Exploring high-resolution targeting till geochemical data in central Lapland, Finland

Charmee Kalubowila[1]*, Pertti Sarala[1], Solveig Pospiech[2] and Peter Filzmoser[3]

[1] *University of Oulu, Oulu Mining School, P.O. Box 3000, Oulu, Finland*

[2] *Helmholtz-Zentrum Dresden-Rossendorf, Helmholtz Institute Freiberg for Resource Technology, Bautzner Landstr. 400, Dresden, Germany*

[3] *TU Wien, Institute of Statistics and Mathematical Methods in Economics, Wiedner Hauptstraße 8-10, Vienna, Austria*

Abstract

The targeting till geochemical dataset of Finland was collected during the 1970s and early 1980s by the Geological Survey of Finland (GTK). The dataset covers central Finnish Lapland and some areas in Ostrobothnia and eastern Finland. A subset of data from central Lapland has been examined in this study. The goal of this study is to present a detailed description of the high-resolution targeting till geochemical data, together with data preprocessing and preliminary data analysis. The geochemical data are considered to be compositional data. Box plots, original concentration maps, and quantile-quantile (Q-Q) plots were used to explore the data and original and centred log-ratio (clr) transformed concentration maps with lithological units, heat maps and principal component analysis (PCA) plots were used as analytical tools. The data analysis revealed mismatches between neighbouring map sheets which were caused by the problems related to till sampling, analytical methods, and the complex levelling problems of elements. A better insight into the data could be obtained by analysing clr-transformed data rather than by analysing raw data. A PCA biplot is a powerful tool for identifying the important patterns in the area and correlating them to the underlying geology.

Keywords: till geochemistry; data analysis; compositional data; PCA; Clr-transformation

*Corresponding author (e-mail: charmee.kalubowila@oulu.fi)

Editorial handling: Antti E. Ojala (e-mail: antti.e.ojala@utu.fi)

1. Introduction

The discovery of new mineral deposits is necessary to meet the growing global demand for mineral resources (Gonçalves et al. 2024). Many challenges are associated with mineral exploration such as the mineral deposits being buried under transported cover. The identification of surficial geochemical anomalies in till during the exploration process is essential in tracing the targeted mineral deposits in these environments. This approach is widely applied in glaciated terrains such as in Finland, where most of the bedrock is covered by glacigenic sediments particularly till (Sarala 2015; Taivalkoski et al. 2024). Till is one of the most appropriate natural sampling media as it is also derived directly from the bedrock (Salminen & Tarvainen 1995). However, drift prospecting using till geochemistry still presents several challenges, including the need to define an anomalous threshold based on local background levels, statistical normalization and levelling of multiple datasets, addressing potential artefacts from the sampling process, and dealing with closed compositional datasets (McClenaghan et al. 2023; Garrett & Sinding-Larsen 1984; Grunsky 2010; Plouffe et al. 2022; Paulen et al. 1998).

A geochemical survey which was carried out in the 1970s and early 1980s by the Geological Survey of Finland (GTK) generated a high-resolution geochemical dataset (targeting till) which covers central Lapland, some areas in Ostrobothnia and eastern Finland (Gustavsson et al. 1979). Moreover, in this study, another regional scale till geochemical dataset (regional till; Salminen & Tarvainen 1995), was used together with the targeting till dataset to explain certain traits such as data distribution and levelling issues between map sheets. It is presumed that the targeting till dataset with dense sampling density has a higher potential of focusing the prospective areas for mineral deposits than the sparser regional till data. However, the usability of targeting till data is restricted due to the poor quality of the analytical method. Thus, extensive data preprocessing is required.

This study was conducted as part of the EU-funded project SEMACRET (Sustainable exploration for orthomagmatic (critical) raw materials in the EU: charting the road to the green energy transition). Thus, this paper, concerns Co, Cr, Cu, Fe, Ni, V, and Ti which are important ore elements in the mafic-ultramafic ore deposits and are listed as critical raw materials by the EU. The objective of this paper is to present a detailed description of the high-resolution targeting till dataset, together with data preprocessing and preliminary data analysis, which would assist in future prospectivity mapping programmes. Summary statistics, concentration distribution maps, boxplots and compositional data analysis (CoDA) techniques are used to explore the data. Moreover, exploratory tools such as quantile-quantile (Q-Q) plots, principal component analysis (PCA) and heat maps are used to understand the behavioural patterns of the chemical elements.

1.1 Targeting till

Targeting till geochemistry survey samples comprise soil (C-horizon) samples collected by GTK along sampling lines in 1971–1983 (GTK n.d.). The objective of the geochemical survey was to map the prevalence of elements in the ground. The samples were collected by 1:20 000 map sheets (Fig. 1a) arranged in a discretionary order, and the sampling areas cover selected areas of central and northern Finland. Most of the samples are composed of till. However, the data include other types of sedimentary material such as silt and sand, pre-glacial weathered bedrock and mixed intermediate sample material containing both till and weathered bedrock. The sampling followed a line sampling grid with the point density of sampling points varies between 6–12 samples/km^2. The line interval was 500–2000 metres, and the point interval was 100–400 metres (Gustavsson et al. 1979). Sampling was done using percussion drilling with a flow-through bit. The sampling depths were between 0.1 and 25.3 metres, with an average sampling depth

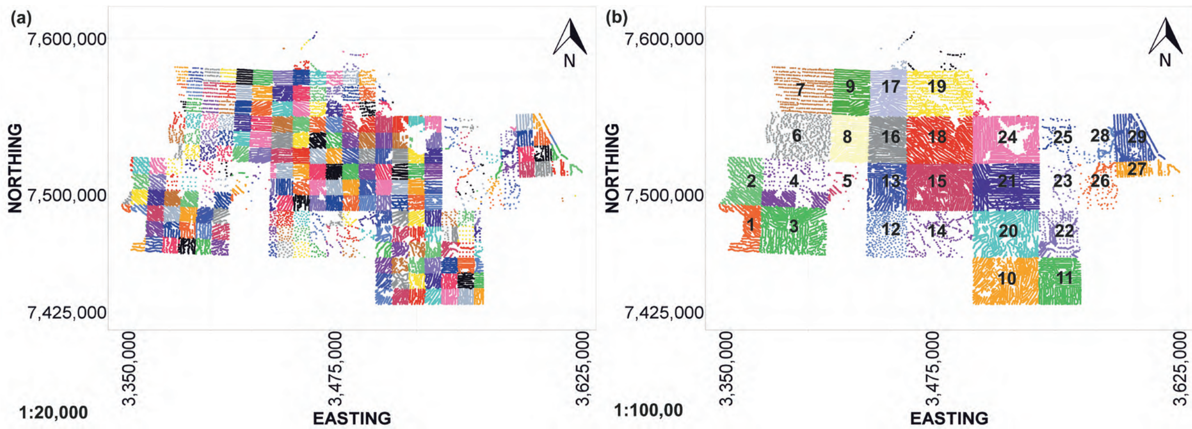

Figure 1. Sampling location map of targeting till a) original sampling model of 1:20 000 scale map sheets b) analytical batches combined from original sampling model following 1:100 000 scale map sheet divisions and map sheet numbers are mentioned on the map sheets.

of 2 metres. A total of 385 000 samples were taken, approximately of which 300 000 are still preserved in the archive. Both the fine fraction (i.e., <0.063 mm) and the coarse fraction (i.e., 0.06–0.25 mm) are preserved. Among these, 191 559 samples were collected from central Lapland.

A size fraction of <0.063 mm was sieved from the original samples, and the concentrations of 17 chemical elements were analysed with an emission quantometer EKV (Danielsson et al. 1959; Danielsson & Sundkvist 1959 a, b). Based on Gustavsson et al. (1979), the samples were collected and analysed continuously on 1:100 000 scale map sheets (Fig. 1b). The analysis capacity with the EKV method was approximately 250 samples per day. Quality control was carried out with the reference samples and repetition of analyses, as well as with duplicate samples taken from every 30th sample point. The total number of control samples was about 15 % of routine samples.

1.2 Regional till

The regional till geochemical dataset is a basic surface geochemical dataset in Finland (Salminen & Tarvainen 1995). The data provide chemical characteristics of till material in different areas, and they can be used to conduct ore exploration and environmental studies. The sampling was done in 1983–1991 with a density of 1 sample / 4 km^2 and dataset covers the entire land area of Finland. 82 062 samples were collected (x3 fractions: fine < 0.063 mm, medium 0.063–0.5 mm and coarse > 2 mm) of which around 95 % are stored in the archive. The samples have been taken from fresh, unoxidized till (horizon-C), in many places below the groundwater table. The sampling depth is approximately 1.5–2 m, and the samples are combined field samples (i.e., subsamples) from 3–5 separate sampling points, located within c. 10 m distance from each other. Estimated coordinates for the sampling points saved in the dataset, have been calculated from the coordinates based on the subsampling locations. The archived samples were the sieved <0.063 mm size fraction, not the original field sample. Furthermore, earlier collected till sample material was used to combine the regional-scale samples in the areas covered by the targeting till dataset. Subsequently, the subsamples were collected from five consecutive sampling points along the line, so that the centre point would be as close as possible to the proposed sampling point in accordance with the sampling strategy (1 sample/4 km^2) of the regional till geochemistry.

A size fraction of <0.063 mm was sieved from each sample for chemical analysis. The analyses were made by inductively coupled plasma atomic emis-

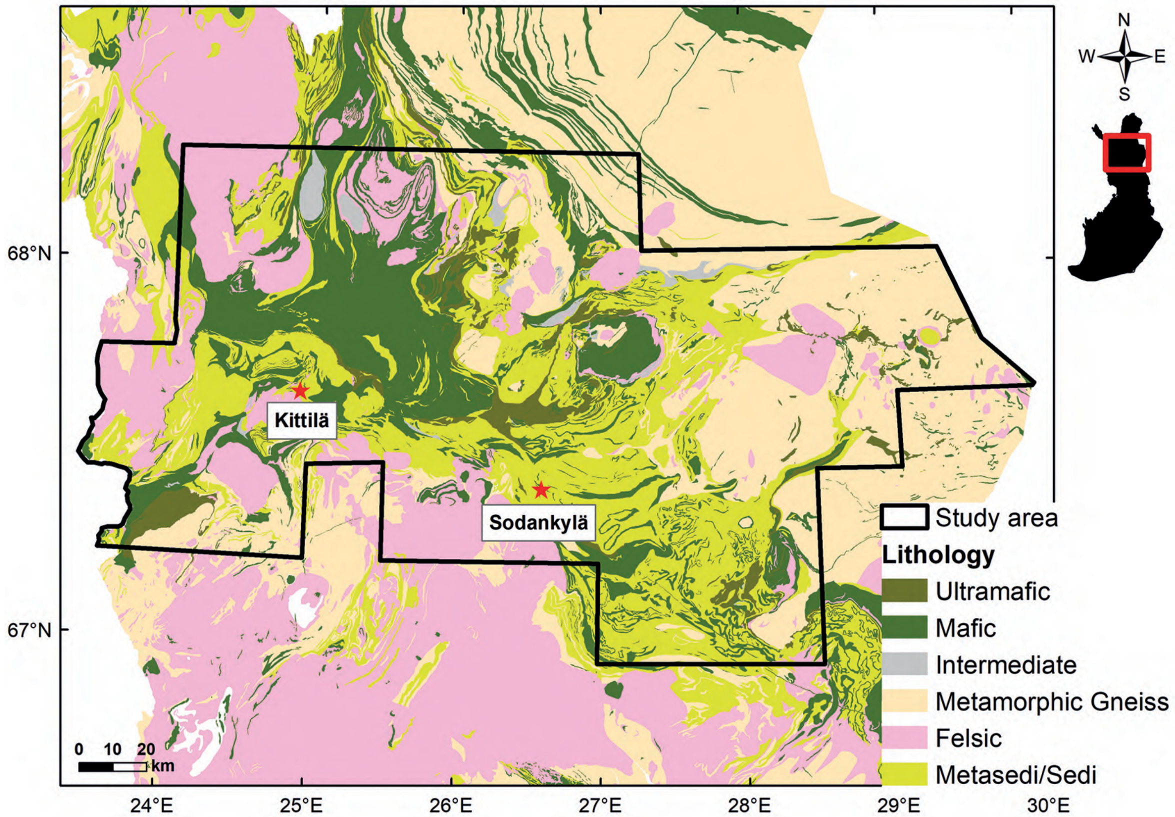

Figure 2. Lithological rock units covered by the sampling points of targeting till data (study area). The bedrock map is constructed using digital map information from the GTK Maankamara: http://gtkdata.gtk.fi/Maankamara/index.html © Geological Survey of Finland (2024, GTK Basic license version 1.1), Bedrock of Finland 1:200 000 (2023–2024), CC-BY-4.0.

sion spectrometry (ICP-AES) after aqua regia-based partial leaching. The data include concentrations of soluble elements for ca. 25 chemical elements.

2. Study area and the geological setting of central Lapland

2.1 Study area

The study area is located in the central Lapland region of northern Finland, which is a part of the Fennoscandian shield. The bedrock in northern Finland mainly consists of Archean granite gneiss–greenstone terrains and Paleoproterozoic schist belts (Hanski & Huhma 2005). The study area contains sampling points which cover almost a rectangular area (Fig. 2). Furthermore, it comprises a part of the Central Lapland Greenstone Belt, Central Lapland Granitoid Complex, Archean gneiss and granite, Archean supercrustal belts and layered mafic–ultramafic intrusions. The lithological units in this area can be observed in the simplified geological map (Fig. 2). In addition, several Ni-Cu-PGE, Cr-V, Ti-V, Fe and Au deposits can be found in this area.

In this paper, map sheets M11 and M15 were selected to explain the data analysis in sections 3 and 4, as the areas covered by those map sheets contain Ni-Cu-PGE and Cr-V deposits. The area in map sheet M15 (Fig. 1b) contains two Ni-Cu-PGE deposits, Kevitsa and Sakatti. Kevitsa

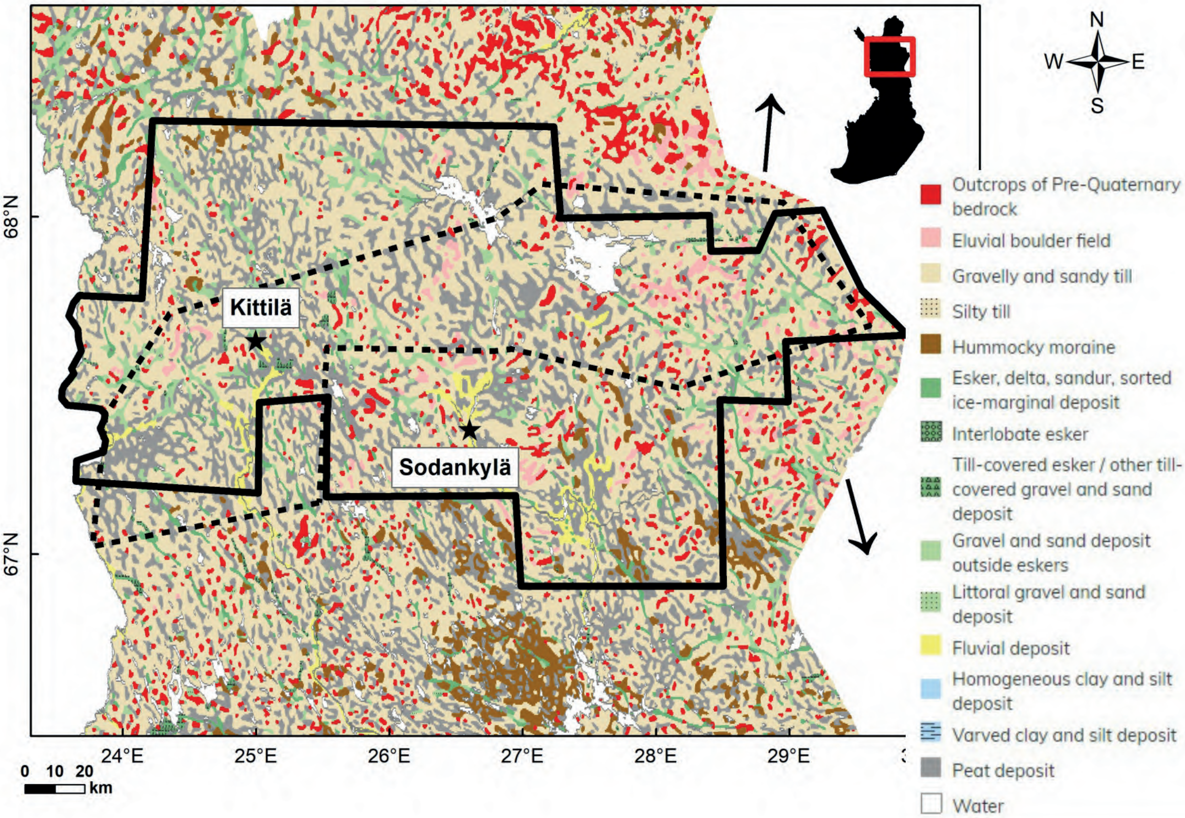

Figure 3. Quaternary geological map of Finland, solid line indicates the study area, and the dotted line marks the ice divide zone. The map is constructed using digital map information from GTK Maankamara: http://gtkdata.gtk.fi/Maankamara/index.html © Geological Survey of Finland (2024, GTK Basic license version 1.1), CC-BY-4.0.

is a Ni mine which produces Cu, Ni, Au, Pt, Pd and contains Co as well. It primarily composed of disseminated sulfidic mineralization, and the ore is comparably richer in Cu than Ni (Santaguida et al. 2015). Moreover, the Kevitsa intrusion contains ultramafic, gabbroic and granophyric areas with magmatic layers (Törmänen et al. 2016). The Sakatti deposit hosts sulfide mineralization that occurs as disseminated sulfides, sulfide veins, and semi-massive to massive sulfides hosted in a large olivine cumulate body (Brownscombe et al. 2015). Additionally, the area covered by map sheet 11 contains the Cr-V bearing layered intrusion, Akanvaara. It hosts several zones that are enriched in different metals. It contains chromitite layers, zones enriched in PGE/Au and vanadium-enriched layers of magnetite gabbro (Karinen & Törmänen 2016).

2.2 Quaternary geology

The bedrock in northern Finland is mostly covered by glacigenic sediments and peat lands. The area is located in the centre of the Fennoscandian ice sheet, where glaciers have existed several times during the Quaternary period. It is also the area in which the last ice-divide zone of the Late Weichselian glaciation was located (Fig. 3). This is characteristically seen in morphology reflecting passive glacigenic features, till stratigraphy representing a long glacial history, and pre-glacial weathered bedrock up to 100 m depth (Sarala & Ojala 2008; Johansson et al. 2011; Sarala et al. 2016). The till stratigraphy is composed of several till beds indicating weak glacial erosion in the area (Hirvas 1991). Based on earlier stratigraphical studies there are, in general, two different basal till beds (Johansson et al. 2011). The younger (i.e. upper) till is formed during the Late Weichselian glacial phase, and the older bed was most probably deposited during the Early and/or Mid-Weichselian glaciations. Above them is a surficial till bed, representing the last melting phase of the glacier. In certain locations, such as within topographic depressions, up to six till beds may be present, which represent glaciations of up to several hundreds of thousand years in age (Hirvas 1991; Putkinen et al. 2020).

One remarkable stratigraphical feature in northern Finland, relates to the late phase of deglaciation. When glacier retreated and ice melted, most of the central Lapland has been in the supra-aquatic position. As the melting progressed, the margin of glacier was divided into smaller ice lobes or tongues which together with variable topography bordered low-land areas to form small, short-lived pro-glacial water bodies, called as ice-lakes (Johansson 2005, 2007). Meltwater, which was released from the glacier transported fine-grained, silty material to the lakes where it deposited on the bottom as a glaciolacustric sediment. This silty sediment layer covers underlying till in the lowest parts in topography. However, its thickness is typically only 0.5–1 m and material are well-stratified. The sampling depth for till samples is typically more than one metre and therefore, the glaciolacustric sediments have not mixed and disturbed the underlying till. Instead, the pre-glacial weathered bedrock, which is commonly found in the central Lapland can be mixed with the lowest till material (Hirvas 1991; Sarala & Ojala 2008; Sarala 2015). In places, the effect of weathered bedrock material in till can be seen as increased or depleted concentrations of elements depending on the bedrock lithology, intensity of weathering, and the level of erosion profile.

3. Materials and methods

The targeting till (refer section 1.1) dataset of Finland is the primary dataset that has been analysed in this paper and a sub-data set from the central Lapland area has been extracted for further analysis (Fig. 2). Moreover, the regional till (refer section 1.1) dataset has been used in some places to enhance the understanding of the reader about the data distribution of the area. Data pre-processing and data analysis were carried out using R (R Core Team 2022). The integration of original and pre-processed data and map plotting was done using the commercial GIS, ArcGIS 10.8.1 a nd ioGAS 7.4.

3.1 Methodology

3.1.1 Preselecting sampling material and summary statistics

Initially, the targeting till dataset was considered and the fine fraction (<0.063 mm) of samples that were collected from the C-horizon of the till cover by using percussion drilling and the test pits were filtered. There were 149 730 sample points after this stage. Summary statistics of the pre-selected area in central Lapland were calculated (Table 1). The table does not contain mean or standard deviation, as for compositional data these measures could be misleading, e.g., when applying classical statistical tests (Reimann et al. 2012).

The data points that contained null data for all element concentrations were removed and after addressing the duplicates, the dataset contained 146 081 samples. This dataset includes analyses of element concentrations in different depths related to each sample point, but most of the samples have been collected from the lowest till. Therefore, the data points taken from the greatest depth concerning each sampling point were considered for further analysis, under the assumption that the deepest samples represent the shortest transportation distance. There are 99 452 sample points after the selection and this data were used for exploring and further analyses of data.

Table 1: Summary statistics for the elements of the targeting till data set (Soil type = till, Layer = C, Sampling method = Percussion drilling and test pits, N = 149,730, Size fraction < 0.06 mm). LDL: lower detection limit; Min: minimum; Q50 = median; Max: maximum; UDL = upper detection limit. All values in ppm.

Element	LDL	Min	Q50	Max	UDL
Ag	0-9.71	11.3	0.92	187	-
Al	0-9910	-16 000	107 000	130 000	131 000-289 000
Ca	0-9990	10 000	11 400	223 000	251 000-841 000
Cr	0-99.9	100	248	4950	5010-9540
Co	0-14.9	15	21.1	498	508-826
Cu	0-9.99	10	72.4	4990	5520-14 000
Fe	0-9980	10 000	64 800	250 000	251 000-1 180 000
K	0-4990	5000	11 400	100 000	101 000-103 000
Mg	0-994	-2770	17 600	133 000	229 000
Mn	0-149	-359	843	2980	3030-4600
Na	0-9990	-487	23 200	257 000	738 000-985 000
Ni	0-19.9	-0.0000009	97.6	3000	3010-4790
Pb	0-49.9	50	25.4	997	1030-1520
Si	0-43 600	-56 600	266 000	449 000	-
Ti	0-988	1000	7540	24 700	-
V	0-149	-174	371	1500	1510-2180
Zn	0-99.9	100	25	1840	-

3.1.2 Data Cleaning

The targeting till geochemical dataset was associated with serious data quality issues, such as detection limit problems, negative values, zeros, and values associated with special symbols. There are some negative values present randomly in the dataset, which may be due to data management or analytical errors. Thus, extensive data cleaning was required. Normally, during data preprocessing, values below the detection limit are replaced with half of the value or 2/3 of the existing value. However, this dataset is a special case, as the lower detection limit (LDL) and the upper detection limit (UDL) of this dataset are not constant; traditional data replacement methods would generate gaps in data distribution plots (ex: normal Q-Q plots). Thus, for this dataset, it was decided to leave those values of less than the LDL unchanged, and the values greater than the UDL were replaced with 1.5 times the existing value. However, zeros and negative values were replaced with ½ of the average LDL concerning the map sheets. If there was no LDL value within the considered map sheet to replace zeros and negative values, they were replaced by half of the lowest positive value in that map sheet.

3.1.3 Replacement based on the nearest neighbour method

Another tested approach was the replacement of zero values with the average of the ten nearest neighbouring values in the targeting till data set. This was done to maximize the usability of the chemical elements for future analyses and the method will specifically help to recover felsic elements such as Al and K as they contain a higher number of zero values. The ´caret´ package in R was used to replace zeros. The ten nearest neighbours that contained concentration values concerning each null sample point were considered. This model used 10 k-nearest neighbours (KNN) regression to make predictions in which the estimated new values were computed as the average of identified 10 k-nearest neighbours. Finally, sample points with null values were replaced by the average of the 10 k-nearest neighbours.

3.1.4 Exploring data and data analysis

Exploring data

Boxplots, original concentration maps, and Q-Q plots were used to explore the data. First, boxplots were used to compare the element distribution of different elements in the study area. Then, original concentration values related to the deepest points of the targeting till dataset were mapped to get an idea about the data distribution characteristics of different chemical elements between different map sheets without applying any data transformation methods. For example, the distribution of the Cr concentration for targeting till data and regional till data in central Lapland was considered. Selecting Cr and using the Cr concentration map of regional till data, were done to give the reader a clear picture of the map sheet levelling issues which prevail in the targeting till dataset. Q-Q plots were used to identify the data distribution between map sheets. The data distributions concerning different elements per individual map sheets were analysed.

Data analysis

a) Selection of elements for log transformation of data

Chemical elements with acceptable data quality should be used when using CoDA-based techniques on the geochemical data sets because log ratios especially with small values, could create artificial outliers. Thus, element selection is crucial before introducing data for such techniques, and the element selection procedure that has been followed is discussed in detail in the results. Original and centred log-ratio (clr) transformed concentration maps with lithological units, heat maps and principal component analysis (PCA) plots were used to analyse the targeting till dataset.

b) Original and clr-transformed elemental concentration maps

The original concentration data were mapped and then transformed using the clr-transformation. The clr-transformed data were then mapped separately using ArcGIS 10.8.1. The maps were subsequently compared with each other (original concentration and clr-transformed values), and the maps were compared to the known lithology and structures of the area. At this point, the chemical compositions of the lithological units and the changes in geochemical elements along different structures were considered.

c) Heat maps

In heat maps, shades of colours are used to represent the two-dimensional (2D) tables of numbers. First, suitable elements concerning a map sheet were selected and the variation matrix was calculated. The variation matrix measures pairwise log ratio variances between components of compositions. In this case, each entry of the variation matrix contains the log ratio variance of the respective element-element log ratios. Finally, the graphical representation of the variation matrix, the heatmap, concerning the selected map sheet was generated. This provides an appropriate tool to visualise which element log ratios are similar throughout the whole data set (small log ratio variance) and which elements exhibit slightly to strongly varying element-element log ratios in the targeting till dataset.

d) Principal component analysis (PCA)

PCA maps were generated using R to assess the correlation between the measured elements, and they were generated with respect to the map sheets. The PCA results were compared to the underlying geology and lithology to examine the association between the geochemical data and the underlying geology.

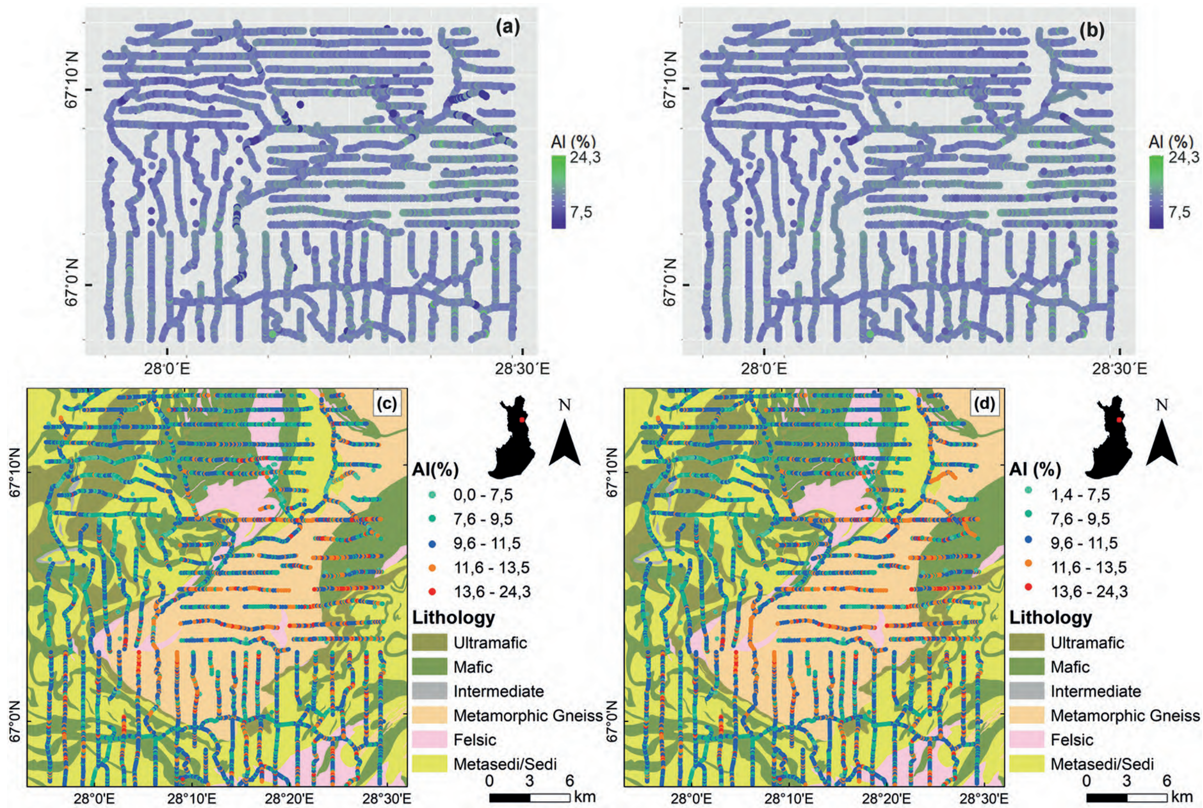

Figure 4. Al concentration maps a) original data with zero values b) data after replacing zero values c) original data with zero values overlaid on the lithological map d) zero-replaced data overlaid on the lithological map. The bedrock map was constructed using digital map information from the GTK Maankamara: http://gtkdata.gtk.fi/Maankamara/index.html © Geological Survey of Finland (2024, GTK Basic license version 1.1), Bedrock of Finland 1:200 000 (2023–2024), CC-BY-4.0. Contains also data from the National Land Survey of Finland, Topographic Database 04/2025, open data CC 4.0.

4. Results

4.1 Replacement based on the nearest neighbour method

It is advantageous to use as many elements as possible when detecting certain unknown mineral deposits. Hence, it would be beneficial for future analyses to recover the important elements of felsic minerals, such as Al and K, to a certain extent instead of removing them completely, even though we are not focusing on those elements in this study. Given the type of material, we assume that the values of the main elements such as, Al or K would not be as low as the LDL values but that these low values are rather an analytical bias resulting from the then-used measurement technique. Therefore, we chose a replacement technique which levels the "missing (zero)" values with values according to neighbouring values. Fig. 4 illustrates the Al concentration of map sheet M11 before (Fig. 4a, c) and after (Fig. 4b, d) replacement with the average of 10 nearest neighbours. When comparing Fig. 4a (original data with zeros) and Fig. 4b (map of zeros replaced with the nearest neighbours), it was noticed that the new Al map in which the zeros were replaced (Fig. 4b, d), does not illustrate the sudden changes of high to low concentration of values as in original data. When comparing the concentrations with the underlying geology before (Fig. 4c) and after replacement (Fig. 4d), the extreme high and low concentration changes within the same geological unit have been smoothed.

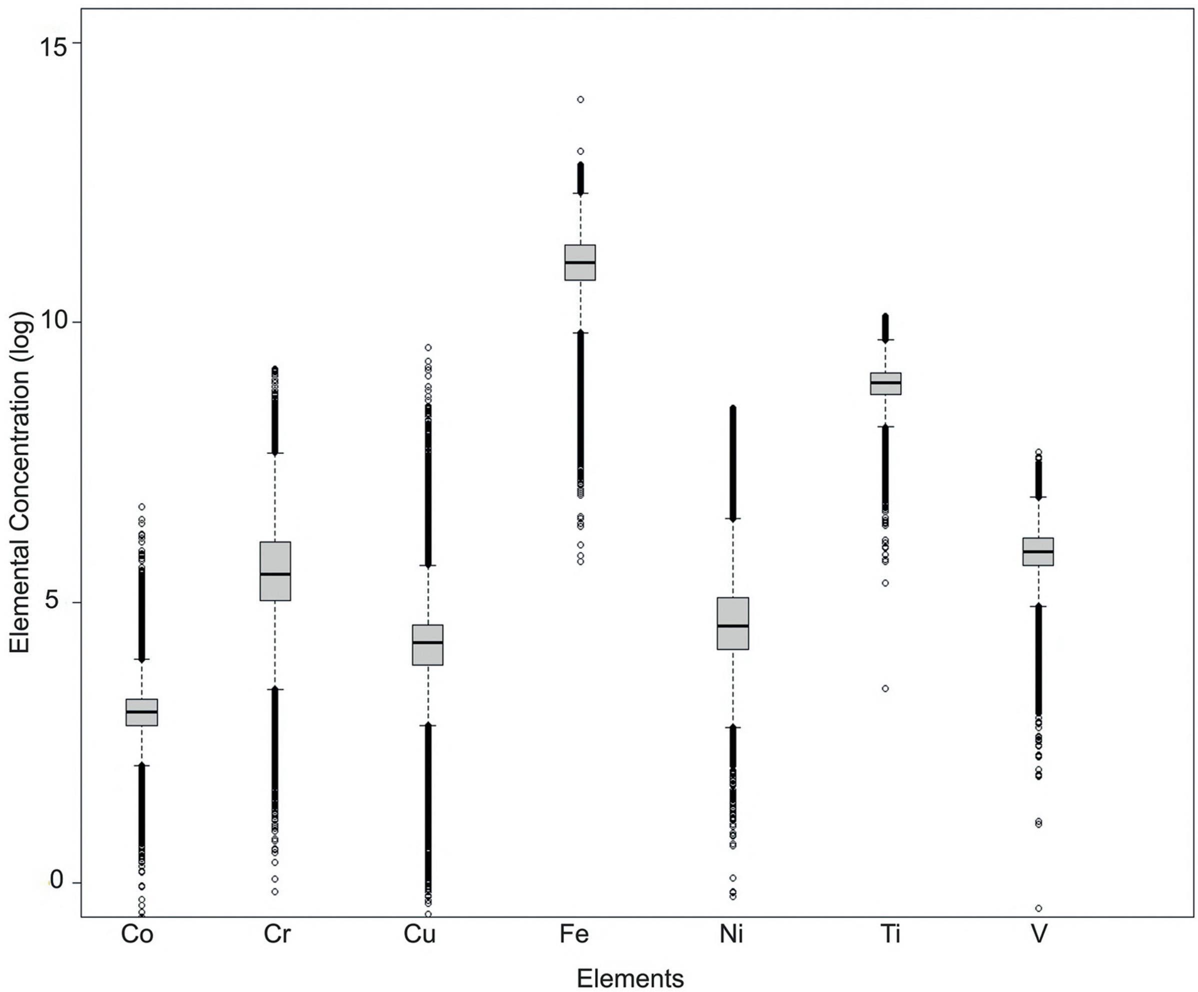

Figure 5. Boxplot of the log-transformed concentrations (ppm) of selected elements from the targeting till data (Till size fraction < 0.06 mm). Original concentration maps

4.2 Exploring data and data analysis

4.2.1 Exploring data

Boxplots, maps representing the original concentrations of the measured elements, and Q-Q plots have been used to get a preliminary idea of the targeting till dataset as well as the distribution of different elements within the study area. This study area has been divided into 29 mapsheets of 1:100 000 scale, as one map sheet is considered to be a single analytical batch.

a) Boxplots

Boxplots of the log-transformed original concentrations of the elements Co, Cr, Cu, Fe, Ni, Ti, and V for the targeting till data sets (<0.063 mm size fraction) are shown in Fig. 5, to compare the distributions before the dataset is clr-transformed. According to Fig. 5, iron shows the highest median and concentration range among all considered elements. The median values of Cu and Ni are closer to each other, while Co shows the lowest median. The area contains a higher concentration of iron, followed by Ti, and it is also rich in elements such as Cu, Ni, V and Co.

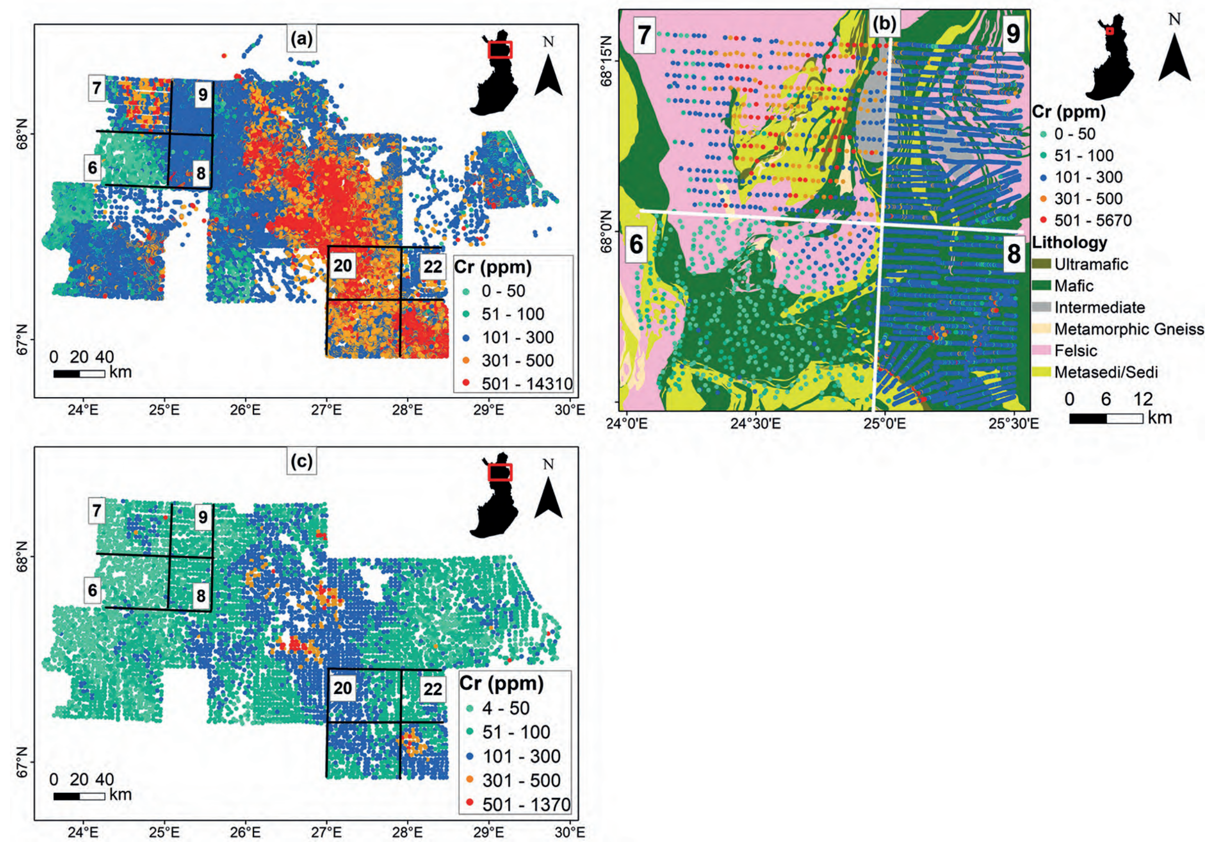

Figure 6. Original concentration of Cr in the central Lapland area for a) targeting till b) map sheets 6-9 overlaid on the lithological map c) regional till. The bedrock map was constructed using digital map information from the GTK Maankamara: http://gtkdata.gtk.fi/Maankamara/index.html © Geological Survey of Finland (2024, GTK Basic license version 1.1), Bedrock of Finland 1:200000 (2023–2024), CC-BY-4.0. Contains also data from the National Land Survey of Finland, Topographic Database 04/2025, open data CC 4.0.

Maps were generated using concentration values which were obtained by the quantometer and ICP-AES methods. In Fig. 6, the Cr concentrations for targeting till (Fig. 6a, b) and regional till (Fig. 6c) are visualised. The Cr concentration map is given to demonstrate the levelling issue between different map sheets. According to the original concentrations of Cr (Fig. 6a) in the targeting till data set in central Lapland, boundaries between some map sheets are detectable. Those map sheet boundaries may be prominent due to map sheet levelling problems (measurement offsets), geological offsets or a combination of those two. The offsets between map sheets 20 and 22 (Fig. 6a), the left side of map sheets 6–7 and the bottom part of 6–8 (Fig. 6b) are due to measurement offsets (a levelling problem). However, the map sheet offsets between map sheets 7–9 may be due to geological offsets, whereas the offsets between map sheets 6–7 may be due to a combination of both. Furthermore, the map sheet levelling issue can be detected in most other elements in targeting till dataset.

C) Quantile-Quantile (Q-Q) plots

Quantile–quantile plots are used to compare two distributions or to compare a sample to a theoretical distribution by matching positioned values such as quantiles. In addition, these plots are useful in identifying differences in location and scale, outliers, and other differences between different distributions (Marden 2004). The distribution of a certain element concerning different map sheets

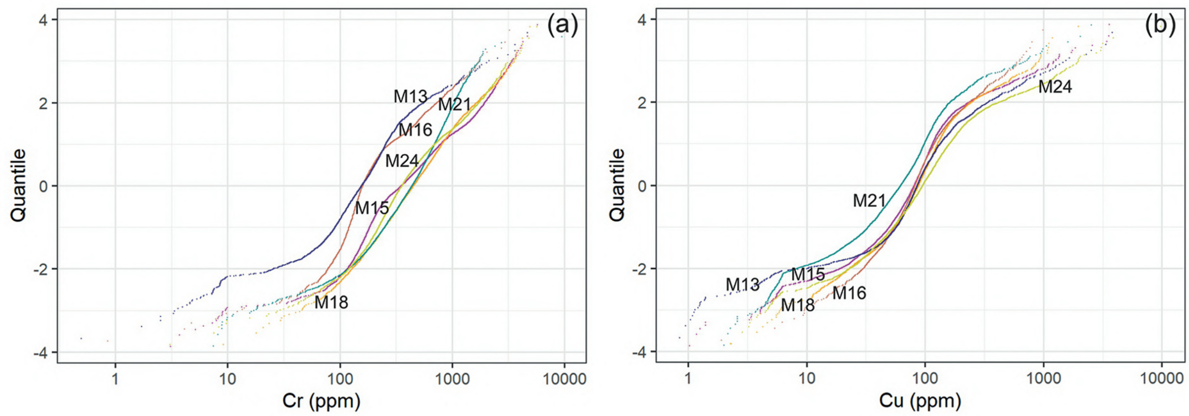

Figure 7. Q-Q plots of the targeting till data for a) Cr and b) Cu across 6 map sheets in the central Lapland area.

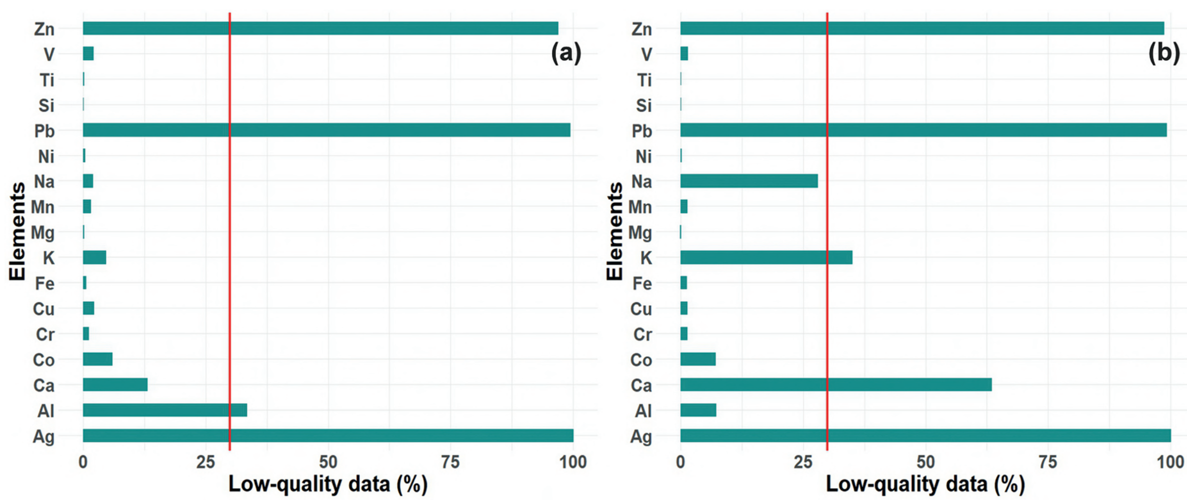

Figure 8. Bar plots showing the percentage of problematic quality data for all elements in map sheets a) M11 and b) M15.

was compared in this study. Altogether, there are 29 map sheets in the targeting till data, and Fig. 7 illustrates the Q-Q plots for the elements Cr and Cu. When considering the Q-Q plots for targeting till, 6 map sheets which lie near each other were selected to interpret the results, as the lithology of the nearby areas is similar to each other than that of remote areas. Therefore, the effect of lithology on the Q-Q plots is likely minor. According to Fig. 7, it is apparent that some map sheets exhibit an entirely different distribution of a given element. For example, the median of Cr for map sheets M13 and M16 is around 200 ppm, and the median of Cr in other map sheets is around 500 ppm (Fig. 7a).

4.2.2 Data Analysis

a) Element selection

As a result of the existing data quality issues in the targeting till geochemical dataset, the percentages of problematic data per map sheet and per element were calculated. This was done based on the 1:100 000 scale map sheets to use as many elements as possible for further analysis. Elements which had more than 30 % quality issues were eliminated from the geochemical data analysis (Gazley et al. 2015).

Fig. 8 shows the percentages of data with problematic quality concerning different elements for the map sheets M11 and M15. Elements with

Table 2: Elements with acceptable data quality across map sheets (√ indicate usable elements)

Map Sheet	Ag	Al	Ca	Co	Cr	Cu	Fe	K	Mg	Mn	Na	Ni	Pb	Si	Ti	V	Zn
1						√	√	√	√	√	√	√		√	√	√	
2						√	√	√	√	√	√	√		√	√	√	
3			√		√	√	√	√	√	√	√	√		√	√	√	
4			√	√		√	√	√	√	√	√	√		√	√	√	
5		√		√	√	√	√	√	√	√	√	√		√	√	√	
6			√	√	√	√	√	√	√	√	√	√		√	√	√	
7			√	√	√	√	√	√	√	√	√	√		√	√	√	√
8		√	√	√	√	√	√		√	√	√	√		√	√	√	
9		√		√	√	√	√	√	√	√	√	√		√	√	√	
10		√		√	√	√	√	√	√	√	√	√		√	√	√	
11			√	√	√	√	√	√	√	√	√	√		√	√	√	
12						√	√	√	√	√	√	√		√	√	√	
13				√	√	√	√	√	√	√	√	√		√	√	√	
14		√		√	√	√	√	√	√	√	√	√		√	√	√	
15		√		√	√	√	√		√	√	√	√		√	√	√	
16		√		√	√	√	√		√	√	√	√		√	√	√	
17		√		√	√	√	√	√	√	√	√	√		√	√	√	
18		√	√	√	√	√	√		√	√		√		√	√	√	
19				√	√	√	√	√	√	√	√	√		√	√	√	
20				√	√	√	√	√	√	√	√	√		√	√	√	
21			√	√	√	√	√	√	√	√	√	√		√	√	√	
22			√		√	√	√	√	√	√	√	√		√	√	√	
23					√	√	√	√	√	√	√	√		√	√	√	
24		√		√	√	√	√	√	√	√	√	√		√	√	√	
25					√	√	√	√	√	√	√	√		√	√	√	
26					√	√	√	√	√	√	×	√		√	√	√	
27					√	√	√	√	√	√	√	√		√	√	√	
28					√	√	√	√	√	√	√	√		√	√	√	
29					√	√	√	√	√	√	√	√		√	√	√	

more than 30 % data quality issues (vertical line) were not considered in the subsequent data analyses. The elements Ag, Pb and Zn were excluded from further analysis, as they contain more than 95 % low-quality data in almost all 29 map sheets. The elements Cu, Fe, Mg, Mn, Ni, Ti, and V can be used for analyses across all map sheets, as they always contain less than 30 % low-quality data. Table 2 provides a summary of these quality control aspects.

b) Original and clr-transformed elemental concentration maps

The targeting till data were clr-transformed and mapped using the ArcGIS 10.8.1 software. Map sheet M15 was selected as it contains two Ni-Cu-PGE deposits. Fig. 9 shows the element distributions of the original data for Ni, Cu and Co (Fig. 9a, c, e) along with the clr-transformed data for the same elements (Fig. 9b, d, f). Elements that contain more than 30 % quality issues (see Table 2) were not considered for the clr-transformation. As an example, when considering the map with clr-transformed data for Ni, a significant improvement is observed compared to the original concentration map. Also, the normal Q-Q plots related to clr-transformed data are more compatible with a normal distribution, in contrast to the Q-Q plot related to the original concentration. When comparing the Ni concentration with the underlying geology, the bedrock related to the area defined by red colour concentration data in Fig. 9a and 9b contains ultramafic rock – komatiite. Thus, the area expected to have high Ni concentration is more well-defined in the clr-transformed data map (Fig. 9b) than in the original concentration data map (Fig. 9a). Moreover, in the clr-transformed

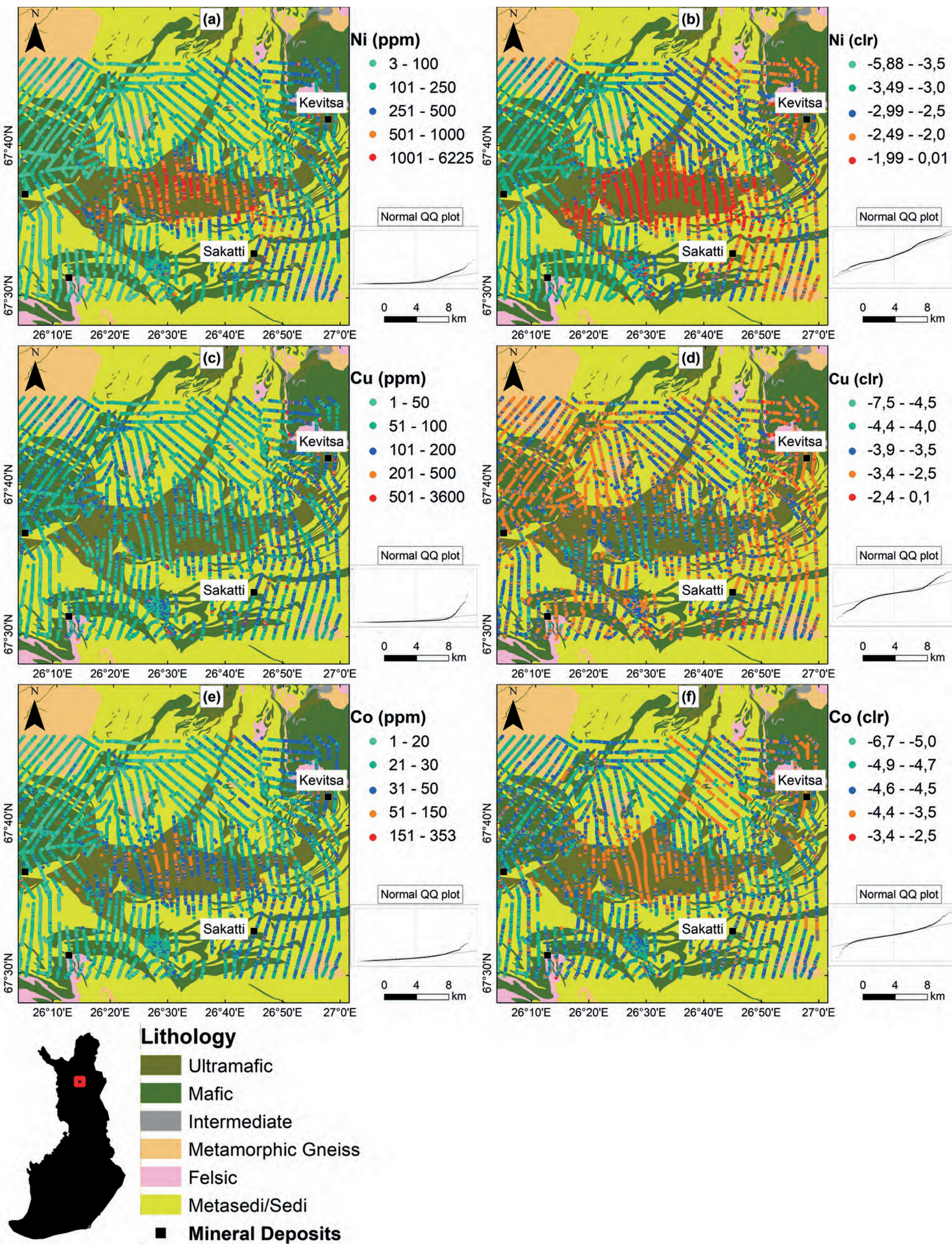

Figure 9. Concentration data for Ni (a) Original, b) clr-transformed), Cu (c) original, d) clr-transformed), Co (e) original, f) clr-transformed). The bedrock map is constructed using digital map information from the GTK Maankamara: http://gtkdata.gtk.fi/Maankamara/index.html © Geological Survey of Finland (2024, GTK Basic license version 1.1), Bedrock of Finland 1:200 000 (2023–2024), CC-BY-4.0.

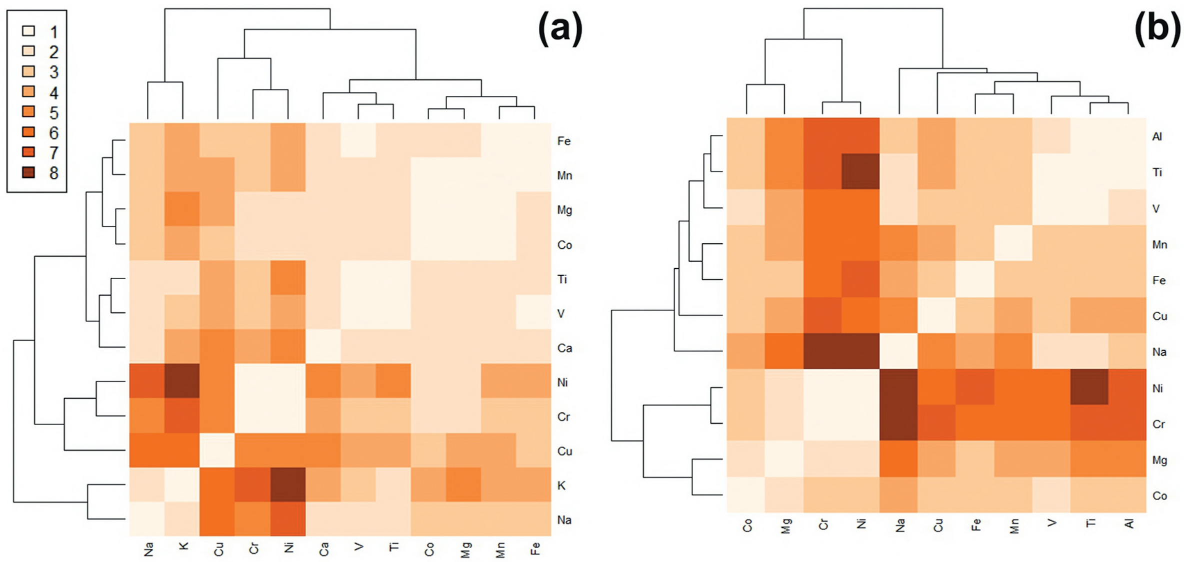

Figure 10. Heat maps for map sheets a) M11 and b) M15.

concentration map, an elevated concentration of Ni can be observed near the Sakatti and Kevitsa Ni-Cu-PGE deposits. Additionally, clr-transformed data for Cu and Co also show improvement compared to the original data (Fig. 9 c-d and e-f).

c) Heat maps

Heat maps for selected map sheets (Fig. 10) can be used to visualise the log ratio variance between different elements. According to the log ratio variances, element groups (Ti, V) and (Ni, Cr) in both heat maps, exhibit stable log ratios throughout the whole dataset (small variance). Hence, it can be assumed that they occur in the same minerals. Mg and Co are also associated in both map sheets with the (Ni, Cr) element group, but less strictly. However, Mn, appears only in map sheet 11 together with the ((Ni, Cr), (Co, Mg)) group, whereas such a trend cannot be observed in map sheet 15. Aluminium is only presented in the map sheet M15 due to data quality and has a very stable log ratio with Ti throughout the whole dataset of map sheet M15. The same type of correlation between these elements could be observed in the respective PCA plots for these map sheets in the next sub-section.

d) Robust principal component analysis (PCA)

According to Filzmoser et al. (2009), the original concentrations of the selected till dataset were transformed to isometric log-ratio (ilr) coordinates to perform robust PCA, and the results were back–transformed to clr-coefficients for interpretation. Fig. 11 shows the results of the 1st and 2nd principal components for the selected areas which were mentioned in the previous section. The principal components 1 and 2 explain 69.9 % and 65.7 % of the variances of the data for map sheets M11 and M15, respectively.

According to Fig. 11a, Ni and Cr are strongly associated with each other. These elements dominate the ultramafic rocks, which are located towards the lower left side of the biplots followed by Mg. However, meta-sedimentary rocks are also located towards the left side. Cu shows a huge variance in log ratios with Na and Ca.

In Fig. 11b, Ni and Cr are strongly associated with each other, similar to map sheet M11, and the elements Ti, Al and V also occur together. However, Cu and Co do not show a significant contribution in the biplot of principal components 1 and 2. As expected, ultramafic and mafic rocks plot towards the upper left side of the biplot,

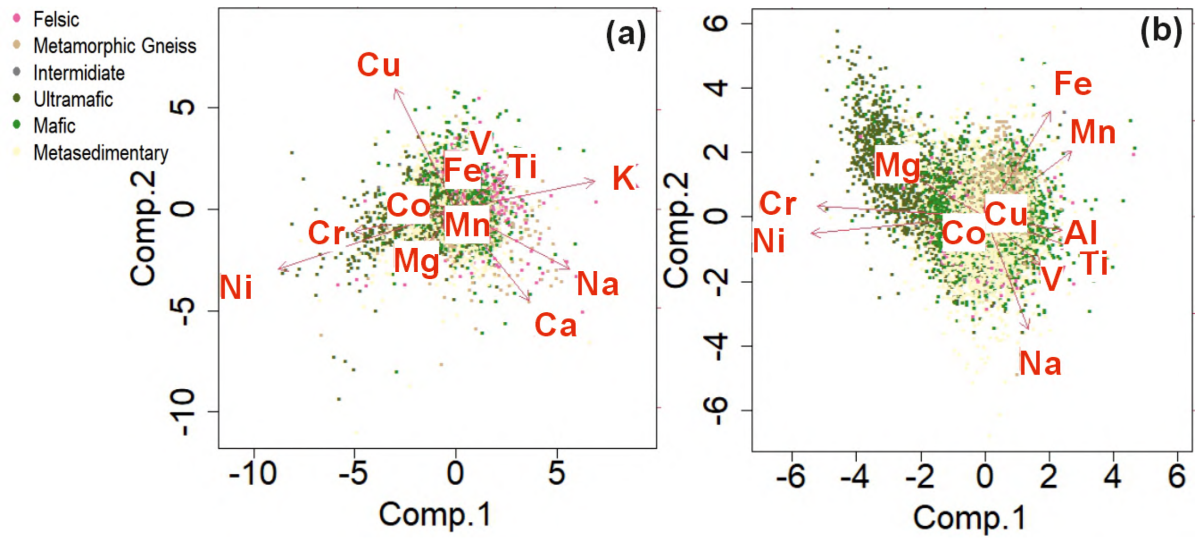

Figure 11. PCA loadings for the first two principal components in the map sheet a) M11 and b) M15 in targeting till.

where a trade-off between Mg versus Na, Al, Ti, and V. The contrasting behaviour is observed for metasedimentary rocks clustered towards the right side of the biplot. However, another cluster of mafic rocks can also be observed towards the right side of the biplot.

5. Discussion

5.1 Exploring data

Two till geochemical datasets, regional and targeting, have been used to study the usability of old data in mineral exploration for some critical metals related to ultramafic rocks in the central Lapland area. In both datasets, the sieved fine fraction of till, i.e. <0.063 mm size fraction was used as an analytical medium. The fine fraction includes both clay-size fraction whose origin is mostly weathered bedrock in central Lapland, and silty-size fraction, consisting of mineral material eroded and pulverized by the glaciers (Lintinen 1995). This fine fraction has been commonly used in geochemical analyses because it represents the most homogenous part of the till matrix and includes both local and distal portions of glacially eroded, transported, and deposited materials (Sarala 2015).

Regional till geochemical dataset has been used since the end of the 1980s as a basic dataset in mineral exploration due to its nationwide coverage. The analytical procedure is based on aqua regia dissolution and ICP-AES elemental analysis. However, the sampling density (1 sample / 4 km^2) is in many cases too sparse, and the sampling depth (about 2 m from the surface) is too shallow for the exploration of covered mineralization, particularly tiny or vein-type occurrences. Instead, in the central Lapland, the targeting till geochemical data collected during 1973–1984 with a density of c. 6–12 samples / km^2 form a much denser and, so far, underutilized media for mineral exploration. However, the detailed examination in this paper shows that there are some limitations of the data which restrict the usability of the whole dataset in mineral exploration.

In the targeting till dataset, the repeatability of the concentrations given by the EKV method at the 95 % confidence level ranges from 20 % to 50 %, depending on the element. The results can, therefore, be considered semi-quantitative (Gustavsson et al. 1979). Furthermore, the element concentration maps of the targeting till dataset show a mismatch between the map sheets, which can be the result of differences caused by a long history of sample collection and sampling strategy (by following the map sheet division), analytical problems, the map sheet levelling problem, geological offsets, or the combined effect of all the above. However, the lack of visibility of map sheet boundaries in the regional till dataset for most map sheets indicate that the mismatch between map sheets in the targeting till dataset should be specifically related to the map sheet levelling problem other than the relation to the underlying lithology. For the targeting till dataset, the data were collected in 1:200 000 map sheets and 1:100 000 map sheets can be considered as one analytical batch. This may be the reason for distinct map sheet boundaries in 1:100 000 map sheets.

In this paper, we removed elements that contained more than 30 % values with problematic data quality. Gazley et al. (2015), Martin-Fernandez et al. (2012) recommend that imputed values less than 30 % are preferred when applying robust techniques. Moreover, Farnham et al. (2002) conducted a study on the chemistry of trace elements in groundwater and demonstrated that their methods yielded poor results when the observations less than LDL exceed 30 %.

Boxplots can be used to gain more insight into the data behaviour, such as by visualising the spread of the data, skewness, existence of outliers, etc. (Reimann 2005). Moreover, Grünfeld (2007) carried out a study to get a better insight into glacial till data in southeastern Sweden by using interactive visualization, spatial and numerical techniques including box plots of different elements. In this paper, we also used box plots to get an idea of the distribution of different elements. According to the results, Fe showed the highest concentration which may be a result of the central Lapland greenstone belt, ultramafic rocks and mafic rocks that locate in the study area. Ultramafic rocks usually contain more than 90 % of mafic minerals rich in Mg and Fe (Downes 2020). According to Bergman et al. (2007), the Fennoscandian shield contains layered intrusions, most of the intrusions with Cr, Ni, Ti, V and/or PGE mineralisation. Therefore, this may be the reason for elevated Ti and V concentrations in the study area in northern Fennoscandia.

Reimann et al. (2011) used the Kola project data set contained 50 chemical elements that were analysed for sample materials collected from 600 sample locations to demonstrate the exploratory data analysis. There were more observations (more than 600) and one of the methods they used to understand the distribution of large data sets was Q-Q plots. Additionally, log-transformed values were used in the plots to reduce the impact of outliers. In this paper, the Q-Q plots reveal significant differences in the targeting till dataset across some map sheets, even among those that are nearby each other. As an example, when considering the Q-Q plot for Cr (Fig. 7a), it contains comparably lower Cr content for map sheets M13 and M16 while displaying higher for other map sheets. This detail goes well with the concentration map of Cr. However, the majority of the underlying rocks related to map sheets 13 and 16 are mafic and ultramafic and thus, those map sheets are also expected to have high Cr medians as the other nearby map sheets. Furthermore, it is not clear why Cr medians in those map sheets are low nevertheless, it can be due to a levelling problem, differences in sampling procedures or till transportation.

5.2 Data analysis

When comparing the concentration maps of the original and the clr-transformed data, it is apparent that the concentration map with clr-transformed data aligns well with the underlying geology (Fig. 9). The clr-transformation helps to overcome issues, such as spurious correlations and bias resulting from the constrained nature of compositional data and this may help to reveal the hidden structures contained in the data.

Sadeghi et al. (2013) used PCA to study the distribution of rare earth elements in Sweden. They noted that the core of PCA is the covariance matrix, in which, the interrelationships between multiple variables are measured. Typically, the first principal component captures the majority of the variance present in the data set. Moreover, Dempster et al. (2013) also used PCA to analyse the geochemical aspects of soil evolved on till in Northern Ireland. In this paper, biplots between the first and second principal components are generated (Fig. 11). The PCA biplot that belongs to the map sheet M11 is shown in (Fig. 11a) and the area is mainly dominated by ultramafic, mafic, gneiss, sedimentary/metasedimentary and granitic rocks (Bedrock of Finland - DigiKP). In the biplot, ultramafic rocks have been clustered along the direction of (Ni, Cr, Mg) and mafic/ultramafic elements (Mg, Ni, Cr, Co) present together. However, according to the heat map (Fig. 10a), the strength of the correlation between these 4 elements varies, whereas elements groups (Ni, Cr) and (Mg, Co) showed a stronger correlation within their respective groups than between the groups. Moreover, elements that are usually associated with felsic rocks, such as Ca, Na, and K, show a close association with each other.

The area of M15, mainly contains ultramafic, mafic, dolerite, gneiss, felsic, granite, metasedimentary, carbonate and intermediate rocks (Bedrock of Finland - DigiKP). In the PCA biplot (Fig. 11b) which represents map sheet M15, the ultramafic rocks have been grouped towards the direction of Mg, followed by mafic rocks. The elements group (Ni, Cr, and Mg) which indicate ultramafic and mafic elements appear together. However, another cluster of mafic rocks could be observed towards the right side of the biplot, where the elements Fe, Ti, V, Al, and Mn dominate. These two clusters may suggest two distinct groups of mafic rocks in the dataset. Besides, one group might be enriched in elements Mg, Cr, and Ni and the second group might be enriched in elements Fe, Ti, and V. Furthermore, sedimentary/metasedimentary rocks are also concentrated towards the bottom right side of the biplot towards the loadings of Na, Ti, V and Al. Typically, sedimentary and meta sedimentary rocks, such as sericite quartzite, sericite paraschist, arkose quartzite, phyllite etc. contain Na and Al. However, according to (Fig. 11b), Ti and V are also associated with Na and Al. Upon examining the heat map (Fig. 10b), it is evident that the strength of correlation between these elements differs. In the heat map, Ti and V have a strong correlation and Al has a strong correlation only with Ti. However, both Ti and V have a weaker correlation with Na than the correlation with each other. The targeted elements Ti and V are mostly associated with the second group of mafic rocks, which dominates the elements Fe, Ti and V rather than ultramafic units. Thus, it is advisable to focus more on elements Ni, Cr, and Mg when exploring orthomagmatic deposits in this area.

6. Conclusions

In this study, the targeting till geochemical dataset that was collected in the 1970s and at the beginning of the 1980s has been analysed. Sampling was carried out by following a 1:100 000 scale map sheet division in central and northern Finland. In the analysis, we used statistical and CoDA-based approaches to analyse data. The analysis of original concentration maps together with Q-Q plots revealed that there are several problems with the data quality that should be considered before using data in mineral exploration. There are mismatches between neighbouring map sheets which are seen as a result of variable quality of geochemical data, differences in the concentrations of elements and levelling problems between map sheets. Those will cause crucial problems when using data as a part of an exploration study or when trying to generate prospectivity maps.

To overcome these problems, a preliminary interpretation of the data was made using summary statistics and boxplot visualizations. Quality issues were then identified using concentration maps and Q-Q plots. Subsequently, the elements with acceptable data quality for further analysis were identified for each map sheet. As explained in the methodology section, CoDA was carried out on individual map sheets separately. During this study, the map sheets in the targeting till data set were analysed separately as they contained sufficient data. However, when merging map sheets for a combined analysis, it is recommended to level the elements' concentrations between map sheets to ensure the reliability and consistency of the analysis.

Better insight into the data could be obtained by analysing clr-transformed data rather than by analysing raw data. The PCA biplot represents the essential variability of the dataset by reducing the dimension of the data. PCA is a powerful tool for identifying the important patterns in the area and correlating them to the underlying bedrock geology. Moreover, heat maps can also be used along with PCA biplots to further interpret the dataset.

Acknowledgements

This work was co-funded by the European Union (SEMACRET, Grant Agreement no. 101057741) and UKRI (UK Research and Innovation). The Geological Survey of Finland (GTK) is thanked for providing the targeting and regional till geochemical datasets in Finland. The authors thank and acknowledge Patricia Puchhammer, Lucija Dujmovic and Markus Raatikainen for their assistance with debugging R codes, and Malcom Aranha for his valuable input on generating geological maps. Additionally, we thank Brandon Datar for assisting with editing.

Authorship contribution statement

Charmee Kalubowila – Conceptualization, Writing, Analytical work, Data curation, Interpretation, Visualization; Pertti Sarala – Conceptualization, Supervision, Writing, Review, Editing; Solveig Pospiech – Supervision, Visualization, Review, Editing; Peter Filzmoser – Supervision, Review, Editing.

References

Bedrock of Finland – DigiKP. Digital map database [Electronic source]. Espoo: Geological Survey of Finland [referred 1.10.2024]

Bergman, S., Weihed, P., Martinsson, O. & Eilu, P., 2007. Geological and tectonic evolution of the northern part of the Fennoscandian Shield. In: Ojala, J., Weihed, P., Eilu, P. & Iljina, M. (eds.), Metallogeny and tectonic evolution of the Northern Fennoscandian Shield: field trip guidebook. Guide book, Geological Survey of Finland, Espoo, 6–15.

Brownscombe, W., Ihlenfeld, C., Coppard, J., Hartshorne, C., Klatt, S., Siikaluoma, J.K. & Herrington, R. J., 2015. The Sakatti Cu-Ni-PGE sulfide deposit in northern Finland. In: Mayer, W.D., Lahtinen, R. & O'Brien, H. (eds.), Mineral deposits of Finland. Chapter 3.7. Elsevier, Amsterdam, 211–252. https://doi.org/10.1016/B978-0-12-410438-9.00009-1

Danielsson, A., Lundgren, F. & Sundkvist, G., 1959. The tape machine – I: A new tool for spectrochemical analysis. Spectrochimica Acta 15, 122–125. https://doi.org/10.1016/S0371-1951(59)80296-4

Danielsson, A. & Sundkvist, G., 1959. The tape machine – II: Applications using different kinds of isoformations. Spectrochimica Acta 15, 126–133. https://doi.org/10.1016/S0371-1951(59)80297-6

Danielsson, A. & Sundkvist, G., 1959. The tape machine – III: Notes on useful corrections in spectrochemical analysis with the tape technique. Spectrochimica Acta 15, 134–137. https://doi.org/10.1016/S0371-1951(59)80298-8

Dempster, M., Dunlop, P., Scheib, A. & Cooper, M., 2013. Principal component analysis of the geochemistry of soil developed on till in Northern Ireland. Journal of Maps 9(3), 373–389. https://doi.org/10.1080/17445647.2013.789414

Downes, H., 2020. Ultramafic Rocks. Reference Module in Earth Systems and Environmental Sciences. https://doi.org/10.1016/B978-0-12-409548-9.12478-9

Farnham, I. M., Singh, A. K., Stetzenbach, K. J. & Johannesson, K. H., 2002. Treatment of nondetects in multivariate analysis of groundwater geochemistry data. Chemometrics and Intelligent Laboratory Systems 60(1-2), 265–281. https://doi.org/10.1016/S0169-7439(01)00201-5

Filzmoser, P., Hron, K. & Reimann, C., 2009. Principal component analysis for compositional data with outliers. Environmetrics: The Official Journal of the International Environmetrics Society 20(6), 621–632. https://doi.org/10.1002/env.966

Garrett, R. G., & Sinding-Larsen, R., 1984. Optimal composite sample size selection, applications in geochemistry and remote sensing. Journal of Geochemical Exploration 21(1-3), 421–435. https://doi.org/10.1016/0375-6742(84)90065-7

GTK, n.d., Till geochemistry sample series, accessed 19 January 2024, https://tupa.gtk.fi/paikkatieto/meta/targeting_till_geochemistry.html.

Gazley, M. F., Collins, K. S., Roberston, J., Hines, B. R., Fisher, L. A. & McFarlane, A., 2015. Application of principal component analysis and cluster analysis to mineral exploration and mine geology. In AusIMM New Zealand branch annual conference, Dunedin New Zealand. 2015, 131–139.

Gonçalves, M. A., da Silva, D. R., Duuring, P., Gonzalez-Alvarez, I. & Ibrahimi, T., 2024. Mineral exploration and regional surface geochemical datasets: An anomaly detection and k-means clustering exercise applied on laterite in Western Australia. Journal of Geochemical Exploration 258, 107400. https://doi.org/10.1016/j.gexplo.2024.107400

Grunsky, E. C., 2010. The interpretation of geochemical survey data. Geochemistry: Exploration, Environment, Analysis 10, 27–74. https://doi.org/10.1144/1467-7873/09-210

Grünfeld, K., 2007. The separation of multi-element spatial patterns in till geochemistry of southeastern Sweden combining GIS, principal component analysis and high-dimensional visualization. Geochemistry: Exploration, Environment, Analysis 7(4), 303–318. https://doi.org/10.1144/1467-7873/06-117

Gustavsson, N., Noras, P. & Tanskanen, H., 1979. Summary: Report on Geochemical mapping methods. Geological Survey of Finland, Report of Investigation 39, 20 p.

Hanski, E. & Huhma, H., 2005. Central Lapland greenstone belt. In: Developments in Precambrian geology (Vol. 14, pp. 139–193). Elsevier. https://doi.org/10.1016/S0166-2635(05)80005-2

Hirvas, H., 1991. Pleistocene stratigraphy of Finnish Lapland. Geological Survey of Finland, Bulletin 354, 123 p.

Johansson, P., 2005. Jääjärvet. In: Johansson, P. & Kujansuu, R. (eds.), Pohjois-Suomen maaperä: maaperäkarttojen 1:400 000 selitys. Summary: Quaternary deposits of Northern Finland – Explanation to the maps of Quaternary deposits 1:400 000. Geological Survey of Finland, Special Publications 46, 127–148.

Johansson, P., 2007. Late Weichselian deglaciation in Finnish Lapland. In: Johansson, P. & Sarala, P. (eds.), Applied Quaternary Research in the Central Part of Glaciated Terrain. Geological Survey of Finland, Special Paper 46, 47–54.

Johansson, P., Lunkka, J. P. & Sarala, P., 2011. The glaciation of Finland. In: Ehlers, J., Gibbard, P.L. & Hughes, P.D. (eds.), Developments in Quaternary sciences 15, 105–116. Elsevier, Amsterdam. https://doi.org/10.1016/B978-0-444-53447-7.00009-X

Karinen, T. & Törmänen, T., 2016. Akanvaara chromitite layers (LLC, LC, ULC and UC). 340–342.

Lintinen, P., 1995. Origin and physical characteristics of till fines in Finland. Geological Survey of Finland, Bulletin 379, 5–73.

Marden, J. I., 2004. Positions and QQ plots. Statistical Science 9(4), 606–614.

Martin-Fernandez, J.A., Hron, K., Templ, M., Filzmoser, P. & Palarea-Albaladejo, J., 2012. Model-based replacement of rounded zeros in compositional data: Classical and robust approaches. Computational statistics and Data Analysis 56, 2688–2704. https://doi.org/10.1016/j.csda.2012.02.012

McClenaghan, M.B., Paulen, R.C., Smith, I.R., Rice, J.M., Plouffe, A., McMartin, I., Campbell, J.E., Lehtonen, M., Parsasadr, M. & Beckett-Brown, C.E., 2023. Review of till geochemistry and indicator mineral methods for mineral exploration in glaciated terrain. Geochemistry: Exploration, Environment, Analysis 23(4). https://doi.org/10.1144/geochem2023-013

Paulen, R.C., Bobrowsky, P.T., Lett, R.E., Bichler, A.J. & Wingerter, C., 1998. Till geochemistry in the Kootenay, Slide Mountain and Quesnel terranes. Geological Fieldwork, 307–319.

Plouffe, A., Kjarsgaard, I.M., Ferbey, T., Wilton, D.H.C., Petts, D.C., Percival, J.B., Kobylinski, C.H. & McNeil, R., 2022. Detecting buried porphyry Cu mineralization in a glaciated landscape: a case study from the Gibraltar Cu-Mo deposit, British Columbia, Canada. Economic Geology 117(4), 777–799. https://doi.org/10.5382/econgeo.4891

Putkinen, N., Sarala, P., Eyles, N., Pihlaja, J., Daxberger, H. & Murray, A., 2020. Reworked Middle Pleistocene deposits preserved in the core region of the Fennoscandian Ice Sheet. Quaternary Science Advances 2, 1–9, 100005. https://doi.org/10.1016/j.qsa.2020.100005

R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/

Reimann, C., 2005. Sub-continental-scale geochemical mapping: sampling, quality control and data analysis issues. Geochemistry: Exploration, Environment, Analysis 5(4), 311–323. https://doi.org/10.1144/1467-7873/03-065

Reimann, C., Filzmoser, P., Garrett, R. & Dutter, R., 2011. Statistical Data Analysis explained applied environmental statistics with R. John Wiley & Sons.

Reimann, C., Filzmoser, P., Fabian, K., Hron, K., Birke, M., Demetriades, A. & GEMAS Project Team, 2012. The concept of compositional data analysis in practice – total major element concentrations in agricultural and grazing land soils of Europe. Science of the total environment 426, 196–210. https://doi.org/10.1016/j.scitotenv.2012.02.032

Sadeghi, M., Morris, G. A., Carranza, E. J. M., Ladenberger, A. & Andersson, M., 2013. Rare earth element distribution and mineralization in Sweden: an application of principal component analysis to FOREGS soil geochemistry. Journal of geochemical exploration 133, 160–175. https://doi.org/10.1016/j.gexplo.2012.10.015

Santaguida, F., Luolavirta, K., Lappalainen, M., Ylinen, J., Voipio, T. & Jones, S., 2015. The Kevitsa Ni-cu-PGE deposit in the Central Lapland greenstone belt in Finland. In: Mayer, W.D., Lahtinen, R. and O'Brien, H. (eds.), Mineral deposits of Finland. Chapter 3.6. Elsevier, Amsterdam. 195–210. https://doi.org/10.1016/B978-0-12-410438-9.00008-X

Salminen, R. & Tarvainen, T., 1995. Geochemical mapping and databases in Finland. Journal of Geochemical Exploration 55 (1-3), 321–327. https://doi.org/10.1016/0375-6742(94)00062-X

Sarala, P., 2015. Surficial geochemical exploration methods. In: Mayer, W.D., Lahtinen, R. and O'Brien, H. (eds.), Mineral deposits of Finland. Chapter 10.1. Elsevier, Amsterdam, 711–731. https://doi.org/10.1016/B978-0-12-410438-9.00027-3

Sarala, P. & Ojala, V. J., 2008. Implications of complex glacial deposits for till geochemical exploration: Examples from the central Fennoscandian ice sheet. In: Stefánsson, Ó. (eds.), Geochemistry Research Advances, Chapter 1. Nova Publishers, New York, pp. 1–29.

Sarala, P., Väliranta, M., Eskola, T. & Vaikutiené, G., 2016. First physical evidence for forested environment in the Arctic during MIS 3. Scientific Reports, 6. https://doi.org/10.1038/srep29054

Taivalkoski, A., Ranta, J.-P., Sarala, P., Nikkola, P., Liu, X., Kalubowila, C., Immonen, N., Gilbricht, S. & Molnár, F., 2024. Mineral exploration using trace element composition of pyrite grains from till: A case study from the Petäjäselkä Au occurrence, northern Finland. Journal of Geochemical Exploration 257, 107359. https://doi.org/10.1016/j.gexplo.2023.107359

Törmänen, T., Karinen, T., Boyd, R., Bjerkgård, T., Nordahl, B. & Schiellerup, H., 2016. The Kevitsa and Sakatti Ni-Cu-PGE deposits. Mineral resources in the Arctic, 351–353.