173	199	155	162	140	124	149	119	112	74	51	30
Parkhead West and Barrowfield	126	156	174	182	157	143	174	197	180	178	177	143	142	139	95	78	23	19
Laurieston and Tradeston	87	63	62	87	158	282	259	230	181	111	150	103	102	85	74	46	16	13
Calton, Galllowgate and Bridgeton	101	79	102	124	213	300	308	201	212	164	161	146	164	137	109	68	36	18
Garrowhill West	125	196	162	146	139	72	165	212	208	162	147	110	102	99	67	63	19	22
Cardonald North	118	129	149	144	121	111	86	109	130	136	116	95	108	94	78	69	50	24
Craigton	96	103	128	126	88	98	115	146	107	130	119	118	131	136	127	89	57	26
Hillington	115	115	137	116	90	115	133	161	141	108	97	79	76	62	63	36	37	18
Greenfield	137	164	157	137	133	145	160	225	166	145	192	147	162	163	117	90	41	24

Intermediate Geography Name	Age																	
	0-4	5-9	10-14	15-19	20-24	25-29	30-34	35-39	40-44	45-49	50-54	55-59	60-64	65-69	70-74	75-79	80-84	85+
Penilee	199	177	220	230	138	111	159	169	180	134	128	129	129	155	112	87	39	23
Kinning Park and Festival Park	99	89	82	97	184	298	230	167	139	109	75	61	67	50	31	31	13	10
Old Shettleston and Parkhead North	85	79	89	92	66	77	97	107	86	65	100	105	109	105	92	75	40	27
Ibrox East and Cessnock	91	81	93	107	171	204	175	150	140	121	104	70	58	56	37	34	13	14
Garrowhill East and Swinton	157	174	175	141	101	132	177	254	208	196	127	89	45	54	37	18	17	3
Barlanark	166	219	198	223	143	118	207	184	175	121	118	99	92	61	47	37	18	11
Gallowgate North and Bellgrove	82	75	92	131	292	234	159	102	117	98	85	71	83	80	55	38	17	10
Carntyne West and Haghill	179	163	156	159	178	195	235	218	213	145	144	116	130	102	76	45	14	8
Ibrox	111	101	127	103	111	96	109	90	91	84	76	75	61	68	46	29	16	11
Dennistoun	87	69	76	106	280	217	215	174	181	121	98	93	56	48	37	37	21	16
Anderston	56	57	51	69	251	230	140	112	95	67	95	63	57	45	33	30	18	5
North Barlanark and Easterhouse South	188	180	215	162	128	94	108	96	123	79	87	49	48	47	47	21	15	11
Easterhouse East	100	131	161	127	122	79	97	131	112	115	79	61	59	47	69	38	19	4
Cranhill, Lightburn and Queenslie South	169	220	249	234	159	121	188	205	213	171	185	149	160	162	112	78	39	24
City Centre West	45	42	43	244	495	301	237	178	120	105	88	77	81	45	51	33	22	16
Govan and Linthouse	175	159	177	179	175	198	210	179	185	187	194	151	150	107	82	48	24	24
Carntyne	57	66	76	75	62	56	85	91	121	97	105	95	118	111	128	96	54	47
City Centre East	88	88	76	457	721	452	322	243	202	172	195	220	224	207	159	113	61	32
Dennistoun North and Alexandra Parade	109	108	126	138	279	237	259	248	198	177	134	137	135	150	104	71	35	32
Drumoyne and Shieldhall	158	179	188	192	144	152	189	199	203	181	175	146	146	150	112	92	50	22
Finnieston and Kelvinhaugh	83	68	83	251	763	448	312	238	165	121	87	71	56	53	42	36	15	9
Central Easterhouse	157	139	161	146	100	89	91	90	99	108	96	62	65	56	51	22	12	6
Craigend and Ruchazie	143	179	204	185	140	139	172	187	147	128	170	176	153	125	92	53	39	24
Hillhead	73	71	80	148	393	401	297	242	192	147	120	103	86	76	65	46	16	23

Intermediate Geography Name	Age																	
	0-4	5-9	10-14	15-19	20-24	25-29	30-34	35-39	40-44	45-49	50-54	55-59	60-64	65-69	70-74	75-79	80-84	85+
Woodlands	124	101	103	241	789	384	282	196	158	152	91	49	67	40	35	24	13	15
Roystonhill, Blochairn, and Provanmill	177	186	192	166	181	170	235	220	184	152	161	135	129	123	100	59	34	12
Riddrie and Hogganfield	141	182	224	201	261	228	224	234	211	164	154	132	130	152	127	88	54	29
Glasgow Harbour and Partick South	40	51	49	67	152	215	195	117	113	82	60	36	54	28	36	13	5	2
Woodside	66	54	47	101	212	178	137	117	84	75	73	67	90	87	71	58	47	21
Sighthill	148	148	132	104	137	180	208	205	147	115	82	87	84	72	35	36	17	10
Partick	46	55	41	84	193	240	185	150	124	101	92	95	58	82	53	60	29	20
Kelvingrove and University	88	63	70	161	633	334	256	199	151	139	140	82	68	37	40	32	13	17
Whiteinch	84	89	78	90	93	156	154	154	123	93	93	69	59	54	50	34	26	10
Garthamlock, Auchinlea and Gartloch	162	158	191	154	110	130	143	164	139	102	81	60	50	35	30	8	6	0
Firhill	169	117	107	128	452	364	322	231	180	143	141	114	115	85	65	47	25	13
Cowlairs and Port Dundas	129	130	122	139	122	175	167	143	140	92	92	90	84	61	58	37	17	9
Partickhill and Hyndland	76	69	62	71	189	195	216	177	155	114	120	91	79	40	36	31	18	19
Blackhill and Barmulloch East	93	110	158	151	112	91	96	121	130	114	96	108	79	103	67	40	18	4
Broomhill	105	109	85	86	102	163	203	176	145	120	157	92	83	61	71	57	34	23
Petershill	186	159	152	157	184	227	223	263	161	136	103	108	85	65	57	36	15	7
Dowanhill	90	84	55	161	276	251	274	204	179	167	142	78	59	48	49	24	13	17
Victoria Park	94	95	119	90	51	57	74	95	104	115	87	68	51	36	59	33	39	21
North Kelvin	103	56	67	115	303	238	204	166	126	115	103	83	62	40	28	23	10	6
Scotstoun South and West	167	142	149	135	133	129	174	153	147	105	98	78	81	73	41	41	26	17
Keppochhill	149	132	202	157	145	137	159	212	138	161	147	144	127	127	87	57	23	8
Scotstoun North and East	172	156	138	119	113	116	127	161	166	152	142	101	89	86	74	50	25	29
Barmulloch	141	149	153	131	148	110	130	139	111	118	91	83	94	113	53	38	23	6
Ruchill	81	96	114	353	360	331	268	175	163	130	113	92	90	81	73	64	44	22
Kelvinside and Jordanhill	139	168	175	160	171	238	253	234	265	222	213	173	107	124	117	89	76	30
Kelvindale	155	143	147	176	214	175	213	237	214	216	217	156	101	105	79	61	40	42

Intermediate Geography Name	Age																	
	0-4	5-9	10-14	15-19	20-24	25-29	30-34	35-39	40-44	45-49	50-54	55-59	60-64	65-69	70-74	75-79	80-84	85+
Springburn	103	136	160	112	101	88	124	154	155	107	111	97	122	157	128	106	34	16
Possil Park	109	129	138	127	146	114	117	148	136	145	155	154	158	163	141	108	43	30
Wyndford	122	119	109	109	117	144	168	170	174	122	140	93	126	137	107	77	25	12
Springburn East and Cowlairs	138	165	116	105	116	135	191	207	160	126	91	97	101	98	90	54	21	14
Robroyston and Millerston	251	225	200	153	128	236	377	360	253	189	109	72	55	51	20	20	12	8
Yoker South	118	111	121	116	111	156	145	180	132	109	95	55	56	49	43	36	27	4
Balornock	89	136	147	184	121	82	110	123	140	90	97	120	134	123	100	59	28	15
Knightswood Park East	96	91	79	82	80	92	104	109	109	106	75	53	68	91	85	64	30	12
Knightswood East	124	105	109	78	97	108	132	163	137	105	94	110	86	117	118	106	35	14
Maryhill West	120	105	101	110	83	95	124	115	113	84	97	64	60	51	35	24	13	10
Milton West	66	119	153	151	117	88	88	113	113	114	127	114	129	117	98	60	44	25
Anniesland East	99	92	72	131	145	165	172	179	141	133	94	82	56	46	56	46	28	23
Knightswood West	66	109	74	75	59	81	106	91	109	85	57	62	89	73	78	32	24	8
Maryhill East	121	98	123	107	90	96	93	111	91	86	96	96	88	84	49	42	21	15
Anniesland West	112	107	151	161	175	161	176	186	168	135	118	122	125	92	98	119	56	38
Knightswood Park West	64	84	83	81	65	75	94	105	115	95	109	95	99	112	108	102	50	43
Yoker North	60	95	117	128	86	85	83	117	132	89	102	88	78	73	53	64	34	18
Milton East	134	154	151	168	106	113	96	134	124	91	126	109	93	94	77	58	29	24
Blairdardie East	153	155	162	163	117	154	191	223	215	158	167	137	156	173	145	109	54	33
Summerston North	110	113	119	171	156	121	113	132	123	135	149	123	113	79	72	52	15	20
Blairdardie West	68	84	113	78	67	40	71	122	114	92	106	57	69	86	57	56	37	17
Drumchapel South	111	139	187	163	103	76	106	92	93	80	96	78	49	51	36	43	28	22
Summerston Central and West	149	160	145	139	95	121	159	187	152	122	119	98	88	71	51	24	15	8
Drumry West	144	141	169	184	107	107	114	104	109	102	71	74	51	33	42	24	14	11
Drumchapel North	171	158	180	146	123	118	100	96	89	89	65	74	53	32	19	19	6	4
Drumry East	118	124	152	150	95	68	84	70	86	99	83	63	63	33	32	29	18	8
Forgewood	216	191	167	163	147	233	270	227	181	135	166	126	106	88	74	67	21	16
Birkenshaw	189	196	201	187	141	119	204	230	238	183	191	117	120	91	64	49	15	8

Intermediate Geography Name	Age																	
	0-4	5-9	10-14	15-19	20-24	25-29	30-34	35-39	40-44	45-49	50-54	55-59	60-64	65-69	70-74	75-79	80-84	85+
Townhead	156	191	247	235	173	182	195	245	203	161	175	156	142	123	98	56	31	18
Glenmavis and Greengairs	132	136	137	140	129	105	155	187	160	147	150	131	103	85	45	25	17	9
Stepps	121	148	163	157	149	116	145	178	210	176	211	154	141	112	87	52	37	18
Gartcosh and Marnock	111	109	102	137	107	84	148	131	122	114	140	102	92	75	47	34	15	20
Chryston and Muirhead	132	138	129	143	116	85	164	179	139	136	131	115	116	85	71	41	25	10
Moodiesburn West	117	110	98	113	105	109	125	118	110	96	106	78	95	84	57	34	22	5
Kilsyth Bogside	111	119	97	107	88	95	119	106	114	95	105	89	63	61	37	36	19	9
Balmalloch	139	163	164	153	132	124	137	164	143	160	149	134	122	103	76	65	30	5
Kilsyth East and Croy	156	144	145	150	133	146	159	165	141	150	161	136	125	131	74	71	40	28
Renfrewshire Rural South & Howwood	86	70	81	74	69	74	120	146	108	94	133	96	76	53	51	27	10	8
Paisley Glenburn West	124	162	164	139	90	86	131	162	142	121	144	113	112	79	79	64	31	16
Paisley Glenburn East	122	137	152	119	120	80	152	138	138	110	119	99	91	66	61	36	16	9
Paisley Foxbar	137	177	190	160	150	138	179	274	207	167	161	142	144	152	143	81	46	10
Johnstone South West	196	207	255	239	122	129	181	220	204	140	151	145	176	151	107	57	26	8
Paisley Dykebar	112	103	95	80	95	103	119	156	111	96	113	101	103	88	62	45	24	15
Paisley South West	142	135	154	188	159	149	168	219	216	227	239	161	133	106	87	37	20	12
Paisley South East	127	123	136	128	117	134	208	185	167	147	172	153	135	183	157	119	59	44
Johnstone South East	132	156	183	125	107	108	154	216	164	139	132	118	99	133	103	65	30	15
Paisley South	119	114	96	98	74	88	158	202	142	112	123	86	70	52	58	40	23	17
Johnstone North West	99	114	112	95	92	96	134	121	114	102	110	89	104	104	54	44	25	12
Paisley West	235	225	203	153	150	200	238	274	216	148	141	84	103	65	75	45	16	10
Paisley East	114	130	161	152	121	116	170	139	143	116	162	111	117	87	77	45	30	12
Johnstone North East	94	111	74	63	70	89	130	156	110	91	91	86	81	66	64	41	29	22
Elderslie and Phoenix	140	155	161	186	148	108	150	216	173	198	228	166	132	141	113	58	43	23
Paisley Central	109	95	101	163	310	347	350	283	222	185	181	118	133	136	81	59	40	42
Paisley North West	81	71	67	85	110	138	142	133	102	98	103	77	94	78	52	51	32	18
Paisley Ralston	126	159	157	140	128	85	100	168	204	215	195	158	142	126	133	94	57	29
Paisley Ferguslie	207	179	234	228	134	160	164	185	135	133	112	65	71	56	46	23	4	3

Intermediate Geography Name	Age																	
	0-4	**5-9**	**10-14**	**15-19**	**20-24**	**25-29**	**30-34**	**35-39**	**40-44**	**45-49**	**50-54**	**55-59**	**60-64**	**65-69**	**70-74**	**75-79**	**80-84**	**85+**
Paisley North East	128	110	129	163	150	202	254	196	200	177	209	149	114	97	145	95	68	40
Linwood South	115	136	142	149	119	109	140	156	142	101	118	126	155	123	134	65	22	13
Linwood North	131	182	207	169	138	104	162	203	184	92	124	118	156	129	56	41	3	6
Paisley Gallowhill and Hillington	155	170	181	133	129	145	198	187	160	143	187	169	178	128	106	79	52	35
Paisley North	87	83	129	108	122	160	209	138	136	114	117	107	112	112	70	53	28	11
Renfrew South	140	134	161	157	153	126	196	160	156	169	197	184	156	164	135	109	35	15
Renfrew East	171	181	216	214	222	154	195	262	246	241	260	185	119	98	71	54	39	19
Renfrew West	177	166	215	199	180	189	248	223	227	201	184	184	127	102	66	49	22	10
Renfrewshire Rural North and Langbank	93	101	116	102	87	68	93	154	164	147	168	149	117	84	94	78	64	48
Renfrew North	103	72	80	89	87	127	152	104	110	109	102	86	75	64	52	40	16	12
Erskine East and Inchinnan	208	234	240	205	120	126	210	287	288	232	195	130	87	64	60	40	36	18
Erskine Central	164	164	186	166	189	182	207	194	189	167	197	173	137	90	70	60	36	19
Chapelton, Glengavel and Sandford	84	81	83	112	71	61	99	104	132	128	110	110	84	64	39	19	13	10
Ashgill and Netherburn	66	104	94	82	65	67	115	99	94	91	75	78	71	58	37	28	21	11
Glassford, Quarter and Allanton	106	97	91	109	85	73	114	117	81	98	114	97	92	54	40	28	24	6
Crosshouse and Lindsayfield	156	145	151	143	168	142	200	185	166	135	145	96	58	42	29	16	7	10
Whitehills West	124	158	159	168	167	125	139	129	135	124	146	103	80	77	57	37	24	4
Greenhills	120	130	148	135	122	121	153	113	105	109	123	100	89	47	22	18	5	6
Mossneuk and Newlandsmuir	75	154	170	146	110	84	107	145	161	133	108	88	40	40	19	14	5	2
Little Earnock	99	81	89	108	90	128	139	121	93	80	99	104	77	51	28	14	10	3
Birniehill, Kelvin and Whitehills East	116	159	210	155	130	109	159	174	175	144	86	97	103	81	67	46	25	2
Nerston and EK Landward Area	78	84	83	90	81	77	122	116	130	135	99	86	90	78	61	29	27	15
Westwood South	145	179	165	161	123	132	203	238	174	140	113	99	131	110	100	56	18	5
The Murray	108	129	120	104	102	103	155	170	130	129	104	80	92	82	135	86	41	21

Intermediate Geography Name	Age																	
	0-4	5-9	10-14	15-19	20-24	25-29	30-34	35-39	40-44	45-49	50-54	55-59	60-64	65-69	70-74	75-79	80-84	85+
Hairmyres and Westwood West	85	108	108	117	90	81	130	138	125	121	111	79	87	79	56	38	15	12
St Leonards South	121	188	165	189	162	158	182	209	202	160	150	175	150	110	87	58	30	6
Westwood East	87	102	92	115	67	77	132	123	157	84	95	82	90	106	90	76	34	11
Thorntonhall, Jackton and Gardenhall	65	107	115	109	66	48	86	124	127	129	120	71	60	29	36	21	9	3
West Mains	85	104	90	99	70	63	86	122	110	112	84	62	63	79	99	85	25	6
St Leonards North	124	176	140	156	109	122	172	214	158	132	123	150	176	136	101	59	25	12
East Mains	54	71	88	97	95	81	91	113	109	103	83	77	94	73	92	78	39	19
Stewartfield West	188	208	148	128	76	83	191	242	204	177	123	86	63	48	57	31	10	10
Calderwood West and Nerston	104	107	127	131	89	69	127	129	147	103	108	82	107	108	101	67	48	15
Stewartfield East	124	138	116	84	47	40	83	182	172	135	97	74	39	30	25	20	9	2
Calderwood Central	98	117	95	89	99	137	171	181	163	109	87	89	121	114	102	82	40	21
Calderwood East	148	154	142	128	111	181	205	176	131	106	78	107	100	97	59	39	15	11
Hamilton Centre and Low Parks	75	66	74	100	107	154	142	129	152	116	147	105	94	81	76	62	39	19
Bothwell South	70	84	101	85	77	69	89	97	123	136	128	92	72	81	47	36	21	25
Fernhill and Cathkin	156	166	166	167	135	131	122	131	136	130	137	104	96	82	77	51	23	10
Whitlawburn and Greenlees	126	138	153	151	162	133	127	150	161	163	158	100	69	36	38	9	18	6
Vicarland and Cairns	149	133	160	142	132	120	129	118	159	147	172	147	106	116	89	80	35	23
Burnside and Springhall	133	149	171	150	139	105	117	148	158	172	174	126	131	136	131	72	41	14
Low Blantyre and Bardykes	96	121	127	125	113	85	99	112	149	113	153	99	64	49	34	17	7	7
Spittal	49	101	97	114	73	73	75	74	89	83	100	84	103	74	59	48	26	9
Cambuslang Central	111	133	118	118	118	75	124	142	139	170	145	113	108	101	74	62	28	16
High Crosshill	85	116	111	118	93	78	85	109	150	124	113	96	74	68	77	48	24	19
Halfway, Hallside and Drumsagard	169	141	128	119	99	169	238	181	168	128	109	70	59	38	29	23	14	3
Westburn and Newton	97	108	119	96	90	96	119	110	110	103	115	78	54	60	34	20	13	3
Bankhead South	111	130	114	124	196	110	124	163	145	131	111	86	82	58	37	25	22	10

Intermediate Geography Name	Age																	
	0-4	5-9	10-14	15-19	20-24	25-29	30-34	35-39	40-44	45-49	50-54	55-59	60-64	65-69	70-74	75-79	80-84	85+
Burgh, Eastfield and Silverbank	114	141	177	161	163	134	184	184	203	226	201	162	140	146	100	91	38	28
Burnhill and Bankhead North	87	108	130	121	98	86	115	126	119	118	123	106	96	108	74	59	29	14
Shawfield and Clincarthill	83	76	91	98	129	118	123	131	142	124	111	98	82	80	77	70	39	32
Farme Cross and Gallowflat North	71	53	46	64	71	118	126	112	121	82	80	64	64	69	75	65	25	27
Blane Valley	134	126	144	128	86	60	94	127	174	173	204	215	145	96	109	58	39	30
IZ One	161	193	170	134	140	144	190	164	168	154	131	119	109	96	76	53	24	14
IZ Two	161	126	125	112	132	168	174	172	140	112	98	86	68	64	57	51	31	28
IZ Three	186	245	217	212	176	185	185	187	181	165	194	186	168	146	104	66	42	23
IZ Four	109	134	129	136	110	134	132	134	137	114	138	158	124	145	139	107	55	24
IZ Five	136	152	181	180	160	162	197	205	204	251	179	153	156	168	116	125	54	27
IZ Six	101	118	76	125	114	111	116	143	123	105	103	107	88	80	87	61	23	14
IZ Seven	220	235	182	174	126	172	242	275	261	182	155	131	95	86	80	45	26	21
IZ Eight	178	221	216	225	170	143	155	192	163	120	144	107	86	84	85	73	40	15
IZ Nine	143	152	165	174	188	210	285	239	229	224	199	125	127	110	104	79	56	39
IZ Ten	116	111	177	203	159	132	141	168	194	177	200	163	124	101	74	43	24	8
IZ Eleven	104	135	162	164	130	109	152	161	178	201	177	146	100	110	85	52	33	13
IZ Twelve	173	167	235	240	163	179	187	247	205	202	206	174	137	110	98	62	38	19
IZ Thirteen	176	219	241	216	172	133	184	193	237	188	165	141	126	115	86	61	38	13
IZ Fourteen	201	191	235	208	158	175	218	224	222	203	212	180	110	78	45	34	11	3
IZ Fifteen	138	198	182	189	126	141	207	208	221	202	200	124	116	88	78	58	40	17
IZ Sixteen	104	119	133	147	128	110	162	175	148	126	135	117	91	85	59	34	25	14

Note: The IG Codes from S02001184 to S02001199 represent IG areas in the West Dunbartonshire region. Since West Dunbartonshire did not provide IG names, they are just labelled as ‘IZ One’.... ‘IZ Sixteen’ for the area covered in this study

Table A.3 – Population distribution of females in Intermediate Geography areas across Greater Glasgow

Intermediate Geography Name	Age																	
	0-4	5-9	10-14	15-19	20-24	25-29	30-34	35-39	40-44	45-49	50-54	55-59	60-64	65-69	70-74	75-79	80-84	85+
Lomond Shore	66	83	83	79	51	60	92	118	112	101	134	110	85	78	82	61	44	32
Auchinairn	113	144	137	146	126	159	209	213	183	142	150	128	132	112	126	101	79	61
Woodhill East	66	110	132	108	66	61	107	131	151	118	100	72	59	29	17	20	8	3
Woodhill West	89	131	180	171	119	103	136	186	211	192	180	128	177	160	70	55	31	20
Westerton East	89	114	94	86	68	54	88	138	146	138	117	97	72	98	85	68	51	49
Bishopbriggs West and Cadder	131	178	189	211	141	131	171	250	248	236	183	183	215	187	160	115	46	51
Westerton West	77	110	76	110	77	67	111	109	133	137	106	108	84	71	61	36	18	15
Kessington East	66	81	68	104	66	44	72	105	131	130	125	102	107	76	84	53	35	28
Bishopbriggs North and Kenmure	138	177	222	163	145	98	170	249	241	184	203	193	230	226	181	98	49	32
Kessington West	76	81	73	105	84	42	70	102	116	117	169	119	119	107	96	106	60	74
Lenzie North	108	165	181	190	130	110	147	224	237	207	247	219	164	169	135	131	66	56
South Castlehill and Thorn	126	169	193	142	106	84	114	194	216	160	176	152	128	99	94	86	33	21
Kilmardinny West	69	78	136	92	58	43	65	108	146	121	129	122	128	109	99	114	108	122
Kilmardinny East	103	92	88	69	63	49	86	147	115	120	130	122	114	85	78	53	42	35
Rosebank and Waterside	108	121	138	138	89	104	137	154	162	131	139	118	105	76	57	31	20	14
Kirkintilloch South	75	85	99	104	75	97	112	162	140	126	99	94	109	80	77	60	38	24
North Castlehill and Thorn	152	172	182	217	155	137	191	222	199	185	176	136	100	74	42	43	27	57
Torrance and Balmore	71	129	127	101	86	48	98	135	162	139	132	72	68	57	51	42	24	15
Keystone and Dougalston	70	102	98	121	89	56	86	124	135	170	184	154	139	138	153	148	111	103
Harestanes	102	86	94	103	90	104	137	149	137	83	113	115	124	130	108	72	45	35
Kirkintilloch West	112	105	80	85	134	167	186	161	147	109	138	129	116	105	102	104	74	68
Barloch	88	97	65	82	59	68	65	118	121	115	119	109	118	113	116	103	66	42
Twechar and Harestanes East	86	85	107	118	98	95	131	133	131	133	131	110	63	58	65	43	18	22
East Clober and Mains Estate	93	140	152	130	84	85	132	153	166	139	133	100	90	104	80	52	36	15
West Clober and Mains Estate	105	128	130	81	63	69	122	153	173	117	102	81	62	44	32	27	11	7

Intermediate Geography Name	Age																	
	0-4	**5-9**	**10-14**	**15-19**	**20-24**	**25-29**	**30-34**	**35-39**	**40-44**	**45-49**	**50-54**	**55-59**	**60-64**	**65-69**	**70-74**	**75-79**	**80-84**	**85+**
Lennoxtown	127	168	131	147	123	147	200	214	188	136	149	100	96	84	70	85	49	56
Eaglesham and Waterfoot	144	158	161	125	91	127	174	203	196	171	173	161	163	118	105	97	60	38
Mearnskirk and South Kirkhill	161	161	181	160	127	91	158	227	202	214	214	196	183	187	172	122	106	88
North Kirkhill	79	107	104	80	72	98	131	159	137	132	122	92	75	81	64	55	23	19
Mearns Village, Westacres and Greenfarm	244	227	204	168	123	157	296	313	267	196	167	110	79	74	64	62	39	49
West Neilston and Uplawmoor	170	212	200	236	146	125	180	275	263	261	235	160	140	129	96	81	44	42
Busby	88	82	100	75	67	132	156	153	141	122	128	88	107	98	116	93	58	68
Whitecraigs and Broom	73	97	127	128	81	38	80	105	161	138	181	155	101	106	89	57	33	22
Clarkston and Sheddens	160	215	237	192	150	122	193	274	282	241	234	171	162	147	107	78	69	48
Crookfur and Fruin	156	208	205	198	141	119	199	241	269	229	213	168	119	106	108	124	107	142
Williamwood	95	99	86	88	83	51	98	130	148	111	107	100	78	69	60	39	20	27
Stamperland	130	144	115	99	61	104	162	201	183	141	117	115	85	88	70	41	27	28
West Arthurlie and North Neilston	150	166	144	153	151	123	195	216	202	170	190	161	175	150	130	103	51	56
Auchenback	120	126	172	141	109	100	149	167	168	140	139	117	143	110	76	62	33	20
Lower Whitecraigs and South Giffnock	82	115	135	131	61	48	84	120	147	159	135	116	91	73	83	64	65	65
South Thornliebank and Woodfarm	108	143	160	151	114	98	120	208	202	153	144	97	117	118	92	71	36	40
Netherlee	124	153	153	128	84	50	114	209	185	191	161	142	114	130	170	129	96	102
Merrylee and Braidbar	117	165	192	174	137	68	122	216	201	214	196	170	140	150	151	123	87	63
North Giffnock and North Thornliebank	86	115	91	103	64	79	124	148	147	125	107	77	97	118	119	133	117	132
Dunterlie, East Arthurlie and Dovecothall	211	182	181	153	196	207	281	273	209	160	167	149	147	144	139	107	90	59
Cross Stobbs	88	117	123	114	83	84	133	157	131	131	146	101	88	90	71	53	33	46
Carmunnock South	80	126	105	107	99	77	103	140	134	117	103	93	111	88	99	84	65	68
Glenwood South	176	147	193	217	173	206	225	243	194	171	145	130	147	132	112	82	45	28
Darnley West	120	140	116	103	100	133	183	203	174	107	63	53	51	36	28	20	13	19

Intermediate Geography Name	Age																	
	0-4	5-9	10-14	15-19	20-24	25-29	30-34	35-39	40-44	45-49	50-54	55-59	60-64	65-69	70-74	75-79	80-84	85+
Glenwood North	146	192	158	180	137	152	216	173	145	103	113	95	72	60	47	39	20	15
Castlemilk	149	166	180	213	165	192	226	255	210	201	188	135	155	156	147	128	90	61
Darnley North	108	89	95	98	100	129	139	154	117	110	56	44	34	59	51	39	38	47
Carmunnock North	65	95	96	114	94	74	95	110	140	151	145	107	77	70	71	44	28	19
Carnwadric West	178	166	177	165	150	181	216	230	175	133	135	134	136	101	105	55	30	23
Muirend and Old Cathcart	131	122	142	136	131	177	171	209	213	184	158	113	133	138	144	147	93	89
Carnwadric East	69	94	115	96	87	84	90	144	93	123	118	114	111	118	88	99	82	60
Kingspark South	163	167	185	134	133	191	224	228	204	150	112	86	92	67	75	56	32	40
Newlands	138	170	227	194	171	124	159	211	234	240	192	165	151	145	125	94	77	55
Nitshill	163	176	232	237	158	129	236	234	221	160	165	146	161	167	79	56	35	70
Merrylee and Millbrae	90	88	102	104	71	81	136	140	147	142	118	97	97	98	99	91	73	84
Cathcart	94	96	109	101	186	272	254	194	161	139	128	92	84	115	133	142	93	68
Kingspark North	159	170	168	156	143	160	207	190	249	182	149	127	88	93	75	69	30	42
Crookston South	130	167	192	215	146	150	164	224	201	119	138	117	126	142	90	61	23	26
Pollokshaws	119	129	109	124	170	203	162	184	168	131	112	130	137	186	173	140	118	139
Mount Florida	100	91	104	131	162	224	228	182	145	104	117	90	94	119	92	90	65	54
Langside	80	67	56	56	139	319	256	147	146	103	84	52	70	76	90	85	73	84
Crookston North	85	94	93	79	78	76	117	151	118	108	99	86	124	161	125	78	48	39
Battlefield	117	101	97	123	197	391	319	251	168	137	120	94	86	99	102	120	80	111
Pollok South and West	172	148	129	143	114	175	208	278	248	171	141	119	119	115	68	67	32	21
Shawlands West	63	46	40	57	160	309	235	153	112	85	84	66	69	87	79	69	66	59
Toryglen and Oatlands	156	185	228	188	180	154	199	271	233	179	167	164	229	224	186	157	124	88
Shettleston South	98	151	123	121	108	96	142	192	214	152	124	100	104	120	112	87	59	40
Maxwell Park	160	171	175	191	192	154	204	228	243	236	223	136	157	146	129	105	78	115
Carmyle and Mount Vernon South	55	63	106	94	86	50	112	140	116	126	88	109	98	108	83	63	36	52
Strathbungo	184	122	100	146	267	445	367	255	204	131	117	77	72	79	63	57	39	57
Govanhill East and Aikenhead	95	89	114	168	125	175	168	158	170	128	102	90	103	111	98	103	73	67
Govanhill West	149	137	114	119	242	343	302	210	172	137	103	89	95	87	82	90	42	51

Intermediate Geography Name	Age																	
	0-4	5-9	10-14	15-19	20-24	25-29	30-34	35-39	40-44	45-49	50-54	55-59	60-64	65-69	70-74	75-79	80-84	85+
Baillieston East	121	123	140	124	103	166	186	187	168	139	147	109	114	112	102	83	53	57
Dalmarnock	99	102	141	128	113	137	144	153	130	123	122	121	124	140	113	100	59	43
Braidfauld	163	177	212	160	148	182	229	232	187	151	138	115	116	163	141	115	72	70
Mosspark	125	127	143	124	115	115	162	185	189	151	133	151	166	212	231	206	127	100
Pollokshields East	177	173	156	164	205	202	236	176	156	91	69	39	43	50	60	38	29	31
Mount Vernon North and Sandyhills	97	109	126	130	84	82	116	156	194	173	154	150	122	116	93	75	38	37
Pollokshields West	179	166	141	155	167	151	192	168	167	147	108	111	100	84	77	69	49	65
Cardonald South and East	121	117	97	98	124	103	151	132	148	100	101	62	78	104	89	102	68	87
Baillieston West	90	116	138	134	122	95	173	162	170	140	178	171	154	176	114	81	36	30
Cardonald West and Central	126	154	187	149	143	142	178	238	205	222	165	178	200	217	197	147	110	125
Gorbals and Hutchesontown	110	108	127	128	190	207	200	207	158	124	126	128	161	176	150	87	73	51
Parkhead East and Braidfauld North	99	126	118	112	118	144	188	164	132	120	108	92	78	76	75	65	34	32
Tollcross	127	131	141	111	138	144	179	183	151	131	107	90	88	110	90	69	58	46
Kingston West and Dumbreck	88	111	68	91	138	140	143	127	132	99	114	91	93	109	132	100	81	111
Shettleston North	98	108	115	95	135	183	200	189	143	141	102	132	150	155	177	149	115	90
Parkhead West and Barrowfield	135	152	154	177	163	177	215	209	202	173	153	135	127	166	129	94	59	70
Laurieston and Tradeston	71	69	71	109	194	218	161	124	97	87	90	91	97	70	71	68	41	36
Calton, Galllowgate and Bridgeton	92	74	89	151	259	284	210	196	142	140	120	116	169	148	124	109	60	54
Garrowhill West	145	154	147	140	112	87	163	275	203	163	142	121	137	110	95	72	62	40
Cardonald North	83	109	115	130	101	108	137	165	166	167	126	130	131	147	158	146	81	60
Craigton	93	119	123	133	90	116	150	173	156	107	120	123	158	209	212	165	114	95
Hillington	114	116	134	123	128	116	173	174	175	115	98	101	78	108	121	76	74	49
Greenfield	133	145	145	146	151	159	187	251	185	140	217	185	234	200	193	128	81	88
Penilee	145	189	213	206	148	149	188	239	214	167	148	181	156	201	178	146	88	90
Kinning Park and Festival Park	99	110	71	99	250	266	174	142	127	105	59	71	69	68	62	48	40	68
Old Shettleston and Parkhead	76	107	100	97	70	103	127	115	115	87	102	123	148	147	148	127	77	74

Intermediate Geography Name	Age																	
	0-4	**5-9**	**10-14**	**15-19**	**20-24**	**25-29**	**30-34**	**35-39**	**40-44**	**45-49**	**50-54**	**55-59**	**60-64**	**65-69**	**70-74**	**75-79**	**80-84**	**85+**
Garrowhill East and Swinton	165	198	184	128	106	147	207	316	212	162	121	79	59	62	56	41	22	21
Barlanark	162	194	214	199	163	184	273	249	205	140	144	105	109	85	72	67	39	36
Gallowgate North and Bellgrove	70	60	84	128	301	215	159	107	107	95	90	72	97	96	85	63	40	29
Carntyne West and Haghill	159	163	173	168	202	227	260	258	224	185	146	95	134	127	99	58	29	44
Ibrox	119	96	107	114	113	115	118	131	99	82	73	59	76	75	62	49	45	47
Dennistoun	77	69	98	128	272	267	190	172	151	93	103	75	74	86	67	56	45	30
Anderston	62	43	47	76	251	205	142	87	82	65	76	60	37	53	40	48	23	15
North Barlanark and Easterhouse South	172	169	171	159	147	149	174	181	159	109	89	64	87	70	53	38	24	21
Easterhouse East	108	139	141	127	121	145	140	153	159	126	96	51	67	73	72	54	28	31
Cranhill, Lightburn and Queenslie South	154	209	210	225	150	172	247	281	232	164	167	178	205	224	152	127	92	60
City Centre West	44	41	37	232	499	238	143	110	70	77	46	45	54	60	50	53	31	24
Govan and Linthouse	165	155	149	178	231	236	217	194	195	171	184	147	127	134	120	86	56	76
Carntyne	46	64	58	70	56	61	84	108	121	105	115	120	147	194	187	174	117	121
City Centre East	75	86	69	728	770	353	194	119	118	98	150	171	166	217	200	203	130	117
Dennistoun North and Alexandra Parade	124	99	114	165	290	263	277	193	198	190	161	145	182	156	154	131	107	140
Drumoyne and Shieldhall	152	156	193	209	186	211	231	260	252	190	174	163	164	208	186	184	121	70
Finnieston and Kelvinhaugh	73	71	67	295	950	467	281	182	141	106	65	66	77	64	66	49	38	29
Central Easterhouse	123	166	157	149	124	135	161	152	141	136	90	83	85	81	83	34	9	16
Craigend and Ruchazie	139	180	150	189	149	198	235	221	187	173	224	193	174	141	105	96	85	66
Hillhead	103	77	61	174	541	407	292	218	171	141	127	117	90	101	75	68	53	50
Woodlands	122	78	108	272	996	356	276	189	139	103	81	53	67	60	52	39	30	28
Roystonhill, Blochairn, and Provanmill	192	201	177	174	191	239	281	229	200	131	137	132	144	135	142	102	67	54
Riddrie and Hogganfield	108	158	201	191	131	123	168	236	209	158	163	148	153	204	186	159	118	93

Intermediate Geography Name	Age																	
	0-4	5-9	10-14	15-19	20-24	25-29	30-34	35-39	40-44	45-49	50-54	55-59	60-64	65-69	70-74	75-79	80-84	85+
Woodside	63	39	43	84	250	179	123	100	72	76	61	76	89	105	117	110	73	83
Sighthill	128	130	132	125	144	219	208	178	121	79	78	80	81	63	64	41	46	43
Partick	66	56	52	67	252	239	182	131	118	107	80	73	91	99	98	92	88	62
Kelvingrove and University	68	56	85	226	884	312	230	164	145	119	109	71	58	55	64	60	51	40
Whiteinch	90	99	73	89	130	172	176	166	119	87	94	54	73	72	69	62	48	38
Garthamlock, Auchinlea and Gartloch	140	175	164	130	104	153	196	205	147	123	89	68	57	48	41	20	9	10
Firhill	122	108	115	180	530	388	295	211	174	149	120	107	98	108	92	89	65	44
Cowlairs and Port Dundas	122	117	115	124	157	198	199	171	149	101	90	82	103	72	68	66	27	55
Partickhill and Hyndland	89	51	49	92	265	231	200	188	141	118	90	79	75	65	103	72	57	68
Blackhill and Barmulloch East	104	128	116	152	95	96	155	178	166	115	134	87	141	106	95	63	34	35
Broomhill	89	80	77	83	117	174	185	174	164	144	151	104	92	114	131	115	114	76
Petershill	151	171	150	164	179	240	239	243	165	111	104	92	78	75	66	42	30	20
Dowanhill	80	75	57	171	338	239	213	188	168	141	117	89	54	57	51	62	49	56
Victoria Park	82	88	98	63	52	48	93	105	138	114	76	60	60	66	77	64	62	58
North Kelvin	71	58	77	140	345	249	219	176	133	121	92	88	59	51	40	25	33	37
Scotstoun South and West	109	162	159	139	182	203	193	172	154	118	87	91	79	65	66	68	49	44
Keppochhill	158	144	171	163	177	205	233	203	182	166	151	135	135	111	101	74	46	47
Scotstoun North and East	152	142	137	131	95	110	161	177	175	145	128	88	108	94	116	94	84	47
Barmulloch	126	116	139	126	167	106	176	166	144	115	106	106	135	124	90	58	45	34
Ruchill	91	91	93	537	498	375	268	216	169	118	126	96	97	92	101	86	70	86
Kelvinside and Jordanhill	148	152	155	229	234	239	236	233	306	211	177	182	114	144	170	156	104	113
Kelvindale	121	142	150	185	260	219	234	265	232	224	229	132	107	122	125	104	93	140
Springburn	108	106	131	128	107	131	201	201	149	101	147	127	196	195	163	122	99	110
Possil Park	122	138	133	180	143	151	150	188	172	167	148	157	195	216	192	169	124	102
Wyndford	118	102	91	123	169	159	148	164	121	98	130	98	132	159	145	100	79	78
Springburn East and Cowlairs	121	147	121	134	172	162	252	211	151	93	86	86	121	155	130	73	41	37
Robroyston and Millerston	269	200	154	133	146	300	421	355	287	165	98	73	59	44	29	28	24	51
Yoker South	128	106	94	118	150	205	210	175	143	113	79	61	67	67	54	35	36	17

Intermediate Geography Name	Age																	
	0-4	5-9	10-14	15-19	20-24	25-29	30-34	35-39	40-44	45-49	50-54	55-59	60-64	65-69	70-74	75-79	80-84	85+
Knightswood Park East	83	106	92	89	83	109	112	108	123	96	68	80	85	112	110	99	73	48
Knightswood East	90	102	107	88	100	112	137	164	115	100	104	117	121	166	169	128	104	99
Maryhill West	108	111	111	115	101	106	153	163	108	89	69	78	82	78	52	43	33	24
Milton West	90	100	136	144	84	77	143	173	149	129	138	138	175	162	150	121	65	44
Anniesland East	98	95	80	122	183	196	193	179	158	115	101	90	75	72	81	92	60	53
Knightswood West	74	76	87	77	57	68	137	120	94	89	66	88	109	107	86	69	39	32
Maryhill East	96	109	108	127	132	99	127	163	120	95	89	110	122	92	85	82	59	36
Anniesland West	112	124	138	130	200	212	228	226	215	155	153	134	170	196	214	197	130	152
Knightswood Park West	65	82	62	58	70	92	93	112	95	101	113	106	131	146	175	178	129	100
Yoker North	91	117	103	94	69	65	118	122	104	118	113	94	93	96	75	108	83	103
Milton East	123	154	169	139	139	137	142	177	161	119	133	119	139	121	92	86	64	57
Blairdardie East	132	131	128	141	129	149	206	247	209	159	164	149	215	235	240	215	124	97
Summerston North	114	122	115	171	135	128	170	162	168	135	164	145	114	135	88	75	45	49
Blairdardie West	50	70	95	82	58	41	80	142	113	106	95	64	73	102	91	100	74	53
Drumchapel South	127	147	169	152	106	130	168	174	170	102	108	86	82	77	58	68	45	58
Summerston Central and West	129	169	134	119	127	139	208	243	171	132	116	111	107	99	50	43	30	30
Drumry West	111	122	165	164	127	133	158	166	146	126	82	95	72	69	49	51	47	11
Drumchapel North	121	141	160	170	107	138	164	143	130	100	78	71	41	40	42	29	18	20
Drumry East	134	133	129	146	112	99	125	130	125	86	93	71	69	63	54	81	50	48
Forgewood	196	179	201	179	190	251	306	244	184	150	155	130	141	111	100	103	64	50
Birkenshaw	172	192	183	191	115	149	211	287	245	219	156	138	126	97	64	46	28	18
Townhead	156	197	230	217	165	167	214	257	237	160	199	153	168	151	119	104	60	41
Glenmavis and Greengairs	142	143	111	118	118	120	158	176	161	156	160	130	113	81	65	58	27	28
Stepps	119	146	180	148	143	112	165	204	205	210	213	182	157	128	99	97	72	32
Gartcosh and Marnock	100	95	101	92	94	99	161	152	122	151	118	123	91	87	67	47	28	26
Chryston and Muirhead	81	132	144	133	107	119	155	188	160	137	160	127	150	104	87	78	44	40
Moodiesburn West	120	100	120	125	139	135	162	151	136	99	116	108	106	83	76	58	41	43
Kilsyth Bogside	98	99	85	85	95	105	146	116	110	90	118	99	66	73	54	55	45	20

Intermediate Geography Name	Age																	
	0-4	**5-9**	**10-14**	**15-19**	**20-24**	**25-29**	**30-34**	**35-39**	**40-44**	**45-49**	**50-54**	**55-59**	**60-64**	**65-69**	**70-74**	**75-79**	**80-84**	**85+**
Kilsyth East and Croy	134	142	133	130	146	154	179	185	155	165	185	152	140	137	130	123	85	81
Renfrewshire Rural South & Howwood	93	78	81	62	64	76	165	133	118	97	117	107	78	51	42	38	28	41
Paisley Glenburn West	124	129	159	138	107	107	170	217	168	143	145	135	117	110	106	90	78	91
Paisley Glenburn East	149	124	148	125	95	133	175	174	160	124	136	102	105	71	80	61	42	17
Paisley Foxbar	144	158	185	183	151	164	204	232	235	159	199	169	187	198	163	98	48	28
Johnstone South West	151	183	210	209	150	155	232	236	196	173	186	177	197	172	118	69	33	28
Paisley Dykebar	83	102	74	79	85	96	140	136	128	90	116	119	108	73	82	58	63	78
Paisley South West	131	148	171	170	165	143	196	219	229	233	236	199	141	111	83	71	42	74
Paisley South East	98	112	124	119	150	189	186	232	186	162	187	168	218	249	255	228	144	133
Johnstone South East	118	154	148	117	148	122	177	211	166	158	119	133	145	200	136	101	66	40
Paisley South	111	112	91	87	70	111	165	190	138	134	121	76	70	69	52	55	34	59
Johnstone North West	118	94	112	119	91	112	149	159	126	101	132	101	121	110	89	71	43	51
Paisley West	199	190	163	149	192	211	292	300	263	149	134	122	96	108	82	60	51	29
Paisley East	106	128	166	169	130	114	186	170	142	128	148	129	107	98	96	70	67	49
Johnstone North East	89	90	79	66	77	109	139	128	106	89	115	105	86	88	90	89	61	83
Elderslie and Phoenix	114	141	180	150	111	112	180	219	246	209	223	168	158	168	134	108	80	69
Paisley Central	121	112	89	166	330	347	272	244	197	172	149	133	157	149	151	147	106	117
Paisley North West	75	87	73	103	157	127	171	120	110	81	95	80	121	96	101	87	69	66
Paisley Ralston	122	149	142	142	123	65	109	219	220	197	204	168	182	173	171	122	69	51
Paisley Ferguslie	202	207	232	189	180	216	266	253	186	121	95	92	86	65	47	29	17	12
Paisley North East	125	141	125	134	171	244	244	223	173	208	186	120	140	188	180	163	129	135
Linwood South	112	131	118	156	136	140	162	188	129	120	132	147	181	180	143	89	44	46
Linwood North	141	167	198	163	121	119	213	233	196	116	155	155	191	131	74	32	15	14
Paisley Gallowhill and Hillington	147	177	184	159	145	194	231	233	182	140	207	201	163	165	167	148	87	98
Paisley North	109	74	83	97	155	211	201	138	139	97	117	118	132	108	72	69	44	62
Renfrew South	114	124	132	134	143	154	172	198	169	193	222	203	200	211	208	139	75	60
Renfrew East	159	166	191	182	186	134	235	263	270	276	275	175	147	91	112	116	63	69
Renfrew West	192	185	193	207	188	235	257	294	222	200	204	186	147	116	93	63	55	31

Intermediate Geography Name	Age																	
	0-4	5-9	10-14	15-19	20-24	25-29	30-34	35-39	40-44	45-49	50-54	55-59	60-64	65-69	70-74	75-79	80-84	85+
Renfrew North	65	65	65	56	101	123	127	129	86	116	98	75	80	74	58	80	46	35
Erskine East and Inchinnan	176	217	239	186	143	129	272	293	293	229	181	150	73	98	73	53	34	40
Erskine Central	156	155	171	177	170	164	225	203	227	211	231	177	128	107	78	66	75	29
Chapelton, Glengavel and Sandford	64	83	96	85	52	52	93	104	132	126	121	107	86	56	35	22	24	17
Ashgill and Netherburn	66	75	107	78	61	70	103	116	99	107	84	78	64	55	52	32	34	32
Glassford, Quarter and Allanton	76	90	88	79	68	75	109	122	121	107	102	96	91	64	48	38	23	29
Crosshouse and Lindsayfield	171	164	167	166	138	141	209	225	181	151	153	93	66	49	28	13	15	5
Whitehills West	98	159	166	192	138	130	164	167	185	153	149	113	88	86	69	68	33	18
Greenhills	127	163	148	180	133	130	151	163	137	122	161	126	62	50	34	29	22	14
Mossneuk and Newlandsmuir	73	115	162	146	109	86	120	188	180	129	120	69	57	42	21	21	4	0
Little Earnock	87	73	98	105	105	160	125	128	113	93	118	95	87	52	40	29	14	6
Birniehill, Kelvin and Whitehills East	124	138	176	180	111	120	188	226	187	130	117	114	107	111	100	65	29	24
Nerston and EK Landward Area	70	84	70	103	76	69	107	112	138	115	109	116	104	77	72	60	38	83
Westwood South	158	184	150	133	150	177	206	230	179	144	129	133	135	140	122	53	24	17
The Murray	101	96	89	84	100	105	156	149	152	126	94	108	117	158	194	135	73	55
Hairmyres and Westwood West	89	110	99	94	100	102	113	146	120	125	91	87	99	87	72	53	31	42
St Leonards South	132	141	157	163	138	133	185	217	211	166	160	200	160	128	105	86	41	28
Westwood East	80	101	87	97	89	78	138	155	145	101	85	101	130	141	146	116	53	35
Thorntonhall, Jackton and Gardenhall	82	91	101	95	59	64	93	132	129	130	120	71	48	37	37	11	7	7
West Mains	73	99	92	85	71	72	137	150	106	103	66	70	87	108	132	85	36	24
St Leonards North	119	124	139	119	122	118	198	193	185	145	146	172	190	173	124	77	67	58
East Mains	65	88	77	77	75	90	105	109	128	86	101	99	105	109	129	116	67	39
Stewartfield West	182	183	125	100	69	102	208	276	214	149	120	98	75	71	55	44	18	38
Calderwood West and Nerston	85	105	94	101	95	93	128	156	156	123	100	91	126	150	152	109	57	37
Stewartfield East	117	128	112	74	59	58	106	191	173	114	89	65	54	34	33	32	4	6
Calderwood Central	92	92	105	78	107	137	178	168	158	103	106	112	160	180	177	99	69	61

Intermediate Geography Name	Age																	
	0-4	5-9	10-14	15-19	20-24	25-29	30-34	35-39	40-44	45-49	50-54	55-59	60-64	65-69	70-74	75-79	80-84	85+
Hamilton Centre and Low Parks	52	78	66	93	118	135	133	128	118	127	132	108	114	97	96	98	85	74
Bothwell South	72	80	94	83	72	76	96	98	136	130	119	93	74	90	67	76	57	50
Fernhill and Cathkin	148	149	160	148	163	180	185	172	176	158	151	106	159	102	116	90	57	30
Whitlawburn and Greenlees	121	126	140	168	168	157	134	171	187	176	154	99	49	54	33	29	22	14
Vicarland and Cairns	146	124	157	165	166	139	154	184	174	171	182	144	141	131	139	135	63	64
Burnside and Springhall	120	144	130	127	135	117	150	186	212	181	171	133	163	197	150	128	77	76
Low Blantyre and Bardykes	105	104	109	110	99	97	122	128	142	140	134	92	70	63	39	25	16	14
Spittal	74	90	111	106	71	81	93	127	122	105	104	105	120	114	102	93	51	32
Cambuslang Central	104	137	123	130	108	87	115	175	161	164	143	120	111	104	96	91	59	41
High Crosshill	75	118	94	95	84	71	94	136	156	152	113	106	95	100	105	77	76	50
Halfway, Hallside and Drumsagard	158	149	119	127	105	227	228	217	174	119	107	73	71	50	44	31	18	21
Westburn and Newton	87	90	117	95	81	153	150	148	128	84	104	96	78	54	52	46	22	14
Bankhead South	107	102	113	126	97	116	152	166	167	144	130	101	92	78	52	67	38	35
Burgh, Eastfield and Silverbank	114	142	157	171	134	133	178	214	248	247	247	169	157	178	174	122	90	67
Burnhill and Bankhead North	87	91	106	126	96	79	153	170	150	121	134	116	138	101	153	102	64	52
Shawfield and Clincarthill	94	92	90	92	125	156	132	144	139	119	127	93	97	111	109	117	89	109
Farme Cross and Gallowflat North	74	57	57	65	97	116	123	129	113	89	91	86	93	100	135	104	90	82
Blane Valley	129	154	161	134	76	58	114	181	198	182	243	190	139	101	111	83	42	47
IZ One	159	138	157	141	135	182	179	221	184	153	140	119	126	119	107	87	78	56
IZ Two	135	143	127	134	175	205	220	181	129	112	94	96	79	94	99	112	83	109
IZ Three	173	174	209	215	192	203	228	223	232	190	211	229	179	182	136	123	90	66
IZ Four	100	111	121	125	101	119	169	145	152	138	170	174	158	211	217	163	118	73
IZ Five	140	157	173	195	196	181	188	214	223	173	205	148	189	186	209	190	132	98
IZ Six	94	105	74	110	97	125	143	150	134	126	135	122	102	124	123	92	49	39
IZ Seven	208	194	207	158	166	196	273	314	282	200	163	129	107	105	85	78	66	53
IZ Eight	156	181	211	230	189	197	238	232	208	172	169	107	118	92	114	130	74	34
IZ Nine	136	139	161	148	184	216	279	251	246	189	175	147	145	153	185	165	150	119

Intermediate Geography Name	Age																	
	0-4	**5-9**	**10-14**	**15-19**	**20-24**	**25-29**	**30-34**	**35-39**	**40-44**	**45-49**	**50-54**	**55-59**	**60-64**	**65-69**	**70-74**	**75-79**	**80-84**	**85+**
IZ Eleven	95	95	150	163	103	89	154	177	199	184	173	141	117	137	125	72	57	38
IZ Twelve	161	175	197	228	163	196	228	244	215	200	227	168	150	143	150	115	89	65
IZ Thirteen	146	174	213	208	178	148	204	234	250	159	150	156	165	151	125	84	58	41
IZ Fourteen	159	207	240	238	189	174	236	272	244	209	245	157	144	72	62	44	20	14
IZ Fifteen	167	158	182	179	144	144	215	243	230	225	183	128	132	122	102	79	77	55
IZ Sixteen	117	152	158	154	83	118	164	214	158	159	146	108	118	117	93	48	61	56

Note: The IG Codes from S02001184 to S02001199 represent IG areas in the West Dunbartonshire region. Since West Dunbartonshire did not provide IG names, they are just labelled as 'IZ One'.... 'IZ Sixteen' for the area covered in this study

Table A.4 – Geometric means of soil metal concentration in each Intermediate Geography area

Intermediate Geography area	**As**	**Cr**	**Cu**	**Ni**	**Pb**	**Se**	**Zn**	**K_2O**
	mg kg^{-1}	mg kg^{-1}	mg kg^{-1}	mg kg^{-1}	mg kg^{-1}	mg kg^{-1}	mg kg^{-1}	wt %
Lomond Shore	10.65	65.55	19.67	16.75	60.41	0.36	52.43	2.01
Auchinairn	6.82	98.71	40.02	48.71	103.34	1.10	138.51	1.46
Woodhill East	8.34	102.93	40.68	32.95	109.21	1.04	129.13	1.26
Woodhill West	8.73	96.96	64.04	47.43	149.91	0.78	160.35	1.19
Westerton East	7.49	87.01	35.64	39.13	90.54	0.47	100.06	1.44
Bishopbriggs West and Cadder	6.20	78.76	34.09	30.99	77.87	0.55	95.91	1.29
Westerton West	6.35	80.25	35.87	36.63	111.28	0.59	102.46	1.39
Kessington East	7.46	82.43	32.10	36.78	102.09	0.64	107.74	1.29
Bishopbriggs North and Kenmure	7.09	82.37	32.65	32.72	89.63	0.72	96.70	1.23
Kessington West	6.75	80.03	28.76	28.82	91.20	0.76	94.87	1.22
Lenzie North	9.04	78.05	41.07	40.02	87.05	0.79	105.50	1.18
South Castlehill and Thorn	7.60	90.56	37.40	33.15	131.04	0.63	122.75	1.32
Kilmardinny West	9.81	106.01	57.61	53.80	151.09	0.98	169.57	1.32
Kilmardinny East	9.27	154.59	76.93	61.16	391.53	1.03	230.93	1.40
Rosebank and Waterside	6.91	93.18	24.97	28.25	70.41	0.82	80.62	1.15
Kirkintilloch South	7.20	91.00	21.50	35.70	60.30	0.90	72.50	1.08
North Castlehill and Thorn	8.21	106.81	38.30	48.27	80.46	0.94	110.36	1.27
Torrance and Balmore	8.98	97.94	30.70	36.26	60.80	1.04	115.26	1.29
Keystone and Dougalston	8.11	87.29	27.95	30.76	72.40	0.77	88.16	1.10
Harestanes	8.10	92.00	123.00	164.30	256.50	1.10	245.10	1.06
Kirkintilloch West	11.86	97.75	28.01	45.54	53.03	0.62	97.16	1.42
Barloch	7.89	115.62	39.10	54.88	96.04	0.96	128.78	1.17
Twechar and Harestanes East	8.63	109.64	30.60	31.35	65.94	1.01	100.15	1.27
East Clober and Mains Estate	8.17	103.81	40.15	48.26	95.98	1.10	129.20	1.14
West Clober and Mains Estate	7.73	113.62	29.99	43.69	57.41	0.88	95.49	1.15
Milton of Campsie	8.99	77.83	25.36	29.58	66.15	0.79	94.28	1.12
Lennoxtown	10.31	82.13	27.81	29.08	74.70	1.26	91.44	0.98
Eaglesham and Waterfoot	6.34	131.04	25.92	51.92	38.84	0.70	117.70	1.50
Mearnskirk and South Kirkhill	5.60	72.79	35.69	38.91	72.04	0.88	150.27	1.61
North Kirkhill	4.89	104.87	40.18	50.15	73.57	0.53	140.09	1.55
Mearns Village, Westacres and Greenfarm	5.05	68.26	23.09	28.25	46.06	0.57	112.05	1.58
West Neilston and Uplawmoor	11.90	134.40	54.46	58.18	122.93	1.21	190.35	1.27
Busby	13.12	103.32	65.14	65.04	149.35	0.84	160.35	1.44
Whitecraigs and Broom	6.73	91.89	42.32	45.65	137.25	0.99	202.07	1.47
Clarkston and Sheddens	6.68	99.27	45.09	44.03	109.55	0.65	141.05	1.24
Crookfur and Fruin	7.16	92.48	34.08	36.87	97.04	0.80	132.47	1.31
Williamwood	8.10	117.69	45.82	41.34	122.45	0.84	154.74	1.17
Stamperland	10.60	106.23	60.96	52.94	141.68	0.96	181.39	1.36
West Arthurlie and North Neilston	12.99	118.38	76.62	54.69	163.81	1.04	197.27	1.22
Auchenback	9.42	118.29	51.00	47.33	119.86	0.75	153.91	1.19
Lower Whitecraigs and South Giffnock	7.15	110.58	36.96	36.15	140.40	0.78	115.15	1.17
South Thornliebank and Woodfarm	8.84	106.18	56.12	47.74	110.76	0.82	191.98	1.41
Netherlee	9.27	166.44	73.67	59.71	125.17	0.80	173.65	1.43
Merrylee and Braidbar	8.29	116.09	55.00	48.22	105.13	0.65	137.71	1.26

Intermediate Geography area	As mg kg^{-1}	Cr mg kg^{-1}	Cu mg kg^{-1}	Ni mg kg^{-1}	Pb mg kg^{-1}	Se mg kg^{-1}	Zn mg kg^{-1}	K_2O wt %
North Giffnock and North Thornliebank	8.70	101.16	61.41	44.06	194.47	0.72	218.43	1.26
Dunterlie, East Arthurlie and Dovecothall	11.17	119.31	63.75	52.22	138.73	1.06	166.38	1.13
Cross Stobbs	13.04	114.30	131.25	63.43	213.35	1.12	301.07	1.14
Carmunnock South	5.94	114.37	35.78	44.82	77.10	0.84	116.63	1.34
Glenwood South	7.98	123.23	92.76	52.39	117.47	0.86	166.40	1.43
Darnley West	9.98	102.88	41.99	40.34	94.13	0.84	122.34	1.30
Darnley East	10.70	111.00	36.80	43.30	119.50	0.50	94.70	1.51
Glenwood North	7.38	125.94	42.47	48.73	96.90	0.79	121.83	1.26
Castlemilk	7.96	113.92	101.01	45.54	137.71	0.79	186.22	1.33
Darnley North	7.34	100.59	39.47	37.80	93.68	0.86	142.93	1.34
Carmunnock North	7.25	119.36	43.35	46.14	130.63	0.92	125.57	1.35
Carnwadric West	7.64	103.22	44.32	40.57	109.45	0.88	176.17	1.38
Muirend and Old Cathcart	6.65	107.67	71.84	47.37	157.86	0.87	172.12	1.29
Carnwadric East	9.07	233.88	44.63	46.04	123.92	0.76	165.35	1.24
Kingspark South	8.06	116.38	55.89	45.79	193.48	1.06	161.89	1.40
Newlands	9.35	139.62	95.87	56.15	215.58	1.04	208.15	1.26
Nitshill	8.66	102.07	46.77	39.53	109.51	0.74	171.35	1.32
Merrylee and Millbrae	8.50	187.00	54.20	45.40	134.00	1.00	147.00	1.27
Cathcart	10.12	209.76	139.72	73.28	244.61	1.20	230.22	1.38
Kingspark North	8.70	167.96	55.75	60.49	158.47	0.92	153.79	1.59
Crookston South	9.04	124.17	40.88	49.82	92.71	1.00	131.07	1.40
Pollokshaws	10.18	128.82	48.22	47.86	128.68	1.05	146.56	1.35
Mount Florida	10.24	243.75	51.03	58.79	116.60	0.91	136.11	1.28
Langside	10.04	148.44	69.43	51.63	241.71	1.20	171.16	1.38
Crookston North	11.23	119.85	44.78	48.89	96.59	0.92	158.81	1.50
Battlefield	9.81	184.91	71.70	54.01	274.02	1.17	190.88	1.30
Pollok South and West	11.21	115.25	50.65	50.11	141.44	1.04	161.77	1.33
Shawlands West	8.20	118.00	53.30	49.20	206.60	1.10	169.90	1.21
Toryglen and Oatlands	9.39	162.55	73.62	57.35	159.58	0.91	184.50	1.32
Shettleston South	14.50	175.74	86.46	63.38	238.65	1.25	371.66	1.32
Maxwell Park	9.42	114.23	62.78	46.64	157.99	0.96	141.23	1.35
Carmyle and Mount Vernon South	11.75	159.94	97.53	53.80	237.79	1.03	305.76	1.22
Strathbungo	9.58	161.92	102.59	60.42	364.64	1.35	274.47	1.28
Govanhill East and Aikenhead	8.86	102.06	99.08	47.17	247.63	0.66	243.25	1.38
Govanhill West	14.00	281.00	116.30	61.90	285.10	1.90	250.40	1.34
Pollok North and East	9.31	113.03	66.94	53.21	134.33	1.05	187.73	1.48
Baillieston East	10.26	109.13	60.47	43.23	172.55	1.09	218.23	1.31
Dalmarnock	9.79	144.59	73.93	44.36	160.44	1.08	183.38	1.43
Braidfauld	9.39	156.46	76.64	60.19	171.79	1.12	266.35	1.45
Mosspark	9.20	111.11	61.78	45.68	178.31	1.03	159.60	1.32
Pollokshields East	9.12	110.50	65.93	42.33	152.90	0.94	162.99	1.33
Mount Vernon North and Sandyhills	7.37	114.38	50.57	31.76	111.51	0.63	189.32	1.35
Pollokshields West	8.40	123.37	53.59	43.77	158.86	1.00	149.05	1.29
Cardonald South and East	7.99	110.27	52.20	46.96	151.33	0.90	141.57	1.54
Baillieston West	7.13	125.55	42.87	38.04	113.12	0.93	141.78	1.26
Cardonald West and Central	9.20	115.42	68.19	56.52	180.52	1.23	201.21	1.51
Gorbals and Hutchesontown	11.56	147.08	84.14	76.08	158.55	0.95	221.03	1.37
Parkhead East and Braidfauld North	10.22	169.87	64.30	48.89	188.84	1.36	210.71	1.32
Tollcross	13.00	131.00	92.10	63.00	295.50	1.30	418.80	1.25
Kingston West and Dumbreck	9.35	110.80	68.04	42.48	222.48	1.09	202.38	1.30
Shettleston North	9.92	102.18	51.45	40.75	120.21	0.73	169.79	1.50

Intermediate Geography area	As mg kg^{-1}	Cr mg kg^{-1}	Cu mg kg^{-1}	Ni mg kg^{-1}	Pb mg kg^{-1}	Se mg kg^{-1}	Zn mg kg^{-1}	K_2O wt %
Parkhead West and Barrowfield	11.54	122.63	89.14	55.98	167.76	1.02	212.22	1.32
Laurieston and Tradeston	8.89	97.81	62.19	39.07	124.88	0.68	150.46	1.36
Calton, Galllowgate and Bridgeton	9.08	92.76	64.13	48.07	196.55	0.79	188.17	1.37
Garrowhill West	7.50	111.82	57.24	45.71	128.37	0.83	181.78	1.32
Cardonald North	8.03	108.98	76.74	50.51	210.05	0.90	201.13	1.47
Craigton	12.12	104.34	65.36	54.19	336.63	1.13	254.91	1.47
Hillington	9.33	111.00	61.39	48.72	188.89	1.08	159.13	1.48
Greenfield	10.06	152.99	82.07	56.77	139.33	1.04	220.08	1.34
Penilee	8.20	101.14	36.81	36.15	87.21	0.76	173.35	1.37
Kinning Park and Festival Park	9.85	96.95	113.29	56.65	260.34	0.90	347.39	1.32
Old Shettleston and Parkhead North	9.35	110.63	138.39	54.99	198.56	0.86	295.51	1.43
Ibrox East and Cessnock	8.42	84.34	57.92	44.26	118.31	0.92	178.09	1.52
Garrowhill East and Swinton	6.25	94.33	32.81	31.07	67.32	0.63	107.53	1.29
Barlanark	8.23	89.13	41.92	33.79	122.61	0.76	148.06	1.19
Gallowgate North and Bellgrove	13.28	92.66	60.25	42.90	179.38	0.80	226.73	1.42
Carntyne West and Haghill	13.23	93.78	77.62	54.35	176.14	1.21	175.89	1.42
Ibrox	10.23	89.26	58.16	44.07	122.31	0.86	148.65	1.49
Dennistoun	8.30	113.00	30.70	41.90	71.60	0.80	108.20	1.26
Anderston	22.27	116.75	219.23	77.52	665.50	0.92	501.65	1.47
North Barlanark and Easterhouse South	5.93	105.82	46.02	38.80	73.39	0.54	115.49	1.48
Easterhouse East	8.21	151.14	63.06	41.44	137.59	0.80	157.09	1.36
Cranhill, Lightburn and Queenslie South	8.10	125.74	74.74	46.37	197.56	0.95	203.74	1.43
City Centre West	9.56	111.24	58.90	33.37	164.13	0.75	204.25	1.38
Govan and Linthouse	10.87	134.83	84.83	54.18	224.56	0.88	202.06	1.35
Carntyne	7.05	123.42	65.86	49.30	119.91	1.10	167.13	1.43
City Centre East	9.11	105.56	58.32	38.07	154.41	0.76	169.56	1.28
Dennistoun North and Alexandra Parade	11.48	110.83	83.17	49.30	170.15	1.16	169.57	1.30
Drumoyne and Shieldhall	9.72	101.93	60.99	47.08	132.64	0.79	162.57	1.47
Finnieston and Kelvinhaugh	7.20	96.21	63.09	30.13	136.73	0.49	337.26	1.26
Central Easterhouse	7.30	97.98	56.05	44.57	144.80	0.99	177.34	1.31
Craigend and Ruchazie	9.95	143.80	69.98	47.63	191.53	1.03	202.49	1.39
Hillhead	9.38	99.04	65.07	45.20	240.37	1.26	202.22	1.29
Woodlands	16.28	115.30	106.23	61.91	348.94	2.16	178.24	1.31
Roystonhill, Blochairn, and Provanmill	12.22	106.19	63.96	45.60	145.00	0.85	178.28	1.38
Riddrie and Hogganfield	11.01	147.55	105.33	59.96	253.33	1.32	263.47	1.33
Glasgow Harbour and Partick South	10.00	108.00	70.80	45.40	172.40	0.80	207.20	1.56
Woodside	11.74	98.42	64.68	44.90	163.98	0.85	156.88	1.24
Sighthill	15.13	104.23	48.39	39.30	103.06	0.84	124.55	1.65
Partick	11.27	99.88	40.56	39.73	203.52	0.90	169.75	1.45
Kelvingrove and University	14.52	102.29	78.00	51.94	472.78	2.50	156.01	1.31
Whiteinch	18.49	142.46	291.75	106.99	299.88	3.39	341.11	1.54
Garthamlock, Auchinlea and Gartloch	9.16	128.61	55.69	43.82	145.39	1.02	154.67	1.33
Firhill	17.90	101.00	91.50	51.90	337.10	1.50	358.10	1.22
Cowlairs and Port Dundas	8.06	94.04	51.34	37.55	111.82	0.73	139.35	1.39
Partickhill and Hyndland	10.02	109.09	88.84	50.33	261.19	1.41	231.83	1.49
Blackhill and Barmulloch East	12.32	115.67	63.15	44.57	169.94	1.11	181.82	1.34
Broomhill	9.70	103.62	71.44	63.71	266.15	1.34	206.68	1.48

Intermediate Geography area	As mg kg^{-1}	Cr mg kg^{-1}	Cu mg kg^{-1}	Ni mg kg^{-1}	Pb mg kg^{-1}	Se mg kg^{-1}	Zn mg kg^{-1}	K_2O wt %
Petershill	11.77	105.88	109.63	64.04	249.96	1.26	259.33	1.45
Dowanhill	12.30	101.00	114.10	63.70	196.30	1.30	282.60	1.29
Victoria Park	12.30	93.06	62.90	52.42	200.11	1.20	177.21	1.32
North Kelvin	14.46	105.09	70.78	57.34	294.03	1.44	213.72	1.40
Scotstoun South and West	18.27	133.94	148.41	87.74	278.81	1.89	295.39	1.31
Keppochhill	9.63	87.31	83.03	39.77	215.49	0.86	344.48	1.41
Scotstoun North and East	12.30	112.42	71.71	52.51	234.55	0.67	222.33	1.27
Barmulloch	10.94	133.84	69.30	48.91	160.38	1.19	209.37	1.39
Ruchill	11.49	100.89	60.00	51.33	219.51	1.19	150.11	1.46
Kelvinside and Jordanhill	10.42	102.36	89.55	59.81	221.30	1.43	225.46	1.41
Kelvindale	10.88	93.49	79.07	51.47	222.58	1.15	192.34	1.35
Springburn	10.27	108.32	73.03	51.15	185.59	1.56	167.15	1.33
Possil Park	9.18	101.36	68.53	44.85	185.72	0.97	161.18	1.37
Wyndford	11.53	92.43	69.29	39.43	131.61	0.90	265.61	1.50
Springburn East and Cowlairs	14.03	98.87	90.02	52.72	271.78	1.02	211.27	1.40
Robroyston and Millerston	9.10	94.63	57.19	46.44	122.23	0.93	141.90	1.41
Yoker South	8.59	98.71	125.96	54.68	182.50	1.36	349.16	1.27
Balornock	7.56	99.60	88.78	43.75	212.15	1.09	141.34	1.39
Knightswood Park East	7.60	95.00	41.20	47.40	146.80	1.00	117.90	1.61
Knightswood East	5.20	78.00	35.10	34.10	138.20	0.60	106.70	1.39
Maryhill West	9.62	123.77	41.32	37.10	114.33	0.84	224.96	1.30
Milton West	8.59	99.99	63.28	50.23	142.37	0.98	126.16	1.24
Anniesland East	8.78	88.74	60.31	39.55	149.62	1.17	126.99	1.26
Knightswood West	16.35	101.38	146.42	74.35	381.38	1.47	306.51	1.33
Maryhill East	29.20	123.28	79.08	80.25	170.84	1.04	166.34	1.23
Anniesland West	8.65	84.31	56.93	43.68	138.28	0.99	161.57	1.32
Knightswood Park West	20.36	118.91	78.48	72.70	267.42	2.03	277.74	1.26
Yoker North	9.23	100.31	64.72	54.53	166.11	0.94	198.87	1.46
Milton East	7.78	89.18	49.31	35.43	143.63	1.06	186.55	1.27
Blairdardie East	7.18	83.58	41.03	36.45	97.18	0.86	118.43	1.31
Summerston North	9.35	80.80	39.94	33.73	103.78	0.99	117.21	1.22
Blairdardie West	7.96	86.11	29.90	36.23	76.49	0.53	93.41	1.28
Drumchapel South	8.21	86.13	51.46	41.56	150.86	0.93	172.15	1.29
Summerston Central and West	9.58	83.68	42.07	37.52	123.78	0.85	122.09	1.29
Drumry West	7.95	79.52	32.42	32.44	79.41	0.58	102.63	1.41
Drumchapel North	7.40	84.45	35.27	36.55	79.95	0.70	117.01	1.30
Drumry East	8.88	95.47	43.65	40.02	117.47	0.93	172.38	1.31
Forgewood	8.40	129.97	52.57	58.05	98.12	0.92	151.26	1.53
Birkenshaw	8.30	88.00	45.20	33.40	111.00	1.10	291.40	1.59
Townhead	8.73	108.14	63.13	46.98	148.78	1.27	169.64	1.07
Glenmavis and Greengairs	11.13	100.44	45.57	35.17	149.69	1.03	167.94	1.12
Stepps	8.09	98.55	45.50	35.49	114.77	0.87	119.57	1.24
Gartcosh and Marnock	7.36	76.30	31.33	22.41	93.02	0.83	83.46	0.92
Chryston and Muirhead	8.61	84.10	32.47	26.95	111.53	0.93	113.78	1.06
Moodiesburn West	11.10	100.00	31.60	26.10	120.50	0.90	97.30	1.15
Kilsyth Bogside	55.30	136.00	48.30	53.60	78.20	6.60	102.30	1.85
Balmalloch	8.09	86.54	20.27	25.11	53.14	0.98	87.13	1.16
Kilsyth East and Croy	8.36	76.22	25.04	28.31	55.59	0.89	86.57	1.13
Renfrewshire Rural South & Howwood	11.15	132.57	48.06	56.86	118.71	1.25	162.36	1.08
Paisley Glenburn West	12.00	109.20	54.83	62.74	96.95	0.80	155.79	1.57
Paisley Glenburn East	12.07	109.61	56.29	54.16	168.09	1.22	199.16	1.26

Intermediate Geography area	As mg kg^{-1}	Cr mg kg^{-1}	Cu mg kg^{-1}	Ni mg kg^{-1}	Pb mg kg^{-1}	Se mg kg^{-1}	Zn mg kg^{-1}	K_2O wt %
Paisley Foxbar	11.44	122.78	64.53	59.93	148.51	1.11	172.40	1.32
Johnstone South West	10.16	142.68	61.38	76.34	94.87	1.17	165.06	1.29
Paisley Dykebar	14.54	119.25	63.72	52.83	150.73	1.11	169.33	1.21
Paisley South West	11.22	107.16	48.09	49.36	129.72	1.00	147.54	1.22
Paisley South East	13.55	114.20	89.54	80.30	316.33	1.41	268.90	1.19
Johnstone South East	11.10	145.01	72.46	79.81	131.92	1.14	203.07	1.20
Paisley South	12.48	103.77	59.75	55.79	135.68	0.81	156.62	1.23
Johnstone North West	12.40	150.51	117.72	93.66	249.64	1.30	433.64	1.28
Paisley West	9.49	106.15	44.84	48.07	99.94	1.11	169.48	1.39
Paisley East	12.83	135.59	67.67	74.11	194.58	1.29	194.03	1.26
Johnstone North East	9.44	103.28	47.92	48.88	86.03	0.80	158.17	1.46
Elderslie and Phoenix	11.69	130.12	74.68	66.39	186.77	1.18	213.69	1.34
Paisley Central	13.67	106.23	80.13	62.44	260.77	1.21	236.90	1.15
Paisley North West	21.67	107.35	59.04	54.63	165.73	1.59	131.26	1.25
Paisley Ralston	12.53	119.57	56.99	63.12	158.90	1.09	180.94	1.44
Paisley Ferguslie	17.30	141.82	104.33	73.23	209.16	1.24	297.75	1.37
Paisley North East	14.99	109.30	96.77	69.81	255.67	1.71	228.80	1.22
Linwood South	11.04	137.93	60.06	79.22	109.99	1.46	182.57	1.58
Linwood North	13.85	127.67	83.55	74.09	198.15	1.55	232.16	1.51
Paisley Gallowhill and Hillington	10.15	109.01	93.25	67.59	131.13	0.83	165.02	1.34
Paisley North	13.16	114.14	70.88	60.95	152.59	1.08	194.74	1.39
Renfrew South	11.71	211.90	84.97	66.58	209.41	0.99	259.37	1.44
Renfrew East	9.69	106.12	60.18	53.83	119.13	0.71	154.81	1.37
Renfrew West	9.25	113.94	64.20	63.68	118.82	0.72	174.15	1.35
Renfrewshire Rural North and Langbank	13.33	126.47	60.79	57.14	141.98	1.23	173.88	1.34
Renfrew North	12.77	112.92	78.33	54.50	173.57	0.80	208.85	1.41
Erskine East and Inchinnan	21.07	97.49	43.64	32.21	136.03	0.63	106.89	1.25
Erskine Central	12.50	86.00	29.40	30.80	55.50	0.70	171.20	1.56
Chapelton, Glengavel and Sandford	6.41	100.33	19.52	24.79	35.12	0.70	79.72	1.27
Ashgill and Netherburn	7.70	111.00	52.70	40.00	79.90	0.80	310.30	1.46
Glassford, Quarter and Allanton	6.56	106.60	26.00	31.49	64.06	0.69	88.64	1.23
Crosshouse and Lindsayfield	9.39	141.56	27.31	55.07	72.32	1.14	108.16	1.33
Whitehills West	7.73	119.13	33.83	44.33	79.60	1.17	100.36	1.48
Greenhills	7.16	124.52	34.75	62.25	49.71	0.85	111.28	1.44
Mossneuk and Newlandsmuir	8.96	112.14	26.63	40.94	51.36	1.00	92.84	1.34
Little Earnock	6.41	102.18	19.64	20.83	52.02	0.68	67.91	1.07
Birniehill, Kelvin and Whitehills East	5.76	109.70	31.46	40.54	54.35	0.79	98.44	1.39
Nerston and EK Landward Area	7.00	125.31	32.93	47.33	71.47	0.85	127.29	1.21
Westwood South	8.85	125.38	36.38	46.06	72.32	1.19	140.26	1.39
The Murray	10.49	127.14	40.83	47.54	77.41	1.41	146.93	1.49
Hairmyres and Westwood West	7.68	115.40	38.82	61.17	76.53	0.84	136.59	1.30
St Leonards South	9.00	109.71	38.96	41.88	74.76	0.99	126.25	1.34
Westwood East	12.63	125.32	49.95	54.05	97.18	1.14	178.23	1.32
Thorntonhall, Jackton and Gardenhall	7.75	111.73	30.92	45.16	57.41	0.74	109.66	1.25
West Mains	10.67	123.04	53.98	51.42	126.58	1.04	193.33	1.35
St Leonards North	6.67	107.26	30.60	40.33	57.62	0.87	97.72	1.26
East Mains	42.20	213.00	362.40	295.40	612.10	1.80	518.20	1.32
Stewartfield West	6.45	132.76	42.39	70.70	51.82	0.55	135.12	1.47
Calderwood West and Nerston	8.61	129.56	52.08	58.77	97.49	1.13	188.16	1.25
Stewartfield East	7.95	140.00	51.80	67.92	64.03	0.92	132.34	1.33
Calderwood Central	8.55	115.84	45.80	52.66	103.83	1.37	191.00	1.43

Intermediate Geography area	As $mg\ kg^{-1}$	Cr $mg\ kg^{-1}$	Cu $mg\ kg^{-1}$	Ni $mg\ kg^{-1}$	Pb $mg\ kg^{-1}$	Se $mg\ kg^{-1}$	Zn $mg\ kg^{-1}$	K_2O wt %
Calderwood East	6.16	116.74	42.53	50.19	81.40	0.86	122.44	1.36
Hamilton Centre and Low Parks	8.54	131.26	66.23	51.79	123.13	1.05	146.19	1.70
Bothwell South	8.84	110.98	31.31	37.32	87.49	0.85	112.91	1.90
Fernhill and Cathkin	7.89	151.44	36.70	50.37	84.57	0.75	129.61	1.33
Whitlawburn and Greenlees	7.56	105.38	37.96	34.44	103.73	0.69	121.04	1.29
Vicarland and Cairns	8.04	117.85	53.76	44.44	122.65	0.76	186.42	1.30
Burnside and Springhall	7.43	120.48	49.60	46.77	122.76	0.66	153.53	1.30
Low Blantyre and Bardykes	11.99	99.29	43.32	37.75	101.63	0.49	153.19	1.48
Spittal	6.78	134.08	36.75	41.54	79.06	0.59	121.05	1.35
Cambuslang Central	11.13	145.39	73.62	64.05	268.01	1.21	268.81	1.22
High Crosshill	8.71	147.88	59.20	55.09	199.83	0.90	192.44	1.44
Halfway, Hallside and Drumsagard	8.39	101.82	43.24	40.44	89.08	0.78	134.05	1.46
Westburn and Newton	10.10	119.72	62.80	56.38	124.49	0.90	236.16	1.42
Bankhead South	8.55	117.61	52.73	54.73	138.00	0.85	151.33	1.43
Burgh, Eastfield and Silverbank	10.09	189.58	77.11	65.88	169.51	0.98	219.34	1.35
Burnhill and Bankhead North	9.23	134.01	73.62	28.98	105.25	0.49	145.53	1.42
Shawfield and Clincarthill	11.91	437.94	80.96	100.63	126.24	0.80	271.80	1.03
Farme Cross and Gallowflat North	11.49	217.10	98.18	79.11	224.19	1.23	312.53	1.29
Blane Valley	4.90	49.00	8.60	5.60	61.30	1.00	22.60	0.42
IZ One	14.68	100.50	68.07	50.98	182.14	0.75	192.26	1.56
IZ Two	8.49	87.98	38.69	35.01	82.13	0.50	112.75	1.47
IZ Three	9.86	89.89	77.96	43.99	152.95	0.79	166.68	1.37
IZ Four	12.48	100.64	54.05	36.45	117.20	0.43	157.12	1.58
IZ Five	9.31	86.42	56.28	42.67	126.53	0.64	134.37	1.48
IZ Six	8.12	80.52	36.69	36.98	102.53	0.75	90.36	1.35
IZ Seven	7.87	86.04	37.91	34.39	74.58	0.53	111.89	1.41
IZ Eight	9.15	98.14	44.58	47.62	111.02	0.85	122.70	1.29
IZ Nine	11.25	89.48	55.13	37.48	146.40	0.55	141.48	1.61
IZ Ten	8.53	98.93	33.45	36.42	98.39	0.95	124.96	1.19
IZ Eleven	9.26	84.64	42.72	35.98	101.56	0.55	117.60	1.73
IZ Twelve	12.16	77.42	39.98	30.58	117.38	0.48	103.72	1.77
IZ Thirteen	8.64	74.55	24.35	24.54	62.24	0.36	65.38	1.83
IZ Fourteen	10.13	95.95	40.45	44.28	119.91	0.35	121.50	2.46
IZ Fifteen	11.00	76.09	35.74	24.90	79.24	0.49	77.36	1.72
IZ Sixteen	8.06	51.81	10.12	10.09	52.54	0.60	36.31	1.34

Note: The IG Codes from S02001184 to S02001199 represent IG areas in the West Dunbartonshire region. Since West Dunbartonshire did not provide IG names, they are just labelled as ‘IZ One’.... ‘IZ Sixteen’ for the area covered in this study

References

ArcGIS Resource Centre, (2010). ArcGIS Resource Centre.
help.arcgis.com/en/arcgisdesktop/10.0/help/index.html#//005p0000000n000000.html
Access date: May 2010.

ATSDR, (2005), Public Health Statement for Nickel, Agency for Toxic Substances & Disease Registry, Department of Health and Human Services, Atlanta.
www.atsdr.cdc.gov/toxprofiles/phs15.html
Access date: December 2009.

BGS, (2010a), Clyde Urban Super Project.
www.bgs.ac.uk/science/landUseAndDevelopment/urban_geoscience/Clyde/.
Access date: June 2010.

BGS, (2010b), Geochemical Baseline Survey of the Environment Project, British Geological Survey, Nottingham.
www.bgs.ac.uk/gbase.
Access date: January 2010.

Bivand, R., Muller, W., G. and Reder, M, (2009), Power calculations for global and local Moran's I. Computational Statistics and Data Analysis, 53: 2859-2872.

Cave, M. R. and Chenery, S., (2010), Measurement of the geogenic contribution to PM10 daily dust samples collected in Northampton, Abstracts of the 27th Society for Environmental Geochemistry and Health (SEGH) European Conference, Galway, Ireland, June 27- July 2, 2010.

CRAN, (2009), The Comprehensive R Archive Network.
cran.r-project.org/
Access date: December 2009.

DEFRA, (2010), Department for Environment Food and Rural Affairs.
www.defra.gov.uk/environment/quality/air/airquality/publications/airpoll/05.htm
Access date: January 2010

Diggle, P. J. and Ribeiro, P. J., (2007), Model-based Geostatistics, Springer Series in Statistics. Springer New York,

EA, (2002a), Contaminants in Soil: collation of toxicological data and intake values for humans, Chromium, R&D Publication Tox 4, Environment Agency, Bristol.

EA, (2002b), Contaminants in Soil: collation of toxicological data and intake values for humans, Lead, R&D Publication Tox 6, Environment Agency, Bristol.

EA, (2009a), Updated Technical Background to the Contaminated Land Exposure Assessment (CLEA) Model. Science Report SC050021/SR3, Environment Agency, Bristol.

EA, (2009b), Using Soil Guideline Values. Science Report SC050021/SGV Introduction, Environment Agency, Bristol.

EA, (2009c), Contaminants in Soil: updated collation of toxicological data and intake values for humans, inorganic arsenic, Science Report: SC050021/TOX 1, Environment Agency, Bristol.

EA, (2009d), Contaminants in Soil: updated collation of toxicological data and intake values for humans, nickel. Science Report: SC050021/TOX 8, Environment Agency, Bristol.

EA, (2009e), Contaminants in Soil: updated collation of toxicological data and intake values for humans, selenium. Science Report: SC050021, Environment Agency, Bristol.

EA, (2010a), Soil Guideline Value reports published using the new approach. Environment Agency, Bristol,
www.environment-agency.gov.uk/research/planning/64015.aspx
Access date: January 2010.

EA, (2010b), Copper and its Compounds, Environment Agency, Bristol,
www.environment-agency.gov.uk/business/topics/pollution/142.aspx.
Access date: January 2010.

EA, (2010c), Zinc and its Compounds, Environment Agency, Bristol,
www.environment-agency.gov.uk/business/topics/pollution/255.aspx.
Access date: January 2010.

Environmental Protection Act Part IIa, (1990), Contaminated Land. HSMO, London.

Ezzaguirre-Garcia, D., Rodriguez-Andres, C., Watt, G.C. and Hole, D., (1999), A study of leukaemia in Glasgow in connection with chromium contaminated land. Journal of Public Health Medicine, 21: 435-438.

Fairbairn, J., Walker, G., Mitchell, G. and Smith, G., (2005), Investigating Environmental Justice In Scotland: Links Between Measures Of Environmental Quality And Social Deprivation, For Scottish and Northern Ireland Forum for Environmental Research (Scottish Executive, Scottish Environmental Protection Agency, Forestry Commission and Scottish Natural Heritage), Edinburgh.

Farmer, J., G. and Jarvis, R., (2009), Strategies for improving human health in contaminated situations: a review of past, present and possible future approaches, Environmental Geochemistry and Health, 31 (2): 227-238.

Fordyce, F., M., Brown, S., E., Ander, B., G., Rawlins, K., E., O'Donnell, K., E., Lister, T., R., Breward, N. and Johnson, C., C., (2005), GSUE: urban geochemical mapping in Great Britain, Geochemistry: Exploration, Environment, Analysis, 5, (4): 325-326.

Fordyce, F. M., Nice, S. E., Lister, T. R., Ó Dochartaigh, B. É., Cooper, R., Allen, M., Ingham, M., Gowing, C., Vickers, B. P. and Scheib, A., In Prep, Urban Soil Geochemistry of Glasgow. Open Report, OR/08/002. British Geological Survey, Edinburgh.

Helsel, D. R., (2005), Nondetects and Data Analysis: Statistics for censored environmental data. Wiley-Interscience, New York.

ISD, (2009), Information Services Division, NHS Scotland,
www.isdscotland.org/isd/6080.html
Access date: December 2009.

ISD (2010), Information Services Division, NHS Scotland,
www.isdscotland.org
Access date: February 2010.

Jarup, L., (2004), Health and environment information systems for exposure and disease mapping, and risk assessment, Environmental Health Perspectives, 112, (9): 995-997.

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

Kibble, A. J. and Saunders, P. J., (2001), Contaminated Land and the Links with Health, In: Hester, R., E. and Harrison, R., M (editors). Assessment and Reclamation of Contaminated Land, Issues in Environmental Science and Technology, The Royal Society of Chemistry, Cambridge, Issue 16, pp 65-84.
Lawson, A. B., (2001), Statistical Methods in Spatial Epidemiology, John Wiley & Sons Ltd, Chichester, England

Lee, D., Ferguson, C., Mitchell, R., (2009), Air pollution and health in Scotland: a multicity study, Biostatistics, 10, (3): 409-423.

Li, H., Calder, C.A. and Cressie, N., (2007), Beyond Moran's I: Testing for spatial dependence based on the spatial autoregressive model, Geographical Analysis, 39, (4): pp357-375.

McConnell, J., (2010), Speech on the Scottish Executive policy on environment and sustainable development, Dynamic Earth Conference, Edinburgh, Feb 18th 2002.
www.spokes.org.uk/oldsite/firstministerspeech.htm
Access date: February 2010.

Mielke, H., W., Berry, K.J., Mielke, P.W., Powell, E.T. and Gonzales, C.R., (2005), Multiple metal accumulation as a factor in learning achievement within various New Orleans elementary school communities, Environmental Research, 97: 67-75.

Mielke, H., W., Zahran, S., Weiler, S., Berry, K., J. and Gonzales, C., (2009), Children's blood lead and standardized test performance response as indicators of neurotoxicity in metropolitan New Orleans elementary schools, NeuroToxicology, 30: 888-897.

Mitchell, R., Sridharan, S., Tunstall, H. and Lawder, R., (2007), An exploratory spatial data analysis approach to understanding the relationship between deprivation and mortality in Scotland, Social Science & Medicine, 65: 1942-1952.

MOE, (2001), Copper in the Environment, Ministry of the Environment programs and initiatives, Ontario, Canada,
www.ene.gov.on.ca/cons/4141e.htm
Access date: December 2009.

Ottesen, R., T., Alexander, J., Langedal, M., Haugland, T. and Hoygaard, E., (2008), Soil pollution in day-care centres and playgrounds in Norway: national action plan for mapping and remediation, Environmental Geochemistry and Health, 30 (6): 623-637.

Plummer, M. and Clayton, D.G., (1996), Estimation of population exposure in ecological studies, Journal of the Royal Statistical Society, Series B 58: 167-182.

Richardson, E.A, Mitchell, R.J., Shortt, N.K., Pearce, J. and Dawson, T., (2010), Developing summary measures of health-related multiple physical environmental deprivation for epidemiological research, Environment and Planning, 42 (7): 1650-1668.

Richardson, S., Stucker, I. and Hemon, D., (1987), Comparison of relative risks obtained in ecological and individual studies: some methodological considerations, International Journal of Epidemiology, 16: 111-120.

RCEP, (1996), Royal Commission on Environmental Pollution Nineteenth Report, Sustainable Use of Soil, HMSO, London,.

School of Geosciences, (2009), Semivariogram Model Parameters.
www.geos.ed.ac.uk/homes/s0198247/variograms.html
Access date: November 2009.

Scottish Executive Environment Group, (2005), Review of Progress on Environmental Justice, December 2005, Scottish Executive, Edinburgh.

Scottish Government, (2005), Investigating Environmental Justice in Scotland: Links Between Measures of Environmental Quality and Social Deprivation, March 2005, Scottish Government, Edinburgh.

Scottish Government, (2011), Intermediate Geography Background Information
www.scotland.gov.uk/Resource/Doc/36496/0026946.pdf
Access date: May 2011.

Scottish Air Quality. (2010), Air Quality Pollutants
www.scottishairquality.co.uk/about.php?n_action=monitoring
Access date: January 2010.

Selinus, O. and Frank, A., (2000), Environmental Medicine. In: Lennart Moller (ed.) Medical Geology, Karolinska Institute, Stockholm, Sweden , pp 164-182.

Shinn, N., J., Bing-Canar, J., Cilas, M., Peneff, N. and Binns, H., J., (2000), Determination of spatial continuity of soil lead levels in an urban residential neighbourhood, Environmental Research Section A, 82: 46-52.

Shohaimi, S., Welch, A., Bingham, S., Luben, R., Day, N., Wareham, N. and Khaw, K-T., (2004), Area deprivation predicts lung function independently of education and social class, European Respiratory Journal, 24: 157-161.

Sierra, (2009), Variogram Models.
www.rsierra.com/DA/node10.html
Access date: December 2009.

SIMD, (2004), Scottish Index of Multiple Deprivation 2004: Technical Report, Published by Scottish Executive, Edinburgh, June 2004.

SIMD, (2009a), Overview and How to Use SIMD.
www.scotland.gov.uk/Topics/Statistics/SIMD
Access date: December 2009.

SIMD, (2009b), Scottish Index of Multiple Deprivation 2009: Technical Report, Published by Scottish Executive, Edinburgh, October 2009.

SIMD, (2009c), Visual Guide to the Methodology Used to Create SIMD 2009.
www.scotland.gov.uk/Topics/Statistics/SIMD/MethodologyVisual2009
Access date: December 2009.

SIMD. (2010), Scottish Index of Multiple Deprivation,
www.scotland.gov.uk/topics/statistics/simd
Access date: November 2009.

SNS, (2010), Scottish Neighbourhood Statistics,
www.sns.gov.uk/Downloads/DownloadHome.aspx
Access date: January 2010.

UK Air Quality Archive, (2007), Air Quality Standards and Objectives
www.airquality.co.uk/standards.php
Access date: January 2009.

Watt, G., Eizaguirre, D., Gillis, C., Hole, D., Stewart, I. and Symington, I., (1991), Assessment of the risk to human health from land contaminated by chromium waste, Report by the Department of Public Health, Glasgow, UK.

Webster, R. and Oliver, M. A., (2001), Geostatistics for Environmental Scientists, John Wiley & Sons, Chichester, New York.

WHO, (1996), Trace Elements in Human Nutrition and Health, World Health Organisation, Geneva.

Wilkins, E., T., (2006), Air pollution aspects of the London fog of December 1952, Quarterly Journal of the Royal Meteorological Society, 80 (344): 267-271.

Wood, S., M., (2006), Generalised Additive Models, Texts in Statistical Science Series, Chapman and Hall/CRC, London.