



Manaaki Whenua
Landcare Research

Determining background soil concentrations of trace elements across New Zealand

Envirolink Grant: 2321-HBRC267

Prepared for: Hawke's Bay Regional Council

August 2023



Determining background soil concentrations of trace elements across New Zealand

Contract Report: LC4324

Jo Cavanagh, Stephen McNeill, Pierre Roudier, Hadee Thompson-Morrison

Manaaki Whenua – Landcare Research

Adam Martin, Rose Turnbull

GNS-Science

Reviewed by:

John Drewry
Senior Researcher
Manaaki Whenua – Landcare Research

Approved for release by:

John Triantafilis
Portfolio Leader – Managing Land & Water
Manaaki Whenua – Landcare Research

Disclaimer

This report has been prepared by Landcare Research New Zealand Ltd for Hawke's Bay Regional Council. If used by other parties, no warranty or representation is given as to its accuracy and no liability is accepted for loss or damage arising directly or indirectly from reliance on the information in it.

© Landcare Research New Zealand Ltd and Hawke's Bay Regional Council 2023

This information may be copied and distributed to others without limitation, provided Landcare Research New Zealand Ltd Hawke's Bay Regional Council are acknowledged. Under no circumstances may a charge be made for this information without the written permission of Landcare Research and Hawke's Bay Regional Council..

Contents

Summary.....	v
1 Introduction	1
2 Background.....	2
2.1 Definitions.....	3
3 Objectives	3
4 Methods	4
4.1 Data	4
4.2 Explanatory variables.....	8
4.3 Statistical modelling.....	11
5 Results.....	17
5.1 Background soil concentrations.....	17
5.2 Comparison with regional background studies.....	36
6 Using background soil concentration information	45
6.1 Legislative context for managing contaminated land	45
6.2 Determination of site-specific background concentrations.....	46
6.3 Waste acceptance criteria.....	47
6.4 Background concentrations and ecological soil guideline values.....	49
6.5 Areas with naturally elevated concentrations.....	58
7 Summary and conclusions.....	58
8 Recommendations.....	60
9 Acknowledgements.....	60
10 References	61
Appendix 1 – Relative influence of individual variables.....	65

Summary

Project and client

- This project revised estimates of background concentrations of naturally occurring trace elements in soils across New Zealand using an extended data set. The project was undertaken for Hawke's Bay Regional Council through Envirolink funding (Envirolink Advice Grant 2321-HBRC267).

Objectives

The specific objectives for this project were to:

- develop revised estimates of background concentrations of naturally occurring trace elements in soils using an extended data set
- compare these revised estimates with existing regional studies of background concentrations (Auckland, Wellington, Christchurch)
- provide guidance on the use of background concentrations including:
 - the use of background soil concentrations to adjust ecological soil guideline values (Eco-SGVs), including when and where this might be appropriate
 - other uses of background concentrations, including ambient vs natural background, and where and when ambient background might be most applicable for managing contaminants (including disposal of materials to clean fills)
 - the site-specific determination of background concentrations
- provide recommendations for any additional monitoring to inform better estimates of background concentrations, including for specific areas (e.g. naturally elevated areas, urban areas).

Methods

- Three data sets were used to develop estimates of background concentrations:
 - New Zealand geochemical baseline data set
 - regional council state of the environment soil quality data
 - regional background studies.
- While up to 65 elements were available from the New Zealand geochemical data set, the focus for this work was a subset of common 'contaminant' elements: arsenic, boron, cadmium, chromium, copper, lead, nickel, and zinc.
- Data from the three data sets were combined and processed to ensure consistent land-use classification, resulting in 2,147 rows of data for further analysis. The land-use classes used were:
 - rural ambient (which combined 'background', 'Forestry', and 'Grazing' that was not 'High Producing Exotic Grassland' using Land Cover Database version 5)
 - grazing on high-producing grassland

- short-rotation cropping
 - perennial horticulture
 - urban
 - wetland.
- Contextual information for each sampling location was obtained from three spatial databases: the pedological S-map, the Fundamental Soils Layers (FSL) in the Land Resource Information System (LRIS), and the digital QMAP geological map of New Zealand. Data from only the most recent sampling of a given site was used for subsequent data analysis. Three data layers were used to provide explanatory variables:
 - parent-material (developed from QMAP and FSL)
 - soil order (developed by combining S-map and the FSL)
 - Land Environments of New Zealand (LENZ, to provide climatic features).
 - The statistical package R was used to assess the influence of land use and the above three explanatory data layers on the key trace elements (arsenic, boron, cadmium, chromium, copper, lead, nickel, and zinc) using a generalised additive model, and to predict rural ambient (background) concentrations across New Zealand.
 - Predicted background concentrations were modelled at a 1 km × 1 km grid, and presented in percentile ranges to reflect the spatial variation and likely concentration ranges across New Zealand, and also to revise previously developed Eco-SGVs.
 - A brief review of the use of background concentrations for land management was also undertaken.

Results and conclusions

- This project has developed nationally consistent estimates of background concentrations of selected naturally occurring trace elements across New Zealand. These estimates are effectively rural ambient concentrations (i.e. data from sites where there is expected to be minimal anthropogenic additions of trace elements to the soil and no difference from true, naturally occurring concentrations). The predicted concentrations are displayed as filled contour plots to enable visualisation of the variation in concentrations across the country, and are also in an online tool.
- The models generally provided reasonable explanatory power for the individual trace elements, with the exception of zinc and arsenic. These models markedly under-predicted concentrations at the higher concentrations, and hence predicted lower (more conservative) background concentrations for the upper percentile estimates (e.g. above 90th percentile). This under-prediction is less important for arsenic, which had a narrow concentration range across the data set. Further model development is required to obtain better estimates for zinc.
- The results from the modelling are summarised in Table S1. Background concentrations vary across the country and are displayed as quantile maps, using the median, 90th, 95th, and 99th percentile as the 'cut-offs' for concentrations within different locations.

Table S1. Summary statistics for predicted rural ambient (background) concentrations (mg/kg) for key trace elements across New Zealand.

Element	Min	5%	Median	Mean	90%	95%	99%	Max
As	0.2	2.2	4.1	4.2	6	7	8	19
B	0.5	2.2	4.6	6.0	12	16	23	83
Cd	0.01	0.02	0.08	0.10	0.20	0.29	0.35	0.58
Cr	2.0	6.8	16	18	25	30	68	765
Cu	3.8	8.5	16	17	24	28	39	76
Ni	1.4	3.8	9	11	14	16	42	590
Pb	1.3	4.8	11	12	17	19	21	30
Zn	11	29	48	48	63	68	80	100

- Spatially explicit concentration data and maps for the individual trace elements are also available at: <https://iris.scinfo.org.nz/layer/114281-pbc-predicted-background-soil-concentrations-new-zealand-h3-resolution-9/>
- The median background concentrations were used to develop the default Eco-SGVs. Where appropriate, background-adjusted Eco-SGVs were also developed.
- The use of background concentration information for other purposes typically involves identifying and setting upper thresholds of background concentrations. The challenges to date have been varying data sets, methods, and spatial coverage, which all lead to different estimates.
- The use of upper thresholds for background concentrations for managing soil contaminants in New Zealand should be re-examined. In particular, clause 5(9) of the National Environmental Standard for Assessing and Managing Contaminants in Soil to Protect Human Health (NES-SC) appears to place an undue emphasis on background concentrations, regardless of the risk associated with the concentration of soil contaminants.
- Background concentrations should only be relevant to consider where naturally occurring or ambient concentrations are anticipated to be *above* risk-based guidelines for the protection of human health or ecological receptors. Similarly, site-specific background determination is only recommended for sites where naturally occurring concentrations are anticipated to be *above* risk-based guidelines.
- It is relevant to determine background concentrations in these cases because it is not reasonable to require remediation to concentrations below naturally occurring, or arguably ambient, concentrations. There is, nonetheless, a requirement that the human health risk associated with any elevated concentrations regardless of whether it is natural or anthropogenic in origin is assessed. However, the NES-SC only applies to a piece of land on which a Hazardous Activities and Industries List (HAIL) activity is occurring, has occurred, or is more than likely to have occurred, and clause 5(9) specifically excludes the ability for the human health risk associated with elevated naturally occurring concentrations to be managed.
- From the current work, areas identified as being at or above the 95th percentile of predicted concentrations are considered to be areas that may contain naturally elevated concentrations that may also vary significantly at smaller scales than

captured in the current study. Small-scale variations may also give rise to naturally elevated concentrations outside of these areas, although that is considered less likely.

- While the setting of clean fill waste acceptance criteria based on the background concentrations relevant to the location of the landfill has some logic in terms of enabling clean fills to be permitted activities, closer evaluation of the establishment and use of clean fills is required to ascertain whether this approach is achieving, or will achieve, the desired outcomes in terms of protecting human health and the environment, or is unduly conservative.

Recommendations

- Further model development should be undertaken for arsenic and zinc to ascertain whether improved predictions can be achieved.
- At a higher level, the use of background soil concentration information in existing policy and regulatory setting should be evaluated to determine whether the intended or optimal outcomes are being achieved (e.g. Is the use of background concentrations appropriate for clean fill waste acceptance criteria? Is clause 5(9) achieving its intended purpose, and what is that purpose?).
- In the context of managing contaminated land, clear guidance needs to be given that background concentrations are only relevant to consider when they are *greater* than any risk-based human health or ecological values and it is considered likely that they are greater than the risk-based values.
- Consideration should be given to the merit in combining the proposed Class 4 and 5 landfills into one class, with waste acceptance criteria based on ensuring protection of the most sensitive receptor (people, soil ecological receptors, groundwater). More stringent criteria – or at least criteria based on leaching or off-site movement – could apply where these landfills might be placed close to waterways or groundwater.
- Further investigations at a finer spatial scale are required to better delineate areas of naturally elevated concentrations and to understand the associated risk to human health or the influence on ecological integrity.

1 Introduction

Background soil concentrations of contaminants are currently used in a number of different contexts for the purposes of managing land in New Zealand. For example, the National Environmental Standard for Assessing and Managing Contaminants in Soil to Protect Human Health (hereafter referred to as the NES-SC) places an increased focus on 'background' soil concentrations, because the standard does not apply 'if contaminants in or on the piece of land of interest are at, or below background concentrations'. However, the NES-SC doesn't define background concentrations, and supporting guidance (MfE 2012) is confusing because it refers to '*naturally occurring ambient* concentrations of the element in the area local to the land' (MfE 2012), which is inconsistent with the definitions typically used to describe naturally occurring and ambient concentrations.

Specifically, *naturally occurring concentrations* are those occurring as a result of natural processes only, whereas *ambient concentrations* include additional concentrations arising from diffuse sources of pollution. Under the NES-SC, background concentrations of arsenic (As) and cadmium (Cd) (defined as the 99th percentile concentration of As and Cd in soils collected from around the country and thought not to have been affected by anthropogenic activities) were considered in the development of the soil contaminant standards for As and Cd, with the rural residential soil contaminant standards for As set so as to not be below this upper threshold concentration (MfE 2011).

For Cd, the background concentration (also defined as the 99th percentile concentration of cadmium in soils collected from around the country and thought not to have been affected by anthropogenic activities) is used to define the trigger value for the first tier of the Tiered Fertiliser Management System for Cadmium (FANZ 2019). Background concentrations are also used to define waste acceptance criteria for clean fills (WasteMINZ 2022) and are an integral component of the development of ecological soil guideline values (Eco-SGVs, Cavanagh & Munir 2019; Cavanagh & Harmsworth 2023).

Previous New Zealand studies have used different pedological and geological groupings to define background concentrations (ARC 2001; URS 2003; Tonkin & Taylor 2006, 2007; McDowell et al. 2013) and don't allow for the assessment of inter-regional similarities, or differences in background concentrations and the factors influencing them. This, in turn, limits the ability to predict likely concentrations in locations for which no data are available.

Provisional national estimates of background concentrations of selected trace elements and organic contaminants were developed through Envirolink Tools Grant C09X1402 for use in the development of Eco-SGVs (Cavanagh et al. 2015). These estimates were considered preliminary because they were based on limited data that were distributed sub-optimally across New Zealand. Through Strategic Science Investment Funding (SSIF), GNS-Science and Manaaki Whenua – Landcare Research have undertaken chemical analyses of over 800 additional samples strategically selected to optimise spatial distribution across New Zealand, with some initial analysis of the data set informing revised estimates of background concentrations.

This project provides an analysis of the extended data set, with additional data (including regional studies on background soil concentrations; e.g. Auckland, Wellington, Canterbury) used to develop nationally consistent estimates of background soil concentrations. The project also addresses specific points raised about the use of background concentrations to manage soil contaminants raised at end-user workshops on the implementation of Eco-SGVs (Cavanagh & Harmsworth 2022).

2 Background

Spatial tools are increasingly used internationally to determine background soils concentrations (e.g. Lado et al. 2008; Diez et al. 2009; Jarva et al. 2010; Cave et al. 2012). Often geostatistical analyses are undertaken and used to define relevant 'domains' or groupings to indicate where background concentrations are similar. Such tools enable the extrapolation of collected data to areas where data have not been collected. Whether the tools are geologically based or soils based appears to depend more on what databases are available as opposed to a rigorous assessment of the factors controlling background concentrations.

For example, in the UK a Soil-Parent Material Model (SPMM) has been developed (Lawley 2009) based on the origin of the parent materials, dominant mineralogy, and texture. This is available at a 1:50,000 scale and was used by Ander et al. (2011), along with mineralisation and historical mining databases, to determine 'normal' (background concentrations) across the UK for the purposes of managing contaminated land. In contrast, Sheppard et al. (2009) used a spatial system based on soil classifications to provide a trace element index to assess the sustainability of Canadian agriculture. Lado et al. (2008) used a geological database, along with various other databases including land cover, nightlights (as a measure of urbanisation), and infrastructure databases, to explain heavy metal concentrations across 26 European countries.

Cavanagh et al. (2015) provided a first attempt to develop nationally consistent estimates of background soil concentrations in New Zealand using existing data collated from different sources. This study evaluated the use of geologically related parameters from several spatial databases and statistical modelling to develop consistent national estimates of naturally occurring concentrations across New Zealand.

Specifically, naturally occurring concentrations were based on a rock-type parameter, 'Chemical4', which provided the ability to predict likely concentrations in locations for which no data are available. Chemical4 was based on *rock-group* from QMAP¹, which is a database of geological maps developed by GNS Science over the period 1993–2012 (Rattenbury & Isaac 2012). Chemical4 subdivided the Miocene and younger sedimentary rocks and sediments (Maui and Pakihi supergroups, Mortimer et al. 2014) of rock-group, forming 72 categories. Chemical4 was used as the basis for generating preliminary background concentration distribution (described by the effective median, 5th and 95th

¹ <http://www.gns.cri.nz/qmap>

percentile estimates) for the individual trace elements for the individual Chemical4 subgroups across most of New Zealand. Predictions for Chemical4 subgroups with few underlying samples ($n < 30$) were considered less reliable and for $n < 10$, unreliable. The current project builds on the approach adopted by Cavanagh et al. (2015), but uses additional variables to help identify factors influencing trace element concentrations across New Zealand.

2.1 Definitions

Numerous terms are used to define the 'background' concentrations of chemical substances in soil, including 'normal', 'typical', 'baseline', 'ambient', 'characteristic', 'natural', 'background', and 'widespread'. These terms are often used interchangeably or can be defined differently in different contexts (see Matschullat et al. 2000, Reimann & Garrett 2005, and Reimann 2005 for more detailed discussion).

For the purposes of this report the following definitions are used.

- *Ambient background*: the concentrations of chemical substances in the environment that are representative of the area surrounding the site that are not attributable to a single identifiable source. This can include contaminants from historical activities and widespread diffuse impacts (e.g. fallout from motor vehicles). This is referred to as 'normal' concentrations in the UK (Defra 2012).
- *Natural background*: the concentrations of naturally occurring elements derived/originating from natural processes in the environment as close as possible to natural conditions, exclusive of specific anthropogenic activities or sources. The term may also be referred to as the 'geochemical background', attributable to mineral content derived from parent materials, other natural processes such as volcanic ash dispersal and geothermal inputs and the influence of soil-forming processes. This is also the definition of 'background' used in *Contaminated Land Management Guideline No. 5* (MfE 2021).
- *Baseline*: the soil concentrations of chemical substances in a specified location at a given point in time. Baseline concentrations are analogous to natural background concentrations where the specified locality is not influenced by diffuse anthropogenic sources, or ambient concentrations when the specified locality is influenced by diffuse anthropogenic sources. In contrast to ambient and natural background concentrations, baseline concentrations also include concentrations in locations known to be influenced by land use (e.g. agricultural land use).
- *Threshold*: the upper limit of background variation (Reimann & Garrett 2005).

3 Objectives

The specific objectives of this project were to:

- develop revised estimates of background concentrations of naturally occurring trace elements in soils using an extended data set

- compare these revised estimates with existing regional studies of background concentrations in soils (Auckland, Wellington, Christchurch)
- provide guidance on the use of background concentrations in soils, including:
 - the use of background soil concentrations to adjust Eco-SGVs, including when and where this might be appropriate
 - other uses of background concentrations, including ambient vs natural background, and where and when might ambient background be most applicable for managing contaminants (including disposal of materials to clean fills)
 - site-specific determination of background concentrations
- provide recommendations for any additional monitoring to inform better estimates of background concentrations, including for specific areas (e.g. mineralised areas, urban areas).

4 Methods

4.1 Data

Three data sets were used to develop estimates of background concentrations:

- New Zealand geochemical baseline data set
- regional council state of the environment soil quality data
- regional background studies.

These are described further below.

4.1.1 New Zealand soil geochemical baseline data set (GNS)

This data set combines soil samples from pre-existing surveys undertaken by GNS or Manaaki Whenua – Landcare Research (Martin et al. 2023a). Over 95% of the samples were collected after 2011; the remaining samples, collected between 1960 and 2011, were included from otherwise under-represented regions of the country.

Samples in this data set were collected by hand, typically using soil corers, augers, or hand trowels to collect a near-surface sample. Sample depths necessarily varied between surveys, but 10 of 20 surveys sampled the top 10 cm, nine sampled the top 20 cm, and one sampled the top 30 cm. Samples were oven dried at 40° C and sieved to retain the sub-2 mm portion.

The surveys determined element concentrations in order to map their variation across the country. A single instrument was used to acquire a suite of 65 elements; inductively coupled plasma mass spectrometry (ICP-MS) on an aqua regia solution were undertaken by Bureau Veritas Minerals Laboratories (BVML) in Vancouver (further details are provided in Martin et al. 2023a). Only the data for As, boron (B), Cd, chromium (Cr), copper (Cu), lead (Pb), nickel (Ni), and zinc (Zn) are considered here.

4.1.2 Regional council state of the environment soil quality monitoring (RC)

Data from regional council state of the environment (SOE) soil quality monitoring programmes were used as an additional data set. These data were recently compiled for national reporting, and Cavanagh et al. (2020) provide a detailed description of the data collation. Briefly, samples collected for soil quality monitoring were typically collected following Hill & Sparling 2009, whereby approximately 25 subsamples (0–10 cm) are collected along a 50 m transect to form a single composite sample. Samples were typically extracted using nitric and hydrochloric acid, followed by analysis by inductively coupled optical emission spectroscopy (ICP-OES) or inductively coupled plasma mass spectroscopy (ICP-MS).

4.1.3 Regional studies on background concentrations (ACWB)

Data from studies to determine background soil concentrations in Auckland (ARC 2001), Wellington (URS 2003), and Canterbury (Tonkin & Taylor 2006, 2007) were used as the final data set. Sampling included single point samples, or composite samples comprising subsamples collected from the centre and points of a grid around a 10–20 m square. Samples were typically extracted using nitric and hydrochloric acid, followed by analysis by ICP-OES, ICP-MS, or atomic absorption spectroscopy (AAS).

4.1.4 Data processing

Data from the three data sets were combined and processed to ensure consistent land-use classification, resulting in 2,146 rows of data for further analysis. A summary of the contribution of the different data sets is shown in Table 1.

Table 1. Number of data points, by land use and data source

Land use	Detailed regional data	GNS sampling	Regional council soil quality data	All sources
Background	284	358	134	776
Exotic Forest	2	32	110	144
Grazing (Dairy and Drystock)	3	318	508	829
Perennial Cropping (Orchards, Vineyards)	0	32	106	138
Short-rotation Cropping	0	81	124	205
Unknown	0	0	5	5
Urban	0	4	35	39
Wetland	0	10	0	10
All NZ	289	835	1,022	2,146

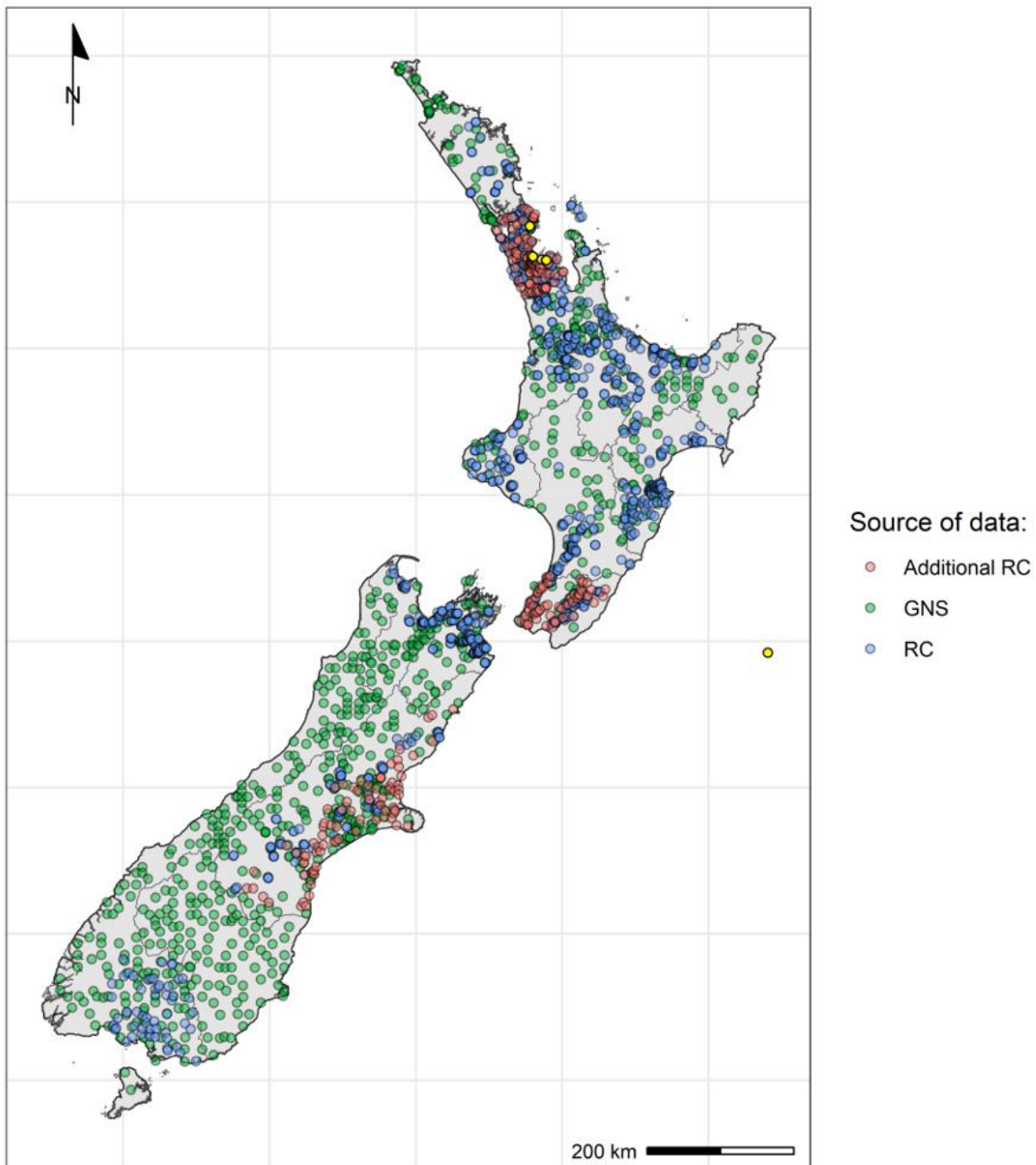


Figure 1. Locations of all data, shown as a function of the source of the data. GNS is the extended dataset, RC is the collated SOE data set, and additional RC refers to the regional background concentration studies. Yellow points indicate those sites that are identified as being in the water; the point located far to the East of New Zealand is evidently faulty.

For the analysis, land uses in Table 1 were further consolidated via spatial analysis to provide a rural ambient class, which combined 'background,' 'Forestry,' and 'Grazing' that was not 'High Producing Exotic Grassland', based on the LCDB version 5 land-use class. A rural ambient concentration was used to recognise the extent of anthropogenic activities across New Zealand, but the land uses included are anticipated to have minimal anthropogenic trace element inputs, and so these rural ambient concentrations are expected to differ little from 'true' naturally occurring concentrations. In an urban environment, ambient trace element concentrations – primarily lead from the historical

use of leaded petrol, and, rarely, other trace elements associated with industrial emission – might reasonably be expected.

Using this approach, 1,074 sites are rural ambient, which is 50.3% of the total number of sites. Figure 2 shows a plot of the rural ambient points alongside sampling points from other land uses.



Figure 2. Distribution of sampling sites based on final land-use category.

4.2 Explanatory variables

The site-based trace element data were complemented by data from three GIS-based polygon spatial databases that regionally and/or nationally delineate areas that are interpreted as having distinct common pedological and/or geological properties. These were:

- parent material
- soil order
- LENZ class.

4.2.1 Parent material layer

The parent material map was generated using a combination of sources. The geological survey of New Zealand, QMAP, was used as a main parent material baseline, using its ROCKCLASS attribute. QMAP² is a database of geological maps developed by GNS Science over the period 1993–2012 (Rattenbury & Isaac 2012). This includes a national seamless GIS component based on the 21 published geological maps at 1:250,000 (Heron 2014).

ROCKCLASS information was combined with additional information from the New Zealand Land Resource Inventory (NZLRI)³ in order to add information about peat, loess, and volcanic ash deposits, which are critical to soil formation but typically not included in QMAP. Some ROCKCLASS attributes were grouped to generalise the parent material classes, with grouped attributes expected to have similar trace element distributions. The parent material classes and descriptions are as follows.

- The 'Ultramafic Igneous' class groups the 'Volcanic', 'Ultramafic Intrusive', and 'Tectonic' classes of QMAP.
- The 'Limestone' class groups the 'Limestone' and 'Chemical Sediment' classes of QMAP.
- The 'Felsic Igneous' class groups the 'Intermediate Intrusive', 'Intermediate Extrusive', 'Felsic Intrusive', and 'Felsic Extrusive' classes of QMAP. Tephra was also included by the different volcanic ash classes mapped as part of the NZLRI.
- The 'Mafic Igneous' class groups the 'Mafic Intrusive' and 'Mafic Extrusive' classes of QMAP.
- The 'Gneiss' class was renamed from the 'Metamorphic' class in QMAP.
- The 'Sediment' class was renamed from the 'Clastic Sediment' class in QMAP.
- The 'Loess' class was sourced from the NZLRI.
- The 'Peat' class was created from a combination of S-map and the NZLRI.
- The 'Schist' class was sourced from the 'Schist' class in QMAP.

² <http://www.gns.cri.nz/qmap>

³ <http://lris.scinfo.org.nz/>

The distribution of parent material classes across New Zealand is shown in Figure 3.

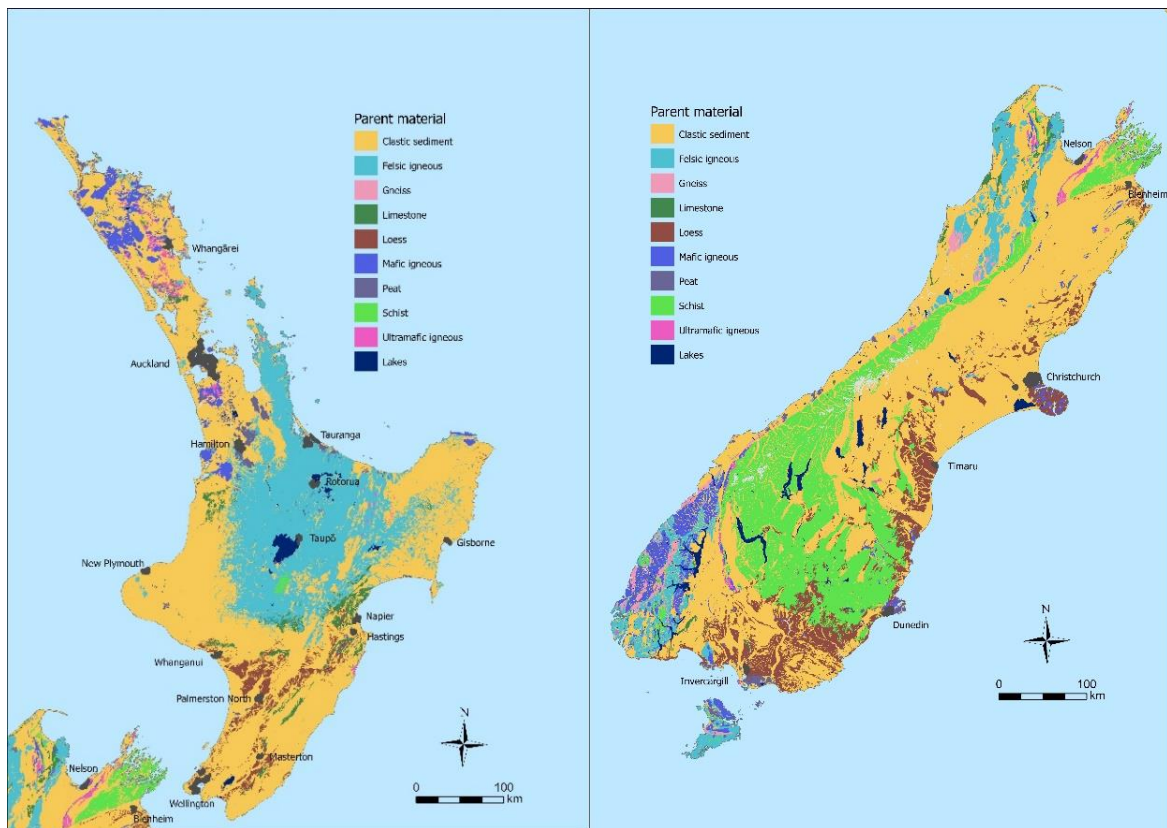


Figure 3. Distribution of parent materials across New Zealand.

4.2.2 Soil order

Determination of soil order was made by combining information from S-map and the Fundamental Soil Layers (FSL) for areas where S-map was not available. S-map⁴ is a spatial database for New Zealand soils designed to provide quantitative soil information for modellers, and it includes links to the National Soils Database (NSD) and Q-Map (Lilburne et al. 2012). Data are mapped at a 1:50,000 scale, or finer in some locations.

The FSL is housed within LRIS (The Land Resources Information System⁵), which is a way for the public to access environmental data held by Manaaki Whenua – Landcare Research. The FSL describes land on the basis of five characteristics (including soil order) and is also mapped at a 1:50 000, scale. The distribution of soil orders nationally is shown in Figure 4.

⁴ <http://smap.landcareresearch.co.nz/home>

⁵ <http://lris.scinfo.org.nz/>

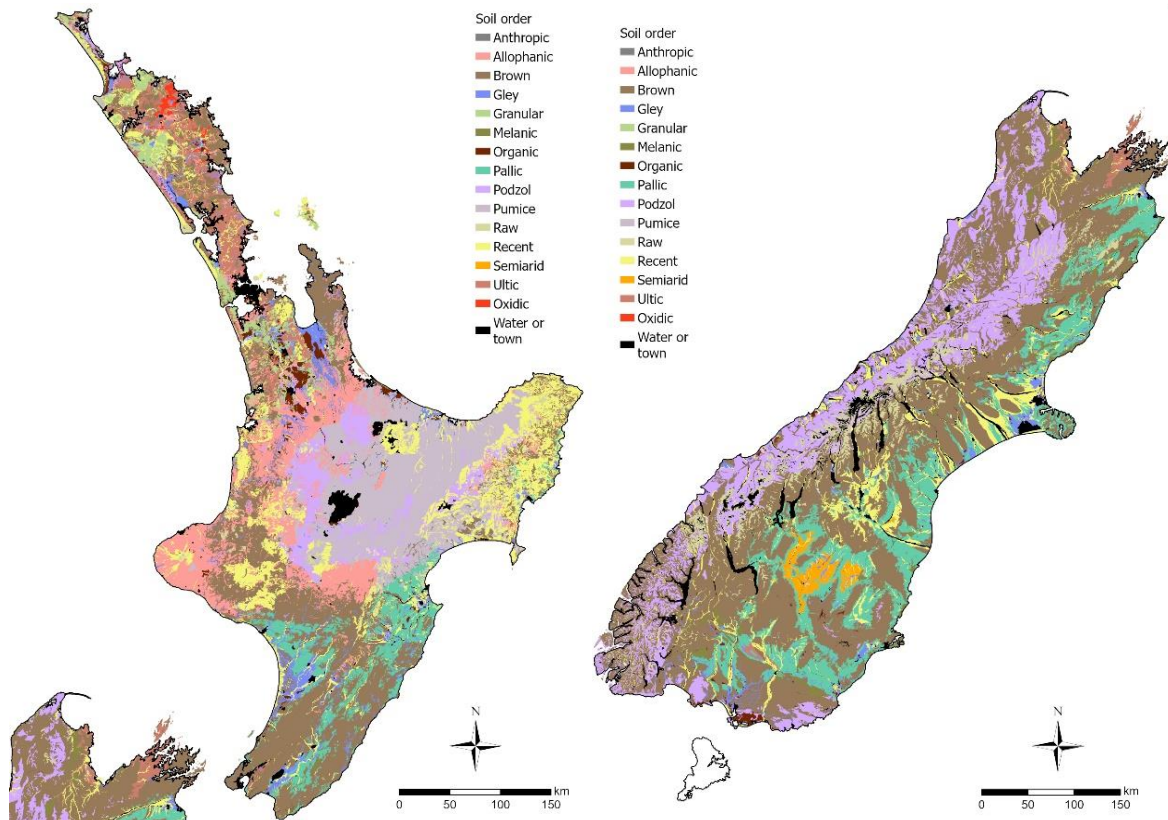


Figure 4. Distribution of soil orders across the North and South Island of New Zealand.

4.2.3 Land Environments of New Zealand

Soil formation (and hence trace element concentrations) can be influenced by climatic conditions. Land Environments of New Zealand (LENZ)⁶ provides a different perspective for considering baseline concentrations of trace elements in soil, describing ecosystems as a function of various aspects of New Zealand's climate, landforms, and soils (Leathwick et al. 2002). Some LENZ classes are associated with distinct soil types and geologies, in particular Ultramafic soils.

At a finer scale, many rare ecosystems⁷ are also associated with distinct geologies, including geothermal areas, and inland and alpine areas with skeletal or poorly developed soils, including boulderfields of calcareous rock, granitic sand plains, and ultrabasic hills (Wiser et al 2013). The distribution of Level 1 LENZ classes across New Zealand is shown in Figure 5.

⁶ <https://www.landcareresearch.co.nz/tools-and-resources/mapping/lenz/>

⁷ <https://www.landcareresearch.co.nz/publications/naturally-uncommon-ecosystems/>

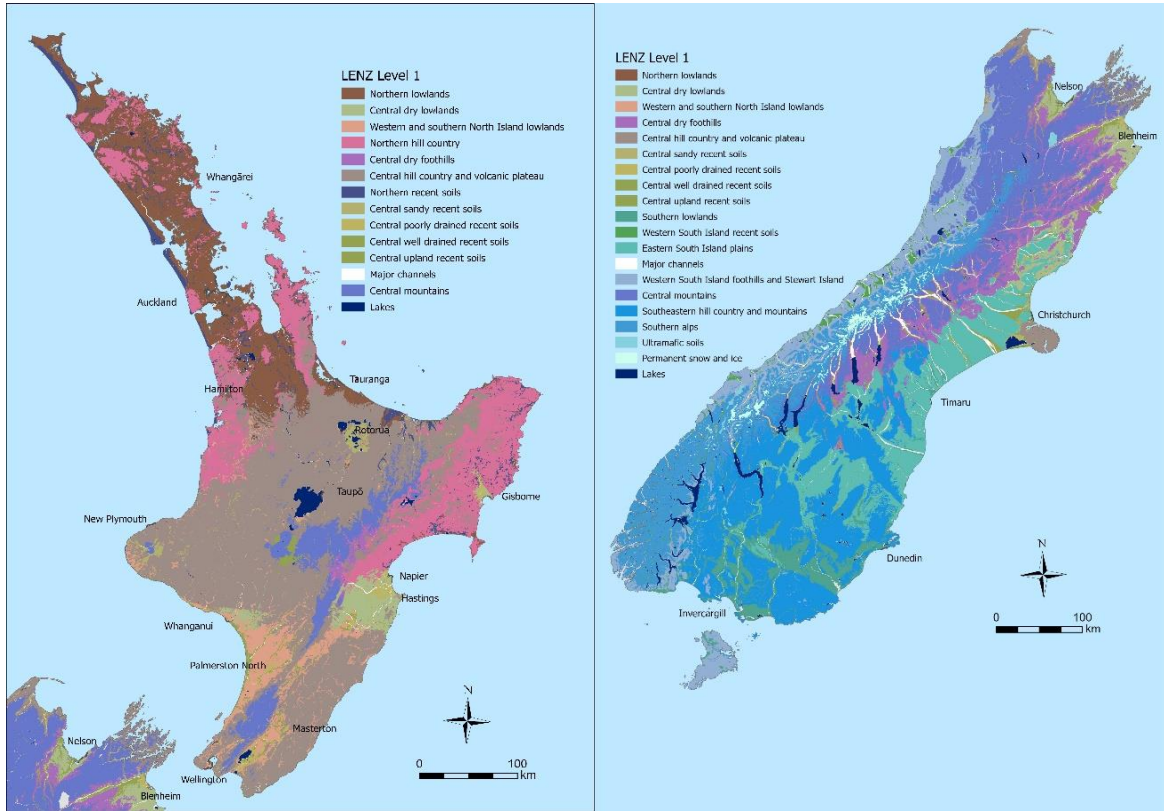


Figure 5. The distribution of Level 1 LENZ classes across the North and South Island of New Zealand.

4.3 Statistical modelling

All statistical analysis was undertaken using R version 4.2.0. The model used was a generalised additive model (GAM) (Hastie & Tibshirani 1990). A GAM is a semi-parametric statistical regression modelling approach in which the response variable (here, the trace element concentration) depends on a combination of covariates, as well as smooth functions of some predictor variables. The model is defined as follows:

$$\log(\mu_{i,j,k,l}(x,y)) = a_0 + \text{LENZ}_i + \text{soilorder}_j + \text{parentmaterial}_k + \text{landuse}_l + h(x,y) + \epsilon$$

where $g(\mu_{i,j,k,l}(x,y)) \equiv \mathbb{E}(Y_{i,j,k,l}(x,y))$ and $Y_{i,j,k,l}(x,y) \sim \text{Tweedie}(\mu_{i,j,k,l}(x,y), \phi)$.

$Y_{i,j,k,l}(x,y)$ is the response variable (the trace element concentration), which is modelled as a Tweedie distribution with mean $\mu_{i,j,k,l}(x,y)$ for LENZ class i , New Zealand Soil Classification (NZSC) soil order j , parent material k , land-use class l , and map location (x,y) . The model includes a smooth function for spatial position ($h(x,y)$). It is additive, but there are implicit interactions that occur spatially, since LENZ, the NZSC soil order, and other covariates all change with spatial location. Finally, ϵ is the uncertainty in the model, which is assumed to be Gaussian, distributed in the scale of the log-scaled mean.

The GAM allows for flexible specification of the dependence of the concentration of the trace element on the covariates, but by specifying the response in terms of the smooth

functions rather than detailed parametric relationships it is possible to avoid the more cumbersome and unwieldy models that would need to be used to present the same model in the form of a generalised linear model (GLM), for instance. This flexibility comes at the cost of the complexity involved in defining the smooth functions, and the computational complexity of deciding how smooth they ought to be (Wood 2017).

It should be noted that the trace element response was modelled as a Tweedie distribution after first investigating using a Gamma distribution. While models using the Gamma distribution were generally satisfactory, small trace element values were often poorly modelled. The Tweedie distribution is a generalisation of the Gamma model, permitting a more flexible fitting to the data, especially for small trace element values.

One difficulty encountered was that an important explanatory variable, NZSC soil order, is missing from both the FSL and S-map for the Auckland region and a small number of other locations. Since the soil order has an important explanatory effect it was important to retain it in the model where it was available. To solve this difficulty, a second model was fitted for the trace element concentration using the same explanatory variables, but omitting NZSC soil order. In this way, predictions are available for areas where the NZSC soil order is available as well as those where it is not.

To some extent the GAM fitted without the NZSC soil order will accommodate the missing variable by including the spatial location term. In regions where the soil order has an important effect but the soil order is not available (e.g Auckland) the spatial effect is increased in importance. However, it is unlikely that all of the information available from the spatial location of the measured points would be sufficient to entirely replace the information from soil order, so some degradation from the model that includes soil order might be expected. The loss in accuracy depends on the trace element being modelled, as well as the complexity of the soil order dependence.

For example, for Zn the percentage of the null deviance explained using a model including soil order was 26.4 %, while the same figure for the model excluding soil order was 20.0 %. Tests conducted at 'overlapping' spatial locations south of Auckland where soil order is available ie where predictions can be made using both models, suggest differences in predicted Zn concentration in the order of 10% of the predicted values between the models. This difference is smaller than the size of the uncertainty in predicted Zn, so the difference between the models is considered acceptable.

The general process used for all trace elements is as follows.

- 1 A Tweedie GAM is fitted using the following form (where Y is the trace element value):

$$\log(Y_{i,j,k,l}(x,y)) = a_0 + \text{LENZ}_i + \text{soilorder}_j + \text{parentmaterial}_k + \text{landuse}_l + h(x,y).$$

- 2 The Cook's distance measure is inspected to identify highly influential outliers. Once influential outliers are removed, the model is re-fitted until there are no further influential points of interest (an example is shown in Figure 6).
- 3 Once a satisfactory model has been obtained with no influential outliers, the residual quantile-quantile plot is inspected to check that the residuals are Gaussian distributed (an example is shown in Figure 6).

- 4 Measured-versus-fitted plots are generated, both in linear- and log-scaled form, and used to identify problems (the log-scaled form is shown for each trace element in the results).
- 5 A spatial plot of residuals is generated to identify spatial clustering of samples with high residual error (Figure 7).
- 6 Once all the above diagnostics have been used to identify and correct problems, a final spatial plot of the predicted log-scaled trace element value is produced and is shown for individual trace elements in results).

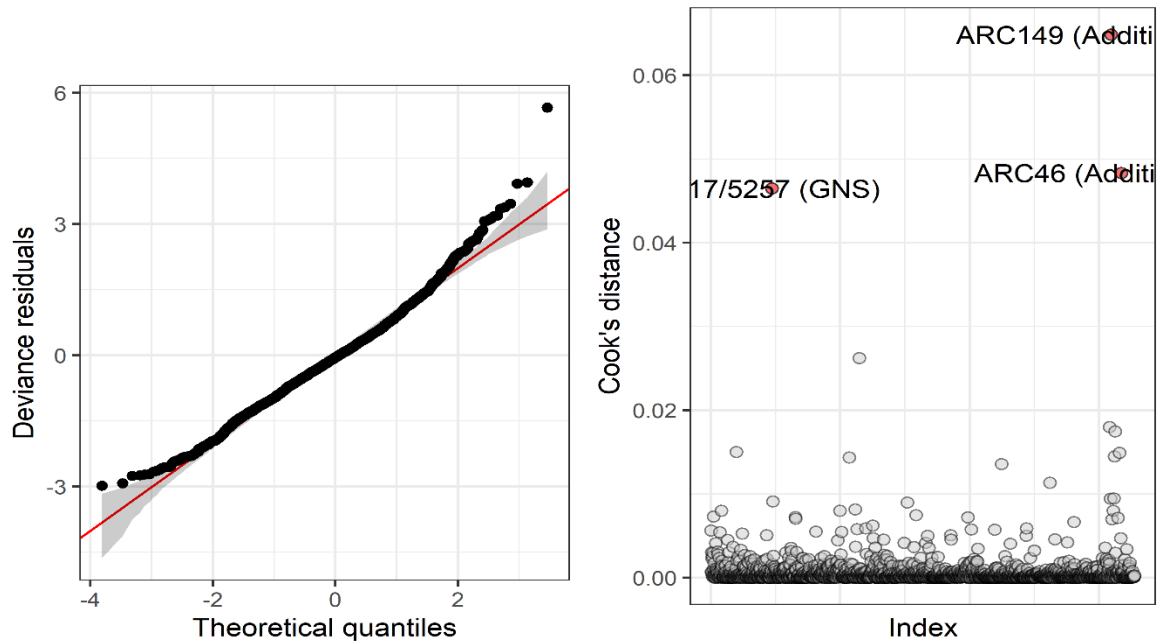


Figure 6. Diagnostic plots for the GAM zinc model. Left: quantile–quantile plot of the deviance residuals, with the 95% confidence region shaded; the red line is the one-to-one line. Right: Cook’s distance plot for all points used in the model fit; the largest three values are highlighted in red, with their identifications shown as labels.

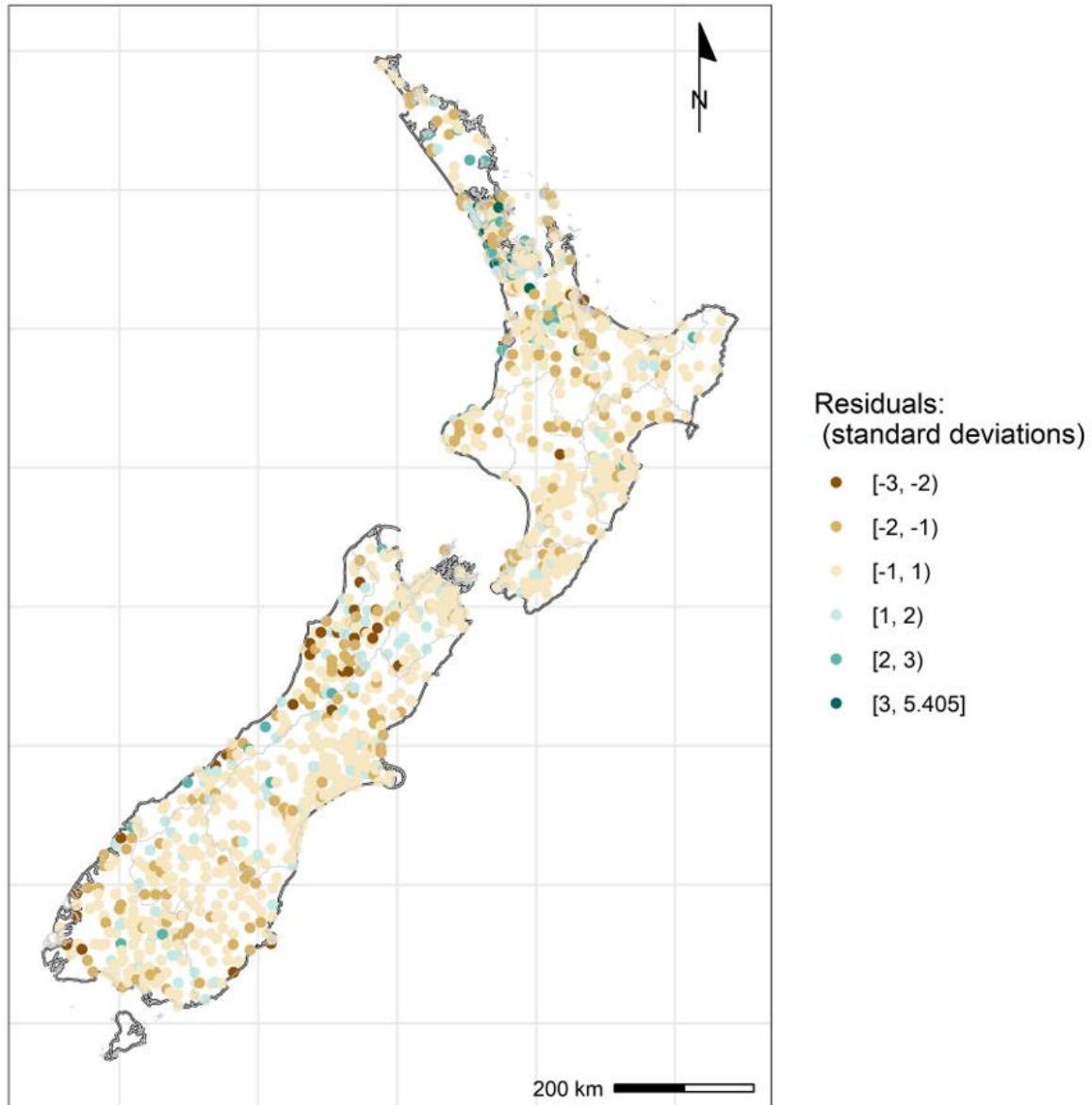


Figure 7. Residual plot of all points used to fit the GAM for zinc. The colour of the point indicates the size of the residual. Extreme positive or negative residuals (red or blue) indicate potential issues of concern.

Typically only a few samples need to be removed from the sample data to generate a satisfactory model (e.g. one without large Cook's distance values). It is notable that the same outliers occur frequently for different elements. Table 2 gives a summary table of results for modelled trace elements, including basic information from the point data (e.g. minimum, median). Also included is the number of influential outliers that were identified, and the maximum value of the Cook's distance for the model after influential outliers have been removed.

Table 2. Summary results for modelled trace elements showing minimum, median, and maximum concentrations (mg/kg) of the data set, and measures of model fit: R² and percentage of the null deviance explained. The last column is the maximum value of the Cook's distance for the model after influential outliers have been removed.

Element	N	Minimum	Median	Maximum	R ²	Percent of null deviance explained	Number of outliers	Max. Cook's distance
As	1,634	0.1	3.8	58	9.37%	20%	1	0.17
B	1,178	0.03	4	840	13%	48.60%	2	10.11
Cd	1,722	0.005	0.12	1.7	55.50%	58.90%	5	0.057
Cr	1,643	2	15	660	49.90%	55.80%	3	0.104
Cu	1,648	1	14	460	16.40%	34.30%	3	0.481
Ni	1,638	0.7	8.8	720	42.70%	51.60%	5	0.084
Pb	1,640	0.4	11	79	28.90%	39.20%	20	0.042
Zn	1,638	4	53	330	20.90%	26.40%	0	0.065

The measures of model quality in Table 2 (R², percentage of null deviance explained) should be interpreted with some care. The definition of R² is specific to the type of model (e.g. linear model, GLM, GAM), and there is no universally accepted definition for models other than the linear model. For the GAM models used here (fitted using the R mgcv package) the value quoted is the adjusted R². This value gives a very broad indication of the quality of the model fit that attempts to account for multiple predictors, but it can be strongly affected by the quality of the fit for very small values and may not be the most appropriate measure by itself. It should be interpreted in conjunction with diagnostic plots (e.g. a measured-versus-fitted plot in Figure 8) and other measures of model quality.

The percentage of the null deviance explained in Table 2 is generally preferred as a metric of GAM model quality. The deviance measures the discrepancy between the measured and fitted values using a definition that is specific to the assumed trace element distribution. The null deviance measures the overall discrepancy when no explanatory variables are included, and the percentage of null deviance explained in Table 2 indicates the proportion of the null deviance that is accounted for when the explanatory variables are included. Like the R², the percentage of null deviance explained is in the range 0–100 %. Values larger than 50% might be considered moderate, while values less than 25–30% would be considered poor, but these are rather broad figures.

The Cook's distance measures the influence of each measurement in the model, which is important for detecting outliers that might strongly affect the regression. Table 2 gives the maximum value of Cook's distance for all measurements used in each model; larger values *might* indicate potential outliers. Generally, the values in Table 2 must be interpreted along with other diagnostic plots for all points. In this report these diagnostic plots are not included, for brevity, but the diagnostic plots for all trace elements were inspected manually.

Generally, the values for the percentage of null deviance explained and the R^2 should be roughly similar. If the value of the R^2 is very small compared with the percentage of null deviance explained (e.g. As in Table 2), this would tend to indicate a problem for very small trace element values, which can be seen in Figure 8 for As. If a single measure of model quality is to be used, the percentage of null deviance explained should be adopted.

For each trace element the following information is provided.

- A summary table is provided that assesses the significance of the various explanatory variables - as the effect of location (included as a spatial effect in the GAM fit) is highly significant for all trace elements, this explanatory variable is not included in each section. In addition to the table, figures showing the measured versus fitted plot of trace element concentration, the ambient trace element concentration, and filled contour plots of trace element concentrations. The summary tables (e.g. Table 3) provide an indication of whether an explanatory variable has a significant effect in the model, but does not indicate which level(s) of that variable are significant. The size of the p-value in these tables should not be given as an indication of the strength of an effect – the variable is either significant or it is not significant, depending on whether the value is less than 0.05 or not, respectively.
- The predicted vs actual concentrations with log axes for all land-use classes are also shown; the error bars are plus and minus one standard error.
- The predicted concentrations for the ambient land use across New Zealand is presented at a 1 km × 1 km resolution. This map also includes important information concerning the model – the number of points, the minimum, median, and maximum, the AIC for the model, the R-squared, and the percentage deviance explained – noting, as discussed above, that the percentage deviance explained is a better measure for the model, since it effectively uses the log-transformed concentration as the response.
- To aid with the application of the predicted estimates in the management of contaminated land or other purposes, the information on predicted concentrations is also presented as contour plots, with contours at the 50, 90, 95, and 99 percentiles of the predicted values. Thus, within each 'band' the concentration is anticipated to fall within, for example min to 50th percentile, or 50th to 90th percentile. A table in the contour maps presents the specific concentrations for each percentile cut-off. This map is presented for the whole of New Zealand, as well as a separate map for the Auckland Region.

5 Results

5.1 Background soil concentrations

5.1.1 Arsenic

Table 3 indicates that LENZ class, parent material, soil order and land use, have a significant influence on determining As concentrations across New Zealand. Further detail on the relative influence of individual variables within each explanatory factor is provided in Appendix 1. There is a relatively low deviation explained (i.e. comparatively poor model fit) for the model (20%, Figure 9 Left). The predicted vs fitted model shows an under-prediction of concentrations at the higher end of measured concentrations – from c. 6 mg/kg (Figure 8). Given the 5th to 95th percentile range across the whole data set is also small (i.e. 0.7–10 mg/kg) and similar to the rural ambient concentrations of 0.4–10 mg/kg, the under-prediction probably falls within sampling and analytical error.

A greater difference is seen at the extreme end of the estimates, where the 99th percentile of the measured data (19.8 mg/kg, $n = 631$) is more similar to the predicted maximum of 18.7 mg/kg than the predicted 99th percentile of 8 mg/kg. The predicted maximum is similar to the 99th percentile background concentration of 17 mg/kg ($n = 372$) determined by MfE (2011).

Predicted spatial estimates of As concentrations in the rural ambient class are presented as log-transformed point (1 km × 1km) estimates (Figure 9 Left) and concentrations based on percentile distribution (Figure 9 Right) show elevated concentrations of As are seen within the Waikato and Otago regions. A higher-resolution image of the area around Auckland is shown in Figure 10.

Table 3. ANOVA table for arsenic, showing the significance of each explanatory variable for the GAM. A Wald test of the significance of term is performed, so the interpretation is like a Type-III ANOVA. The p-values for each term are calculated using the Bayesian estimated covariance matrix of the parameter estimators.

Variable	df	F	p-value
lenz	19	2.14	0.003
soilorder	13	6.63	< 0.001
parentmaterial	8	8.97	< 0.001
landuse_ambient	5	6.6	< 0.001

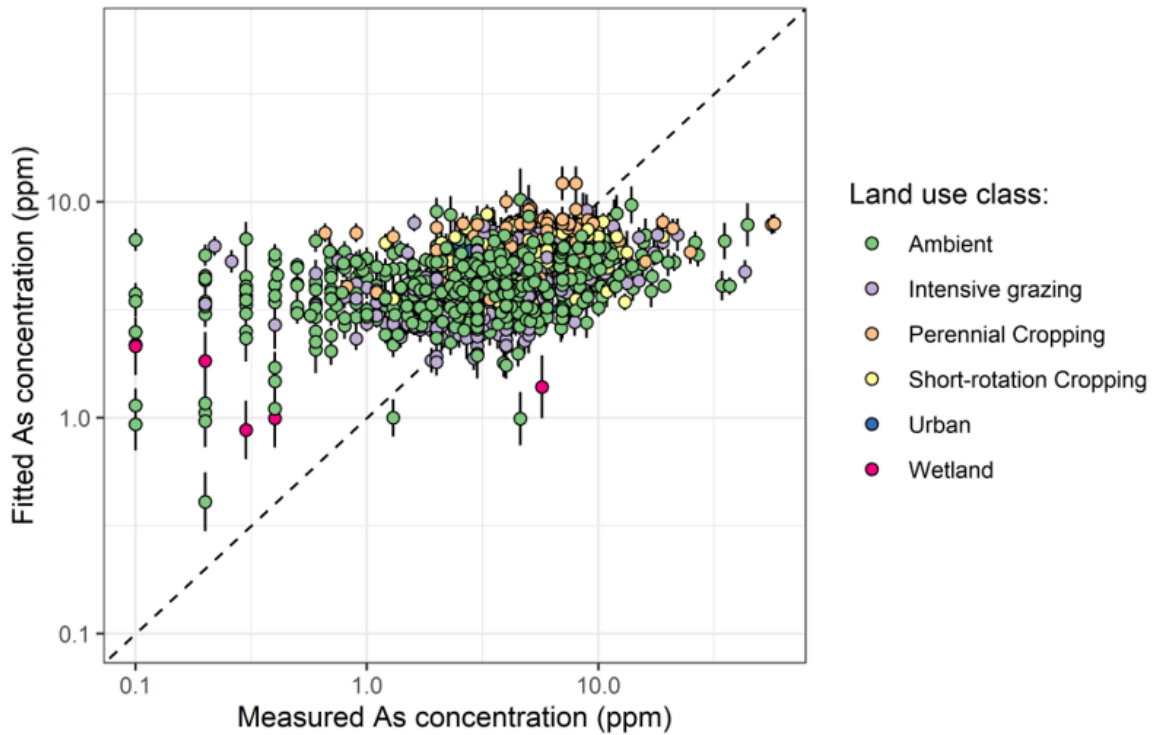


Figure 8. Measured versus fitted plot of arsenic concentration, with both axes log scaled. The points are coloured by the land-use class.

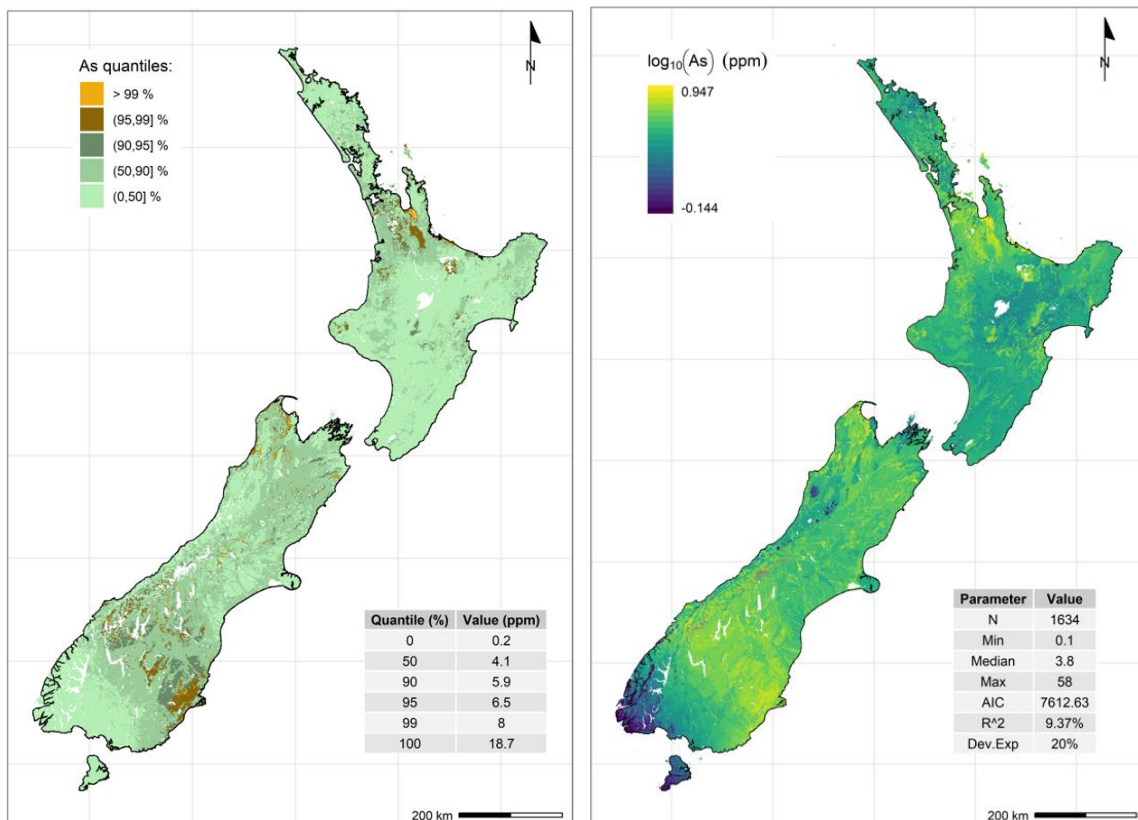


Figure 9. Left: Predicted log-transformed ambient arsenic concentration for all locations in New Zealand. The calculations are based on a grid with 1,000 m spacings. A table summarising the fitting data set is also included. Right: Filled contour plot of arsenic concentrations, with contours at the 50, 90, 95, and 99 percentiles of the predicted values. The table in the lower right shows the values associated with common quantiles, including the 0 and 100% quantiles (representing minimum and maximum estimates).

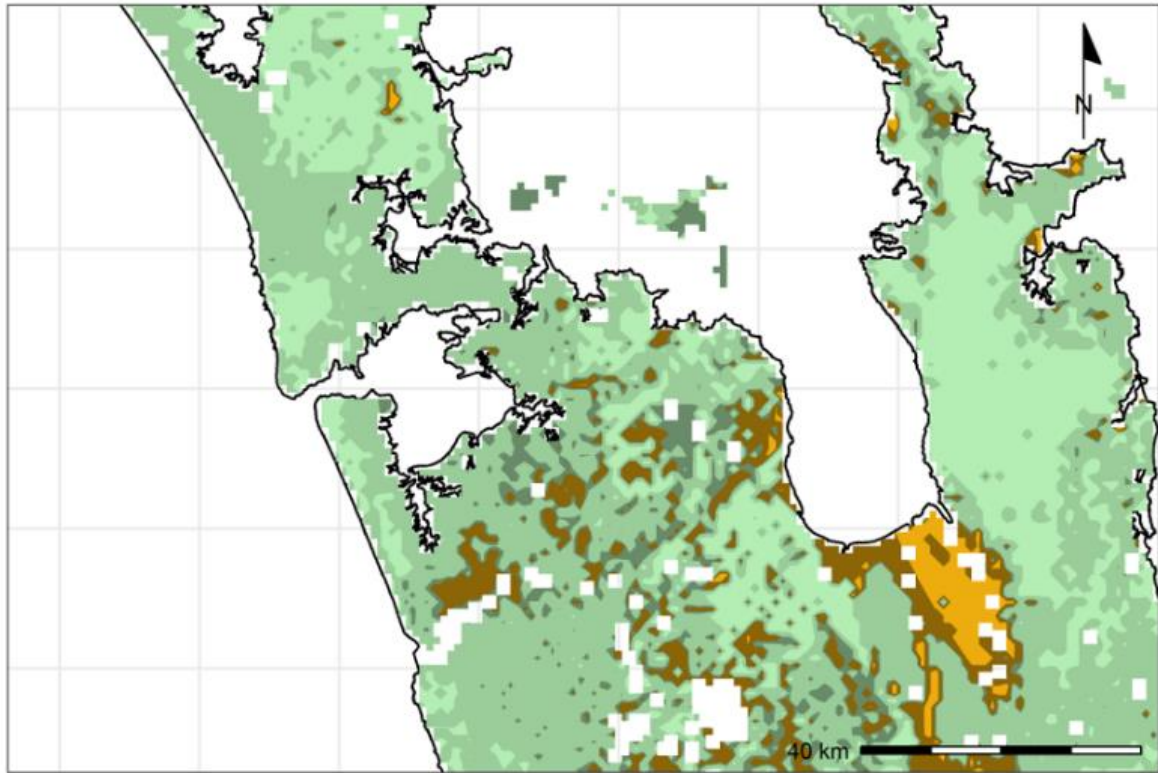


Figure 10. Filled contour plot of arsenic concentrations in the Auckland region, with contours at the 50, 90, 95, and 99 percentiles of the predicted values. Refer to Figure 9 for legend. White spaces indicates missing co-variate data.

5.1.2 Boron

Table 4 indicates that LENZ class, parent material, and soil order, has influence on determining B concentrations across New Zealand. Further detail on the relative influence of individual variables within each explanatory factor is provided in Appendix 1. Overall there is a reasonable fit of the data, as shown by the measured vs fitted model (Figure 11), and c. 49% of deviance is explained by the model (Figure 12).

Predicted spatial estimates of B concentrations in the rural ambient class are presented as log-transformed point (1 km × 1 km) estimates (Figure 12 Left) and concentrations based on percentile distribution (Figure 12 Right). Elevated concentrations (>99th percentile of predicted concentrations) of B are seen across the upper Waikato, Coromandel, Auckland and Northland regions, with elevations also seen along the East Coast of Hawke’s Bay and Gisborne. A higher-resolution image of the area around Auckland is shown in Figure 13.

Table 4. ANOVA table for boron, showing the significance of each explanatory variable for the GAM. A Wald test of the significance of term is performed, so the interpretation is like a Type-III ANOVA. The p-values for each term are calculated using the Bayesian estimated covariance matrix of the parameter estimators.

Variable	df	F	p-value
lenz	19	2.2	0.002
soilorder	13	2.3	0.005
parentmaterial	8	11.22	< 0.001
landuse_ambient	5	4.05	0.001

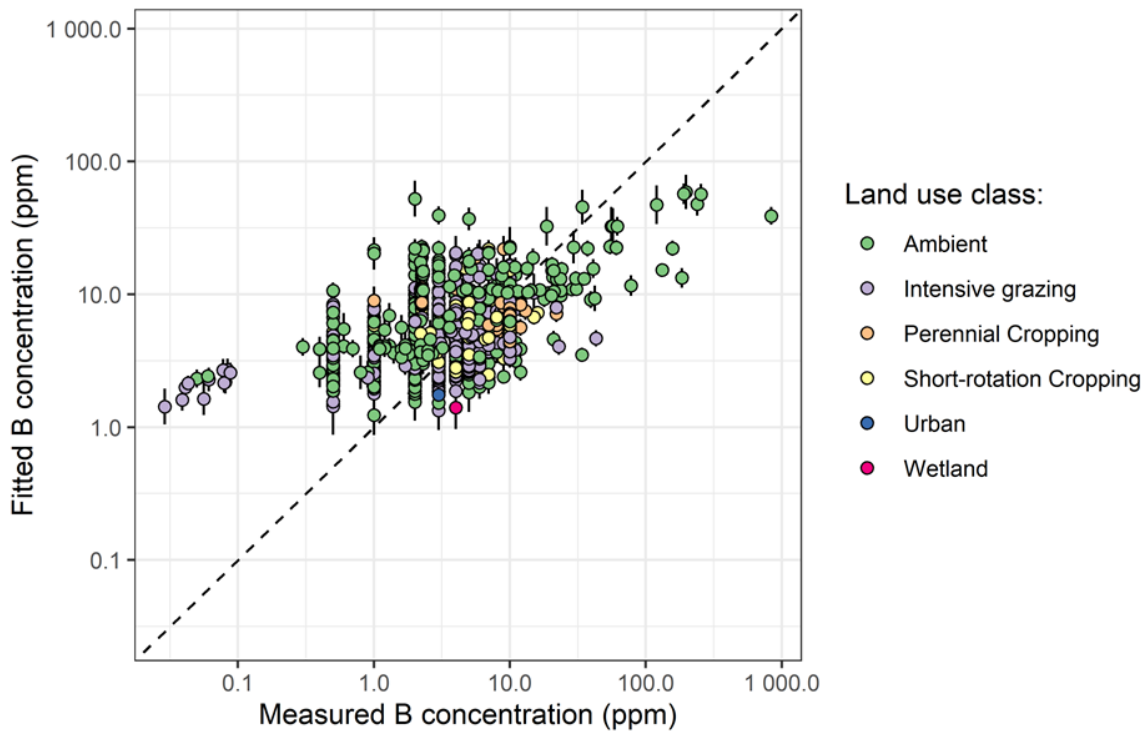


Figure 11. Measured versus fitted plot of boron concentration, with both axes log scaled. The points are coloured by land-use class.

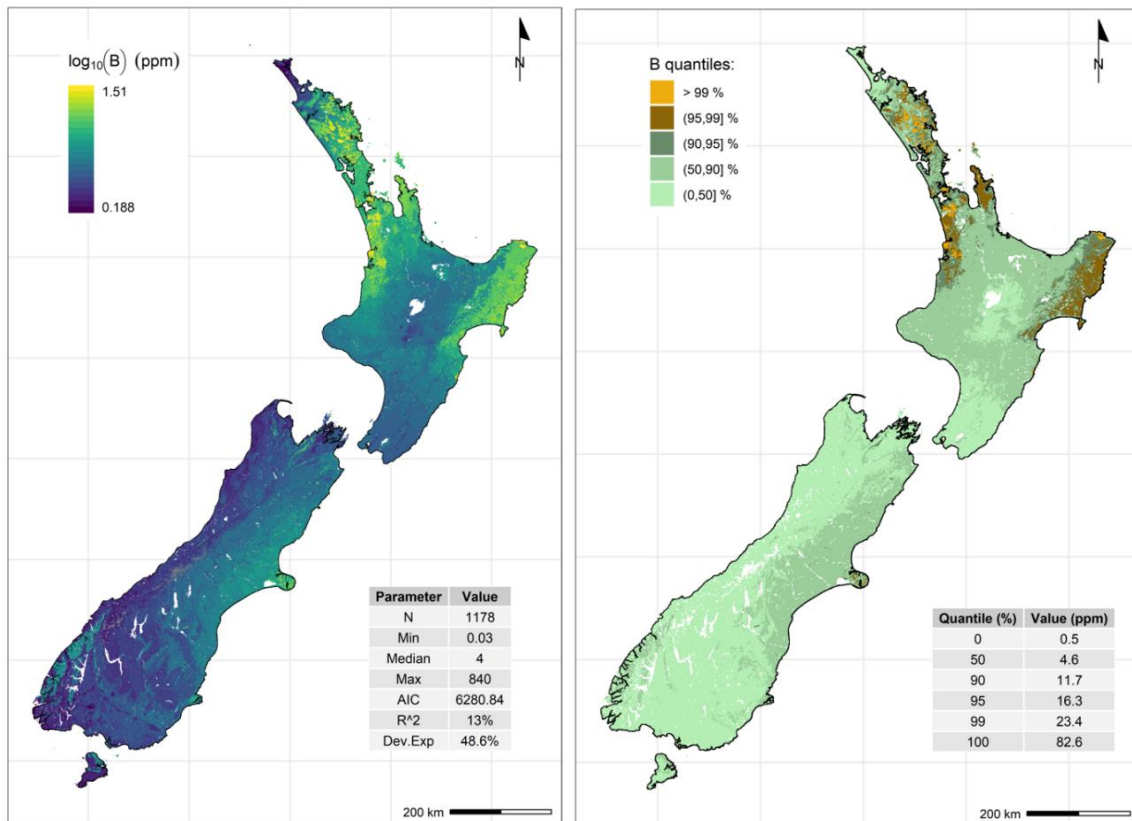


Figure 12. Left: predicted log-transformed ambient boron concentrations for all locations in New Zealand. The calculations are based on a grid with 1,000 m spacings. A table summarising the fitting data set is also included. Right: filled contour plot of boron concentrations, with contours at the 50, 90, 95, and 99th percentiles of the predicted values. The table in the lower right shows the values associated with common quantiles, including the 0 and 100% quantiles (representing minimum and maximum estimates).

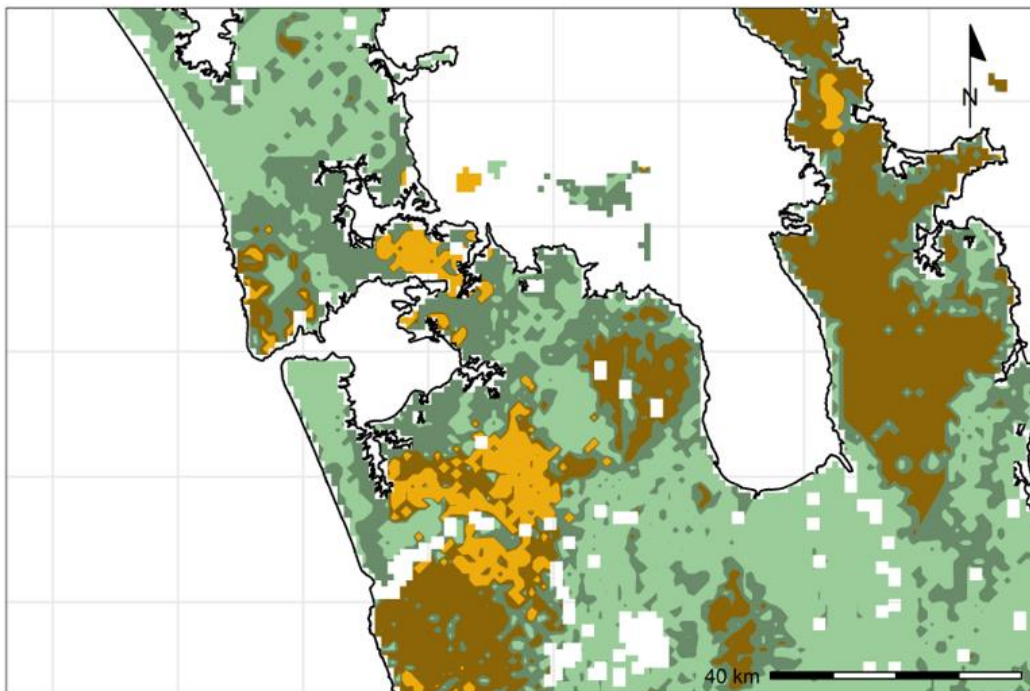


Figure 13. Filled contour plot of boron concentrations in the Auckland region, with contours at the 50, 90, 95, and 99th percentiles of the predicted values. Refer to Figure 12 for legend. White spaces indicates missing co-variate data.

5.1.3 Cadmium

Table 5 indicates that LENZ class, land use, soil order, and parent material, have a significant influence on determining Cd concentrations across New Zealand. Further detail on the relative influence of individual variables within each explanatory factor is provided in Appendix 1. Overall, there is a reasonable fit of the data, as shown by fitted vs measured (Figure 14), and c. 59% of the null deviance is explained by the model (Figure 15). The 5th–95th percentile range in Cd concentration across the whole data set is comparatively small (0.01 to 0.78 mg/kg), with a smaller range for the rural ambient of 0.01 to 0.28 mg/kg.

The 99th percentile of the measured rural ambient data (0.48 mg/kg, $n = 698$) falls between the predicted 99th percentile (0.4 mg/kg) and the predicted maximum of 0.6 mg/kg for the rural ambient. The latter is similar to the 99th percentile background concentration of 0.65 mg/kg ($n = 486$) determined by MfE (2011).

Predicted spatial estimates of Cd concentrations in the rural ambient class are presented as log-transformed point (1 km × 1 km) estimates (Figure 15 Left) and concentrations based on percentile distribution (Figure 15 Right). Higher concentrations are observed in the Taranaki and Waikato regions, which are also recognised to have higher Cd concentrations as a result of historical phosphate fertiliser use (Abraham 2020). A higher-resolution image of the area around Auckland is shown in Figure 16.

Table 5. ANOVA table for cadmium, showing the significance of each explanatory variable for the GAM. A Wald test of the significance of term is performed, so the interpretation is like a Type-III ANOVA. The p-value for each term are calculated using the Bayesian estimated covariance matrix of the parameter estimators.

Variable	df	F	p-value
lenz	19	7.81	< 0.001
soilorder	13	10.24	< 0.001
parentmaterial	8	4.72	< 0.001
landuse_ambient	5	69.95	< 0.001

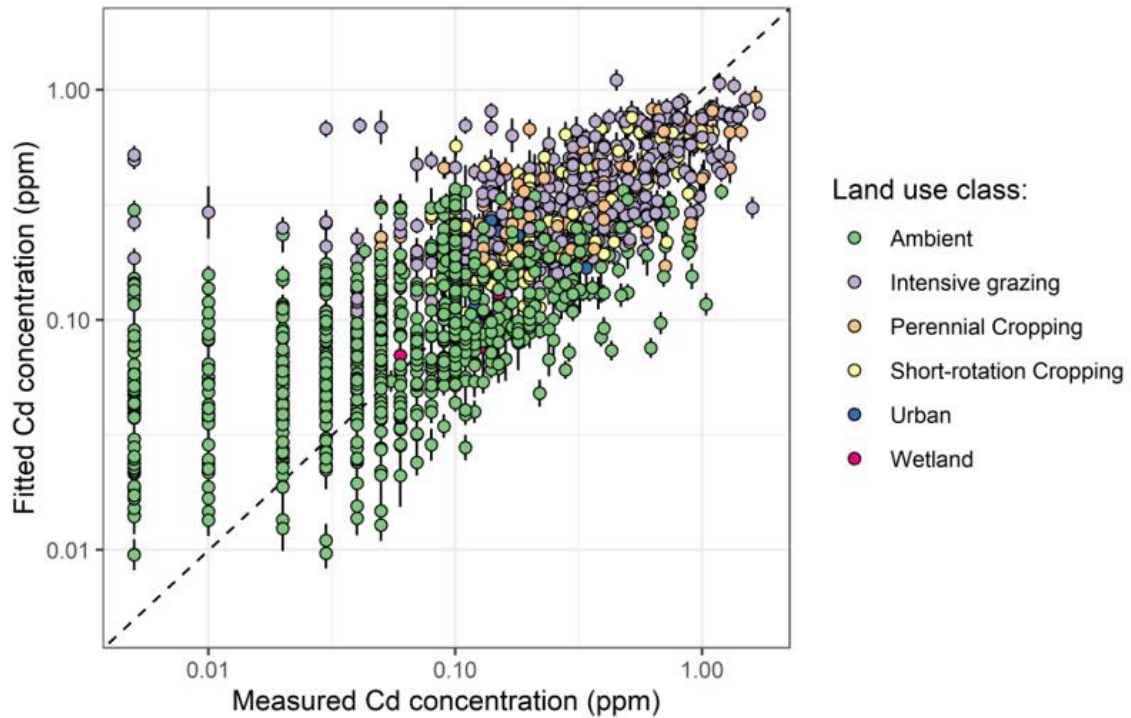


Figure 14. Measured versus fitted plot of cadmium concentration, with both axes log scaled. The points are coloured by the land-use class.

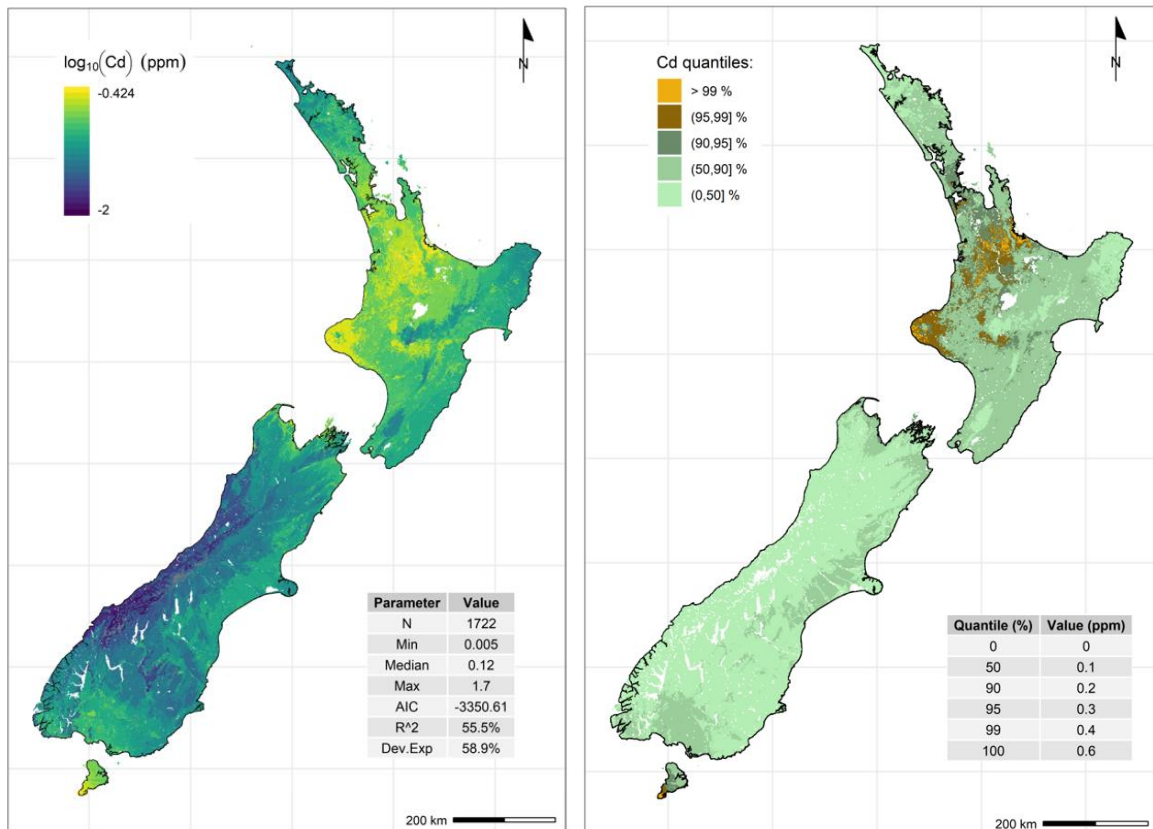


Figure 15. Left: predicted log-transformed ambient cadmium concentrations for all locations in New Zealand. The calculations are based on a grid with 1,000 m spacings. A table summarising the fitting data set is also included. Right: filled contour plot of cadmium concentrations, with contours at the 50, 90, 95, and 99th percentiles of the predicted values. The table in the lower right shows the values associated with common quantiles, including the 0 and 100% quantiles (representing minimum and maximum estimates).

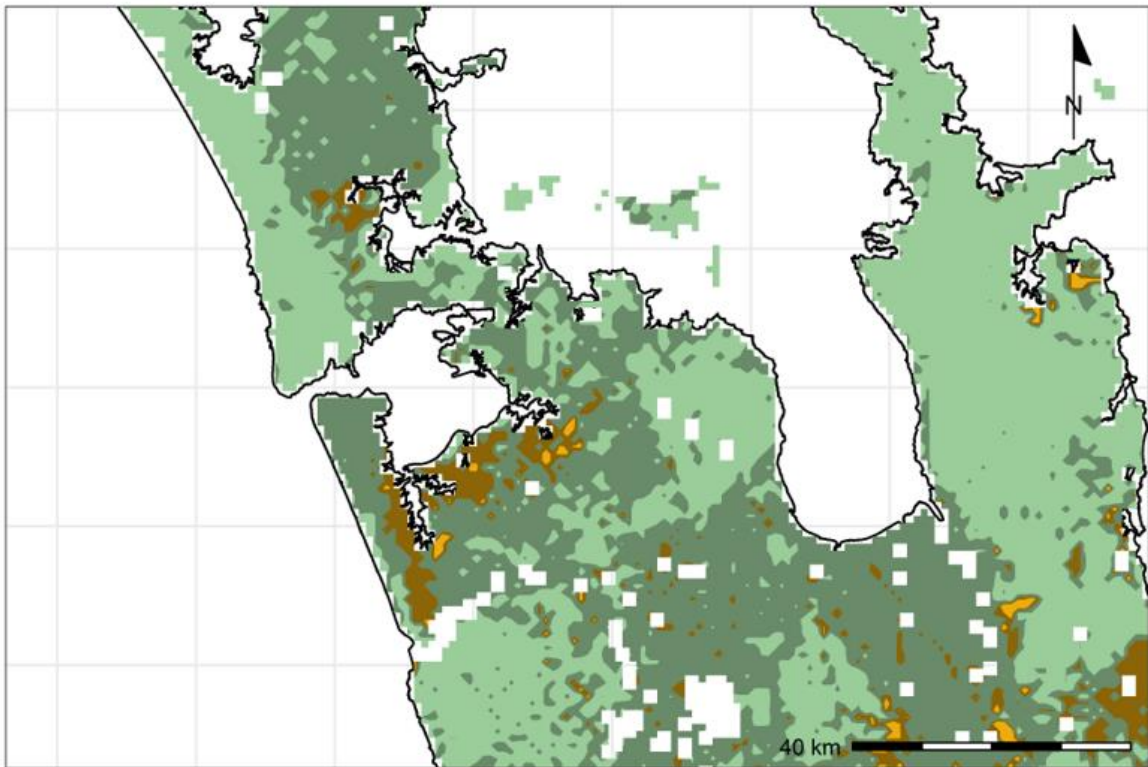


Figure 16. Filled contour plot of cadmium concentration in the Auckland region, with contours at the 50, 90, 95, and 99th percentiles of the predicted values. Refer to Figure 15 for legend. White spaces indicates missing co-variate data.

5.1.4 Chromium

Table 6 indicates LENZ class, soil order, parent material, and land use have a significant effect on Cr concentrations. Further detail on the relative influence of individual variables within each explanatory factor is provided in Appendix 1 and shows a marked influence of ultramafic igneous parent material. Overall, there is a reasonable fit of the data, as shown by fitted vs measured (Figure 17), and c. 56% of the null deviance explained by the model (Figure 18).

Predicted spatial estimates of Cr concentrations in the rural ambient class are presented as log-transformed point (1 km × 1 km) estimates (Figure 18 Left) and concentrations based on percentile distribution (Figure 18 Right). Elevated concentrations are found in Northland and Auckland, across the top of the South Island around Golden Bay, the Richmond mountain range near Nelson, as well as areas in Fiordland. Smaller areas of elevated concentrations are also shown in the Taranaki region. A higher-resolution image of the area around Auckland is shown in Figure 19.

Table 6. ANOVA table for chromium, showing the significance of each explanatory variable for the GAM. A Wald test of the significance of term is performed, so the interpretation is like a Type-III ANOVA. The p-value for each term are calculated using the Bayesian estimated covariance matrix of the parameter estimators.

Variable	df	F	p-value
lenz	19	2.46	< 0.001
soilorder	13	9.08	< 0.001
parentmaterial	8	31.89	< 0.001
landuse_ambient	5	17.56	< 0.001

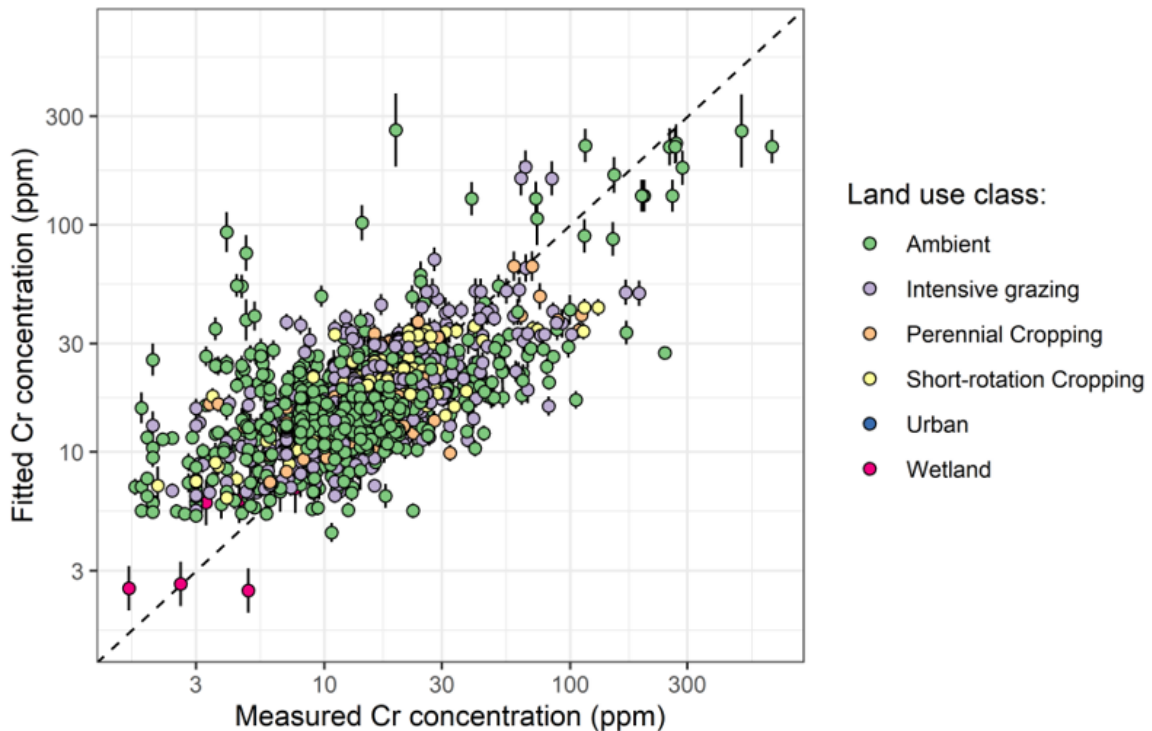


Figure 17. Measured versus fitted plot of chromium concentration, with both axes log scaled. The points are coloured by land-use class.

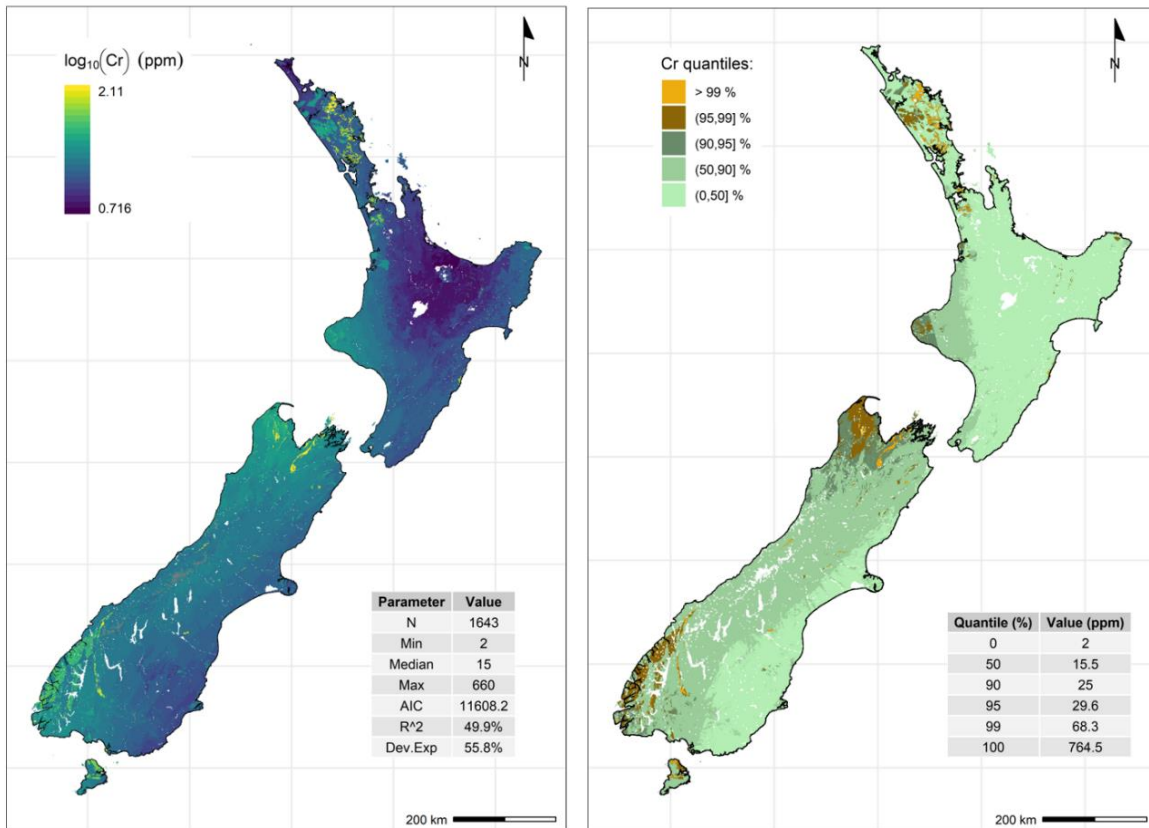


Figure 18. Left: predicted log-transformed ambient chromium concentrations for all locations in New Zealand. The calculations are based on a grid with 1,000 m spacings. A table summarising the fitting data set is also included. Right: filled contour plot of chromium concentrations, with contours at the 50, 90, 95, and 99th percentiles of the predicted values. The table in the lower right shows the values associated with common quantiles, including the 0 and 100% quantiles (representing minimum and maximum estimates).

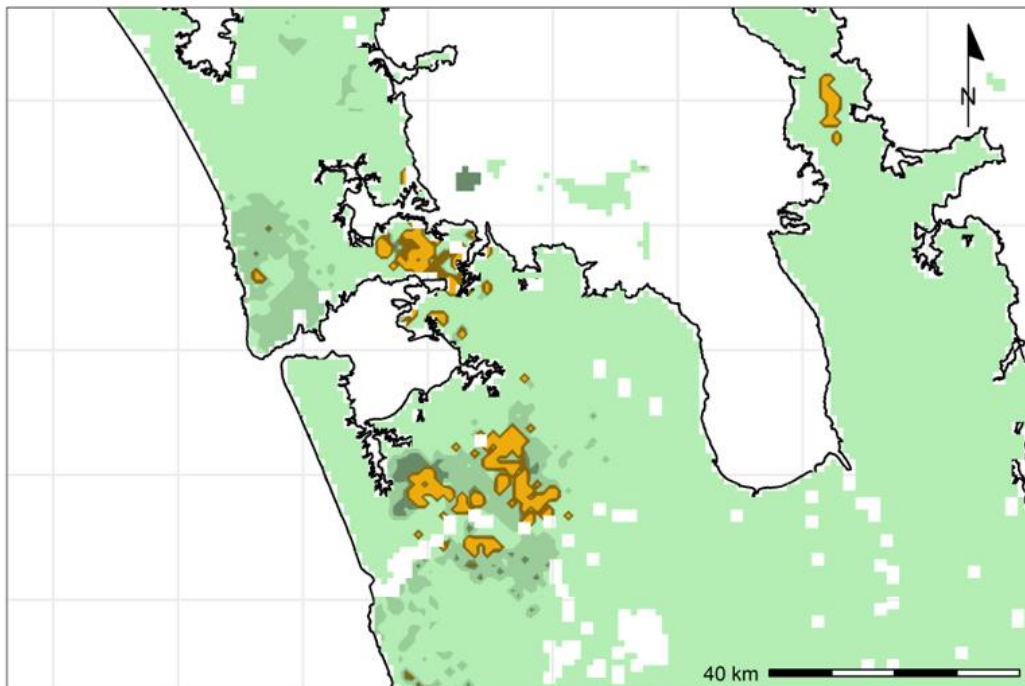


Figure 19. Filled contour plot of chromium concentrations in the Auckland region, with contours at the 50, 90, 95, and 99th percentiles of the predicted values. Refer to Figure 18 for legend. White spaces indicates missing co-variate data.

5.1.5 Copper

Table 7 indicates that LENZ class, land use, parent material, and soil order, has a significant influence on determining Cu concentrations across New Zealand. Further detail on the relative influence of individual variables within each explanatory factor is provided in Appendix 1. Overall, there is a reasonable fit of the data, as shown by fitted vs measured (Figure 20), and c. 34% of the null deviance is explained by the model (Figure 21).

Predicted spatial estimates of Cu concentrations in the rural ambient class are presented as log-transformed point (1 km × 1 km) estimates (Figure 21 Left) and concentrations based on percentile distribution (Figure 21 Right)). Higher concentrations (>99th predicted concentrations) are found around the Taranaki region, in particular, with elevated concentrations also observed in Waikato and Fiordland. The elevated concentrations in Taranaki are also supported by the findings of a previous study (Percival & Sutherland 2002). A higher-resolution image of the area around Auckland is shown in Figure 22.

Table 7. ANOVA table for copper, showing the significance of each explanatory variable for the GAM. A Wald test of the significance of term is performed, so the interpretation is like a Type-III ANOVA. The p-value for each term are calculated using the Bayesian estimated covariance matrix of the parameter estimators.

Variable	df	F	p-value
lenz	18	5.13	< 0.001
soilorder	13	4.76	< 0.001
parentmaterial	8	7.49	< 0.001
landuse_ambient	5	29.53	< 0.001

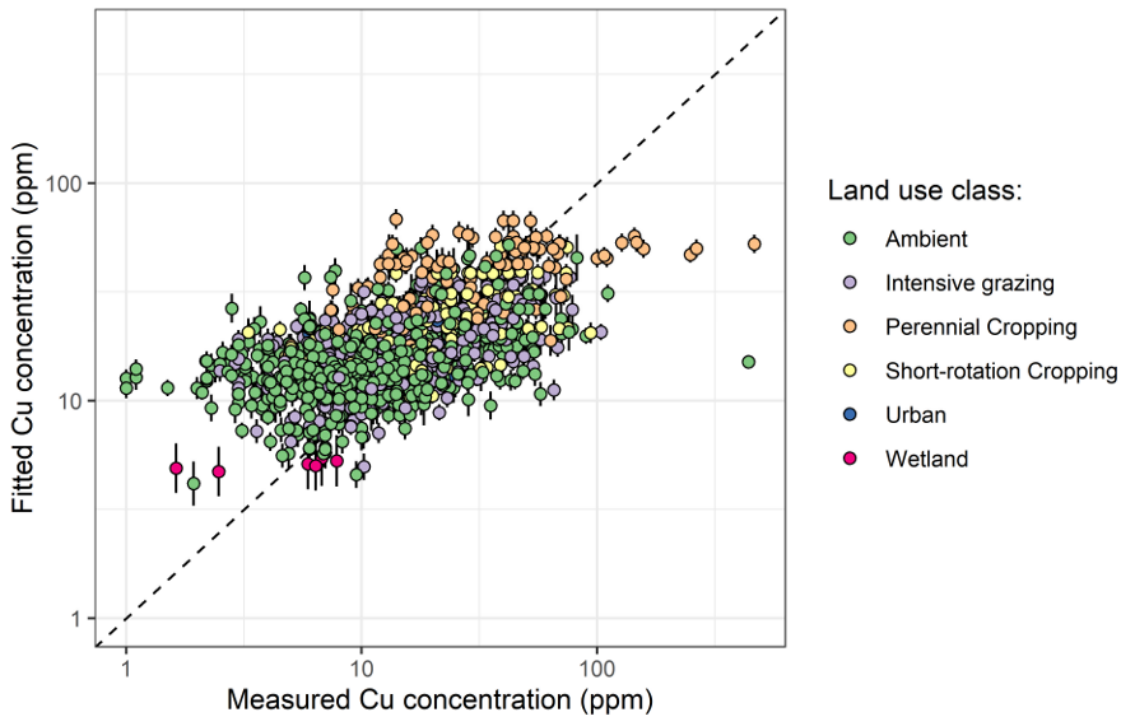


Figure 20. Measured versus fitted plot of copper concentrations, with both axes log scaled. The points are coloured by land-use class.

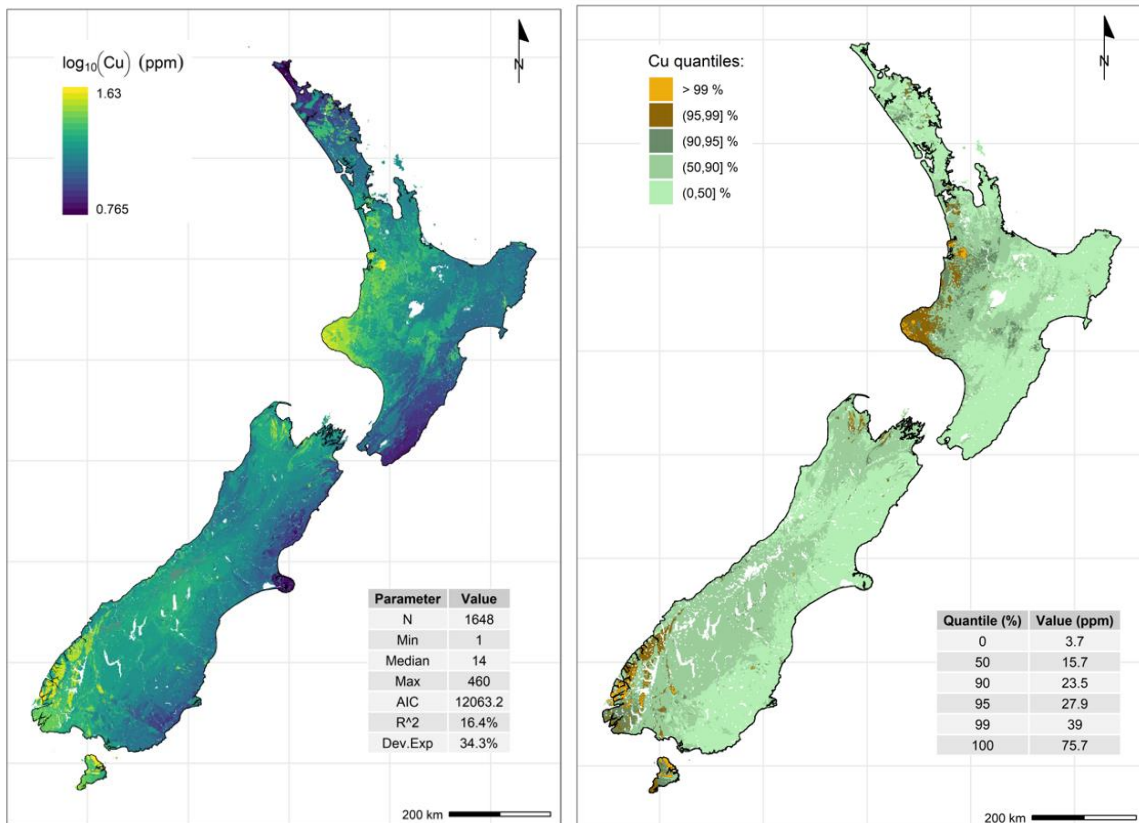


Figure 21. Left: predicted log-transformed ambient copper concentrations for all locations in New Zealand. The calculations are based on a grid with 1,000 m spacings. A table summarising the fitting data set is also included. Right: filled contour plot of copper concentrations, with contours at the 50, 90, 95, and 99th percentiles of the predicted values. The table in the lower right shows the values associated with common quantiles, including the 0 and 100% quantiles (representing minimum and maximum estimates).

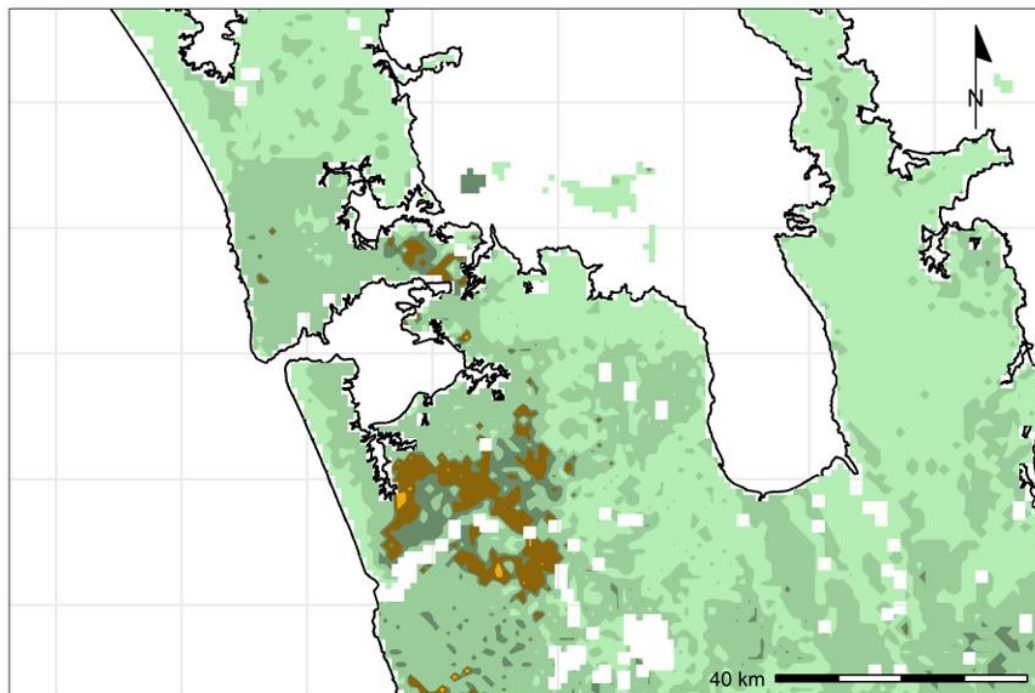


Figure 22. Filled contour plot of copper concentrations in the Auckland region, with contours at the 50, 90, 95, and 99th percentiles of the predicted values. Refer to Figure 21 for legend. White spaces indicates missing co-variate data.

5.1.6 Lead

Table 8 indicates that LENZ class, soil order, parent material, and land use have a significant influence on Pb concentrations across New Zealand. Further detail on the relative influence of individual variables within each explanatory factor is provided in Appendix 1. Overall there is a reasonable fit of the data, as shown by fitted vs measured (Figure 23), and c. 39% of the null deviance is explained by the model (Figure 24).

Predicted spatial estimates of Pb concentrations in the rural ambient class are presented as log-transformed point (1 km × 1 km) estimates (Figure 24 Left) and concentrations based on percentile distribution (Figure 24 Right). Higher concentrations are found in the mid-Canterbury region, and in the Waikato region, although overall the predicted concentration range is relatively small, with the 5th–95th percentile range of 4.8–19 mg/kg and a predicted maximum concentration of 30 mg/kg. A higher-resolution image of the area around Auckland is shown in Figure 25.

Table 8. ANOVA table for lead, showing the significance of each explanatory variable for the GAM. A Wald test of the significance of term is performed, so the interpretation is like a Type-III ANOVA. The p-value for each term are calculated using the Bayesian estimated covariance matrix of the parameter estimators.

Variable	df	F	p-value
lenz	18	3.98	< 0.001
soilorder	13	11.1	< 0.001
parentmaterial	8	7.14	< 0.001
landuse_ambient	5	13.71	< 0.001

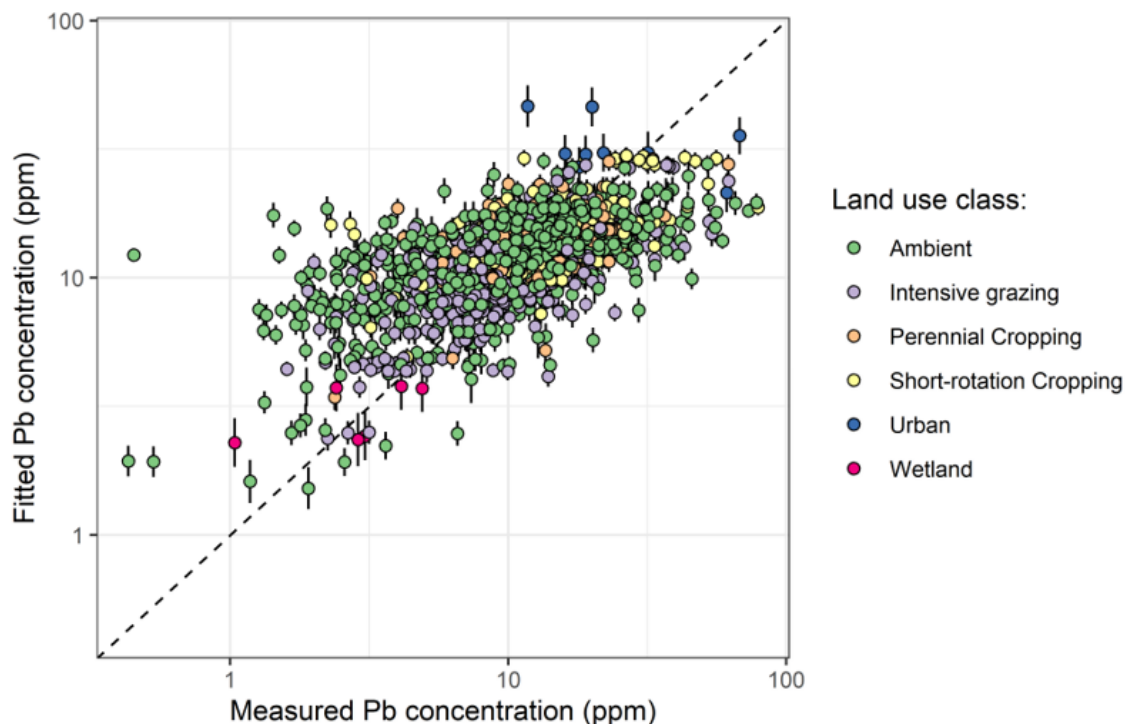


Figure 23. Measured versus fitted plot of lead concentrations, with both axes log scaled. The points are coloured by land-use class.

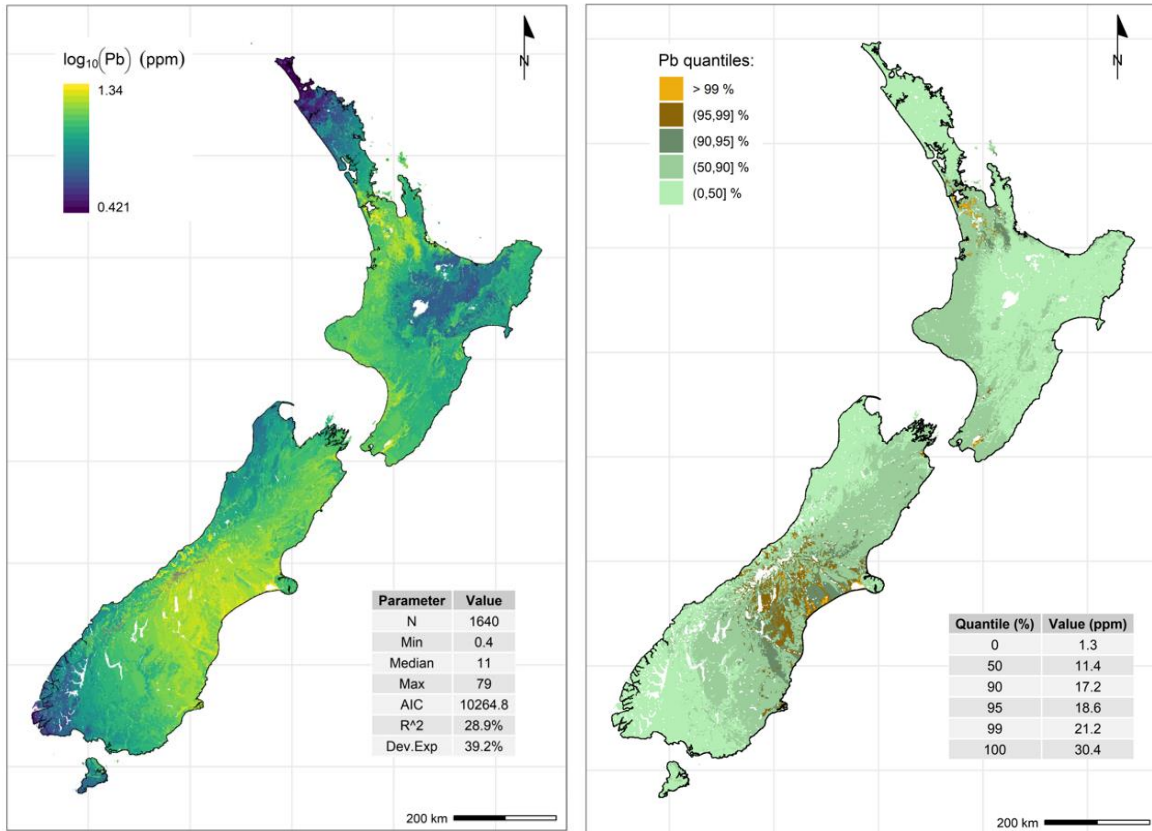


Figure 24. Left: predicted log-transformed ambient lead concentrations for all locations in New Zealand. The calculations are based on a grid with 1,000 m spacings. A table summarising the fitting data set is also included. Right: filled contour plot of lead concentrations, with contours at the 50, 90, 95, and 99th percentiles of the predicted values. The table in the lower right shows the values associated with common quantiles, including the 0 and 100% quantiles (representing minimum and maximum estimates).

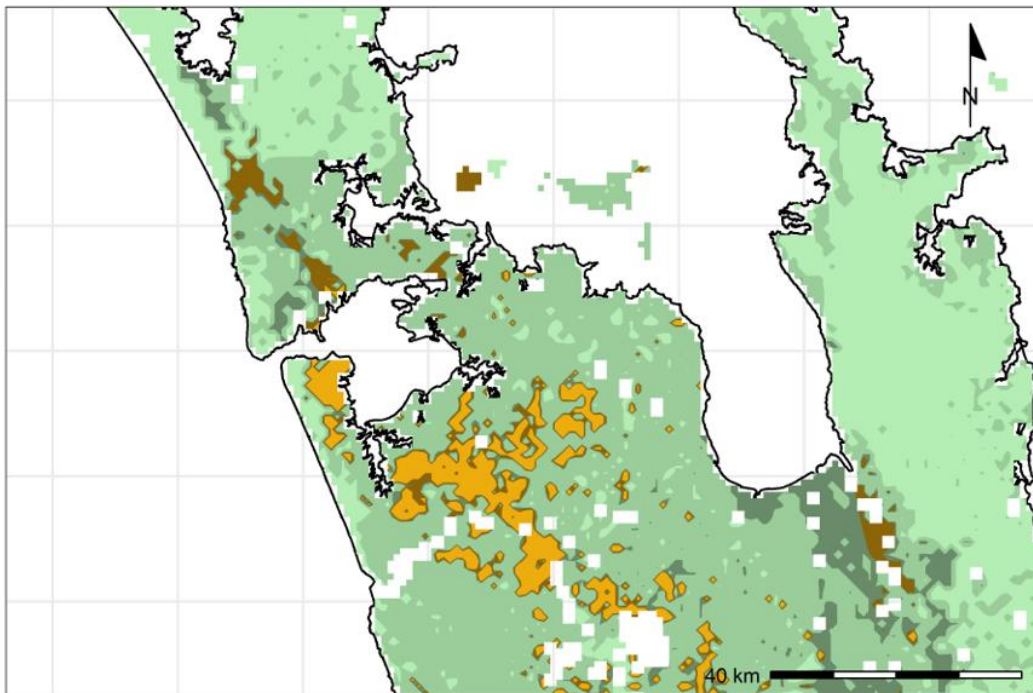


Figure 25. Filled contour plot of lead concentrations in the Auckland region, with contours at the 50, 90, 95, and 99th percentiles of the predicted values. Refer to Figure 24 for legend. White spaces indicates missing co-variate data.

5.1.7 Nickel

Table 9 indicates that LENZ class, soil order, parent material, and land use have a significant influence of parent material on Ni concentrations. Further detail on the relative influence of individual variables within each explanatory factor is provided in Appendix 1 and shows a marked influence of ultramafic igneous parent material. Overall, there is a reasonable fit of the data, as shown by fitted vs measured concentrations (Figure 26), and c. 52% of the null deviance is explained by the model (Figure 27).

Predicted spatial estimates of Ni concentrations in the rural ambient class are presented as log-transformed point (1 km × 1km) estimates (Figure 27 Left) and concentrations based on percentile distribution (Figure 27 Right). Similar to Cr, elevated concentrations of Ni are found in Northland and Auckland, across the top of the South Island around Golden Bay, and in the Richmond mountain range near Nelson, as well as areas in Fiordland. A higher-resolution image of the area around Auckland is shown in Figure 28.

Table 9. ANOVA table for nickel, showing the significance of each explanatory variable for the GAM. A Wald test of the significance of term is performed, so the interpretation is like a Type-III ANOVA. The p-value for each term are calculated using the Bayesian estimated covariance matrix of the parameter estimators.

Variable	df	F	p-value
lenz	18	3.12	< 0.001
soilorder	13	5.48	< 0.001
parentmaterial	8	39.66	< 0.001
landuse_ambient	5	11.71	< 0.001

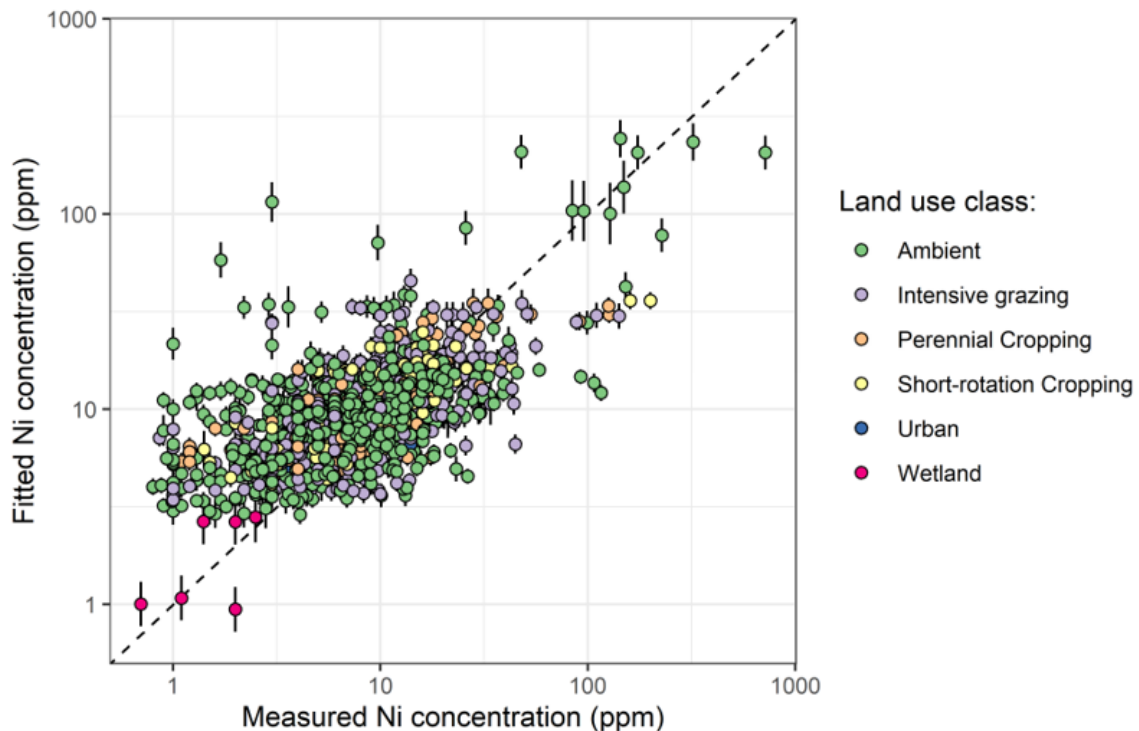


Figure 26. Measured versus fitted plot of Ni concentrations, with both axes log scaled. The points are coloured by land-use class.

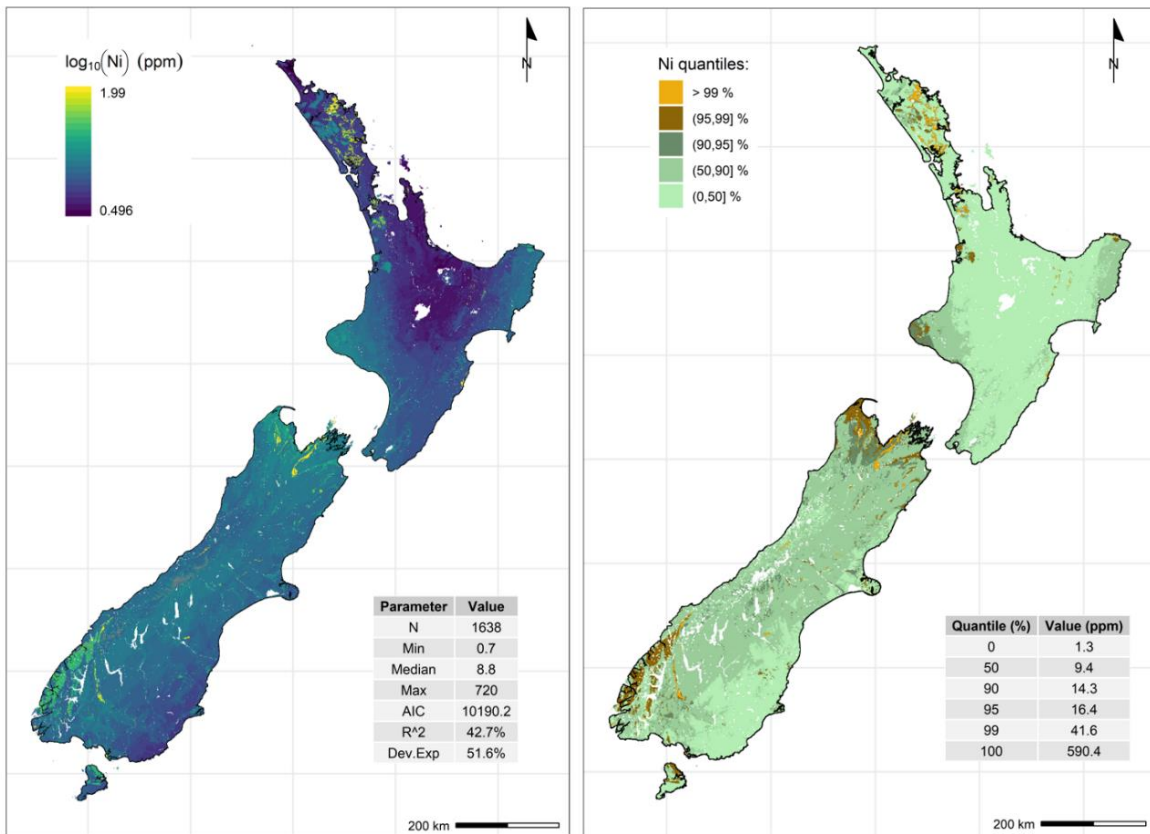


Figure 27. Left: predicted log-transformed ambient nickel concentrations for all locations in New Zealand. The calculations are based on a grid with 1,000 m spacings. A table summarising the fitting data set is also included. Right: filled contour plot of nickel concentrations, with contours at the 50, 90, 95, and 99th percentiles of the predicted values. The table in the lower right shows the values associated with common quantiles, including the 0 and 100% quantiles (representing minimum and maximum estimates).

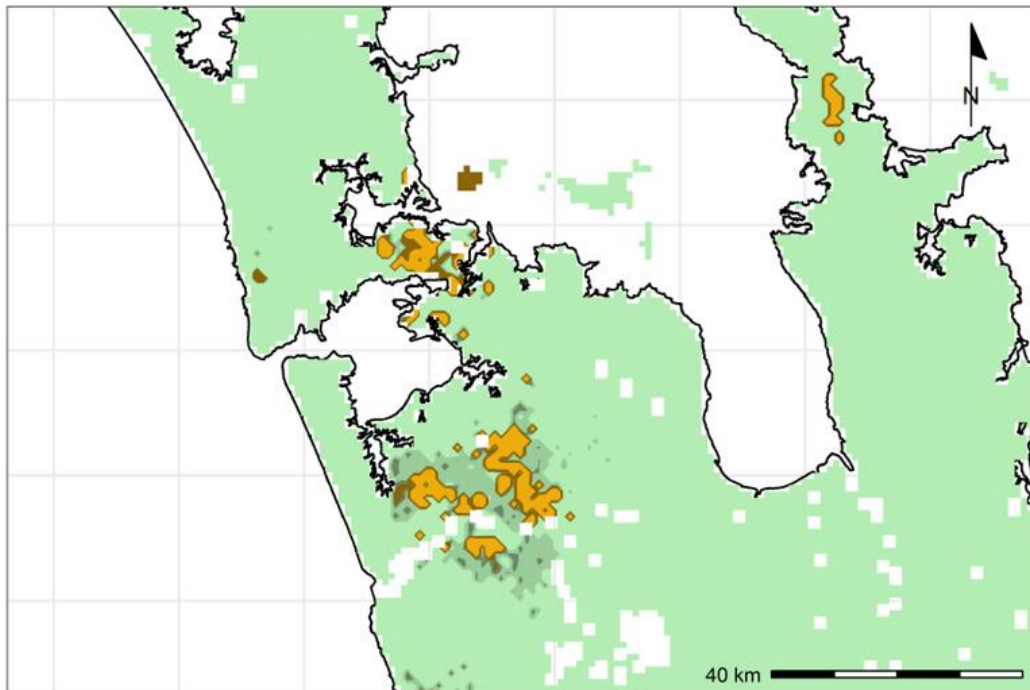


Figure 28. Filled contour plot of nickel concentrations in the Auckland region, with contours at the 50, 90, 95, and 99th percentiles of the predicted values. Refer to Figure 27 for legend. White spaces indicates missing co-variate data.

5.1.8 Zinc

Table 10 indicates that LENZ class, soil order, and land use, have a significant influence on determining Zn concentrations across New Zealand, but parent material does not appear to have a significant effect. Further detail on the relative influence of individual variables within each explanatory factor is provided in Appendix 1. The influence of location is perhaps reflected in the relatively low deviance (i.e. comparatively poor model fit) explained for the model (26%, Figure 30). The measured vs fitted model also shows the model under-predicts concentrations compared to measured concentrations at the higher end (>c. 70 mg/kg). Further analysis of the data, including assessing the influence of low and high concentration values, did not improve the fit. Given also the comparatively wide range of concentrations across the data set, some other unidentified factors may have a greater influence on Zn concentrations than those currently modelled.

Predicted spatial estimates of Zn concentrations in the rural ambient class are presented as log-transformed point (1 km × 1 km) estimates (Figure 30 Left) and concentrations based on percentile distribution (Figure 30 Right). Higher concentrations (>99th predicted concentrations) are found around the Taranaki region, in particular, and are supported by the findings of a previous study (Percival & Sutherland 2002). A higher-resolution image of the area around Auckland is shown in Figure 31.

Predicted concentrations are typically lower than has been reported for Auckland, in particular (ARC 2001; Martin et al. 2023b). These studies found that Zn concentrations in volcanic soils were elevated compared to non-volcanic soils. While the magnitude of this variation differs between the studies, we note that the 'volcanic' grouping used in these studies effectively comprises grouping three of our current parent material classes: Felsic Igneous, Mafic Igneous, and Ultramafic Igneous. Further discussion is provided in section 5.2.1.

Table 10. ANOVA table for zinc, showing the significance of each explanatory variable for the GAM. A Wald test of the significance of term is performed, so the interpretation is like a Type-III ANOVA. The p-value for each term are calculated using the Bayesian estimated covariance matrix of the parameter estimators.

Variable	df	F	p-value
lenz	17	3.83	< 0.001
soilorder	13	5.56	< 0.001
parentmaterial	8	1.45	0.17
landuse_ambient	5	12.18	< 0.001

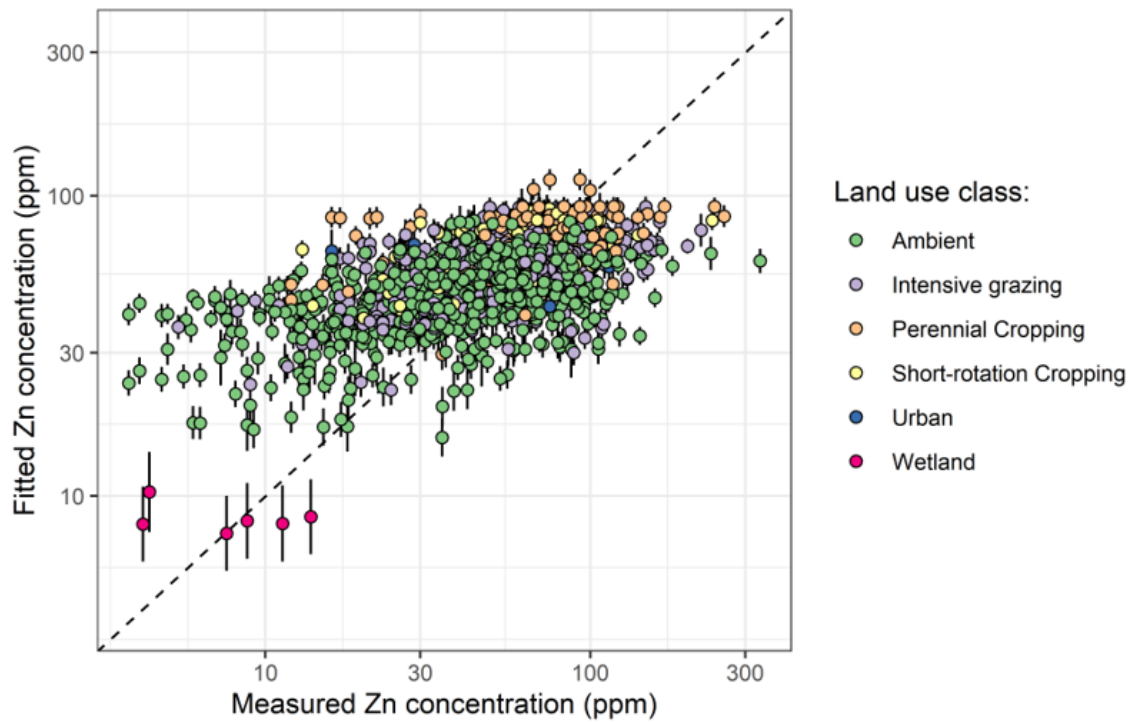


Figure 29. Measured versus fitted plot of zinc concentration, with both axes log scaled. The points are coloured by land-use class.

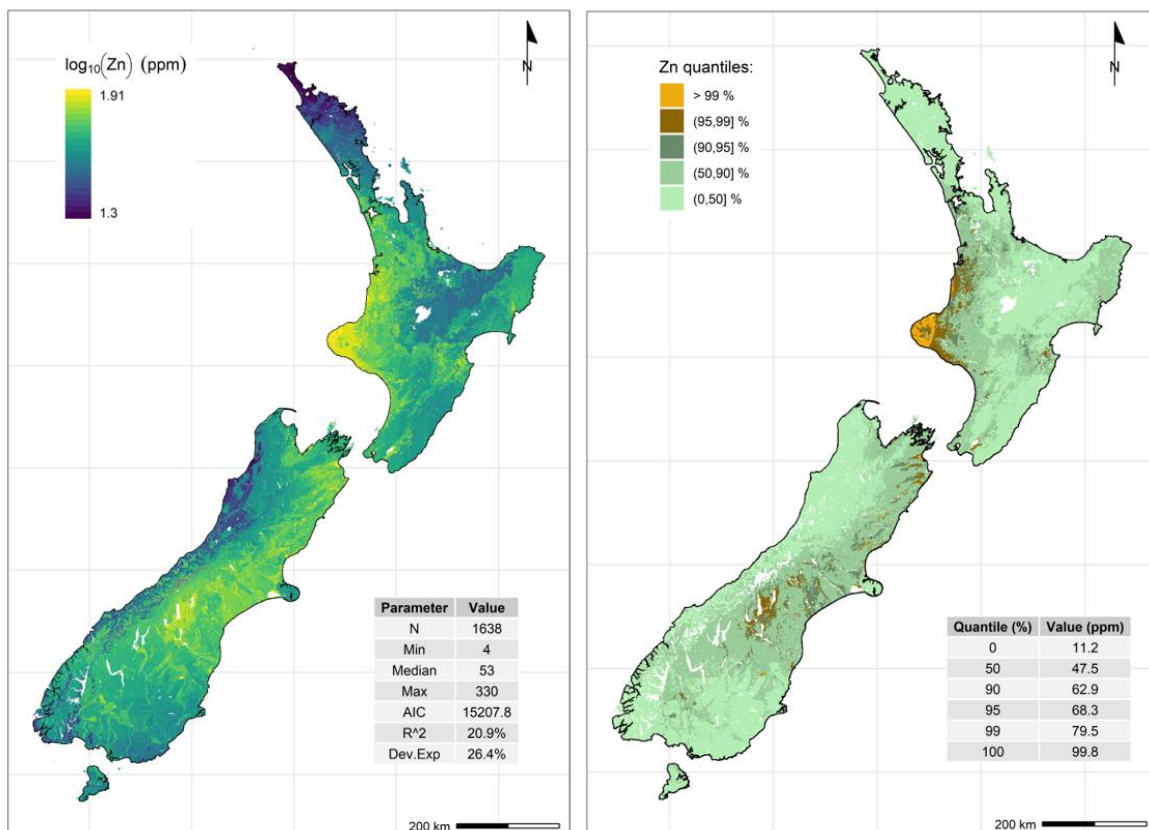


Figure 30. Left: predicted log-transformed ambient zinc concentrations for all locations in New Zealand. The calculations are based on a grid with 1,000 m spacings. A table summarising the fitting data set is also included. Right: filled contour plot of zinc concentrations, with contours at the 50, 90, 95, and 99 percentiles of the predicted values. The table in the lower right shows the values associated with common quantiles, including the 0 and 100% quantiles (representing minimum and maximum estimates).

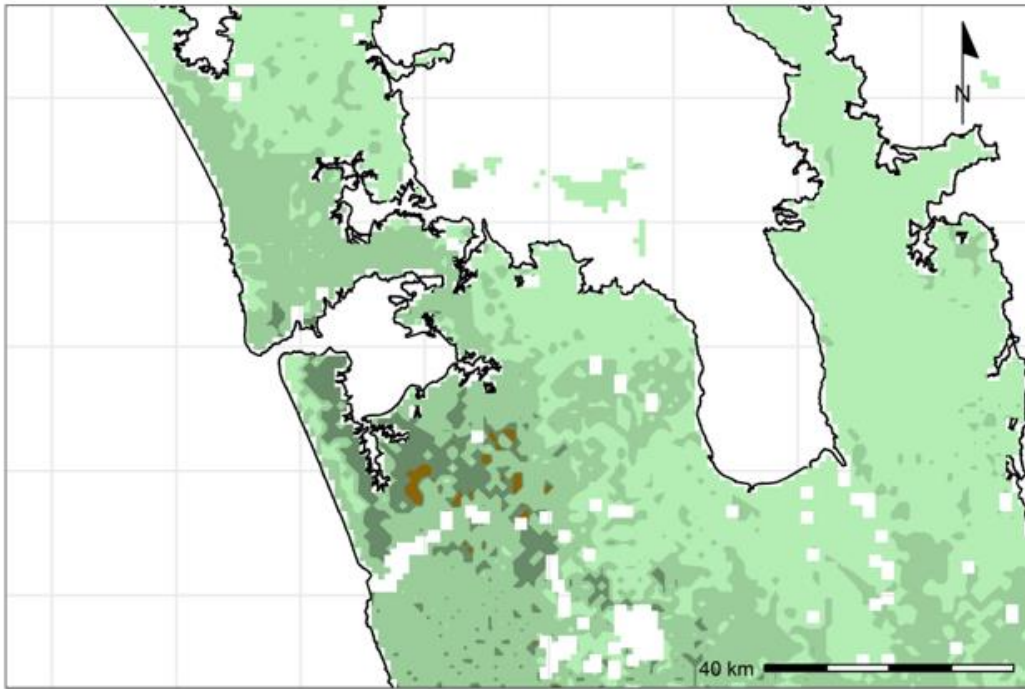


Figure 31. Filled contour plot of zinc concentrations in the Auckland region, with contours at the 50, 90, 95, and 99th percentiles of the predicted values. Refer to Figure 30 for legend. White spaces indicates missing co-variate data.

5.1.9 Summary

Table 11 provides a summary of the predicted rural ambient background concentrations for selected trace elements across New Zealand. The modelled estimate provides greater spatial differentiation of the variation of in rural ambient concentrations than has been afforded by previous studies. The spatially explicit concentration data and maps for the individual trace elements are also available at: <https://iris.scinfo.org.nz/layer/114281-pbc-predicted-background-soil-concentrations-new-zealand-h3-resolution-9/>

Table 11. Summary statistics for predicted rural ambient concentrations (mg/kg) for selected trace elements

Element	Min	5%	Median	Mean	90%	95%	99%	Max
As	0.2	2.2	4.1	4.2	6	7	8	19
B	0.5	2.2	4.6	6.0	12	16	23	83
Cd	0.01	0.02	0.08	0.10	0.20	0.29	0.35	0.58
Cr	2.0	6.8	16	18	25	30	68	765
Cu	3.8	8.5	16	17	24	28	39	76
Ni	1.4	3.8	9	11	14	16	42	590
Pb	1.3	4.8	11	12	17	19	21	30
Zn	11	29	48	48	63	68	80	100

5.2 Comparison with regional background studies

Regional background studies for Auckland (ARC 2001), Wellington (URS 2003), and Canterbury (Tonkin & Taylor 2006, 2007) have, not surprisingly, formed the basis for background concentrations in those regions. These studies have used different pedo-geological units to define background concentrations, making it difficult to compare across regions and with the predictions developed in the current project. As noted in section 4.1, the current modelled estimates include these individual studies, as well as additional data. A high-level visual evaluation is undertaken here to provide a comparison between the predicted concentrations in the current study and the individual regional studies.

5.2.1 Auckland

Investigations to determine background concentrations in the Auckland region were carried out over 1999–2001 and are reported in ARC 2001. Soil samples (0–15 cm) were collected as a single point sample from 91 locations in May 1999 on soil derived from the major lithological units (eight) of the region. Fifteen sites were revisited in October 2001 to confirm the validity of outlier data obtained during the 1999 survey. The location of the individual sampling sites and the approximate extent of key lithological units are shown in Figure 32.

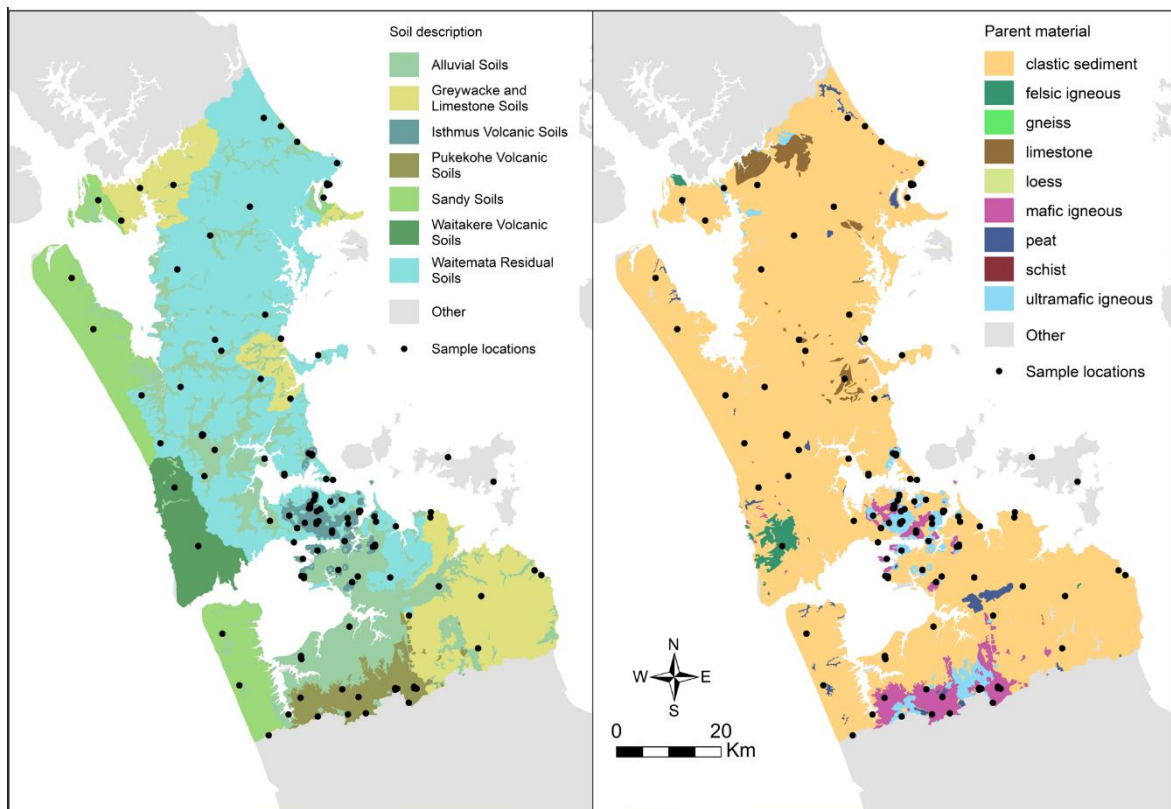


Figure 32. Location of samples used to determine background concentrations in Auckland in ARC 2001. Left: the base layer is the 'Soil types' layer provided by Auckland Council, which approximates the lithological groupings used in ARC 2001. Right: the distribution of the parent material groupings used in the current study for the same region.

Soils were grouped to provide trace element concentration ranges for volcanics and non-volcanics (Table 12), although it was recognised that individual soil types within the groups may be geochemically different (ARC 2001). For As, Cd, and Pb no difference was observed for background concentrations for volcanic and non-volcanic soils, although it was noted for Pb that some specific lithologies (Mt Smart volcanics) may require additional analyses to confirm background concentrations. For all other elements, concentration ranges are overlapping, although higher for soils derived from volcanic geologies than from non-volcanic origins (Table 12).

As noted earlier, the 'volcanic' grouping used in this study effectively comprises grouping three of our current parent material classes (Felsic Igneous, Mafic Igneous, and Ultramafic Igneous) (see also Figure 32), which may account for some difference between our predicted concentrations and those determined in ARC 2001. In general, our predicted concentrations up to the 99th percentile fall within the range for non-volcanic soils or the entire range (As, Cd, Pb) in ARC 2001, with the maximum value falling into the volcanic range, where applicable. The exceptions were Cr and Ni, for which our 99th percentile estimates fell within the volcanic range while maximum estimates were outside this range. ARC 2001 estimates of the maximum background concentration of Cd are higher than our maximum predicted concentration.

Table 12. Comparison of background concentration ranges (mg/kg) for volcanic and non-volcanic soils from ARC 2001, with predicted concentrations from the current study

Trace element	Soil grouping					
	Non-volcanic ^a	Volcanic ^a	Min ^b	95th percentile ^b	99th percentile ^b	Max ^b
<i>n</i> ¹	51–54	38–42				
As	0.4–12		0.2	7	8	19
B	2–45	<2–260	0.5	16	23	83
Cd	<0.1–0.65		0.01	0.29	0.35	0.58
Cr	2–55	3–125	2.0	30	68	765
Cu	1–45	20–90	3.8	28	39	76
Ni	0.9–35	4–320	1.4	16	42	590
Pb	<1.5–65		1.3	19	21	30
Zn	9–180	54–1160	11	68	80	100

^a Number of samples varied slightly for different elements. Source: ARC 2001.

^b Predicted concentrations developed in this study.

It is also relevant to compare the spatial distribution of the predicted concentrations to provide a comparison.

- Modelled estimates for As concentration (Figure 10) across the Auckland region (were typically up to the 90th percentile (5.9 mg/kg), with higher concentrations (up to the 99th percentile, 8 mg/kg) in the proximity of the Pukekohe Volcanic soils (as shown in Figure 32).

- For B, predicted concentrations were typically at or above the 90th percentile (12 mg/kg), with volcanic soils showing clear elevations (>23 mg/kg) (Figure 13).
- For Cd, predicted concentrations were typically at or above the 90th percentile (0.2 mg/kg), with elevated concentrations (up to the 99th percentile, 0.35 mg/kg) occurring in the alluvial soils to the south of Manukau Harbour (Figure 16).
- The predicted concentrations for Cr and Ni show a similar distribution, with volcanic soils (Figures 19 & 28), particularly the Pukekohe and Isthmus volcanic soils, showing elevated concentrations (>99th percentile, >68 mg/kg and >42 mg/kg respectively).
- Predicted Cr and Ni concentrations in non-volcanic areas were typically up to the median estimates of 16 and 9 mg/kg.
- Predicted Cu concentrations across Auckland were typically up to the predicted 90th percentile concentration (24 mg/kg), with volcanic soils predicted to have concentrations up to or above the 95th percentile of 28 mg/kg (Figure 22).
- Predicted Pb concentrations showed no clear association with volcanic or non-volcanic areas, consistent with the observation of ARC 2001 of no differentiation between volcanic and non-volcanic soils (Figure 25). Concentrations were typically up to the 90th percentile concentration of 17 mg/kg, with some areas further elevated above the 99th percentile (21 mg/kg). These concentrations are lower than the range reported in ARC 2001 of up to 65 mg/kg.
- Finally, for Zn, predicted concentrations indicate that most of the Auckland region falls within the 50th to 90th percentile concentration range of 48 to 63 mg/kg (Figure 31). Concentrations are predicted to be elevated in the proximity of Pukekohe volcanic soils up to the 99th percentile (80 mg/kg). These concentrations are lower than the those reported in ARC 2001 and Martin et al. 2022. Martin et al. (2022) suggest the inter-quartile range for Zn in volcanic soils is between 80 and 140 mg/kg, indicating the majority of soils will fall below 140 mg/kg, although some soils are still likely to have higher concentrations.

5.2.2 Wellington

The development of initial estimates for background soil concentrations in the Wellington region is described in URS 2003 and briefly summarised here.

Five main soil types were identified based on a review of the major parent rock formations soils information using the 1:250,000 Geological Maps of New Zealand, 1:1,000,000 North Island Soil Map of New Zealand, and Regional Soil Plan for the Wellington Region. In each of these main soil type areas (Figure 33), specific sample sites were chosen to gain a relatively even spatial distribution of soils best representing background (non-anthropogenically influenced) soil quality, and for ease of access, and included some sites that had been previously sampled through the SOE monitoring programme.

Surficial (0–150 mm depth) soil samples were collected from 40 locations during June 2003. The collected samples were analysed for: As, soluble B, Cd, Cr, Cu, Pb, Hg, Ni, Zn, total petroleum hydrocarbons (TPHs), and polycyclic aromatic hydrocarbons (PAHs). URS (2003) noted that further soil sampling was required to improve statistical robustness before the data could be considered fully representative of background soil quality for the Wellington Region. Only trace element concentrations are discussed further here.

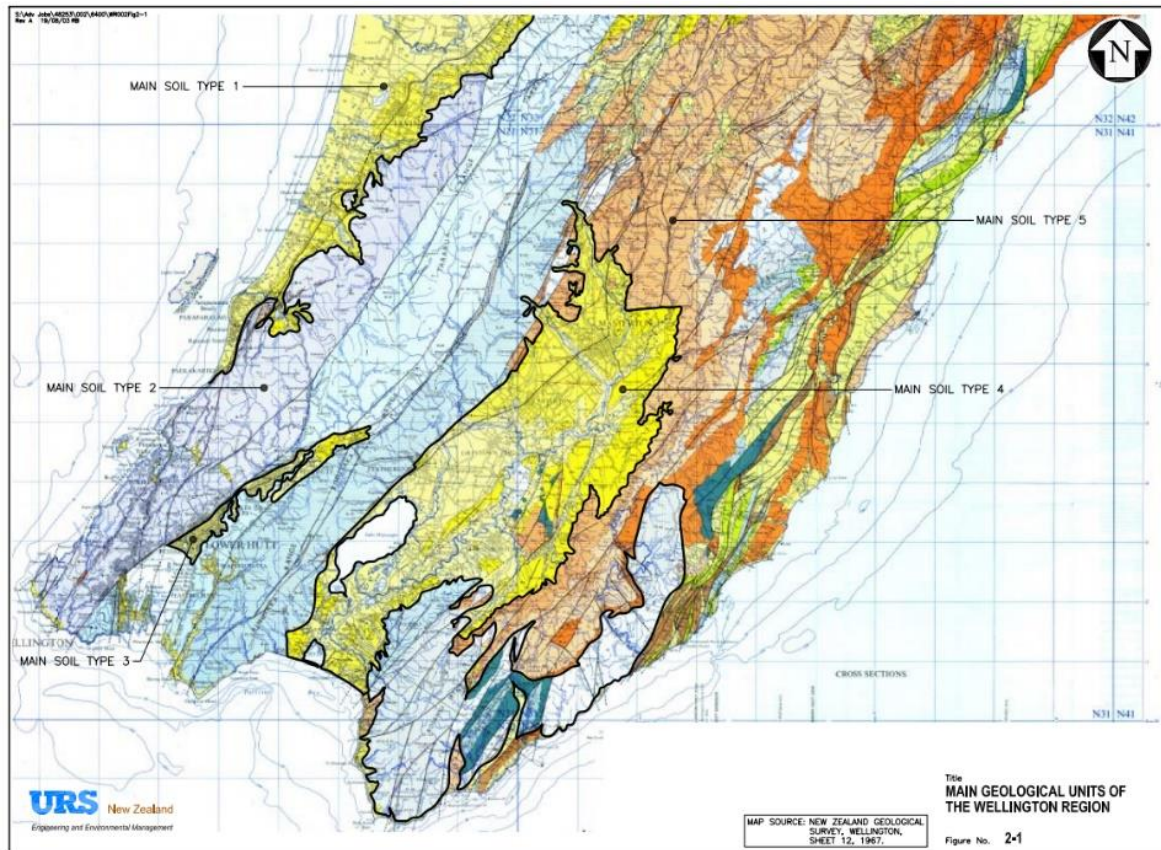


Figure 33. Approximate location of the main five soil types identified for determining background concentrations in the Wellington region: type 1 – sandy soils, type 2 – Greywacke soils, type 3 – Hutt alluvium, type 4 – Wairarapa alluvium, type 5 – Mudstone/siltstone. (Source: URS 2003)

A summary of the background concentration ranges for trace elements in the main soil types in the Wellington region, as determined by URS (2003), is shown in Table 13, along with the predicted concentrations for key elements.

- Predicted concentrations for As across most of the Wellington region are ≤ 50 th percentile, with some Wairarapa alluvium soils ≤ 90 th percentile (Figure 9). There is no differentiation between most soil types in the reported background concentrations, which are all within the 99th percentile predicted concentrations.
- Predicted concentrations for Cd in some greywacke soil areas are ≤ 50 th percentile, while other areas are mostly ≤ 90 th percentile (Figure 15). This is generally consistent with measured background concentrations, all of which fall within the 90th percentile estimates.
- Predicted concentrations for Cu show concentrations across most of the region to be ≤ 50 th percentile, although there is slight elevation up to 90th percentile concentrations around Lake Wairarapa (Wairarapa alluvium soils, Figure 21). While Wairarapa alluvium soils do have reported background concentrations ≤ 90 th percentile estimates, so do Hutt alluvium and Mudstone/siltstone soils. Greywacke soils have the highest upper range of reported background concentrations, placing them ≤ 95 th percentile model estimated concentrations, inconsistent with the mapped predictions.

- Predicted concentrations of Cr and Ni showed very similar distributions, with most areas across Wellington (Sand, Hutt alluvium and Mudstone/siltstone soils) having concentrations \leq 50th percentile (Figures 18 & 27). Some Greywacke soils and some Wairarapa alluvium soils are predicted to have higher Cr and Ni concentrations: \leq 90th percentile, with small areas of Wairarapa alluvium soils \leq 95th percentile for Ni.
- For Cr, the predicted concentrations are consistent with the reported background concentrations for most soils, but reported background concentrations for Hutt alluvium soils fall within \leq 90th percentile rather than the predicted \leq 50th percentile. For Ni, sand soils do have reported background concentrations \leq 50th percentile, but for other soil types the upper reported concentrations are slightly higher than predicted concentrations.
- Predicted Pb concentrations were typically between the 50th and 90th percentile across the region, while some concentrations around Lake Wairarapa (Wairarapa alluvium) were $>$ 99th percentile concentration (Figure 24). Predicted Pb concentrations were markedly lower than the upper concentration range reported for all soil types in URS 2003.
- Finally, Zn concentrations were predicted to be \leq 50th percentile across most of the region, with some areas in sand, Wairarapa alluvium, and Greywacke soils \leq 90th percentile, and some other areas in sand \leq 95th percentile and \leq 99th percentile in some Wairarapa alluvium (Figure 30). Most of these concentrations fall within, although are lower than, the upper range of the reported background concentrations. For example, Hutt alluvium soils have the highest reported background concentrations, contrasting with model estimates. The upper ranges of reported background concentrations for three soil types exceed the maximum model estimate, while the remaining two fall within the \leq 99th percentile range (Table 13).

Table 13. Comparison of background concentration ranges (mg/kg) in the main soil types in the Wellington region, as determined by URS 2003, with predicted concentrations developed in the current study.

Soil type (Main soil type #) ^a	As	Cd	Cr	Cu	Ni	Pb	Zn
Sand (1)	<2–7	<0.1–0.1	7–12	4–10	4–9	4.5–180.0	28–79
Greywacke (2)	<2–7	<0.1–0.1	6–16	3–25	4–13	5.9–78.6	24–105
Hutt alluvium (3)	2–7	<0.1–0.2	9–18	5–19	5–14	16.7–73.3	38–201
Wairarapa alluvium (4)	2–7	<0.1–0.2	11–21	7–19	6–21	9.4–34.0	44–121
Mudstone/siltstone (5)	<2–4	<0.1–0.2	8–15	6–19	5–13	10.8–38.1	31–72
percentile estimate ^b							
Median	4.1	0.08	16	16	9	11	48
90th	6	0.20	25	24	14	17	63
95th	7	0.29	30	28	16	19	68
99th	8	0.35	68	39	42	21	80
Max	19	0.58	765	76	590	30	100

^a URS 2003

^b This study.

5.2.3 Canterbury

Background concentrations of selected trace elements were determined in Canterbury's major soil groups in studies in 2006 and 2007 (Tonkin & Taylor 2006, 2007), and are available online.⁸ Soils were classified using the New Zealand genetic soil classification system, which is no longer used, as this was the prevalent classification scheme available in electronic data sets available at the time: the Land Resource Inventory and Canterbury Soils data sets. A total of 90 sample sites were distributed across these soil groups: 17 in the Christchurch urban area and 73 throughout the rest of Canterbury (Figure 34). From these samples, concentrations of As, B, Cd, Cr, Cu, Pb, Mn, Hg, Ni, and Zn were measured in mg/kg.

Level 1 gives the maximum concentration values of the above trace elements measured in each soil group. Level 2 concentrations are also available, which provides the maximum concentration values of the above trace elements measured in each soil group, plus half the interquartile range (buffer). A two-tiered approach was recommended to account for the limited data used to determine the background concentrations (Tonkin & Taylor 2006). The intent was for the approach to enable sites clearly below background to be designated as not contaminated, while retaining some flexibility in assessing sites within the upper limits of the proposed background concentrations.

⁸ <https://opendata.canterburymaps.govt.nz/datasets/593db381b6b04bcf8b6f01dc53d91954/about>

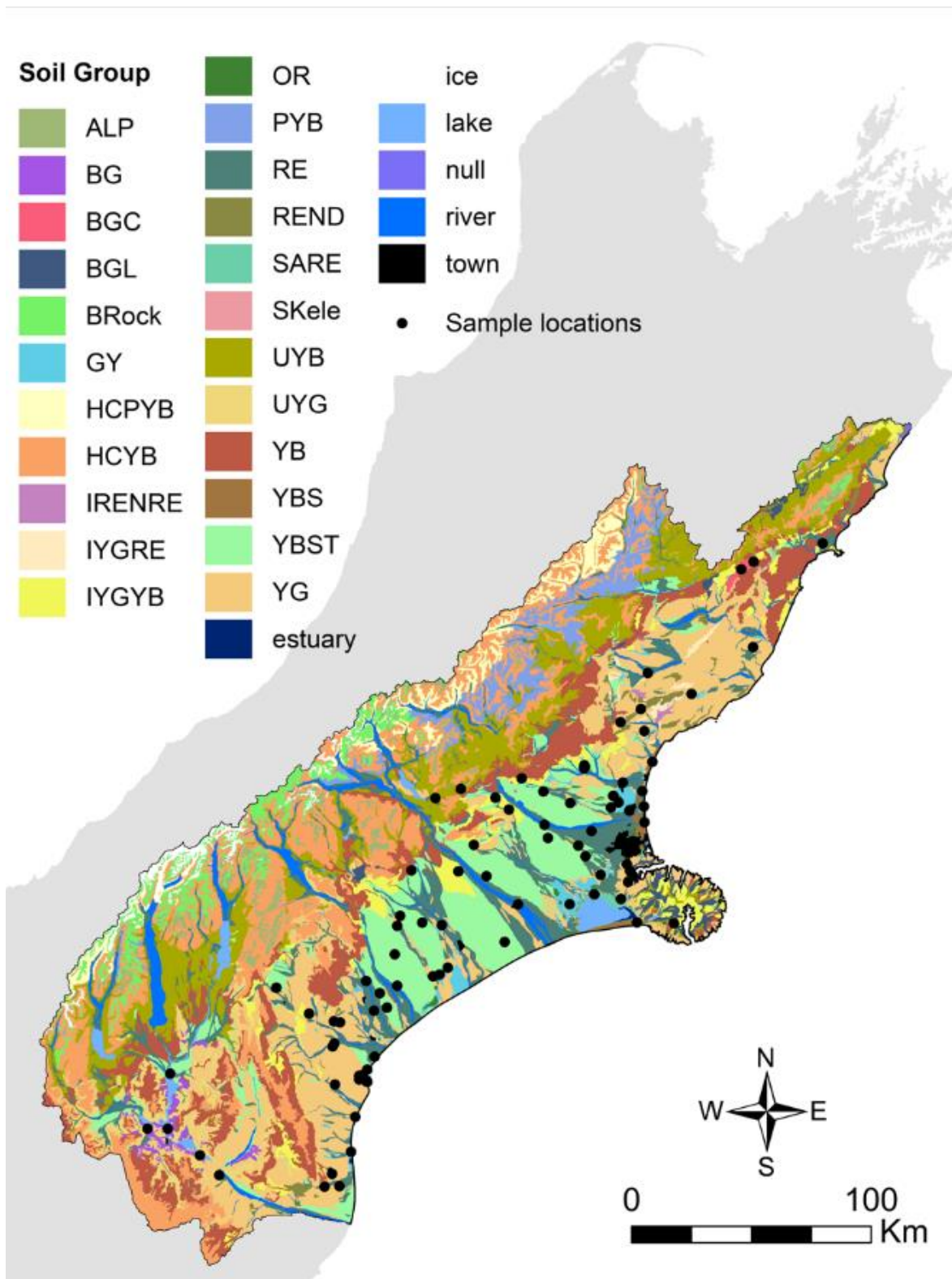


Figure 34. Sample site location for background concentration studies, and distribution of soil groups across Canterbury.

The reported maximum background concentrations are shown in Table 14. Model predictions indicate that small pockets of soils across Canterbury may have concentrations >99th percentile, although most areas fall within either the 50th percentile or 90th percentile concentrations for all trace elements; the exception is Pb, which has elevated concentrations predicted across Canterbury (Figure 24).

- Reported background concentrations of As for Rendzina soils are markedly higher and exceed the maximum model estimate. Reported As concentrations for Gley (regional and Christchurch urban), Recent (regional and Christchurch urban), and Organic (urban) soils are >99th percentile model estimate. Saline Gley recent soils (regional and Christchurch urban) are between the 95th and 99th percentile estimates, while other soil types are ≤95th percentile estimates for As.
- Predicted concentrations for Cd are mostly ≤90th percentile across the region (Figure 15). The maximum reported background concentrations for Gley soils fall within the 95th percentile estimates, while Organic and Rendzina soils have Cd concentrations ≤99th percentile estimates.
- For Cu, some small areas are predicted to have concentrations >99th percentile, but most of the region is predicted to have concentrations ≤90th percentile, with the majority of this area region with concentrations ≤50th percentile (Figure 21). This is largely consistent with reported background concentrations, as most are within 50th percentile estimates, although reported maximum concentrations for Organic and Recent soils are ≤90th percentile, and BG/BGL (Brown Granular Earths/Brown Granular loams) soils fall within the 95th percentile estimates.
- Predicted concentrations for Cr show some small areas as >95th percentile, but concentrations for most of the region are ≤90th percentile, with approximately half the region (closer to the coast and including the Canterbury plains) being ≤50th percentile (Figure 18). These estimates are slightly lower than the maximum reported background concentrations for key soil groups on plains and along the east of the region – including Recent, Yellow Brown Earth, Yellow Brown Stony Soils, and Intergrade, which are ≤90th percentile. The maximum background concentrations for Rendzina soils are higher, but fall within ≤95th percentile predicted concentrations.
- Predicted concentrations show areas where Pb concentrations are >99th percentile in mid-Canterbury, and most areas have concentrations >90th percentile (Figure 24). Nonetheless, the model appears to under-predict Pb concentrations in Canterbury, with reported background concentrations for eight soil types exceeding the maximum model estimate and three falling between the 99th percentile and maximum model estimates.
- The predicted concentrations of Ni show small areas with concentrations >95th percentile, but most areas ≤90th percentile (Figure 27). Reported background concentrations show most soils are within 95th percentile estimates, excluding Recent (both regional and Christchurch urban) and BGC/BGL soils, which fall within 99th percentile estimates.
- Predicted Zn concentrations show some small areas with concentrations >99th percentile, and other areas with concentrations ≤99th percentile or ≤95th percentile, with the majority of areas are ≤90th percentile (Figure 30). As noted earlier, the model under-predicts Zn concentrations at the higher end, and reported background concentrations for four soils (BGC/BGL, Gley (Christchurch urban), Recent (Christchurch urban) and YGE (yellow-grey earths) (Timaru urban)) exceed the maximum model estimate for Zn, while Recent (regional) and Saline Gley Recent (Christchurch urban) are >99th percentile model estimates. Intergrade (regional)

soils are >95th percentile estimates, while other soil types fall within 95th percentile estimates.

Table 14. Reported maximum background concentrations for soils from the Christchurch region (all in mg/kg)

Area	Soil group	<i>n</i>	As	Cd	Cr	Cu	Ni	Pb	Zn
Regional ^a	BG/BGL	4	5.1	0.20	22.5	27.3	20.7	17.2	116.0
	Gley	6	8.7	0.24	16.8	15.5	13.4	17.8	65.6
	Intergrade	8	6.1	0.12	24.5	15.2	15.3	27.5	69.8
	Organic	3	2.9	0.34	13.8	18.4	7.2	25.9	53.5
	Recent	18	11.5	0.18	20.8	18.8	19.0	37.4	86.5
	Rendzina	4	36.9	0.31	26.4	9.5	15.9	16.7	57.4
	Saline Gley Recent	4	6.8	0.09	13.2	12.2	9.6	44.4	47.3
	Yellow Brown Sand	4	3.4	0.06	11.0	7.1	8.7	31.9	50.7
	Yellow Brown Stony	14	5.8	0.10	18.3	10.2	12.8	18.7	64.0
	Yellow Grey Earth	15	4.6	0.11	15.6	11.5	11.6	18.8	62.4
	Yellow Brown Earth	2	4.2	0.04	18.9	10.1	16.1	11.9	43.1
Christchurch urban ^a	Gley	6	10.6	0.20	18.5	23.3	15.6	34.9	138.0
	Organic	2	13.2	0.11	12.4	13.3	11.7	40.9	63.3
	Recent	8	15.3	0.20	19.0	17.7	16.6	101.0	149.0
	Saline Gley Recent	2	7.5	0.06	22.1	10.2	14.1	31.2	87.7
	Yellow Brown Sand	4	5.6	0.10	15.4	8.8	11.7	22.3	54.9
Timaru urban ^a	YGE	5	3.3	0.11	12.4	13.1	8.6	56.0	122.0
Model predictions ^b	Percentile								
	Median		4.1	0.08	16	16	9	11	48
	90th		6	0.20	25	24	14	17	63
	95th		7	0.29	30	28	16	19	68
	99th		8	0.35	68	39	42	21	80
	Max		19	0.58	765	76	590	30	100

^a Source: Tonkin & Taylor 2007

^b This study.

5.2.4 Summary

When comparing model predictions with upper estimates of background concentrations from the regional studies, the model performed reasonably well for most trace elements, although upper concentrations of Pb and Zn were consistently under-predicted. It was, however, challenging to compare data that were based on different groupings of results. All individual regional studies acknowledged the limited data on which background

concentrations were determined, which is where modelled results, based on a larger data set, should be more robust. Nonetheless, there are also challenges with modelling the 'extreme' end of the concentration range.

6 Using background soil concentration information

As noted in the introduction, background soil concentration information is used in different contexts in New Zealand. With the exception of the use of background concentrations for the development of Eco-SGVs, the use of background concentrations fundamentally revolves around determining some upper level of concentration, either nationally, regionally, or for specific purposes (e.g. landfill waste acceptance criteria).

The upper thresholds currently used vary, and include the maximum of the range in specific studies (ARC 2001; Tonkin & Taylor 2007; URS 2003), the 99th percentile concentrations of collated data sets (MfE 2011), and the maximum concentration plus half the interquartile range for a given data set (Level 2 background, Tonkin & Taylor 2007). These estimates also vary because they have been developed for different locations, geographical extents, pedo-geological units, and sample numbers.

There is much discussion in the literature about the statistical analysis of geochemical data, including the determination of upper limits and outliers or anomalies, the shape of the data, and identifying different 'populations' of data it may be appropriate to statistically summarise (e.g. Reiman & Garrett 2005; Reiman et al. 2005; Mrvić et al. 2011; US EPA 2002). However, rather than focusing on the technical aspects, it is useful to consider the specific contexts in which background concentration information is used in New Zealand.

6.1 Legislative context for managing contaminated land

In the National Environmental Standard for Managing Contaminants in Soil for the Protection of Human Health (NES-SC), background concentrations are used in Clause 5(9), which states: 'These regulations do not apply to a piece of land ... about which a detailed site investigation exists that demonstrates that any contaminants in or on the piece of land are at, or below, background concentrations'.

For this clause to apply, there is the requirement that the land must have been identified as having, having had, or is more likely than not to have had a HAIL (Hazardous Activities and Industries List) activity on it. However, this clause appears to often be interpreted as then indicating that the NES does apply to land with soil concentrations considered to be *above* background (i.e. it's an indication a HAIL activity must have occurred), even if below any applicable human health criteria (i.e. Soil Contaminant Standard) or environmental guideline.

At this point it is also worth considering the definition of contaminated land under the Resource Management Act (RMA) and the Natural and Built Environment Bill (NBE). Under the RMA, 'contaminated land' means land that has a hazardous substance in or on it that has, or is reasonably likely to have, significant adverse effects on the environment. Under

the NBE, 'contaminated land' means land where a contaminant is present in concentrations that exceed an environmental limit or pose an unacceptable risk to human health or the environment. In both cases there is clear reference to adverse effect or unacceptable risk. Thus, even if concentrations of trace elements are above background, if they are below any relevant risk-based human health standard or Eco-SGV, then no further action should be required.

In a different context, under the NBE a natural hazard includes 'soil that contains concentrations of naturally occurring contaminants that pose an ongoing risk to human health.' Logically, 'naturally occurring contaminants' in this context is intended to mean background concentrations, and, in particular, naturally elevated background concentrations. Therefore, clause 5(9) is contradictory to the purpose of the NES-SC (protection of human health), because excluding land at or below background concentration doesn't necessarily ensure protection of human health, particularly where those concentrations are naturally elevated.

It is worthwhile noting that none of the predicted 99th percentile background concentrations are greater than the most conservative soil contaminant standard (rural residential). However, the rural residential value for As is based on a previous determination of background concentration for As.

Finally, in US guidance background concentrations only become relevant after it is determined that soil concentrations are greater than risk-based screening values. If concentrations are determined to be below any background concentration, then no further action (i.e. remediation) is required.⁹ Similarly, background concentrations may be used to set remedial goals when background concentrations *are greater than* risk-based remedial goals. The rationale is that it is not reasonable to expect clean-ups to achieve concentrations less than soil background.

6.2 Determination of site-specific background concentrations

There are numerous documents that provide guidance on determining site-specific or local background concentrations (e.g. US EPA 2002; Diamond et al. 2009; ISO 2011; ITRC¹⁰; EPA Victoria 2018). Once again, it is useful to take a step back from the technical detail to clarify the purpose of undertaking site-specific determination.

From the perspective of managing risk associated with soil contaminants, there seems little value in undertaking a site-specific determination of background concentrations *unless* these concentrations are anticipated to be *above* risk-based guidelines for the protection of human health or ecological receptors (including ground and surface water). Assessment in this case is to inform what might be appropriate remediation goals, specifically, it is unreasonable to require remediation goals to be below naturally occurring

⁹ https://sbr-1.itrcweb.org/using-soil-background-in-risk-assessment/#4_1

¹⁰ <https://sbr-1.itrcweb.org/>

background concentrations, rather the focus should be on how any associated risk from these naturally occurring concentrations might need to be managed.

A rural ambient concentration has been used to generate background concentrations here to recognise the extent of anthropogenic activities across New Zealand, but the land uses included are anticipated to have minimal anthropogenic trace element inputs, and so these rural ambient concentrations are expected to differ little from 'true' naturally occurring concentrations. In an urban environment, ambient trace element concentrations – primarily lead from the historical use of leaded petrol, and, rarely, other trace elements associated with industrial emission – might reasonably be expected.

Arguably, more of a challenge arises when considering background concentrations of some organic contaminants, notably polycyclic aromatic hydrocarbons and per- and polyfluoroalkyl substances.

Note that it is irrelevant to determine background concentrations at a site for the purposes of disposal to clean fill unless the site is located in the same region/geographical area as the clean fill, as the waste acceptance criteria for the clean fill should be based on background concentrations relevant to its location (see also section 6.3).

Where a site-specific determination of background concentration is considered appropriate, other factors such as the size of the site, and the ability to identify a sufficient number of locations that have not been affected by anthropogenic activity (to get sufficient samples to robustly determine background concentrations for the site in question) should also be considered. A minimum of 30 samples is recommended to characterise background concentrations of a given pedo-geological area (ISO 2012), although this depends on the area under consideration: it may be appropriate to analyse 30 samples for a regionally based characterisation, and a smaller number for a specific site investigation. In this instance, 7 to 10 samples may be appropriate, as used by some US EPA jurisdictions (Diamond et al. 2009).

6.3 Waste acceptance criteria

The *Technical Guidelines for Disposal to Land* (WasteMINZ 2022) provides technical guidance on siting, design, construction, operation, and monitoring for disposal to land, and classifies landfills into five types:

- Class 1 Landfill – municipal solid waste landfill or industrial waste landfill
- Class 2 Landfill – construction & demolition landfill or industrial waste landfill
- Class 3 Landfill – managed fill
- Class 4 Landfill – controlled fill
- Class 5 Landfill – clean fill.

In the context of background soil concentrations, Class 5 clean fill is of most relevance. Materials placed within a Class 5 clean fill are intended to be inert, and the regional soil background levels for trace elements should be adopted as the basis for acceptance of materials for these sites. The waste acceptance criteria for a Class 5 clean fill are based on

the accepted background concentrations for inorganic elements within the intended catchment of the site, and provide for trace concentrations of a limited range of organic compounds.

It should also be noted that approaches used by regional councils for clean fill criteria have been variable, based either on background concentrations alone; or on a combination of background concentrations and Eco-SGVs and protecting human health (e.g. Cavanagh 2021, 2013); or on concentrations that are not lower than the 95th percentile of the regional background and not exceeding the lower of protective thresholds for the most sensitive receptor (i.e. the lower of human health or ecological thresholds) (Waikato Regional Council 2022).

The current study provides information at a regional level that could be used in the absence of a more specific determination of background concentrations. As also demonstrated by the current study, background concentrations vary across New Zealand. In this context the use of national background levels that are taken as the 99th percentile of the available data set (suggested as default values for As and Cd in WasteMINZ 2022) could result in marked elevations in concentration compared to regional backgrounds. In other words, background concentrations used as clean fill criteria should reflect the regional background. Some specific regional background studies (Auckland, Wellington, Canterbury) are available, and section 5 provides a comparison of the results from the current project with those studies.

While setting clean fill waste acceptance criteria based on the background concentrations relevant to the location of the landfill has some logic in terms of enabling clean fills to be permitted activities, closer evaluation of the establishment and use of clean fills is required to ascertain whether this approach is achieving / will achieve the desired outcomes in terms of protecting human health and the environment, or is unduly conservative. Conceptually (and as outlined in WasteMINZ 2022), waste acceptance criteria should be developed by considering:

- the protection of human health from direct contact (or inhalation of volatiles)
- the potential for leaching into groundwater (including that used for drinking-water)
- organisms living in or on the soil (ecological receptors)

and should be based on the most sensitive receptor. These criteria should be appropriate to the landfill's construction, the nature of the wastes, and the potential future land use (i.e. unrestricted or restricted to certain land uses).

Cavanagh (2021) also raises the question about the relevance of having separate classes for Class 4 (controlled fill) and Class 5 (clean fill), given that both landfill types are not intended to impose any restrictions on future land use. Thus, nationally, it would be relevant to consider whether there is merit in combining the proposed Class 4 and 5 landfills into one class, with waste acceptance criteria based on ensuring protection of the most sensitive receptor (people or ecological receptors). More stringent criteria – or at least criteria based on leaching or off-site movement – could apply where these landfills might be placed close to waterways or groundwater.

6.4 Background concentrations and ecological soil guideline values

Soil guideline values for the protection of ecological receptors (microbes, invertebrates, plants, wildlife, livestock; Eco-SGVs) provide a first assessment of the potential negative environmental effects associated with contaminants (Cavanagh & Munir 2019). Eco-SGVs for naturally occurring contaminants (i.e. metals and metalloids) have been developed using the 'added-risk' approach (Cavanagh & Munir 2019). This approach considers that soil biota are adapted to the naturally occurring concentrations of potential contaminants, and that it is the 'added' anthropogenic component that drives toxicity responses. The magnitude of the acceptable 'added' concentration is determined by assessing toxicity data, with the 'added concentration limit' (ACL) added to the background concentration (see Cavanagh & Munir 2019 for further details). In turn, the added-risk approach allows for variation in the Eco-SGVs based on variation in naturally occurring background concentrations. Cavanagh and Harmsworth (2022) further evaluate the application of Eco-SGVs and outline the use of Eco-SGVs based on three levels of protection: 95%, 80%, and 60% of species (based on chronic toxicity data).

In the first instance, the median background concentration is used to determine the Eco-SGVs (Cavanagh & Munir 2019). Thereafter some pragmatism is required to determine when it is appropriate to modify the Eco-SGVs based on background concentrations to avoid overly complex application of the Eco-SGVs, and to ensure they still afford an appropriate level of protection. This judgement has been made by considering both the percentile range and the proportional contribution of the natural background concentration to the Eco-SGV, but also on the relevance of allowing further adjustment for lower levels of protection.

Overall, we recommend that background concentration adjustment only be acceptable for the 95% protection values, partly because background concentrations generally comprise a small proportion of the 80% and 60% protection values, but also because it was not considered appropriate to enable further adjustment of values associated with lower levels of protection. For the 95% protection values, background adjustment was only considered relevant for those contaminants for which the difference between median concentration and the upper percentiles was > 10 mg/kg, and where background concentration comprised > 10% of the Eco-SGV.

As noted earlier, percentile background concentrations are the percentile of predicted concentrations. A summary of the upper percentile predicted background concentrations for the individual trace elements is shown in Table 15. The default Eco-SGVs are based on the median concentration, with the remaining bolded values showing those concentrations for which it is considered acceptable to adjust the 95% protection level Eco-SGVs. The revised Eco-SGVs for the individual trace elements are provided in the following sections, with further detail on the application of Eco-SGVs available in Cavanagh & Harmsworth 2023.

Table 15. A summary of relevant statistics for the range of predicted background (rural ambient) concentrations of selected trace elements. Bolded values are those to be used for developing Eco-SGVs, including for background adjustment of the 95% protection-level Eco-SGVs.

Element	Median	90th	95th	99th *	Max
As	4.1	5.9	6.5	8.0	18.7
B	4.6	12	16	23	83
Cd	0.08	0.2	0.29	0.35	0.58
Cr	16	25	30	68	765
Cu	16	24	28	39	76
Ni	9	14	16	42	590
Pb	11	17	19	21	30
Zn	48	63	68	80	100

* It is recommended that the 99th percentile be used as a default value for these areas initially. Where there is recognised to be significant local small-scale elevation in background concentrations, it may be appropriate to undertake site-specific determination of background concentrations.

6.4.1 Arsenic

The predicted background (rural ambient) concentrations for As are summarised in Table 16, with the spatial variation in concentrations across New Zealand shown in Figure 9. Given the small range in predicted ambient concentrations for As, it is not recommended that Eco-SGVs be adjusted for background concentrations, except in areas where there are recognised to be significant small-scale elevations in naturally occurring concentrations. In these cases, site-specific determination of background concentrations is probably required. The revised Eco-SGVs based on updated median background concentrations and the derived added concentration limits determined for As by Cavanagh and Munir (2019) at the different protection levels are shown in Table 17. See Cavanagh & Harmsworth 2023 for more details on applying the Eco-SGVs.

Table 16. Summary of predicted background (rural ambient) concentrations for arsenic. The bolded value shows the concentration used to derive Eco-SGVs

Element	Min	Median	90th	95th	99th*	Max
As	0.22	4.1	5.9	6.5	8.0	18.7

* It is recommended that the 99th percentile be used as a default value for these areas initially. Where there is recognised to be significant local small-scale elevation in background concentrations, it may be appropriate to undertake site-specific determination of background concentrations.

Table 17. Eco-SGVs for arsenic based on median background concentration and added contaminant limits (ACLs) at three protection levels. See Cavanagh & Harmsworth 2023 for more details on applying the Eco-SGVs.

Protection level ^a	Median background concentration ^b (mg/kg)	ACL ^c _(EC30) (mg/kg)	Eco-SGV ^d (mg/kg)
95%	4.1	15	20
80%		55	60
60%		144	150

^a These protection levels equate to the non-food production land (95%), residential/recreational area (80%), and commercial/industrial area (60%) land uses from Cavanagh & Munir 2019.

^b It is recommended that the median be used as a default value. Where there is recognised to be significant local small-scale elevation in background concentrations, it may be appropriate to undertake site-specific determination of background concentrations.

^c From Cavanagh & Munir 2019.

^d Values have been rounded.

6.4.2 Boron

The predicted background (rural ambient) concentrations for B are summarised in Table 18, with the spatial variation in concentrations across New Zealand shown in Figure 12. Based on the range in rural ambient concentrations for B, background adjustment for Eco-SGVs based on total B concentrations is recommended in areas identified as being above the 95th percentile of modelled estimates. ACLs for B are based on both total and hot-water-soluble B (HWS-B, Table A4), although the contribution of background HWS-B is considered to be negligible (Cavanagh & Munir 2019), so Eco-SGVs based on HWS-B do not change. Values for Eco-SGVs based on revised predicted median background concentrations for the different protection levels are shown in Table 19, with background-adjusted 95% protection level Eco-SGVs shown in Table 20. See Cavanagh & Harmsworth 2023 for more details on the application of Eco-SGVs.

Table 18. Summary of predicted background (rural ambient) concentrations for boron. Bolded concentrations are the values used to develop Eco-SGVs.

Element	Min	Median	90th	95th	99th*	Max
B	0.6	4.6	12	16	23	83

* It is recommended that the 99th percentile be used as a default value for these areas initially. Where there is recognised to be significant local small-scale elevation in background concentrations, it may be appropriate to undertake site-specific determination of background concentrations.

Table 19. Eco-SGVs for boron based on median background concentration, and added contaminant limits (ACLs) expressed as total boron and hot-water-soluble boron (HWS-B) concentrations at three protection levels

Protection level ^a	Median background concentration (mg/kg)	ACL ^b _(EC30) (mg/kg)	Eco-SGV ^c _(EC30) (mg/kg)	HWS-B Eco-SGV ^d _(EC30) (mg/kg)
95%	4.6	9.7	14	7
80%	4.6	17	22	14
60%	4.6	21	26	17

^a These protection levels equate to the non-food production land (95%), residential/recreational area (80%), and commercial/industrial area (60%) land uses from Cavanagh & Munir 2019.

^b From Cavanagh & Munir 2019.

^c Values have been rounded.

^d Based on hot-water-soluble boron concentrations; the contribution of background HWS-B is considered to be negligible.

Table 20. Background-adjusted 95% protection-level Eco-SGVs for boron based on the 95th and 99th percentile predicted background concentrations and the ACL for total boron^a

Background concentration percentile	Background concentration (mg/kg)	ACL ^b _(EC30) (mg/kg)	Eco-SGV ^c _(EC30)
95th	16	9.7	26
99th ^d	23	9.7	33

^a The contribution of background HWS-B is considered to be negligible, so Eco-SGVs associated with HWS-B do not vary with background concentration.

^b See Table 19.

^c Values have been rounded.

^d It is recommended that the 99th percentile be used as a default value for these areas initially. Where there is recognised to be significant local small-scale elevation in background concentrations, it may be appropriate to undertake site-specific determination of background concentrations.

6.4.3 Cadmium

The predicted background (rural ambient) concentrations for Cd are summarised in Table 21, with the spatial variation in concentrations across New Zealand shown in Figure 15. Given the low rural ambient concentrations for Cd, background concentrations are not used to derive Eco-SGVs. Also, given that Eco-SGVs based on providing protection for biomagnification are lower than those based on total Cd, these are the recommended Eco-SGVs to use. See Cavanagh & Harmsworth 2023 for more details on the application of Eco-SGVs.

Table 21. Summary of predicted background (rural ambient) concentrations for cadmium.

Element	Min	Median	90 th	95 th	99 th	Max
Cd	0.01	0.1	0.2	0.29	0.35	0.58

Table 22. Eco-SGVs for cadmium based on median background concentration and added contaminant limits (ACLs), based on total cadmium, allowing for protection for biomagnification at three protection levels

Protection level ^a	Median background concentration (mg/kg)	ACL ^b _(EC30) (mg/kg)	Eco-SGV ^c (mg/kg)	ACL ^b _(EC30BM) (mg/kg)	Eco-SGV ^c _{BM} (mg/kg)
95%	0.1	4.8	4.8	1.5	1.5
80%	0.1	17	17	12	12
60%	0.1	40	40	33	33

^a These protection levels equate to the non-food production land (95%), residential/recreational area (80%), and commercial/industrial area (60%) land uses from Cavanagh & Munir 2019.

^b From Cavanagh & Munir 2019.

^c Values have been rounded.

BM = protective from exposure via biomagnification.

6.4.4 Chromium

The predicted background (rural ambient) concentrations for Cr are summarised in Table 23, with the spatial variation in concentrations across New Zealand shown in Figure 18. Based on the range in rural ambient concentrations for Cr, background adjustment for Eco-SGVs is recommended in areas identified as being above the 95th percentile of modelled estimates. Where there are recognised significant small-scale elevations in naturally occurring concentrations, it may be appropriate to undertake site-specific determination of background concentrations. Values for Eco-SGVs based on revised predicted median background concentrations for the different protection levels are shown in Table 24, with background-adjusted 95% protection-level Eco-SGVs shown in Table 25. See Cavanagh & Harmsworth 2023 for more details on the application of Eco-SGVs.

Table 23. Summary of predicted background (rural ambient) concentrations for chromium. Bolded concentrations show the values used to develop Eco-SGVs.

Element	Min	Median	90 th	95 th	99 th *	Max
Cr	1.96	16	25	30	68	765

* It is recommended that the 99th percentile be used as a default value for these areas initially. Where there is recognised to be significant local small-scale elevation in background concentrations, it may be appropriate to undertake site-specific determination of background concentrations.

Table 24. Eco-SGVs for chromium based on median background concentrations and added contaminant limits (ACLs) at three protection levels

Protection level ^a	Median background concentration (mg/kg)	ACL ^b (EC30) (mg/kg)	Eco-SGV ^c (EC30) (mg/kg)
95%	16	184	200
80%	16	382	400
60%	16	641	660

^a These protection levels equate to the non-food production land (95%), residential/recreational area (80%), and commercial/industrial area (60%) land uses from Cavanagh & Munir 2019.

^b From Cavanagh & Munir 2019.

^c Values have been rounded.

Table 25. Summary of background-adjusted 95% protection Eco-SGVs based on the 95th and 99th percentile predicted background concentrations and the added contaminant limit (ACL) for chromium

Background concentration percentile	Background concentration (mg/kg)	ACL ^a (EC30) (mg/kg)	Eco-SGV ^b (EC30) (mg/kg)
95th	30	184	215
99th ^c	68	184	250

^a See Table 24.

^b Values have been rounded.

^c It is recommended that the 99th percentile be used as a default value for these areas initially. Where there is recognised to be significant local small-scale elevation in background concentrations, it may be appropriate to undertake site-specific determination of background concentrations.

6.4.5 Copper

The predicted background (rural ambient) concentrations for Cu are summarised in Table 26, with the spatial variation in concentrations across New Zealand shown in Figure 21. Based on the range in rural ambient concentrations for Cu, background adjustment for Eco-SGVs is recommended in areas identified as being above the 95th percentile of modelled estimates. Where there are recognised significant small-scale elevations in naturally occurring concentrations, it may be appropriate to undertake site-specific determination of background concentrations.

There were sufficient toxicity data to derive added contaminant limits for three reference soils, with the values for the sensitive soils recommended as default values (Table 27, see Cavanagh & Munir 2019 for further details). Values for Eco-SGVs based on revised predicted median background concentrations for the different protection levels are shown in Table 28, with background-adjusted 95% protection-level Eco-SGVs shown in Table 29. See Cavanagh & Harmsworth 2023 for more details on the application of Eco-SGVs.

Table 26. Summary of predicted background (rural ambient) concentrations for copper. Bolded concentrations show the values used to develop Eco-SGVs.

Element	Min	Median	90th	95th	99th *	Max
Cu	3.8	16	24	28	39	76

* It is recommended that the 99th percentile be used as a default value for these areas initially. Where there is recognised to be significant local small-scale elevation in background concentrations, it may be appropriate to undertake site-specific determination of background concentrations

Table 27. Added concentration limits (ACL) derived for copper using LOEC/EC30 toxicological endpoints for aged contamination, the typical, sensitive, and tolerant New Zealand reference soils, and three protection levels

Protection level (%)*	ACL _(EC30) Typical soil (mg/kg)	ACL _(EC30) Sensitive soil (mg/kg)	ACL _(EC30) Tolerant soil (mg/kg)
95%	108	55	90
80%	197	120	412
60%	339	250	600

* These protection levels equate to the non-food production land (95%), residential/recreational area (80%), and commercial/industrial area (60%) land uses. (Source: Cavanagh & Munir 2019)

Table 28. Eco-SGVs for copper based on median background concentration and added contaminant limits developed for the three New Zealand reference soils at three protection levels

Protection level (%)	Median background concentration (mg/kg)	Eco-SGV ^a _(EC30) typical soil	Eco-SGV ^a _(EC30) sensitive soil ^b	Eco-SGV ^a _(EC30) tolerant soil
95%	16	110	95	135
80%	16	245	190	350
60%	16	430	330	640

^a Values have been rounded.

^b Suggested default Eco-SGV.

Table 29. Summary of background-adjusted 95% protection values Eco-SGVs for the three New Zealand reference soils, based on the estimated 95th and 99th percentile ambient concentrations

Percentile background concentration	Background concentration (mg/kg)	Eco-SGV ^a _(EC30) typical soil	Eco-SGV ^a _(EC30) sensitive soil ^b	Eco-SGV ^a _(EC30) tolerant soil
95th%	28	125	110	150
99th%	39	135	120	160

^a Values have been rounded.

^b Suggested default Eco-SGV.

6.4.6 Lead

The predicted background (rural ambient) concentrations for Pb are summarised in Table 30, with the spatial variation in concentrations across New Zealand shown in Figure 24. Given the small range in predicted ambient concentrations for Pb, it is not recommended that Eco-SGVs be adjusted for background concentrations. Values for Eco-SGVs based on revised predicted median background concentrations for the different protection levels are shown in Table 31. Given that Eco-SGVs based on providing protection for biomagnification are lower than for total Pb for the lower protection levels (80%, 60%), these are the recommended Eco-SGVs for use at those protection levels. See Cavanagh & Harmsworth 2023 for more details on the application of the Eco-SGVs.

Table 30. Summary of predicted background (rural ambient) concentrations for lead. Bolded concentrations show the value used to develop Eco-SGVs.

Element	Min	Median	90th	95th	99th	Max
Pb	1.3	11	17	19	21	30

Table 31. Eco-SGVs for lead based on median background concentration and added contaminant limits (ACLs) based on total lead, and allowing for protection for biomagnification at three protection levels

Protection level ^a (%)	Median background concentration (mg/kg)	ACL ^b _(EC30) (mg/kg)	Eco-SGV ^c _(EC30) (mg/kg)	ACL ^b _(BM) (mg/kg)	Eco-SGV ^c _{BM} (mg/kg)
95%	11	275	290		NA
80%	11	1,276	1,290	918	900
60%	11	3,049	3,060	2,541	2,500

^a These protection levels equate to the non-food production land (95%), residential/recreational area (80%), and commercial/industrial area (60%) land uses from Cavanagh & Munir 2019.

^b From Cavanagh & Munir 2019.

^c Values have been rounded.

BM = protective from exposure via biomagnification, recommended for use to account for secondary poisoning at high concentrations of Pb.

6.4.7 Zinc

The predicted background (rural ambient) concentrations for Cu are summarised in Table 32, with the spatial variation in concentrations across New Zealand shown in Figure 30. Based on the range in rural ambient concentrations for Cu, background adjustment for Eco-SGVs is recommended in areas identified as being above the 95th percentile of modelled estimates. Where there are recognised significant small-scale elevations in naturally occurring concentrations, it may be appropriate to undertake site-specific determination of background concentrations.

There were sufficient toxicity data to derive added contaminant limits for three reference soils, with the values for the sensitive soils recommended as default values (Table 33, see Cavanagh & Munir 2019 for further details). Values for Eco-SGVs based on revised predicted median background concentrations for the different protection levels are shown in Table 34, with background-adjusted 95% protection-level Eco-SGVs shown in Table 35. See Cavanagh & Harmsworth 2023 for more details on the application of the Eco-SGVs.

Table 32. Summary of predicted background (rural ambient) concentrations for zinc. Bolded concentrations show the values used to develop Eco-SGVs.

Element	Min	Median	90th	95th	99th*	Max
Zn	11.2	48	63	68	80	100

* It is recommended that the 99th percentile be used as a default value for these areas initially. Where there is recognised to be significant local small-scale elevation in background concentrations, it may be appropriate to undertake site-specific determination of background concentrations.

Table 33. Added concentration limits (ACLs) derived for zinc using LOEC/EC30 toxicological endpoints for aged contamination for the typical, sensitive, and tolerant New Zealand reference soils

% protection*	ACL _(EC30aged) typical (mg/kg)	ACL _(EC30aged) sensitive (mg/kg)	ACL _(EC30aged) tolerant (mg/kg)
95%	152	131	203
80%	273	236	361
60%	463	404	597

* These protection levels equate to the non-food production land (95%), residential/recreational area (80%), and commercial/industrial area (60%) land uses.

Source: Cavanagh & Munir 2019.

Table 34. Eco-SGVs for zinc based on median background concentration and added contaminant limits developed for the three New Zealand reference soils at three protection levels

Value name (% protection)	Median background concentration (mg/kg)	Zn Eco-SGV ^a _(EC30) typical soil	Zn Eco-SGV ^a _(EC30) sensitive soil ^b	Zn Eco-SGV ^a _(EC30) tolerant soil
95%	48	200	180	250
80%	48	320	285	410
60%	48	510	450	645

^a Values have been rounded.

^b Suggested default Eco-SGV.

Table 35. Background-adjusted 95% protection values Eco-SGVs for zinc for the three New Zealand reference soils based on the estimated 90th, 95th and 99th percentile ambient concentrations

Percentile background concentration	Background concentration (mg/kg)	Zn Eco-SGV ^a _(EC30) typical soil	Zn Eco-SGV ^a _(EC30) sensitive soil ¹	Zn Eco-SGV ^a _(EC30) tolerant soil
90th	63	215	195	265
95th	68	220	200	270
99th	80	230	210	280

^a Values have been rounded.

^b Suggested default Eco-SGV.

6.5 Areas with naturally elevated concentrations

As noted above, the NBE defines a natural hazard as including ‘soil that contains concentrations of naturally occurring contaminants that pose an ongoing risk to human health’. This, on its own, places greater emphasis on identifying areas with naturally elevated concentrations and then assessing the risk these concentrations may pose to human health in these locations. In the first instance, areas predicted to have concentrations >95th percentile in the current study would be target locations to commence further evaluation.

From an ecological perspective, areas with naturally elevated concentrations can have unique ecosystems that have adapted to these elevated concentrations (i.e. will have high value for ecological integrity, a fundamental pillar of the NBE). A further consideration of the influence of areas with naturally elevated concentrations is the extent to which naturally elevated concentrations may be present in areas outside those expected to have naturally elevated trace element concentrations, such as the Dun mountain mineral belt, where soils are naturally elevated in Ni, Cr, and, to a lesser extent, Cu (Cavanagh 2021). These elevations may represent the ‘downstream’ influence of geological materials derived from this mineral belt through erosion and pedogenic processes: such fine-scale movement of trace elements is unlikely to have been captured in the current work.

Similarly, significant small-scale variations in As concentrations, suspected to be natural in origin, are known to occur in the Waikato region (pers. comm., M. Begbie, Waikato Regional Council). Thus finer-scale investigations are required to better delineate areas of naturally elevated concentrations and to understand the associated risk to human health or influence on ecological integrity.

7 Summary and conclusions

This project has developed nationally consistent estimates of background concentrations of selected naturally occurring trace elements across New Zealand. These estimates are effectively rural ambient concentrations (i.e. data from sites where there is expected to be

minimal anthropogenic additions of trace elements to the soil, and which are not expected to differ from true, naturally occurring concentrations). The rural ambient class includes indigenous vegetation/conservation sites, forestry, and grazing on low-producing grassland, as determined from LCDB5.

Statistical modelling was undertaken using three explanatory layers – a parent material layer, a soil order layer, and a land environment layer that allows for the inclusion of climatic variables – to develop predictions for background (rural ambient) concentrations across New Zealand. The predicted concentrations are displayed as filled contour plots to enable visualisation of the variation in concentrations across the country, and they are also presented in an online tool available at: <https://iris.scinfo.org.nz/layer/114281-abc-predicted-background-soil-concentrations-new-zealand-h3-resolution-9/>.

The models generally provided reasonable explanatory power for the individual trace elements, with the exception of Zn and As: the models markedly under-predict concentrations at the higher concentrations, and hence predict lower (more conservative) background concentrations for the upper percentile estimates (e.g. above the 90th percentile). This under-prediction is less important for As, which had a narrow concentration range across the data set. Further model development is required to obtain better estimates for Zn.

These revised background concentrations were used to update Eco-SGVs developed by Cavanagh and Munir (2019). Median background concentration are used to develop the default Eco-SGVs, with background-adjusted Eco-SGVs also developed where there was a sufficient change in predicted background concentrations.

The use of background concentration information for other purposes typically involves identifying and setting upper thresholds. The challenge has been varying data sets, methods, and spatial coverage, which lead to different estimates. Greater clarity on the outcomes desired to be achieved by the current use of background concentration information in the assessment and management of contaminated land, and as waste acceptance criteria, is needed to ascertain whether these outcomes are being realised. In particular, clause 5(9) of the NES-SC appears to place an undue emphasis on background concentrations, regardless of the risk associated with the concentration of soil contaminants.

Background concentrations should only be relevant to consider where naturally occurring or ambient concentrations are anticipated to be *above* risk-based guidelines for the protection of human health or ecological receptors. Similarly, site-specific background determination is only recommended for sites where naturally occurring concentrations are anticipated to be *above* risk-based guidelines. In these cases it is relevant to determine background concentrations, because it is not reasonable to require remediation to concentrations below naturally occurring, or arguably ambient, concentrations.

There is, nonetheless, a requirement that the risk associated with any elevated concentrations, regardless of whether it is natural or anthropogenic in origin is assessed. However, the NES-SC (which only applies to a piece of land on which a HAIL activity is occurring, has occurred, or is more than likely to have occurred) and clause 5(9) specifically exclude the ability for this risk to be managed.

The model predictions help to identify areas where naturally elevated concentrations may occur. In particular, these areas are considered to be within those areas at and above the 95th percentile of model estimates – there is potential for marked small-scale variations in background concentration in these regions. In these areas, further investigations at a finer spatial scale are required to better delineate areas of naturally elevated concentrations and to understand the associated risk to human health or influence on ecological integrity. Small-scale variations may also give rise to naturally elevated concentrations outside of these areas, although that is considered less likely.

8 Recommendations

Based on the current study we make the following recommendations.

- Further model development should be undertaken for As and Zn to ascertain whether improved predictions can be developed.
- At a higher level, the use of background soil concentration information in existing policy and regulatory setting should be evaluated to determine whether the intended or optimal outcomes are being achieved (e.g. Is the use of background concentrations appropriate for clean fill waste acceptance criteria? Is clause 5(9) achieving its intended purpose, and what is that purpose?).
- In the context of managing contaminated land, clear guidance needs to be given that background concentrations are only relevant to consider when concentrations are *greater* than any risk-based human health or ecological values, and it is considered likely that background concentrations are greater than the risk-based values.
- Consideration should be given to the merit in combining the proposed Class 4 and 5 landfills into one class, with waste acceptance criteria based on ensuring protection of the most sensitive receptor (people, soil ecological receptors). More stringent criteria – or at least criteria based on leaching or off-site movement – could apply where these landfills might be placed close to waterways or groundwater.
- Further investigations at a finer spatial scale are required to better delineate areas of natural elevated concentrations and to understand the associated risk to human health or the influence on ecological integrity.

9 Acknowledgements

Thanks to Graham Curwen and particularly James Ardo (Manaaki Whenua – Landcare Research) for undertaking GIS and geospatial data management support including developing the predicted background concentrations explorer web-page. Thanks also to Andre Egar (Manaaki Whenua – Landcare Research) and Jonathan Caldwell (Waikato Regional Council) for reviewing earlier drafts of the report, and Ray Prebble and Kate Boardman for editorial and formatting support. This study was undertaken with additional funding support from Manaaki Whenuas' Strategic Science Investment Fund.

10 References

- Abraham E. 2020. Cadmium in New Zealand agricultural soils. *New Zealand Journal of Agricultural Research*. 63:202–219.
- Ander EL, Cave MR, Johnson CC, Palumbo-Roe B 2011. Normal background concentrations of contaminants in the soils of England: available data and data exploration. British Geological Survey Commissioned Report, CR/11/145.
- ARC (Auckland Regional Council) 2001. Background concentrations of inorganic elements in soils from the Auckland Region. Technical Publication No. 153. Auckland, Auckland Regional Council.
- Cavanagh JE, 2013. Determining background soil concentrations of contaminants for managing land. Envirolink Advice Grant 1251-MLDC83 for Marlborough District Council. Lincoln, Manaaki Whenua – Landcare Research.
- Cavanagh JE 2021. Managing potentially contaminated soils in the Tasman District. Envirolink Advice Grant 2147-TSDC178 for Tasman District Council. Lincoln, Manaaki Whenua, Landcare Research
- Cavanagh JE, Harmsworth G 2023. An implementation framework for ecological soil guideline values. LC Contract Report: LC4311. Envirolink Tools Grant: C09X2206. Lincoln, Manaaki Whenua – Landcare Research.
- Cavanagh JE, McNeil S, Arienti C, Rattenbury 2015. Background soil concentrations of selected trace elements and organic contaminants in New Zealand. Landcare Research Report 2440 for Envirolink Tools Grant C09X1402.
- Cavanagh JE, Munir K 2019. [Updated] Development of soil guideline values for the protection of ecological receptors (Eco-SGVs): technical document. Manaaki Whenua – Landcare Research Report for Envirolink Tools Grant C09X1402.
- Cavanagh JE, Whitehead B, Vickers S, Stevenson B 2020. Collating soil quality and trace element State of Environment monitoring data. Landcare Research Report 3803 for Envirolink Advice Grant 2050-HBRC254.
- Cave MR, Johnson CC, Ander EL, Palumbo-Roe B 2012. Methodology for the determination of normal background contaminant concentrations in English soils. Nottingham, UK, British Geological Survey (CR/12/003).
- DEFRA (Department for Environment, Food and Rural Affairs) 2012. Environmental Protection Act 1990: Part 2A Contaminated Land Statutory Guidance. April 2012. UK, HM Government. Available online at: <https://www.gov.uk/government/publications/contaminated-land-statutory-guidance> (accessed July 2023).
- Diamond D, Baskin D, Brown D, Lund L, Najita J, Javandel I 2009. Analysis of background distributions of metals in the soil at Lawrence Berkeley National Laboratory. Report prepared for U.S. Department of Energy under Contract DE-AC02005CH11231.
- Diez M, Simón M, Martín F, Dorransoro C, Garcia I, Van Gestel CAM 2009. Ambient trace element background concentrations in soils and their use in risk assessment. *Science of the Total Environment* 407: 4622–4632.

- Environmental Protection Authority Victoria. 2021. Proposed methodology for deriving background level concentration when assessing potentially contaminated land. Melbourne, Victoria.
- FANZ (Fertiliser Association of New Zealand) 2019. Tiered Fertiliser Management System for Soil Cadmium For the management of soil cadmium accumulation from phosphate fertiliser applications. Wellington, Fertiliser Association of New Zealand.
- Hastie TJ, Tibshirani RJ. 1990. Generalized Additive Models. London, Chapman & Hall.
- Heron DW [custodian] 2014. Geological map of New Zealand 1:250,000. GNS Science geological map 1. Lower Hutt, New Zealand, GNS Science.
- Hill RB, Sparling GP 2009. Soil quality monitoring. In: Land and soil monitoring: a guide for SoE and regional council reporting. Hamilton, Land Monitoring Forum. Pp. 27–88.
- ISO (International Standards Organisation) 2011. Soil quality – Guidance on the determination of background values. ISO 19258:2011-09.
- Jarva J, Tarvainen T, Reinikainen J, Eklund M 2010. TAPIR – Finnish national geochemical baseline database. *Science of the Total Environment* 408: 4385–4395.
- Lado R, Hengl L, Reuter HI 2008. Heavy metals in European soils: a geostatistical analysis of the FOREGS Geochemical database. *Geoderma* 148: 189–199.
- Lawley 2009. The Soil-Parent Material Database (SPM-v4): A User Guide. Nottingham, UK, British Geological Survey (OR/08/034). 45pp.
- Leathwick J, Morgan F, Wilson G, Rutledge D, McLeod M, Johnston K. 2002. Land Environments of New Zealand: A Technical Guide. Wellington, Ministry for the Environment.
- Lilburne LR, Webb TH, Hewitt, AE, Lynn IH, de Pauw B 2012. S-map database manual. Lincoln, Landcare Research.
- Martin AP, Lim C, Kah M, Rattenbury MS, Rodgers KM, Sharp EL, Turnbull RE. 2023b. Soil pollution driven by duration of urbanisation and dwelling quality in urban areas: An example from Auckland, New Zealand. *Applied Geochemistry* 148: 105518.
- Martin AP, Turnbull RE, Roudier P, Cavanagh J, Rattenbury MS, Rodgers KM, et al., 2023a. Geochemical Atlas of Aotearoa New Zealand. GNS Science Report 2023/[23]). doi:10.21420/[P9BK-7016]. Lower Hutt, GNS Science.
- Matschullat J, Ottenstein R, Reimann C 2000. Geochemical background – can we calculate it? *Environmental Geology* 39: 990–1000.
- McDowell RW, Taylor MD, Stevenson BA 2013. Natural background and anthropogenic contributions of cadmium to New Zealand soils. *Agriculture, Ecosystems and Environment* 165: 80–87.
- MfE (Ministry for the Environment) 2011. Methodology for deriving standards for contaminants in soil to protect human health. Wellington, Ministry for the Environment.
- MfE (Ministry for the Environment) 2012. Users' Guide National Environmental Standard for Assessing and Managing Contaminants in Soil to Protect Human Health. Wellington, Ministry for the Environment.

- Ministry for the Environment. 2021. Contaminated land management guidelines No 5: Site investigation and analysis of soils (Revised 2021). Wellington: Ministry for the Environment.
- Mortimer N, Rattenbury MS, King PR, Bland KJ, Barrell DJA, Bache F, et al. 2014. High-level stratigraphic scheme for New Zealand rocks. *New Zealand Journal of Geology and Geophysics* 57: 402–419.
- Mrvić V, Kostić-Kravljanac L, Čakmak D, Sikirić B, Brebanović B, Perović V, et al. 2011. Pedogeochemical mapping and background limit of trace elements in soils of Branicevo Province (Serbia). *Journal of Geochemical Exploration* 109: 18–25.
- Percival H, Sutherland A 2002. Maps of total soil concentrations (background levels) of chromium, copper, lead, nickel, vanadium, and zinc in the Taranaki Region. Landcare Research Report LC0102/152 prepared for Taranaki Regional Council.
- Rattenbury MS, Isaac MJ 2012. The QMAP 1:250 000 Geological Map of New Zealand project. *New Zealand Journal of Geology and Geophysics* 55: 393–405.
- Reimann C, Filmozer P, Garrett RG 2005. Background and threshold: critical comparison of methods of determination. *Science of the Total Environment* 346: 1–16.
- Reimann C, Garrett RG 2005. Geochemical background: concept and reality. *Science of the Total Environment* 350: 12–27.
- Tonkin & Taylor 2006. Background concentrations of selected trace elements in Canterbury soils. Environment Canterbury Report No. R07/1.
- Tonkin & Taylor Ltd 2007. Background concentrations of selected trace elements in Canterbury soils. Addendum 1: Additional samples and Timaru specific background levels. Report R07/1/2 for Environment Canterbury.
- URS 2003. Determination of common pollutant background soil concentrations for the Wellington Region. Report prepared for Greater Wellington Regional Council.
- US EPA (United States Environmental Protection Agency) 2002. Guidance for comparing background and chemical concentrations in soil for CERCLA sites. EPA 540-R-01-003, OSWER 9285.7-41. US, United States Environmental Protection Agency.
- Wiser SK, Buxton RP, Clarkson BR, Hoare RJB, Holdaway RJ, Richardson SJ, et al. 2013. New Zealand's naturally uncommon ecosystems. In Dymond JR ed. *Ecosystem services in New Zealand – conditions and trends*. Manaaki Whenua Press, Lincoln, New Zealand. Pp. 49–61.
- Wood SN 2017. *Generalized Additive Models: An Introduction with R*. 2nd edition. London, Chapman & Hall.
- Waikato Regional Council 2022. Standard operating policies for defining cleanfill acceptance criteria. WRC Doc # 10581789. Hamilton, Waikato Regional Council
- WasteMINZ 2022. Technical guidelines for disposal to land (Land disposal technical guidelines).
https://www.wasteminz.org.nz/files/Disposal%20to%20Land/TG%20for%20Disposal%20to%20Land_12Oct22_FINAL.pdf

Appendix 1 – Relative influence of individual variables

The following figures present the relative influence or partial effect of individual variables of landuse, LENZ, parent material and soil order. For example, for landuse_ambient, the reference is "Ambient", and the partial effect is the relative change in As concentration when the landuse_ambient is changed from "Ambient" to another level (e.g. "Urban"), after the effect of all other variables is accounted for. The marker is the nominal value for the reference; the error bars are the 95% range. A dashed line is drawn through the reference level for each explanatory variable. Those variables that do not cross the dashed line are significantly different to the reference level.

In the following graphs LENZ is presented as the letter assigned to each level 1 category – the descriptors are presented in Table A1. Figures 3-5 show the distribution of parent material, soil order and LENZ category across New Zealand.

Table A1. LENZ level 1 categories. Source (Leathwick et al 2002).

LENZ Level 1 Category
A - Northern Lowlands
B - Central Dry Lowlands
C - Western and Southern North Island Lowlands
D - Northern Hill Country
E - Central Dry Foothills
F - Central Hill Country and Volcanic Plateau
G - Northern Recent Soils
H - Central Sandy Recent Soils
I - Central Poorly Drained Recent Soils
J - Central Well-drained Recent Soils
K - Central Upland Recent Soils
L - Southern Lowlands
M- Western South Island Recent Soils
N - Eastern South Island Plains
O - Western South Island Foothills and Stewart Island
P - Central Mountains
Q - Southeastern Hill Country and Mountains
R - Southern Alps
S - Ultramafic Soils
T - Permanent Snow and Ice

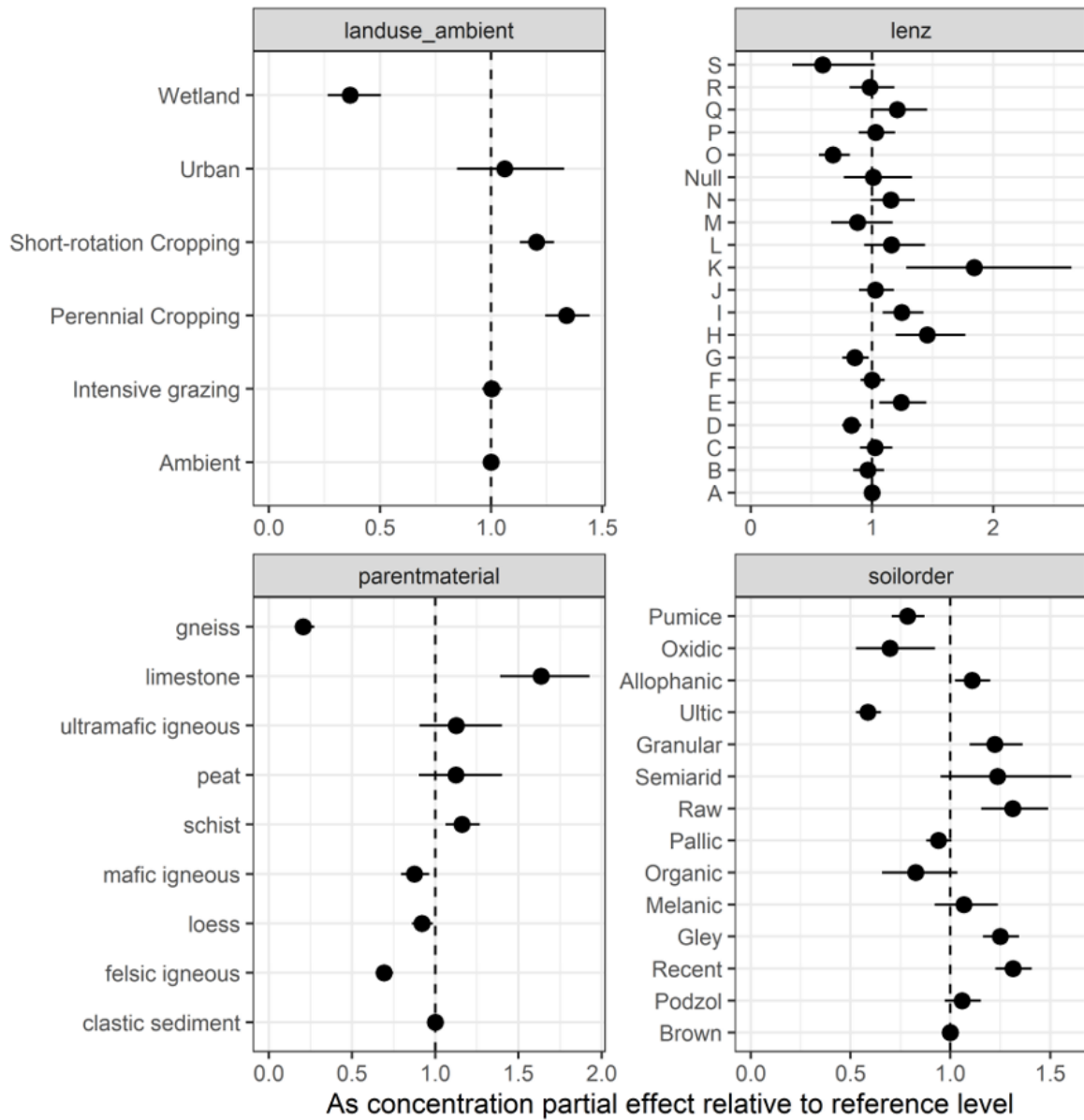


Figure A1. Partial effect of arsenic for each explanatory variable with respect to the reference for that explanatory variable, after the effect of all other explanatory variables is accounted for by the GAM. The marker is the nominal value; the error bars are the 95% range. A dashed line is drawn through the reference level for each explanatory variable. Those variables that do not cross the dashed line are significantly different to the reference level.

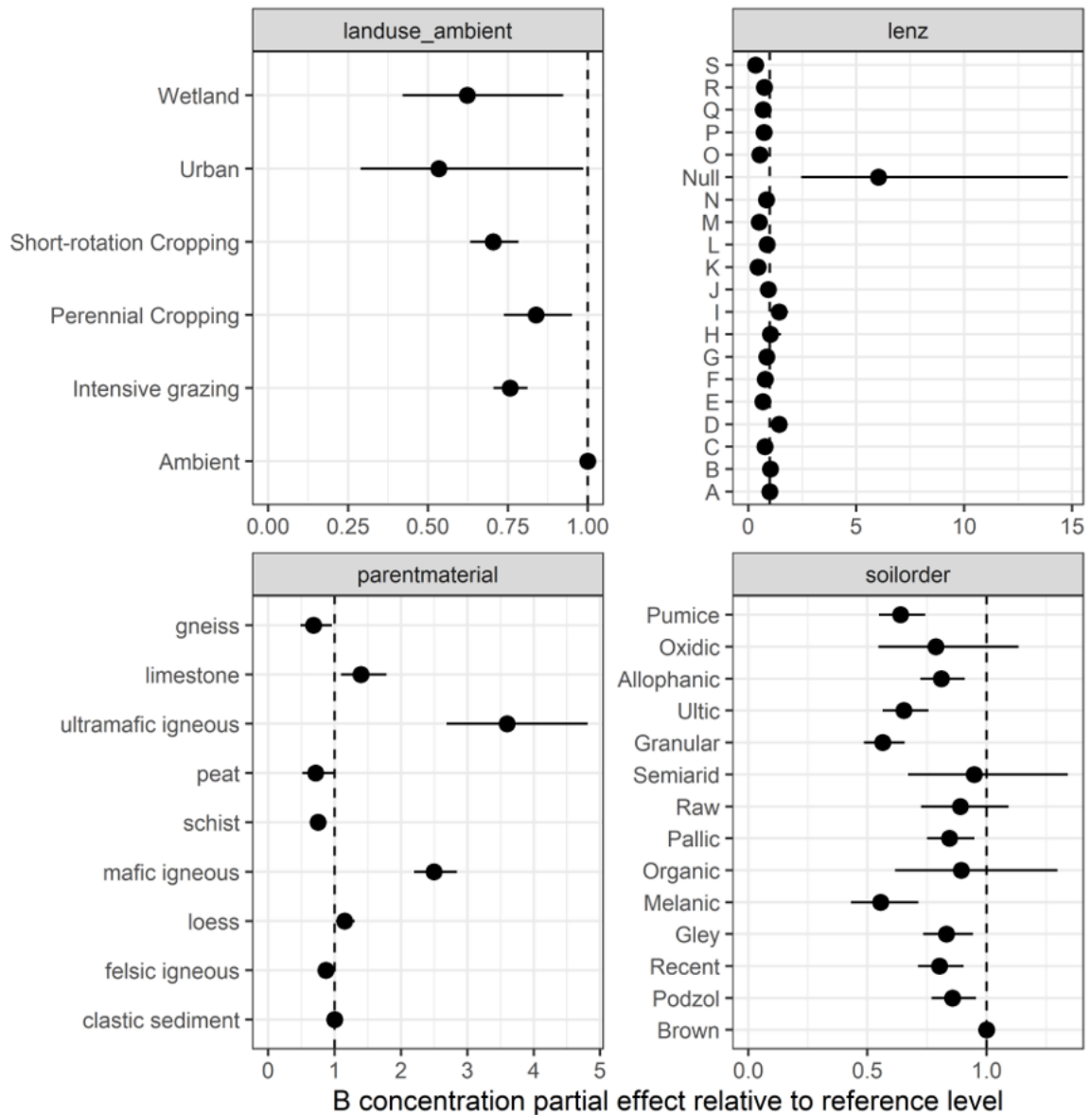


Figure A2. Partial effect of boron for each explanatory variable with respect to the reference for that explanatory variable, after the effect of all other explanatory variables is accounted for by the GAM. The marker is the nominal value; the error bars are the 95% range. A dashed line is drawn through the reference level for each explanatory variable. Those variables that do not cross the dashed line are significantly different to the reference level

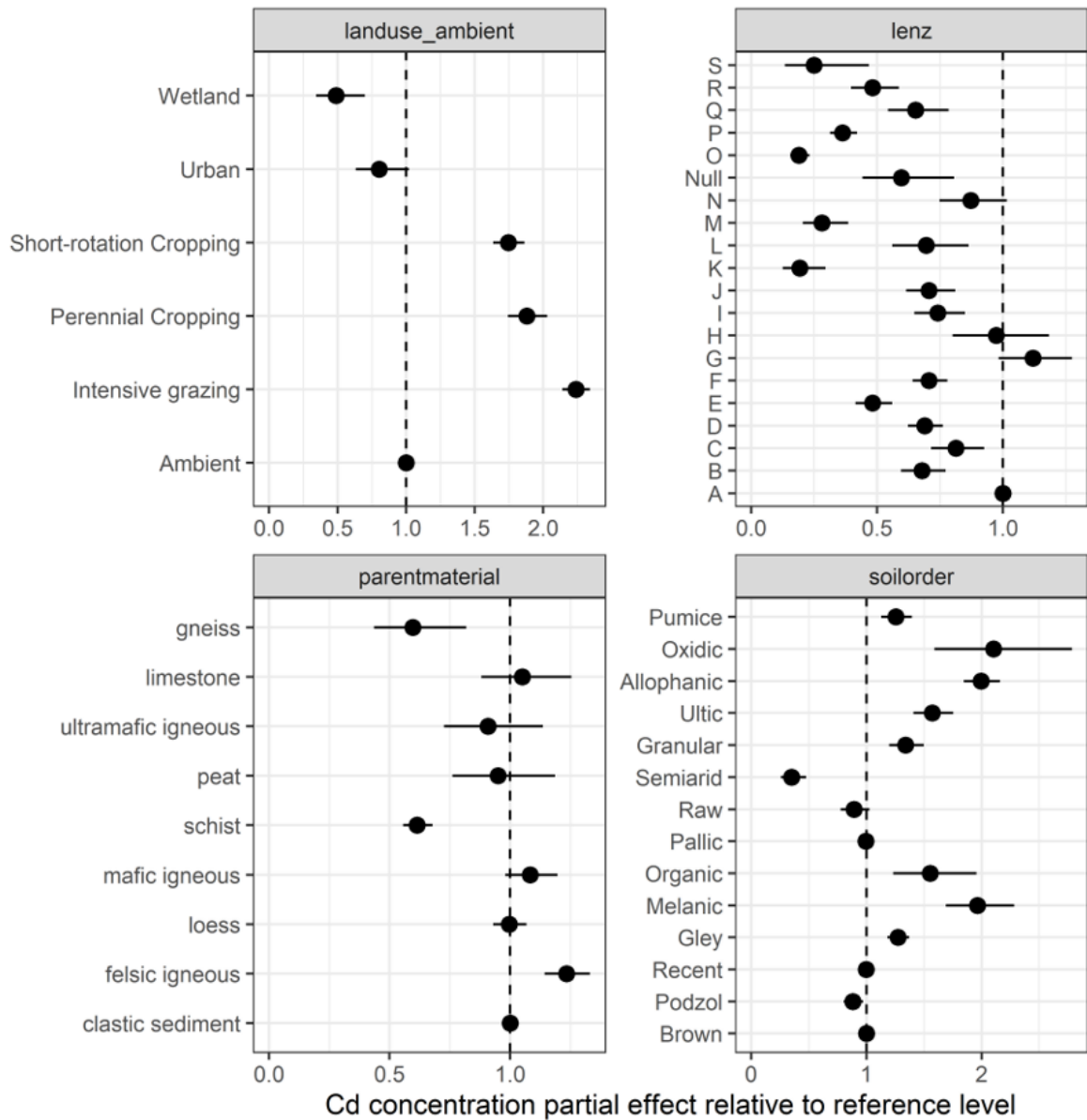


Figure A3. Partial effect of cadmium for each explanatory variable with respect to the reference for that explanatory variable, after the effect of all other explanatory variables is accounted for by the GAM. The marker is the nominal value; the error bars are the 95% range. A dashed line is drawn through the reference level for each explanatory variable. Those variables that do not cross the dashed line are significantly different to the reference level

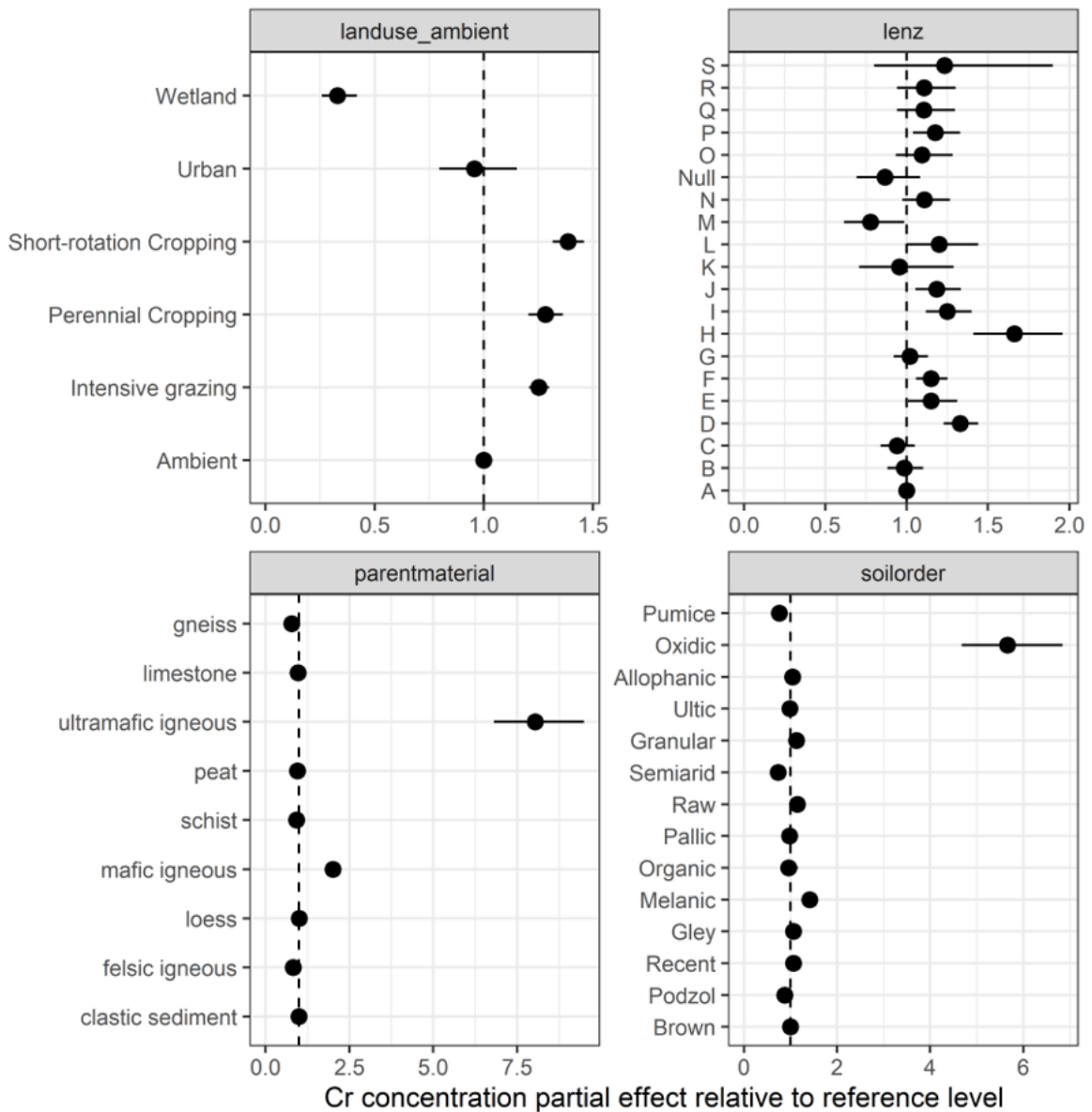


Figure A4. Partial effect of chromium for each explanatory variable with respect to the reference for that explanatory variable, after the effect of all other explanatory variables is accounted for by the GAM. The marker is the nominal value; the error bars are the 95% range. A dashed line is drawn through the reference level for each explanatory variable. Those variables that do not cross the dashed line are significantly different to the reference level.

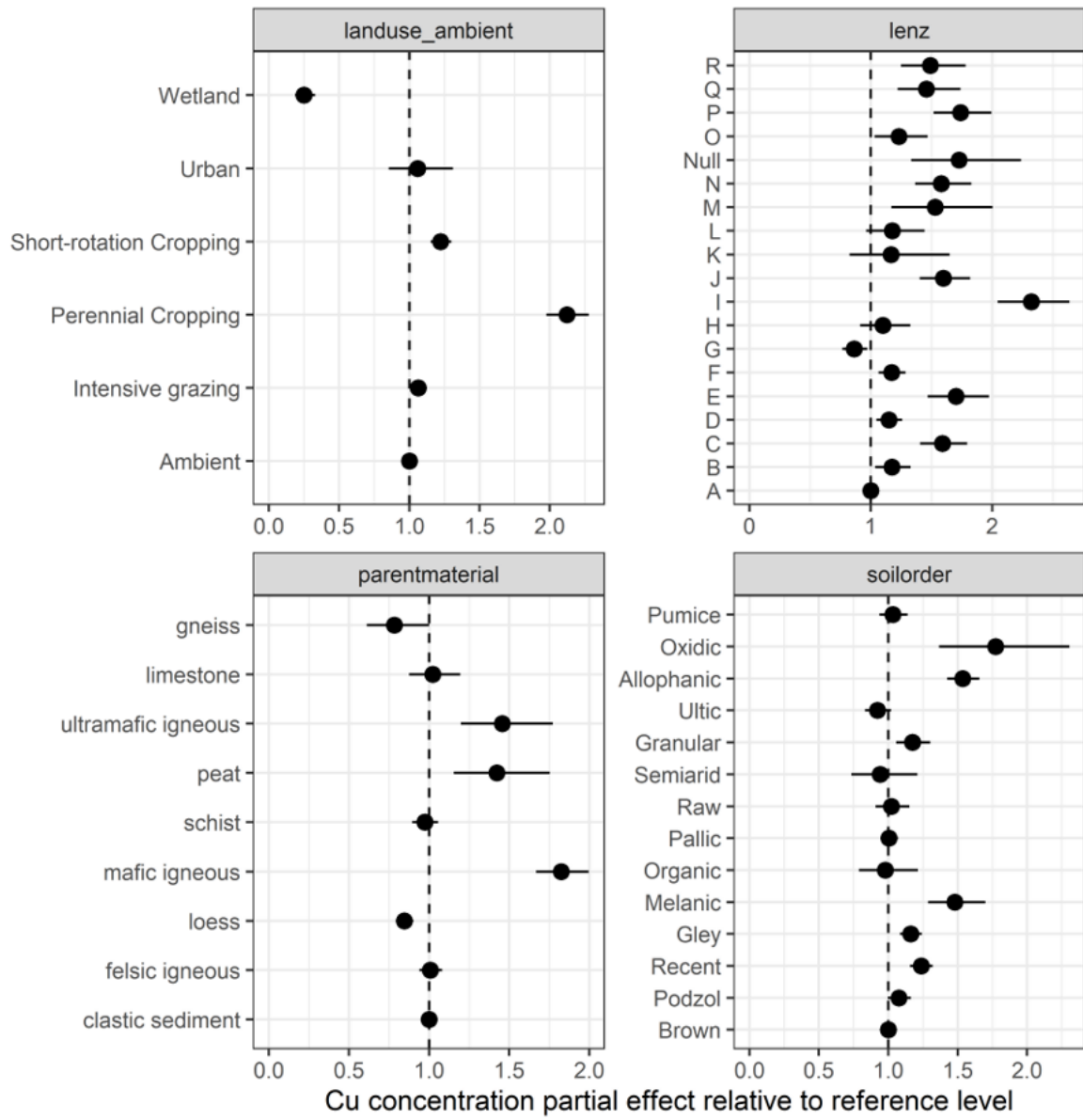


Figure A5. Partial effect of copper for each explanatory variable with respect to the reference for that explanatory variable, after the effect of all other explanatory variables is accounted for by the GAM. The marker is the nominal value; the error bars are the 95% range. A dashed line is drawn through the reference level for each explanatory variable. Those variables that do not cross the dashed line are significantly different to the reference level.

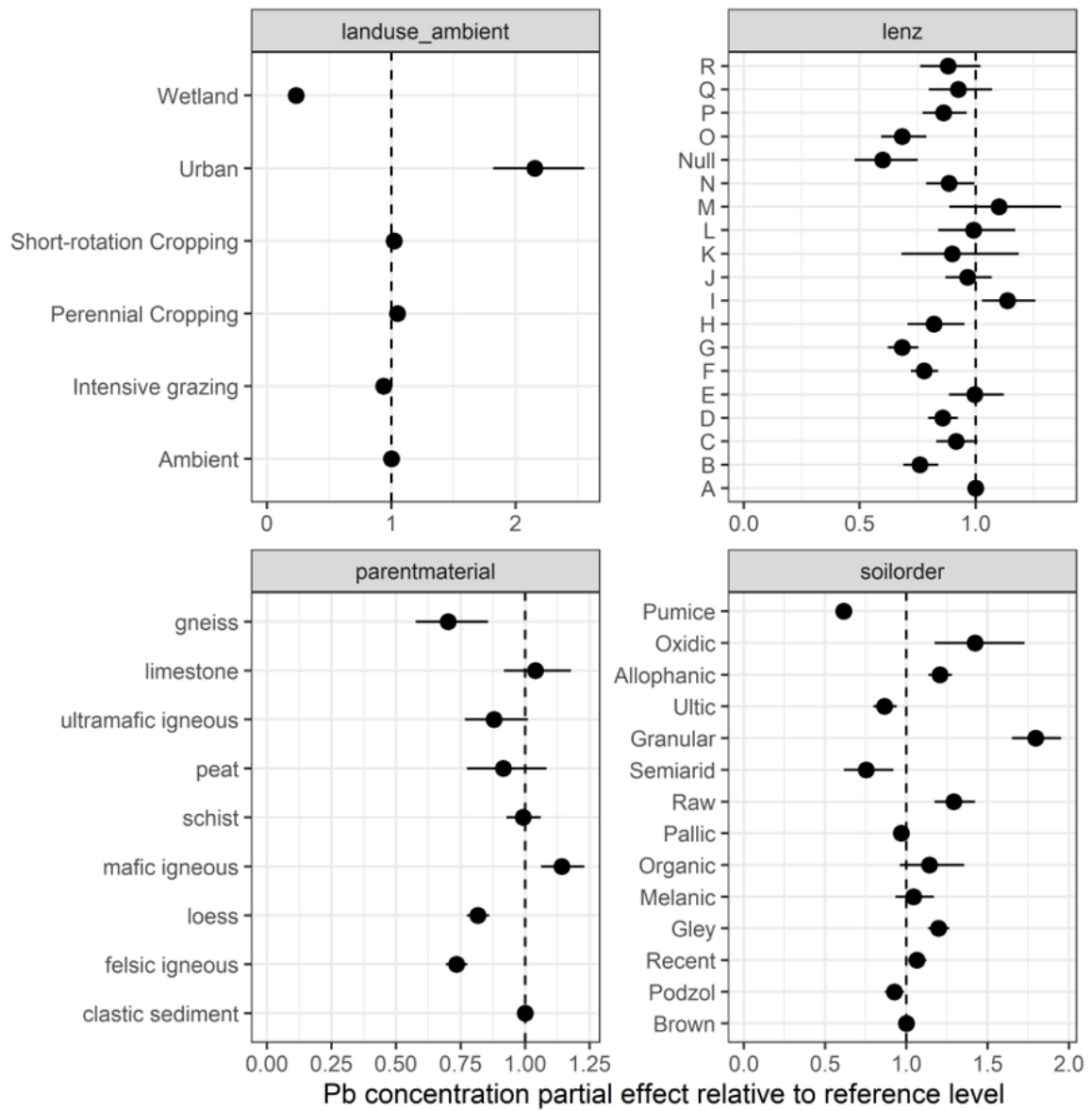


Figure A6. Partial effect of lead for each explanatory variable with respect to the reference for that explanatory variable, after the effect of all other explanatory variables is accounted for by the GAM. The marker is the nominal value; the error bars are the 95% range. A dashed line is drawn through the reference level for each explanatory variable. Those variables that do not cross the dashed line are significantly different to the reference level.

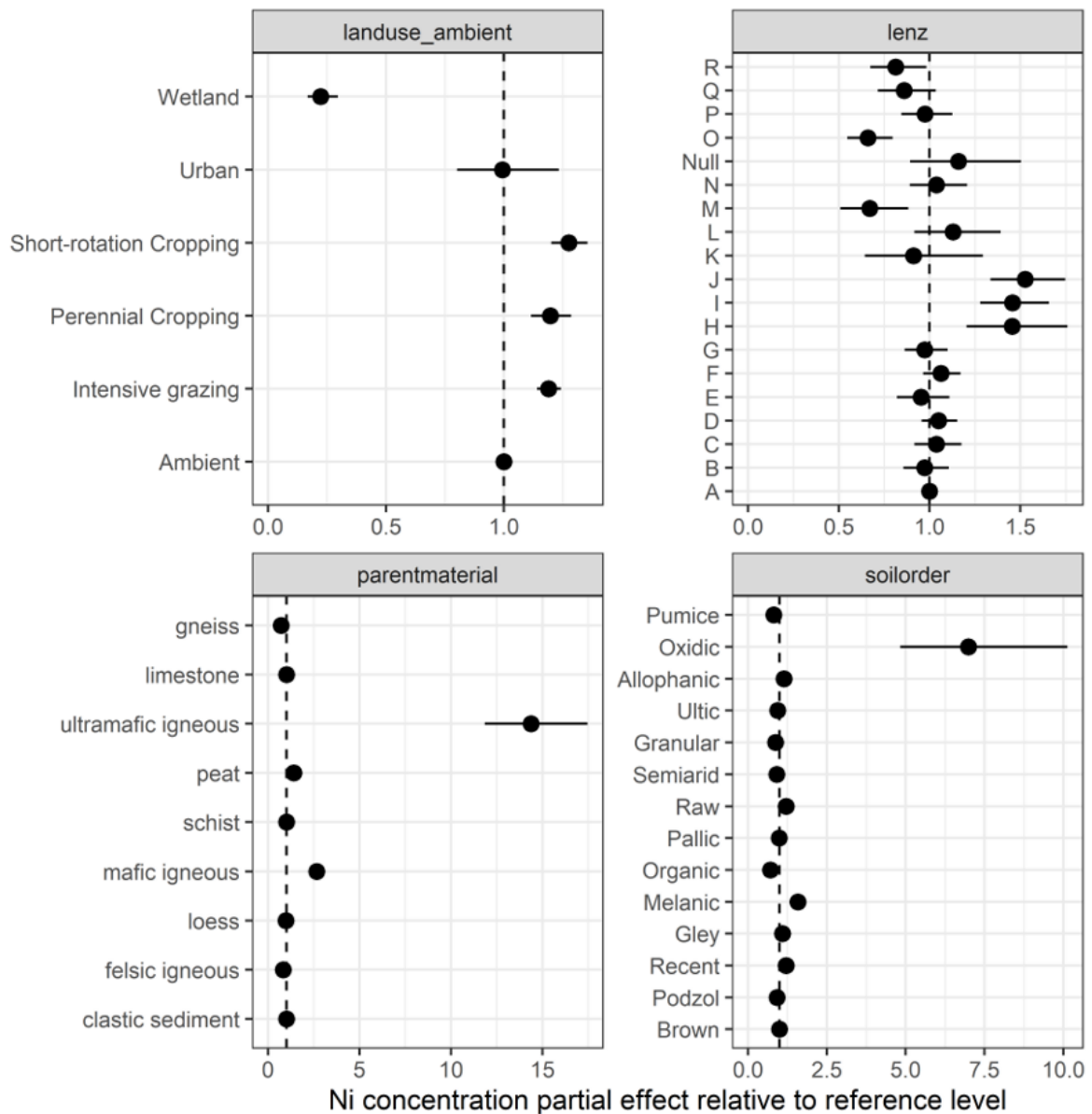


Figure A7. Partial effect of nickel for each explanatory variable with respect to the reference for that explanatory variable, after the effect of all other explanatory variables is accounted for by the GAM. The marker is the nominal value; the error bars are the 95% range. A dashed line is drawn through the reference level for each explanatory variable. Those variables that do not cross the dashed line are significantly different to the reference level.

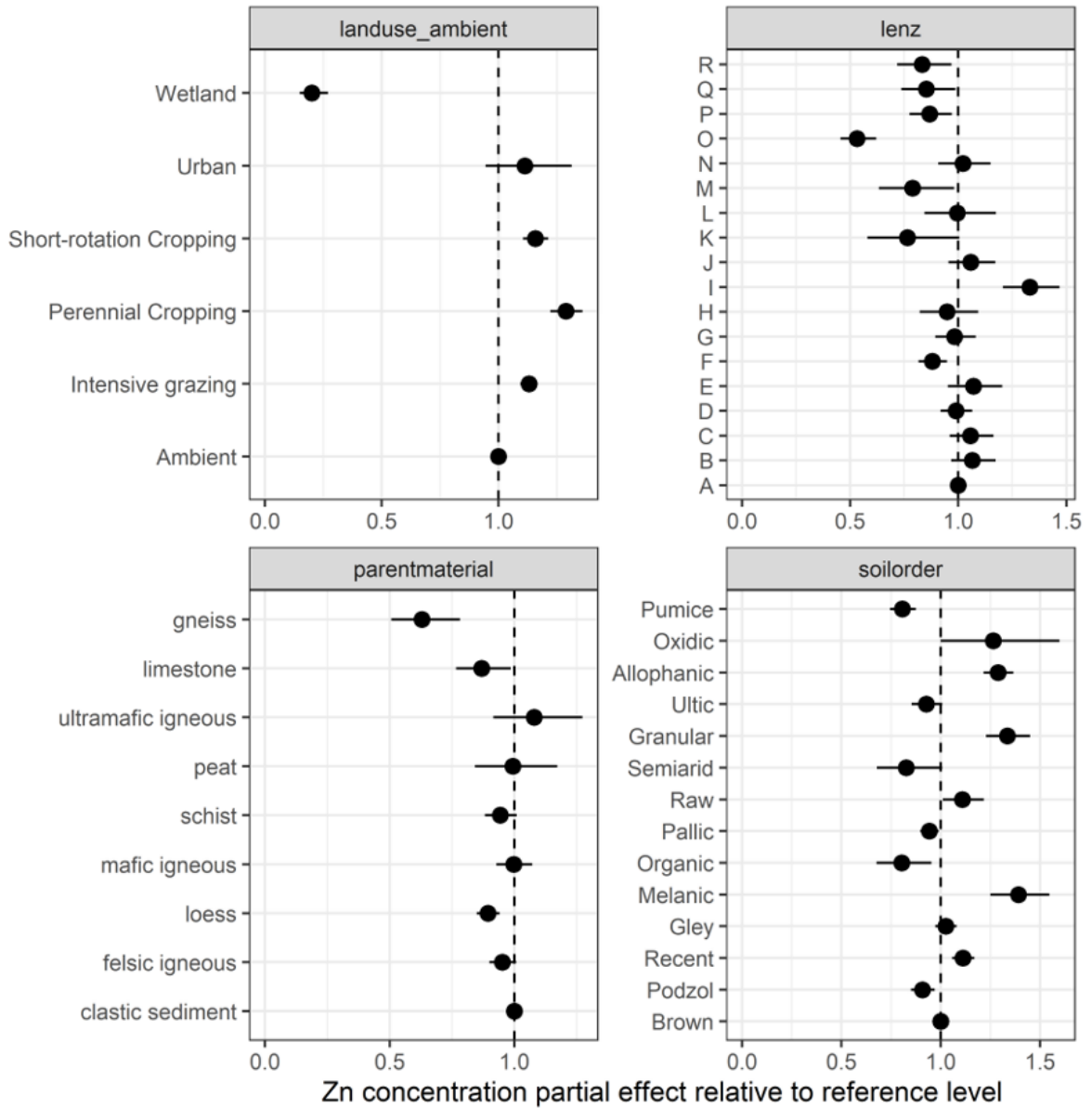


Figure A8. Partial effect of zinc for each explanatory variable with respect to the reference for that explanatory variable, after the effect of all other explanatory variables is accounted for by the GAM. The marker is the nominal value; the error bars are the 95% range. A dashed line is drawn through the reference level for each explanatory variable. Those variables that do not cross the dashed line are significantly different to the reference level.