National Research FLAGSHIPS
Minerals Down Under

CSIRO

An Investigation Using SiroSOM for the Analysis of QUEST Stream-Sediment and Lake-Sediment Geochemical Data

Stephen J. Fraser and Jane H. Hodgkinson
Geoscience BC - Report 2009-14
CSIRO Report Number: EM MDU P2009 / 983
September 2009

www.csiro.au

Geoscience BC

Enquiries should be addressed to:
Stephen Fraser
Stream Leader, Enhancing Knowledge from Drilling
Minerals Down Under National Research Flagship
CSIRO
QCAT, 1 Technology Court
Pullenvale, Queensland 4069
Australia
Tel: +61 (0)7 3327 4544
Fax: +61 (0)7 3327 4566
Email: stephen.fraser@csiro.au

Copyright and Disclaimer

Important Disclaimer

EXECUTIVE SUMMARY

CSIRO Exploration & Mining (CEM) has used a Self-Organizing Map (SOM) approach to analyze a database of stream and lake sediment geochemical analytical results compiled and collected as part of the Geoscience BC's QUEST Project.

The study's objective was to use CSIRO's implementation of the Self-Organizing Map (SiroSOM) to identify patterns and establish relationships amongst the various stream sediment and lake sediment geochermical data that may be indicative of geochemical dispersion related to mineralization. Anomalous samples and some spatially-coherent responses for some elements result from the study.

Sample point data, extracted from the levelled and imputed elemental grids produced by Barnett and Williams (2009) were used as inputs to this study. The levelled and imputed grids were intersected by the sample point locations and the relevant grid (elemental) values subsequently assigned to the sample point location and number. The input data set consisted of some 15,020 samples each with 42 elemental values.

There is a multitude of ways that SOM can be used to analyse and assist with knowledge extraction and creation, and only those relevant to this study are detailed. It was never our intent to document and describe all the significant occurrences and distributions for each of the 42 available elements; however, we have demonstrated our approach to target prospective Au, Cu, Mo and Ni samples, so that these procedures can be repeated by the reader as required.

Various "anomalous" or prospective samples have been identified; and, the spatial context and coherence of such samples is critical for their assessment. Sample locations are displayed colour-coded by either the K-means cluster to which their SOM-derived BMU "node" has been assigned, or by the magnitude of their quantization-error, which is used to identify the locations of outliers. The locations and distributions of samples highlighted by this study need to be assessed in the context of known mineral occurrences and with regards to the structural evolution and metallogenesis of the Quesnellia Terrane.

Because this study has been undertaken to promote resource assessment and exploration, we have endeavoured to present our results in a manner so that all potential users of these data can understand and benefit from them. A series of digital maps and ASCII digital data outputs are provided, so readers can use the results in their own analysis software, and display them in their spatial information system of choice.

Contents

FIGURES

1. INTRODUCTION

CSIRO Exploration & Mining (CEM) has undertaken a Self-Organizing Map (SOM) analysis of stream and lake sediment geochemical data from the QUEST Project area to identify samples that exhibit "anomalism" that may be associated with mineralization. The study area (Figure 1) is over the Quesnellia Terrane of central British Columbia, which has potential to host copper and gold porphyry-style mineralization, but prospective areas are variably covered by Quaternary deposits of glacial sands and gravels. The Quest Project includes areas of Mountain Pine Beetle (MPB) infestation, and it is anticipated that successful mineral exploration and mining ventures resulting from Geoscience BC initiated projects will provide economic and development opportunities for MPB affected areas.

1.1 Input Data Background

As part of a previous geochemical QUEST project, publicly available regional geochemical data collected via various federal and provincial geoscience initiatives over central British Columbia (for example see, Lett, 2005; Lett and Bluemel, 2006) were compiled into a digital data base: the MPB Data Repository (Jackaman, 2007). This data repository contains multi-element analytical results from over 30,000 samples collected during various regional bark, lake, till and stream surveys. More recently, a further 5000 samples were added to the available data resulting from the reanalysis of archived sample pulps, from these earlier surveys, and, by some 2200 new samples collected to fill holes and gaps in the data coverage (Jackaman, 2008 (a), (b), (c); Jackaman and Balfour, 2008). While the now available, expanded geochemical data set over the QUEST area is substantial, its integration, analysis and interpretation is a significant challenge. Different analytical methods with various detection limits and levels of precision have been used over the 30 years these survey data were collected.

Despite the above challenges, Barnett and Williams (2009) have produced uniform grids over a common area for 42 of the elements available in the combined MPB data set. These authors have developed a pragmatic approach to overcome the inherent issues in such a combined data set and the procedures they have employed are thoroughly documented in their report. Their procedure essentially involves three steps: (1) selection of a preferred analytical method and detection limit from the available analytical methods for each element; (2) re-levelling and blending of adjacent geochemical surveys for the selected method/element combination; and (3) imputation of any missing data values (on a site basis) to ensure that the spatial coverages of all selected elements were comparable. Table 1 details the 42 elements, the preferred analytical method, the detection limits, and the measurement units for the geochemical values used to create the grids.

Barnett and Williams (opt cit) demonstrated the validity of their approach by creating a model that relates their resulting elemental grids to mapped geology in areas of known geological outcrop and by predicting the "unknown geology" in areas of till and recent sediment cover. The spatial coherence of transition zones from known to unknown geology provides evidence of the success of their approach.

For this study we have used as input, sample point data extracted from the levelled and imputed elemental grids produced by Barnett and Wilson (2009). The levelled and imputed grids were intersected by the sample point locations and the relevant grid (elemental) values subsequently assigned to the sample point location and number.

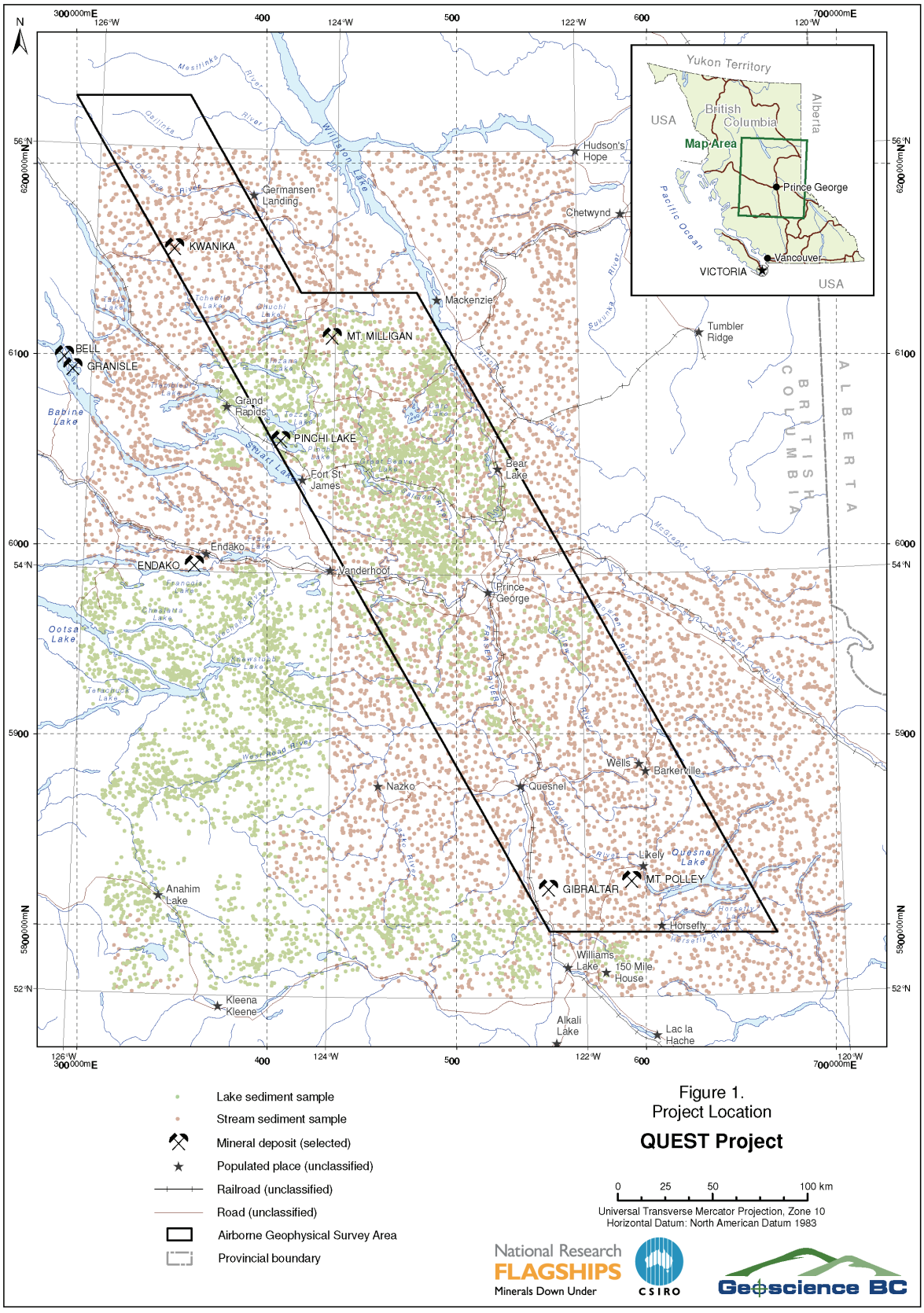


Figure 1. Location of the QUEST SOM Geochemical Project Area.

Table 1: **The elements, analytical method, detection limits and measurement units used to create the levelled and imputed grids, from which the point data used in this study were derived (after Barnett and Wilson, 2009).**

ELEMENT	Method	DL	UNIT	ELEMENT	Method	DL	UNIT
ALUMINUM	ICPMS	0.01	PCT	MANGANESE	ICPMS	1	PPM
ANTIMONY	INAA	0.1	PPM	MERCURY	ICPMS	5	PPB
ARSENIC	INAA	0.5	PPM	MOLY	ICPMS	0.01	PPM
BARIUM	INAA	100, 50	PPM	NICKEL	ICPMS	0.1	PPM
BISMUTH	ICPMS	0.02	PPM	PHOSPHORUS	ICPMS	0.001	PCT
BROMINE	INAA	0.5	PPM	POTASSIUM	ICPMS	0.01	PCT
CADMIUM	ICPMS	0.01	PPM	RUBIDIUM	INAA	5	PPM
CALCIUM	ICPMS	0.01	PPM	SAMARIUM	INAA	0.5, 0.1	PPM
CERIUM	INAA	10, 5	PPM	SCANDIUM	INAA	0.5, 0.2	PPM
CESIUM	INAA	0.5	PPM	SELENIUM	ICPMS	0.1	PPM
CHROMIUM	INAA	5, 20	PPM	SILVER	ICPMS	2	PPB
COBALT	INAA	5	PPM	SODIUM	INAA	0.1, 0.02	PCT
COPPER	ICPMS	0.01	PPM	STRONTIUM	ICPMS	0.5	PPM
GALLIUM	ICPMS	0.2, 0.1	PPM	SULPHUR	ICPMS	0.02	PCT
GOLD	INAA	2	PPB	TERBIUM	INAA	0.5	PPM
HAFNIUM	INAA	1	PPM	THALLIUM	ICPMS	0.02	PPM
IRON	INAA	0.2	PCT	THORIUM	INAA	0.5, 0.2	PPM
LANTHANUM	INAA	5, 2	PPM	TITANIUM	ICPMS	0.001	PCT
LEAD	ICPMS	0.01	PPM	URANIUM	INAA	0.2	PPM
LUTETIUM	INAA	0.2	PPM	VANADIUM	ICPMS	2	PPM
MAGNESIUM	ICPMS	0.01	PCT	ZINC	ICPMS	0.1	PPM

1.2 SiroSOM Background

CSIRO Exploration & Mining (CEM) has an ongoing interest in tools and techniques that assist in the integrated analysis, interpretation and visualization of geoscience data sets, especially those that are stored within a geographic information system (GIS). For some time, CEM has been promoting the use of self organizing map (SOM) technology for the analysis of exploration data sets, as a data mining tool to assist in exploratory data analysis and knowledge discovery (Fraser and Dickson, 2007).

The SOM procedures and algorithms are described in detail by Kohonen (2001); however, a brief description follows. If one can represent all data points as vectors in a data-space defined by the variables (element concentrations), the SOM procedure provides a non-parametric mapping (regression) that transforms these high dimensional, nonlinearly-related vector items to a typically two-dimensional (2D) representation. This procedure attempts to maintain the relationships between the best-matching unit vectors (BMU) in nD multidimensional space and their 2D representation as nodes[1] on the "map". The resulting, typically rectilinear, self-organized map is a 2D representation of the n-dimensional data space, which attempts to preserve and show the relationships and complexity of the input data. The SOM procedure essentially provides an un-supervised "self-organization", which allows a visual representation of the relationships present in the multivariate data space. Also this "self-organization" occurs at a variety of levels. The samples that are represented by each node on the map in effect form a

[1] The term "best-matching unit" (BMU) refers to a vector in nD space that was trained to represent a group of samples in a data set. The term BMU node refers to the cell on the self-organized map that has been assigned the characteristics of a particular BMU vector. These terms are largely interchangeable.

grouping of statistically similar samples; however, other more traditional clustering techniques (we typically use K-means) can then be used to cluster the various characteristics of nodes themselves into self-similar patterns.

SOM procedures are used in a range of applications, but they have had significant impact in the fields of data exploration (Kaski, 1997) and data mining (Vesanto, 2000).

Geologists in the exploration and mining industries are increasingly challenged by the joint integration and interpretation of ever-increasing amounts of new and historic, spatially-located data (geochemistry, geophysics, geology, mineralogy, elevation data, etc.). Because these data can be gathered faster than they can be interpreted, the availability of GIS and similar software packages, has, in some respects, compounded, rather than reduced this problem. Computational data mining methods, such as SOM, can assist in the analysis and understanding of such data sets. SOM is an ideal technique to assist in the analysis of spatially-located stream and lake sediment geochemical data.

There is an increasing acceptance of SOM in the petroleum industry to assist with calibration and interpretation of well-logs and seismic data (e.g., Strecker and Uden, 2002; Briqueu et al., 2002). There have also been demonstrations of the technique for the analysis of broader earth sciences data (Sliwa et al., 2003; Penn, 2002, 2004; Zhou et al., 2005; and, Fraser and Dickson, 2005, 2007; Mikula et al., 2008; Fraser and Hodgkinson, 2008).

The SOM software developed by CSIRO Exploration & Mining (SiroSOM) is a specific implementation of the SOM algorithms, which is aimed at providing geoscientists with access to new methods for determining the intricate relationships within and between multiple, spatially-located and complex data sets. The analysis and visualization provided by SOM has the potential to be significant to both regional and mine-based mineral exploration activities, as we seek to understand those relationships associated with mineralization, identify prospective zones, and discover new ore bodies. Fraser and Dickson (2007) provide more detail on the SiroSOM processing.

2. PROJECT OBJECTIVE & DELIVERABLES

The objective of this study was to use SiroSOM to seek patterns and establish relationships amongst the various stream sediment and lake sediment[2] parameters, to highlight anomalous samples or spatially-significant responses that may indicate geochemical dispersion indicative of mineralization.

The typical output from a SiroSOM procedure is an ASCII file, which is essentially, the input data file with extra appended columns that contain a number of the resulting SOM-derived parameters specific for each sample. Various other maps and digital data files will be included as digital data to accompany this report either on CD ROM or an associated web-page.

[2] It was planned to include the parameters from the water surveys in this study; however the irregular and varied nature of these measurements would present significant challenges. Water parameters were not included in the 42 variables selected by Barnett and Williams (2009) so such parameters are not included in this combined study.

3. SOM ANALYSIS PROCEDURES AND OUTPUTS

Data analysis using SiroSOM is an interactive process, which typically involves a range of approaches and strategies that can be used depending on the objective of an analysis and the complexity of the data. There are numerous ways that a SOM analysis can be used to assist with knowledge extraction and creation.

The point data set extracted from the levelled and imputed grids consist of some 15,020 samples each with 42 elemental values. The elemental values were in log10 format. This was considered an advantage as the first step in the SiroSOM processing would have been to undertake this transformation.

After some initial experimentation a "self-organized map" size of 35 x 29 was chosen as being appropriate to display the overall variability of the input data. We have used a toroidal projection surface and the various "self-organized map" displays wrap from top-to-bottom and from side-to-side.

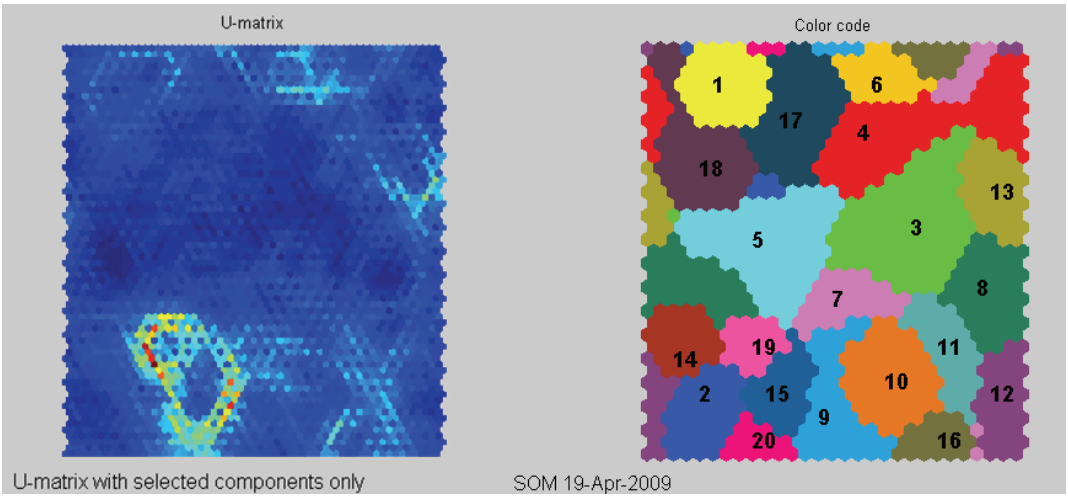


Figure 2. (a) U-matrix representation of the self-organized map that results from the analysis of the levelled and gridded QUEST geochemical data; and, (b) A K-means clustering of the map "nodes" shown on the U-matrix plot. The cluster numbers are shown.

The U-matrix representation of the "map[3]" is a display that is intended to summarize the structure, yet illustrate the complexity inherent in a data set. The nodes on the U-matrix representation are coloured using a colour-temperature scale to indicate how similar (blue) or dissimilar (red) they are to adjacent nodes. To assist in this process, every second row and column on the U-matrix consists of dummy nodes that are coloured to show the average "similarity" between adjacent real nodes. The "real nodes" are coloured according to their average "similarity" to their adjacent real nodes. At one level, the U-matrix can be considered broadly analogous to an elevation model, with blue "plains of similarity" separated by hotter temperature nodes forming walls or areas of increasing diversity. Figure 2(a) is the U-matrix representation of the self-organized map, resulting from our analysis of the QUEST area geochemical input data.

[3] The self-organized "map" is typically a two-dimensional rectilinear representation of a multivariate data set; it is not a spatial representation of samples or other distributions.

Appendix 1 details the co-ordinate system and numbering convention for identifying nodes or best matching units (BMUs) on the resulting self-organized map. Appendix 2 describes the content of the ASCII output files that can be used for further analysis or to display the SiroSOM processed data.

The ACSII output files ("QUEST_out_data.csv" and "QUEST_out_map.csv") are included in the digital data that accompany this report. Appendix 3 contains a catalogue of the digital data (ASCII files and map pdfs) that should accompany this report.

Each node on the self-organized map corresponds to a BMU data vector, which represents a group of similar input samples. We have found there are advantages for visualization and interpretation purposes in clustering these BMU vectors using a traditional clustering approach such as K-means (Scheffe, 1953; Xu and Wunsch, 2005). Figure 2(b) shows the results of a K-means clustering of the map nodes. A Davies-Bouldin (1979) analysis of the node data revealed that 20 was a suitable, natural segmentation of these data, so consequently in the K-means analysis, 20 classes were determined. A detailed comparison of Figures 2 (a) and (b) shows how the cluster boundaries conform to patterns evident on the U-matrix. In this study the K-means analysis essentially breaks up the data into segments or domains of self-similar geochemical responses. Appendix 4 contains "Box and Whisker" plots of the 42 elements for each of the 20 K-means classes.

The patterns evident on the U-matrix, and hence the distribution of the K-means clusters, are controlled by the underlying contributions of the components (variables) that constitute the samples used in the analysis. Figure 3 displays the 42 elemental component plots that sit behind the "self-organized map". These plots display the elemental BMU vector values for each node on the map, and are displayed using a colour-temperature scale that denotes low-range values as blue and high-range values as red.

Figure 4 shows a nearest-neighbour gridded, spatial map of the samples, colour-coded by the K-means cluster assignment of the SOM nodes to which the samples belong. This map represents a domaining of the samples that broadly reflects the local "upstream" geology and stream transport conditions. (Geoscience BC; QUEST Project Map 2009-14-01 is a more detailed version of this map).

To assist with interpretation and to create a product that can be used immediately to target areas of interest, "Cluster-normalized" maps for each of the 42 elements have been produced. For a given sample, the "Cluster-normalized" value for a given element has been calculated by subtracting the mean of the sample values for that cluster, from the sample value, then dividing that result by the standard deviation of that cluster. A set of these "Cluster-normalized" maps, and the file from which they were derived - "QUEST_out_data_Cluster _Normalized.csv" - have been included in the digital data that accompany this report.

The "Quantization Error" (QER) is another SOM-derived parameter that can assist in the interpretation of a dataset. In a SOM analysis, each sample vector is assigned to, and represented by a BMU vector, (which is displayed as a node on the self-organized map). The QER is a measure of how similar a sample is to the BMU vector that represents it. The QER is determined for each sample and the larger the QER value, the more of an outlier (anomalous) that sample is. It is important to understand that large QER values (anomalism) may be caused

by either an increased or reduced presence of particular elements (variables); hence care is needed in its interpretation. In general, the QER allows the identification of anomalous trends that cross geological boundaries. The process tends to compensate for changes in the thresholds of anomalism for individual elements in different geological background chemistries. Figure 5 shows the spatial distribution of each sample's QER value. (Geoscience BC; QUEST Project Map 2009-14-02 is a more detailed version of this map).

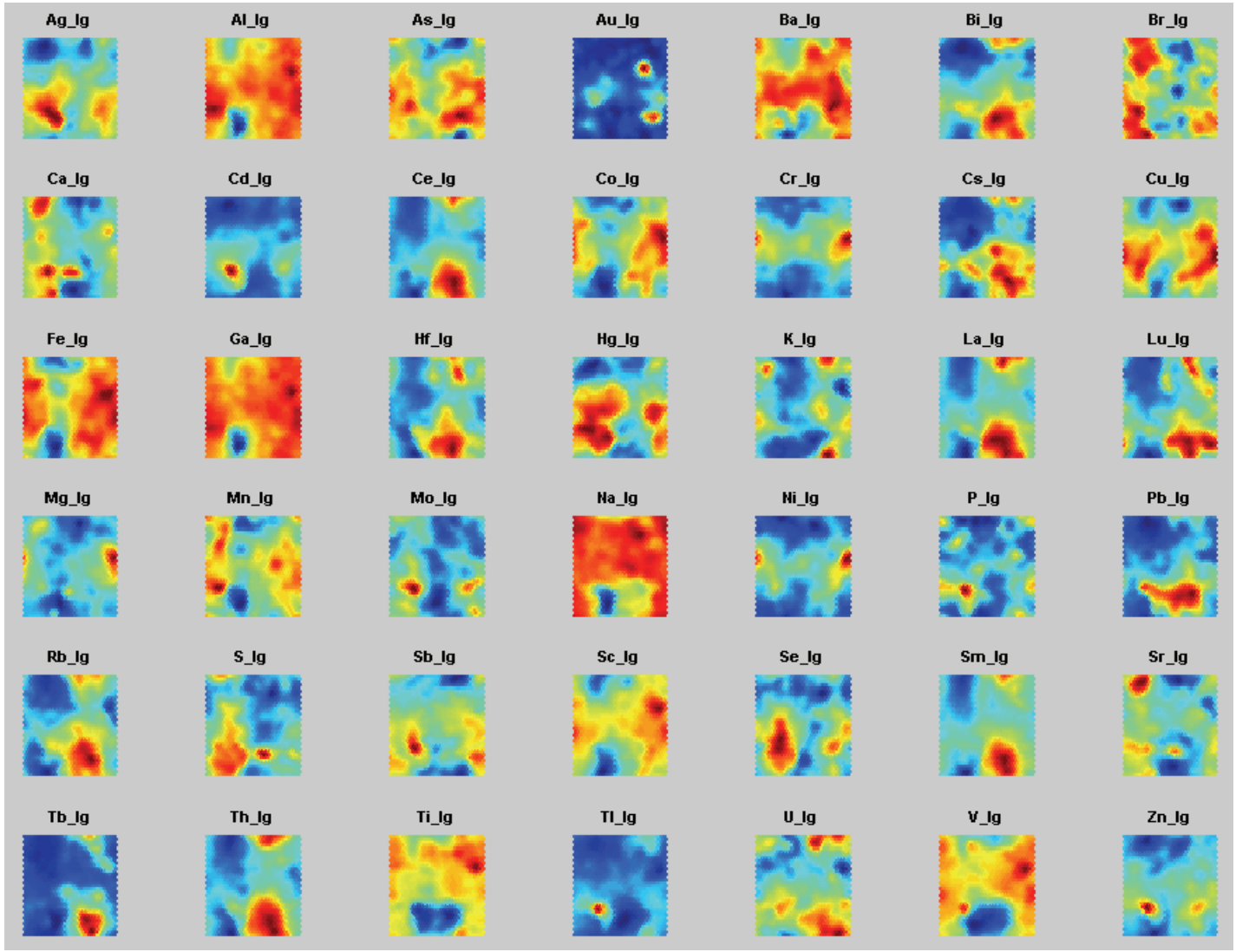


Figure 3. Component Plots of the elemental contributions for the nodes of the self-organized map shown on Figure 2.

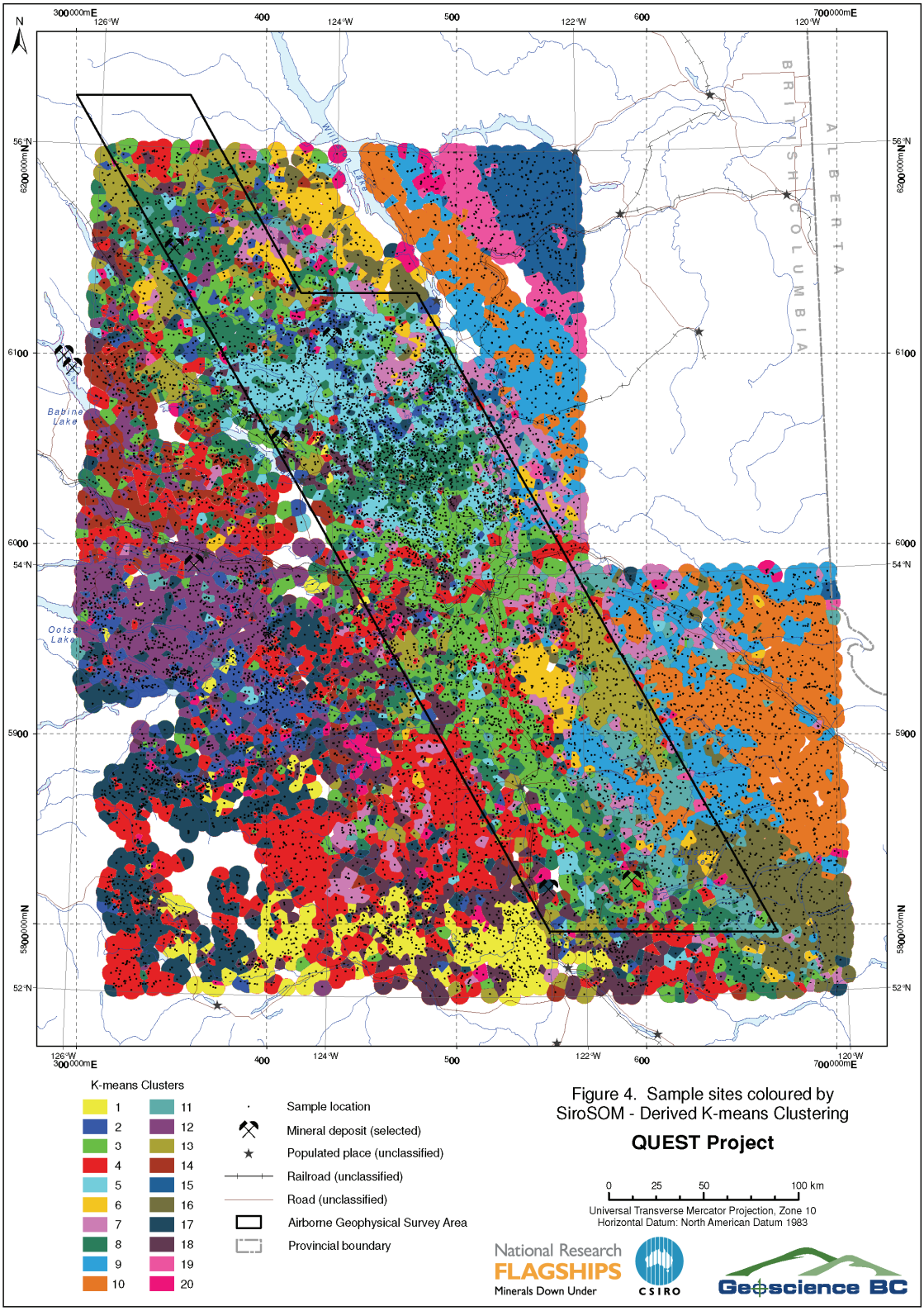


Figure 4. **Sample sites coloured by the SiroSOM-derived K-means cluster of the nodes to which they belong. Point data were gridded using a Nearest-Neighbour interpolation (courtesy of P. Kowalczyk).**

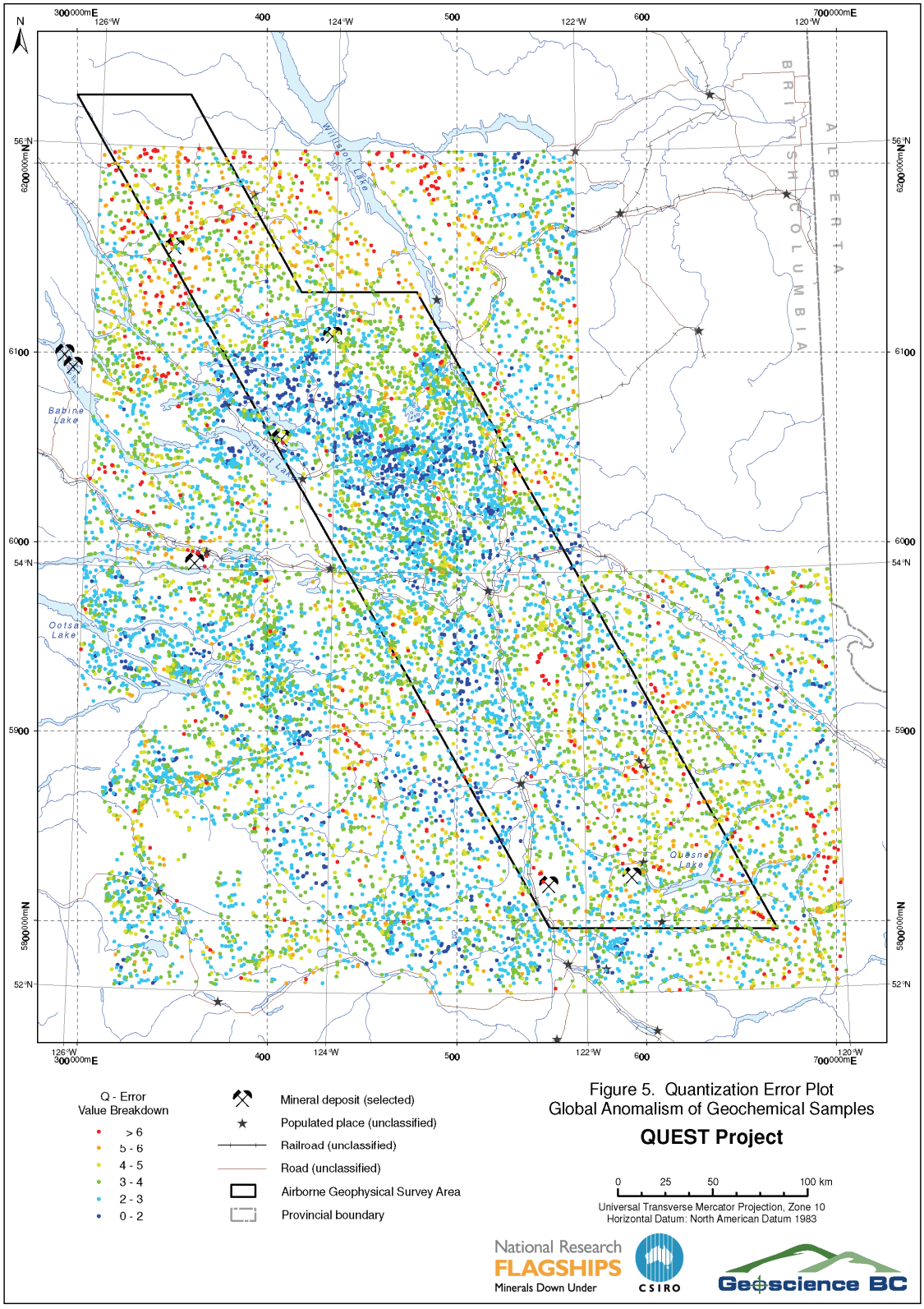


Figure 5 Quantization Error Map: Global Anomalism of Geochemical Samples.

4. NON-EXHAUSTIVE INTERPRETATION OF SOM RESULTS

4.1 General

Output data from a SOM analysis can be used in various ways to extract information and knowledge from a geochemical data set. Some of these approaches are detailed in the following sections of this chapter.

It is beyond this study's scope to document and describe exhaustively all the significant occurrences and distributions for each of the 42 elements resulting from this SiroSOM analysis. Our intent is to describe and demonstrate the processes used to target anomalous or prospective samples so that the procedures can be repeated by the reader as required. The following sections will provide examples that concentrate on Au, Cu, Mo and Ni.

As previously indicated, various digital ASCII files are available with this report; and these allow further post-SOM processing analysis and visualization to be undertaken. For specific procedures detailed in the following sections, ASCII output files have also been provided to assist with follow-up and location of highlighted samples etc.

4.2 "Cluster-Normalized" Element Anomaly Maps

During the SOM analysis, field samples were assigned to SOM-derived BMU nodes on the resulting self-organized map. These BMUs subsequently were "clustered" by a K-means analysis to yield 20 classes. Samples were then assigned membership to clusters according to the cluster that their BMU has been assigned. These SOM-derived, K-means clusters have been used to highlight "anomalism" within the geochemical data set.

(Geoscience BC has produced a series of 42 "Cluster-Normalized" element anomaly maps with samples normalized to the mean and standard deviation of the K-means cluster to which a sample's BMU belongs (Geoscience BC QUEST Project Maps: 2009:14:002 - 2009:14: 044). The file from which these maps were derived - "QUEST_ out_data _Cluster _Normalized.csv" - is included in the digital data that accompany this report.)

4.2.1 Gold Anomalism Map

Figure 6 shows the "Gold Anomalism" map resulting from the above procedure. This map shows sample sites colour-coded in terms of the magnitude of the difference a sample value has, compared to the average value of the cluster to which that sample belongs. Anomalous values occurring in coherent spatial regions or forming sediment trains indicate areas where some form of follow-up sampling and further reconnaissance is recommended. (Geoscience BC; QUEST Project Map 2009-14-06 is a more detailed version of this map).

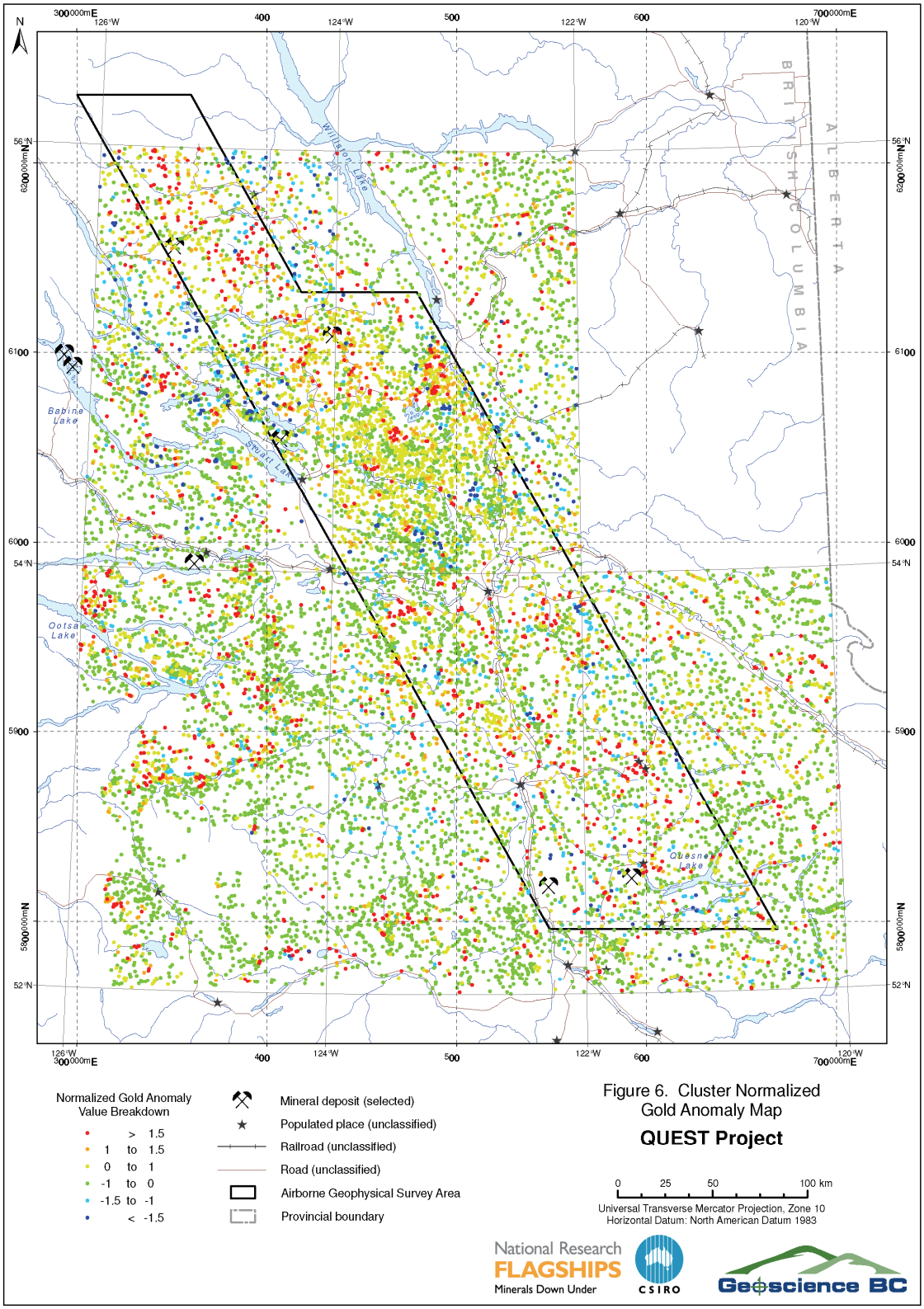


Figure 6. **Cluster-Normalized Gold "Anomalism" Map. Samples are normalized to the mean and standard deviation of the K-means cluster to which they belong.**

4.3 Element Associations with SOM-Derived K-means Clusters

Cross-plots of elemental values for each BMU node can be used to highlight samples, trends, processes and relationships amongst the variables in a data set analysed via SOM. If the nodes on the cross-plots are colour-coded based on their K-means cluster assignment, the cross-plots also highlight clusters, or particular BMU nodes that may be of significance.

4.3.1 Au verses Ag Cross-Plot

The Au verses Ag cross-plot is shown on Figure 7. This figure shows the BMU nodes from the self-organized map, coloured by the K-means cluster assignments as shown on Figure 2(b). The nodes are sized in proportion to the number of samples from the initial data set that report to the node.

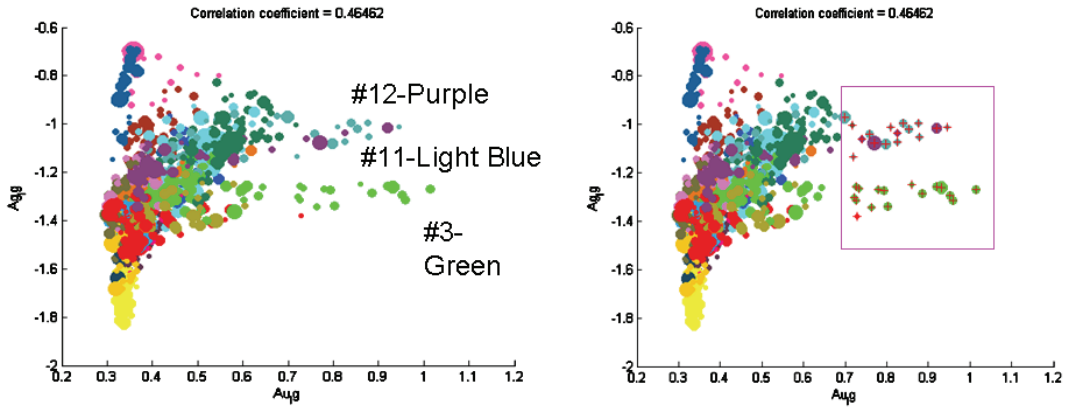


Figure 7. **Au vs Ag Cross-Plot. The graph on the left shows the cluster-labelling of the nodes associated with the highest Au values. The locations of samples belonging to those higher-value Au nodes within the box on the right-hand plot have been spatially-located in Figure 8.**

In Figure 7, the nodes with the highest Au values plot on two trends: one trend with lower Ag values contains nodes from Cluster #3 (green) and one node belonging to Cluster #4 (red); and another trend with higher associated Ag values containing nodes associated with Clusters #11 (light blue) and #12 (purple). The spatial contexts of samples belonging to those higher-value Au nodes within the box on the right hand plot are shown in their respective cluster-colours on Figure 8. Samples belonging to Cluster #3 (green) occur throughout the study area. The samples of Cluster #11 (light-blue) are concentrated in the south-east of the study area,, with other representative samples scattered throughout the area. The samples of Cluster #12 (purple) are concentrated on the central western area, but some samples also occur on the eastern side. The three samples of Cluster #4 (red) occur on the western side of the area. (An ASCII file containing information on the samples plotted on Figure 8 - "Au vs Ag Box_QUEST_out_data.csv" - has been included in the digital data that accompany this report).

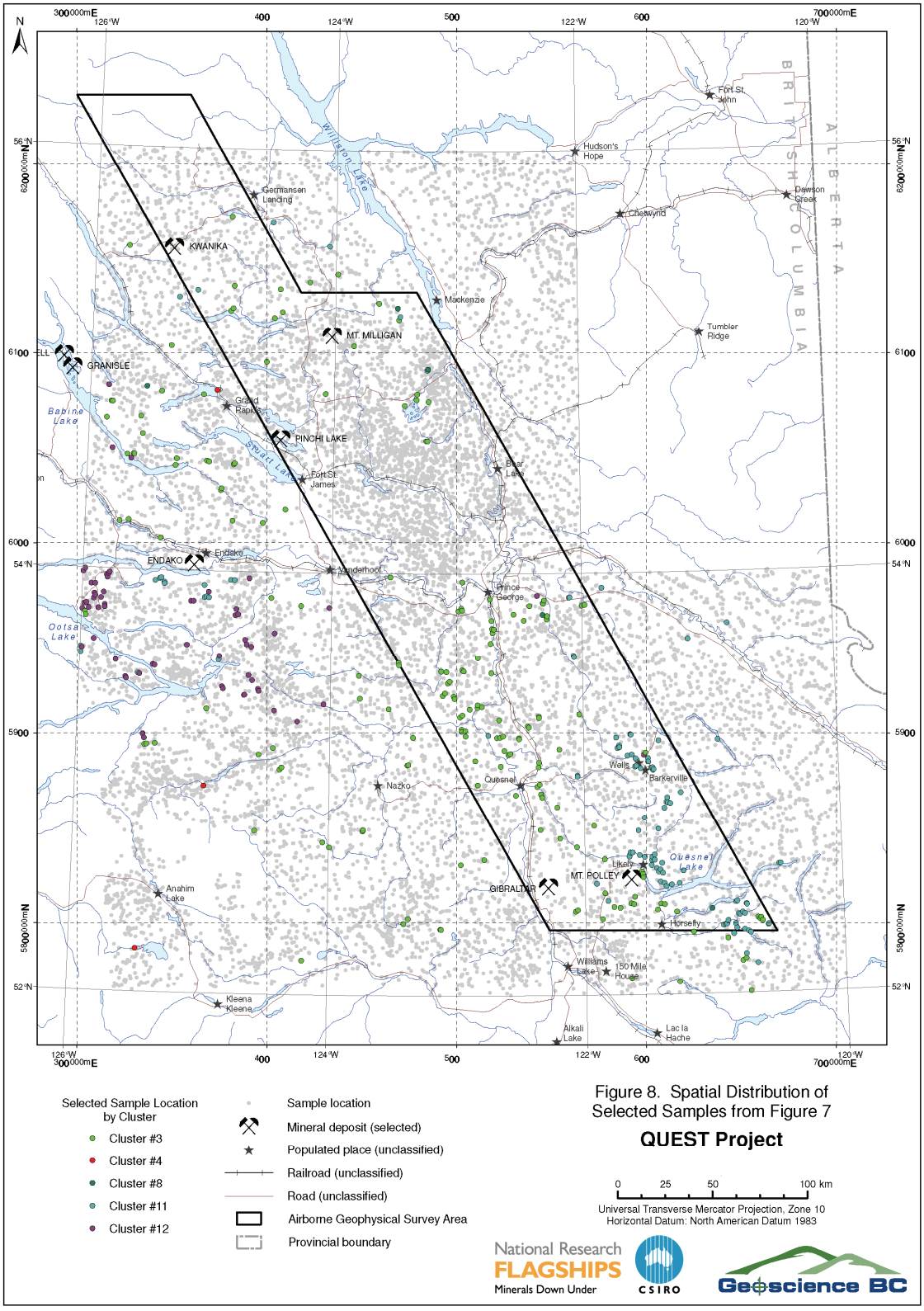


Figure 8. **Spatial distribution of samples belonging to the elevated Au nodes within the box shown on the right-hand plot of Au vs Ag Cross-Plot on Figure 7.**

4.3.2 Cu verses Ni Cross-Plot

The Cu verses Ni cross-plot is shown on Figure 9. The display on this figure is similar to that shown for the Au verses Ag Cross-Plot on Figure 7.

On Figure 9, the nodes with the highest Ni values belong to Cluster #13 (olive green); whereas the nodes containing the highest Cu values belong to Cluster #8 (teal blue). The spatial locations of samples belonging to those higher-value Ni and Cu nodes within the boxes outlined on the right hand plot are shown on Figure 10.

On Figure 10, samples containing elevated Cu belong to Cluster #8 (teal blue); and those with elevated Ni belong to Cluster #13 (olive green). The spatial distribution of samples that belong to the highlighted nodes forms zones predominately in the north-west and south-east corners of the study area, apart from some elevated Ni samples in the centre of the map area.

(Two ASCII files containing information on the samples plotted on Figure 10 - "Cu vs Ni CuBox_QUEST_out_data.csv" and "Cu vs Ni NiBox_QUEST_out_data.csv"- have been included in the digital data that accompany this report).

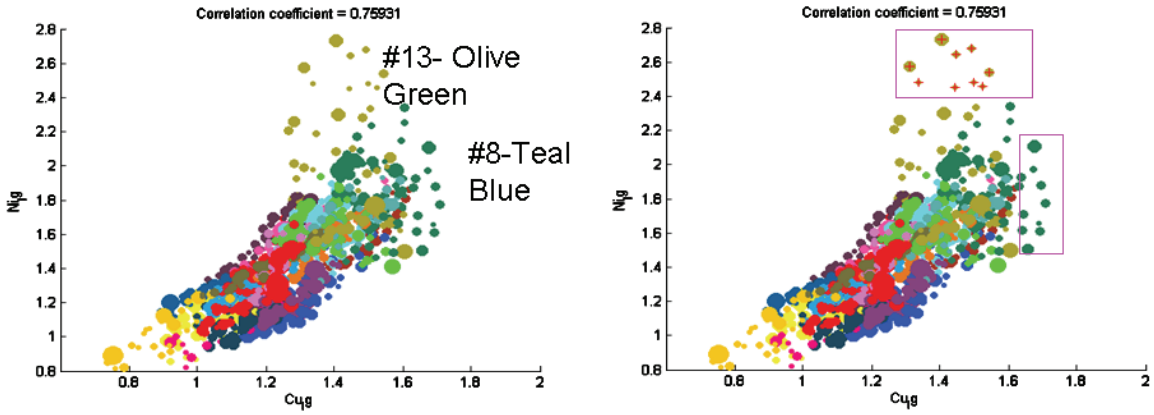


Figure 9. **Cu vs Ni Cross-Plot. The graph on the left shows the cluster-labelling of the nodes associated with the highest Ni (Cluster #13) and Cu (Cluster #8) values. The spatial locations of samples belonging to those higher-value Ni and Cu nodes within the boxes outlines on the right hand plot are shown on Figure 10.**

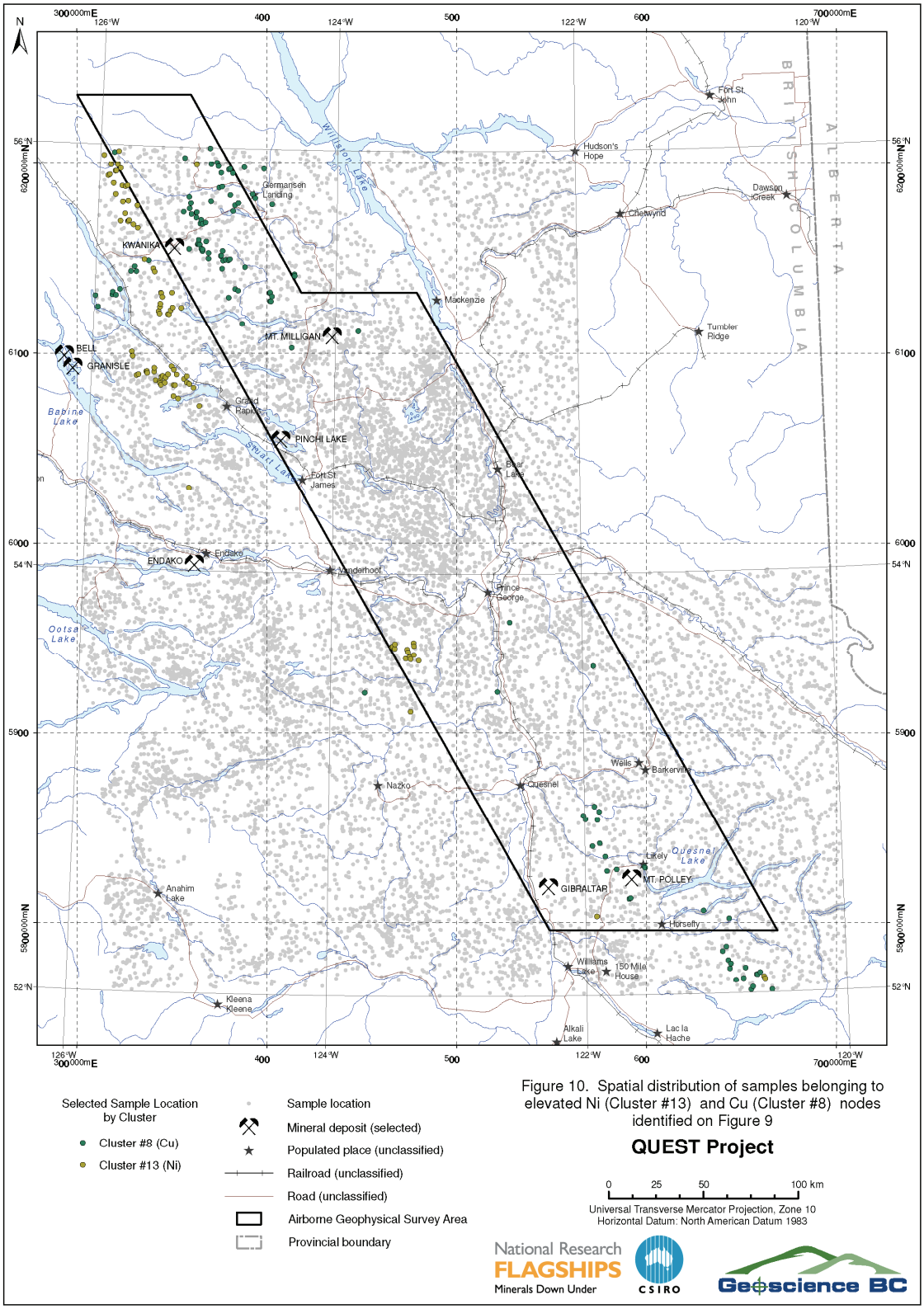


Figure 10. Spatial distribution of samples belonging to the nodes with elevated Ni and Cu values within the boxes shown on the right-hand plot of Figure 9. Samples belonging to Cluster #8 (Cu) are teal blue; and those belonging to Cluster #13 (Ni) are olive green.

4.3.3 Cu verses Mo Cross-Plot

The Cu verses Mo cross-plot is shown in Figure 11. This figure's display is similar to those of the previous cross-plots shown on Figures 7 and 9.

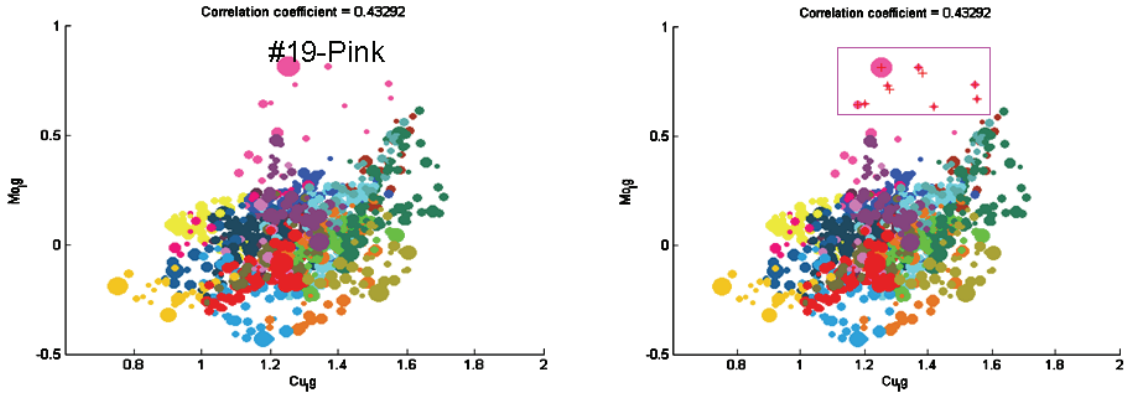


Figure 11. Cu vs Mo Cross-Plot. The graph on the left shows the cluster-labelling of the nodes associated with the highest Mo (Cluster #19) values. The spatial locations of samples belonging to those higher-value Mo nodes within the box outlined on the right hand plot are shown in Figure 12.

On Figure 11, the nodes with the highest Mo values belong to Cluster #19 and are coloured pink. The spatial locations of samples belonging to those higher-value Mo nodes within the box on the right hand plot are shown on Figure 12.

An ASCII file containing information on the samples plotted on Figure 12 - "Cu vs Mo Box_QUEST_out_data.csv"- has been included in the digital data that accompany this report.

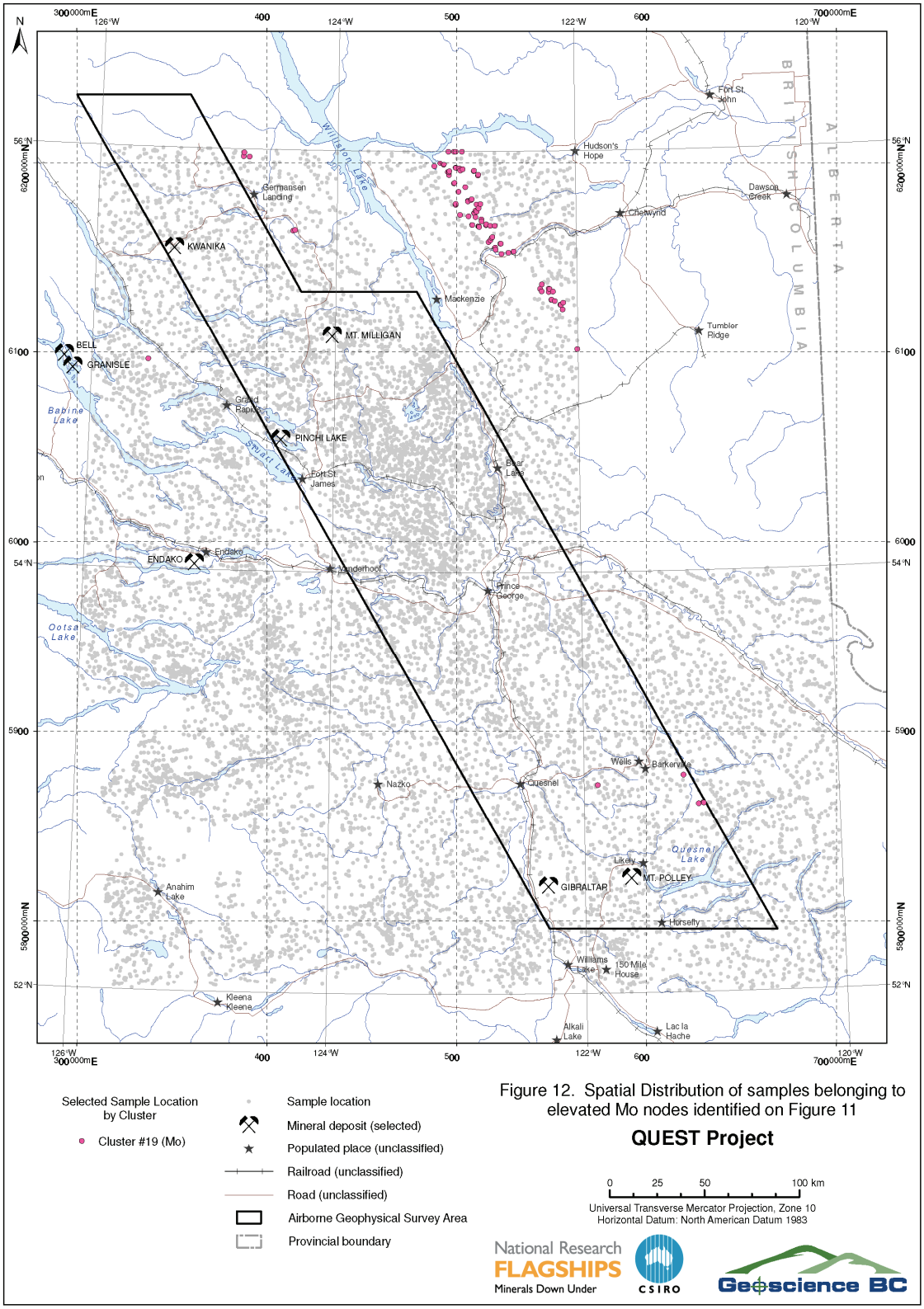


Figure 12. Spatial distribution of samples belonging to the nodes with elevated Mo values within the box shown on the right-hand plot of Figure 11.

4.4 Analysis of Component Plot Nodes

The component plots shown on Figure 3 display the relative contribution of the element in question for each node on the self-organized map. These are coloured using a temperature scale from low-values in blue, to high-values in red, across the range of available BMU node values. Because these nodes are arranged on the "map" in a fashion that retains (and reflects) the structure and topology of the samples belonging to the input data set, the patterns of the sample values on the nodes also reflect this structure and topology. The component plots may, however, also provide insight as to the primary controls behind the structure and composition of the self-organized map and hence, that of the input data set.

4.4.1 Analysis of the Au Component Plot Nodes

Figure 13(a) shows the Au component plot as extracted from Figure 3. (Nodes with elevated Au values are displayed in hotter-temperature colours.) Six populations of Au values were visually identified and ranked from #1 to #6 (Figure 13 (b)). Samples belonging to selected nodes from these six populations (Figure 13 (c)) were extracted to separate and a combined ASCII output files as tabulated below.

Table 2. ASCII files produced for the Six Selected Au Node Populations.

AU_1_ QUEST_out_data.csv
AU_2_ QUEST_out_data.csv
AU_3_ QUEST_out_data.csv
AU_4_ QUEST_out_data.csv
AU_5_ QUEST_out_data.csv
AU_6_ QUEST_out_data.csv
AU_1_6_ QUEST_out_data.csv

Spatial plots for each of the six populations are shown on Figure 14.

Readers who want to follow-up these results should import the above files into their own spatial analysis systems, then assess the significance of their locations with respect to known mineral occurrences, prior mining activity and the readers own models for mineralization in the area.

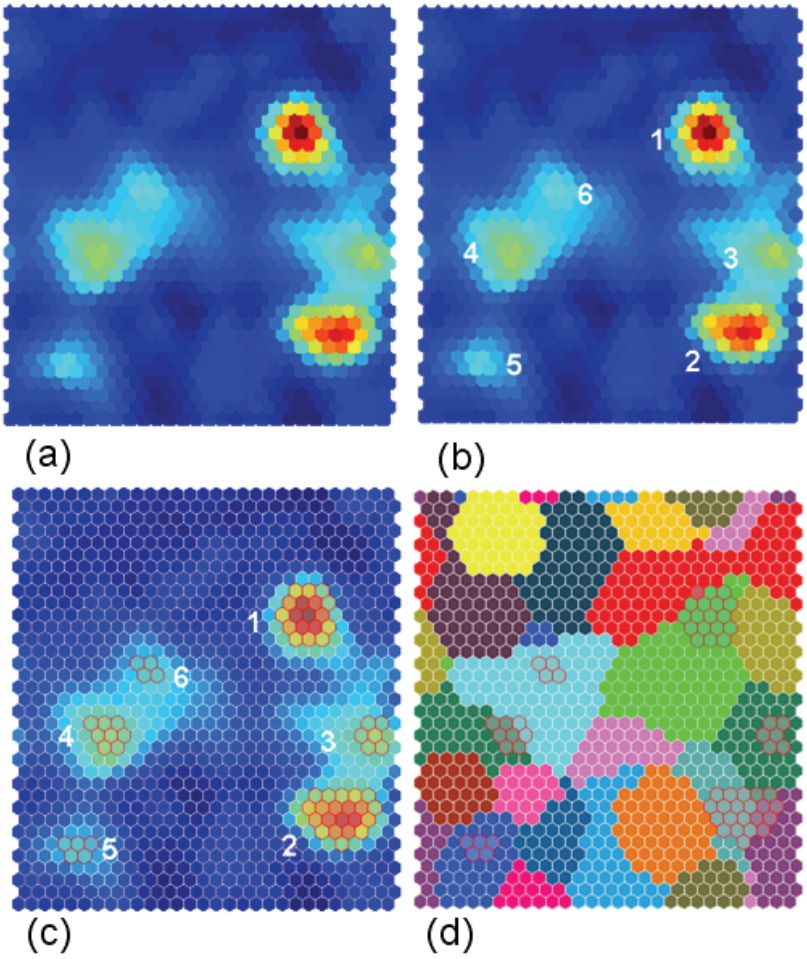


Figure 13. **(a) The Au Component Plot as shown on Figure 3. (b) Six Au populations were identified and ranked #1-#6, with #1 being the highest ranked. (c) Samples belonging to the selected nodes, identified with red outlines, were output to ASCII files. (d) The selected nodes superimposed over the K means clustering of the "map".**

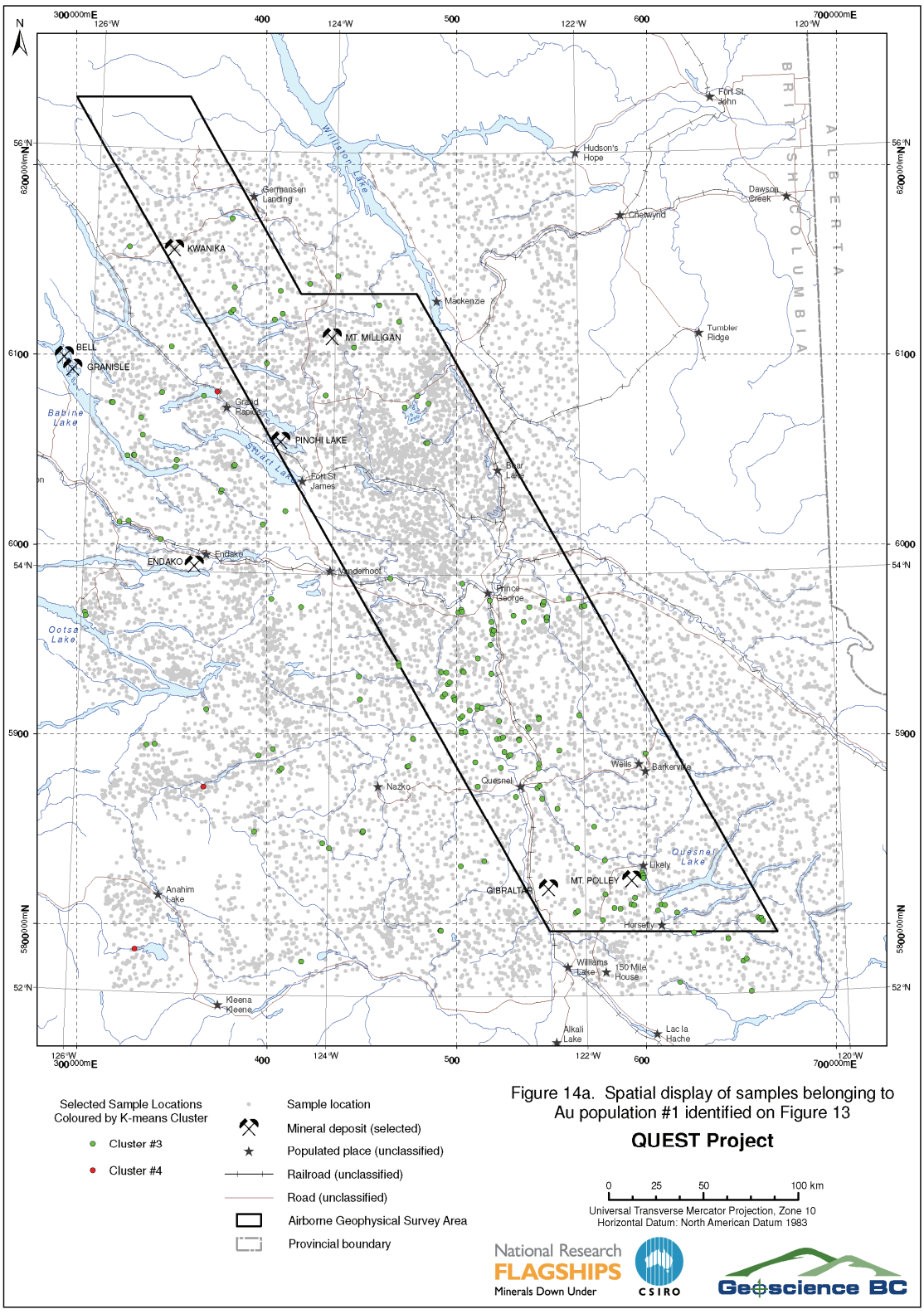


Figure 14a. Spatial displays of samples belonging to Au population #1 that is identified on Figure 13.

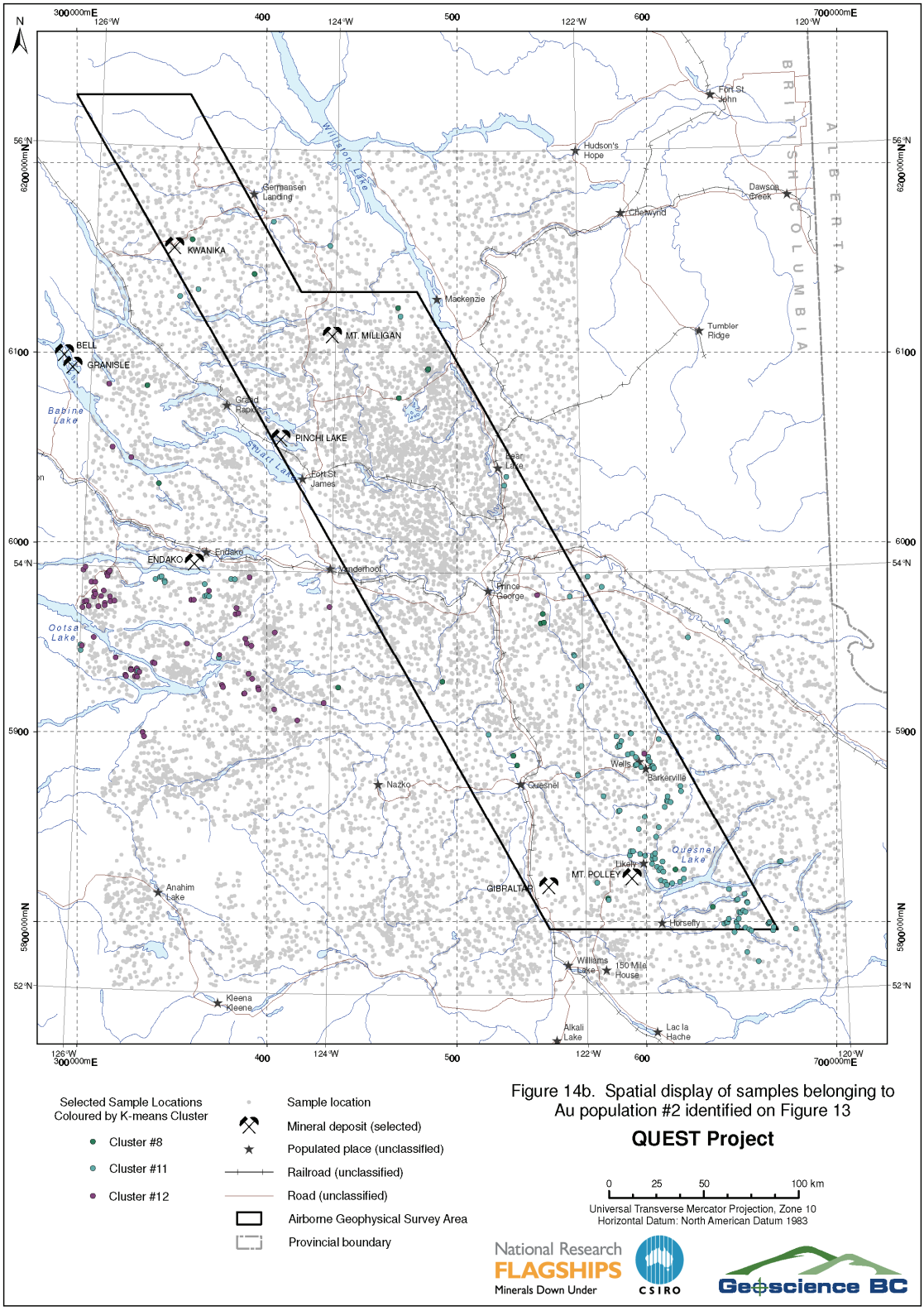


Figure14b. Spatial displays of samples belonging to Au population #2 that is identified on Figure 13.

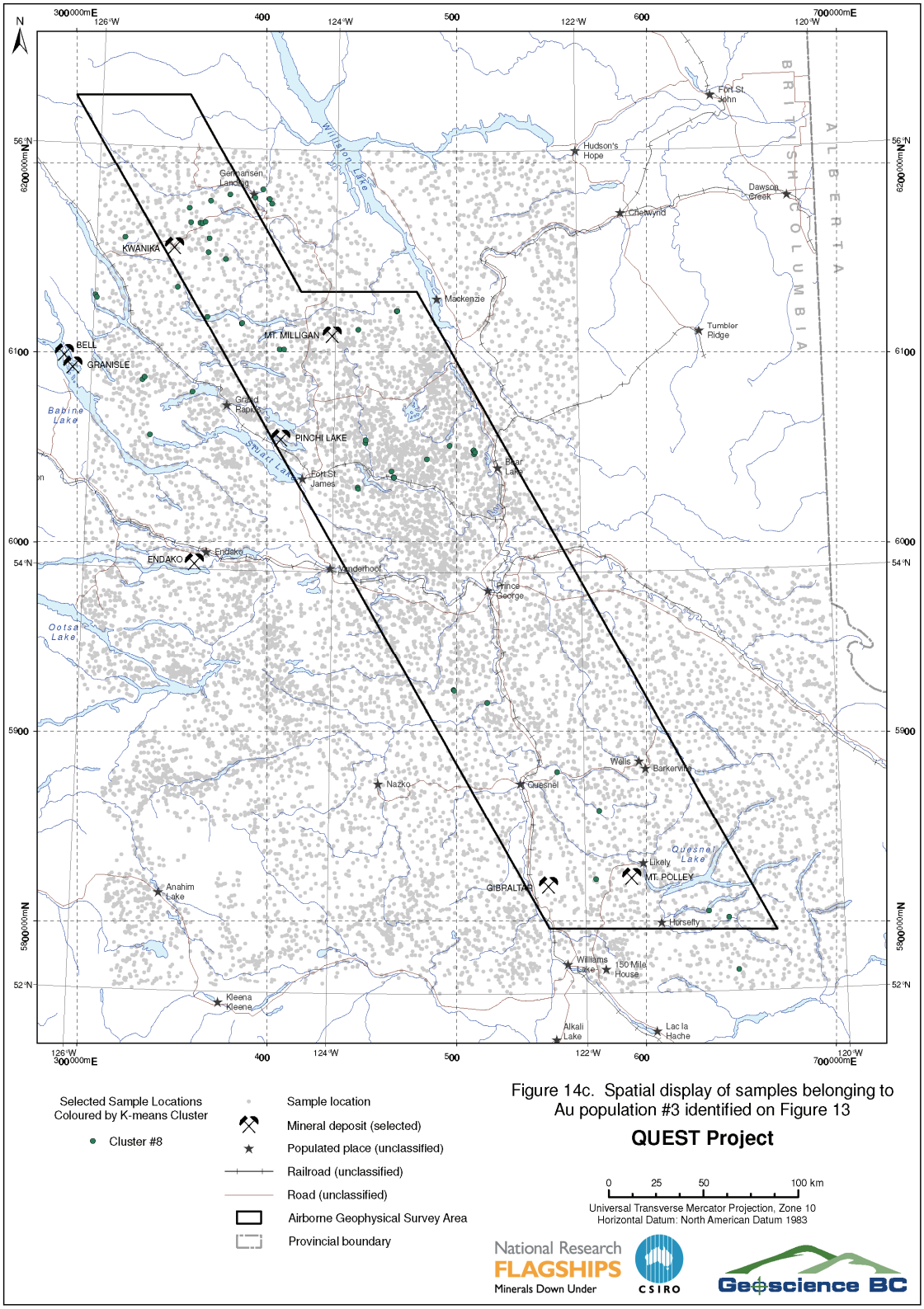


Figure 14c. Spatial displays of samples belonging to Au population #3 that is identified on Figure 13.

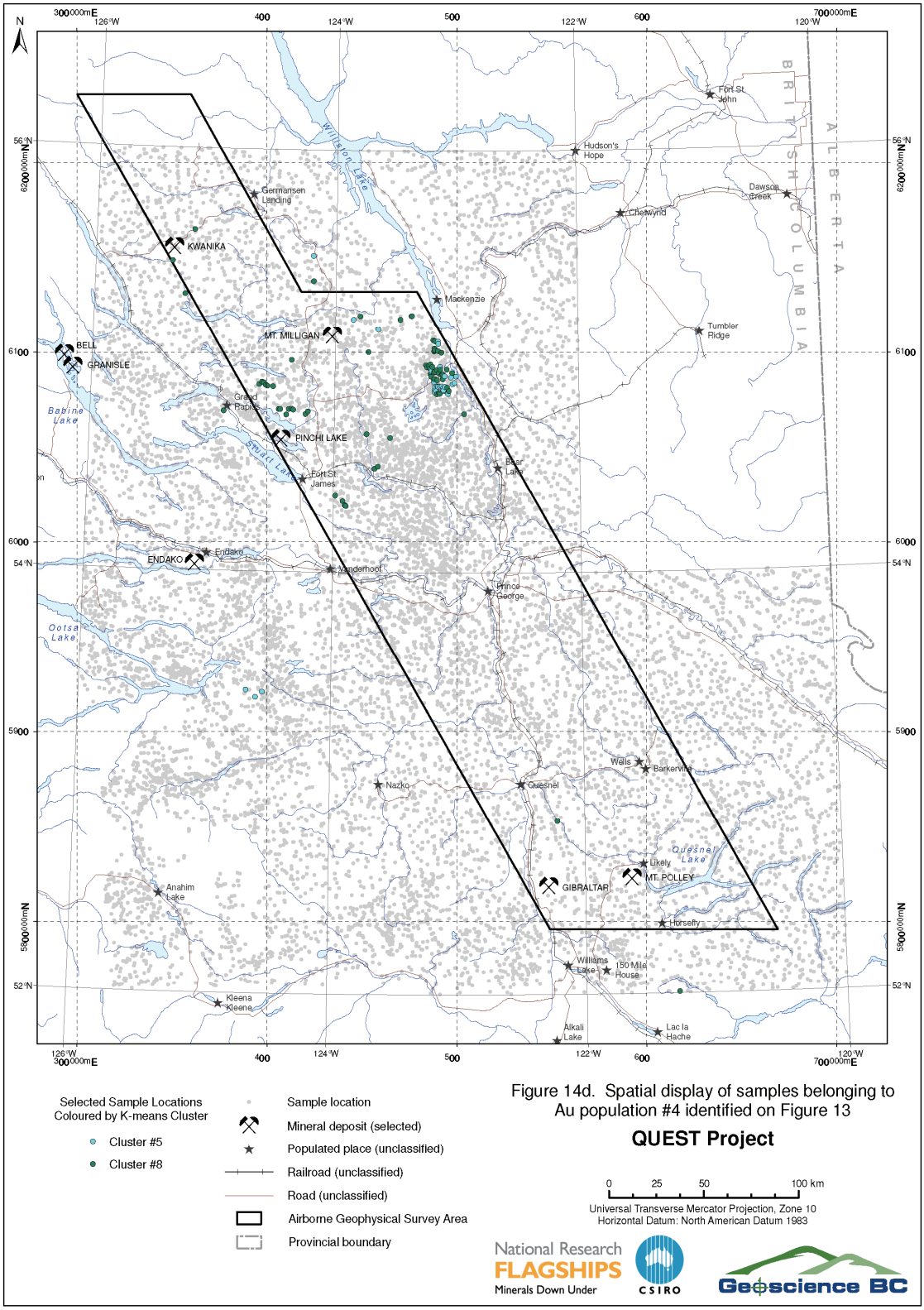


Figure 14d. Spatial displays of samples belonging to Au population #4 that is identified on Figure 13.

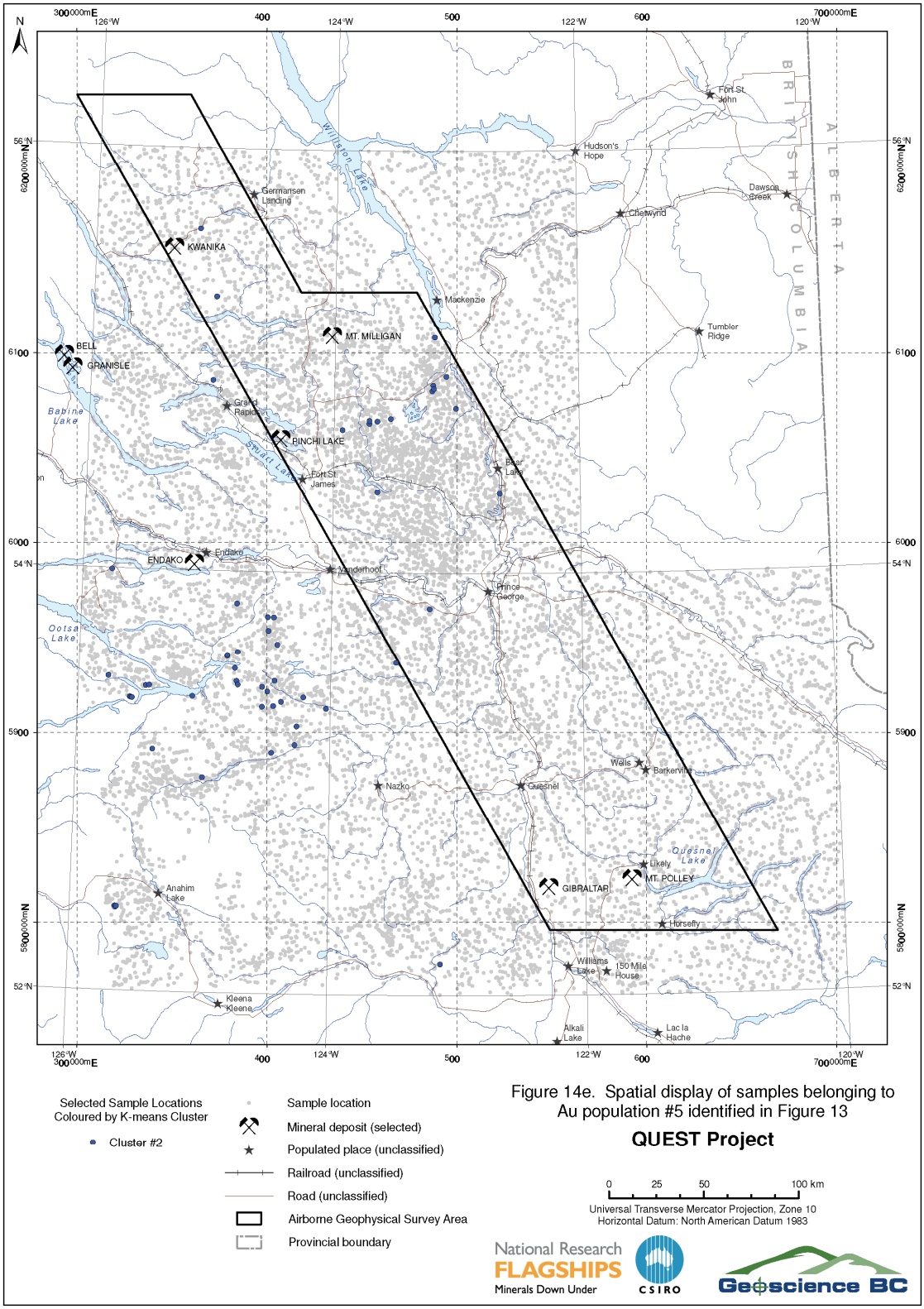


Figure 14e. Spatial displays of samples belonging to Au population #5 that is identified on Figure 13.

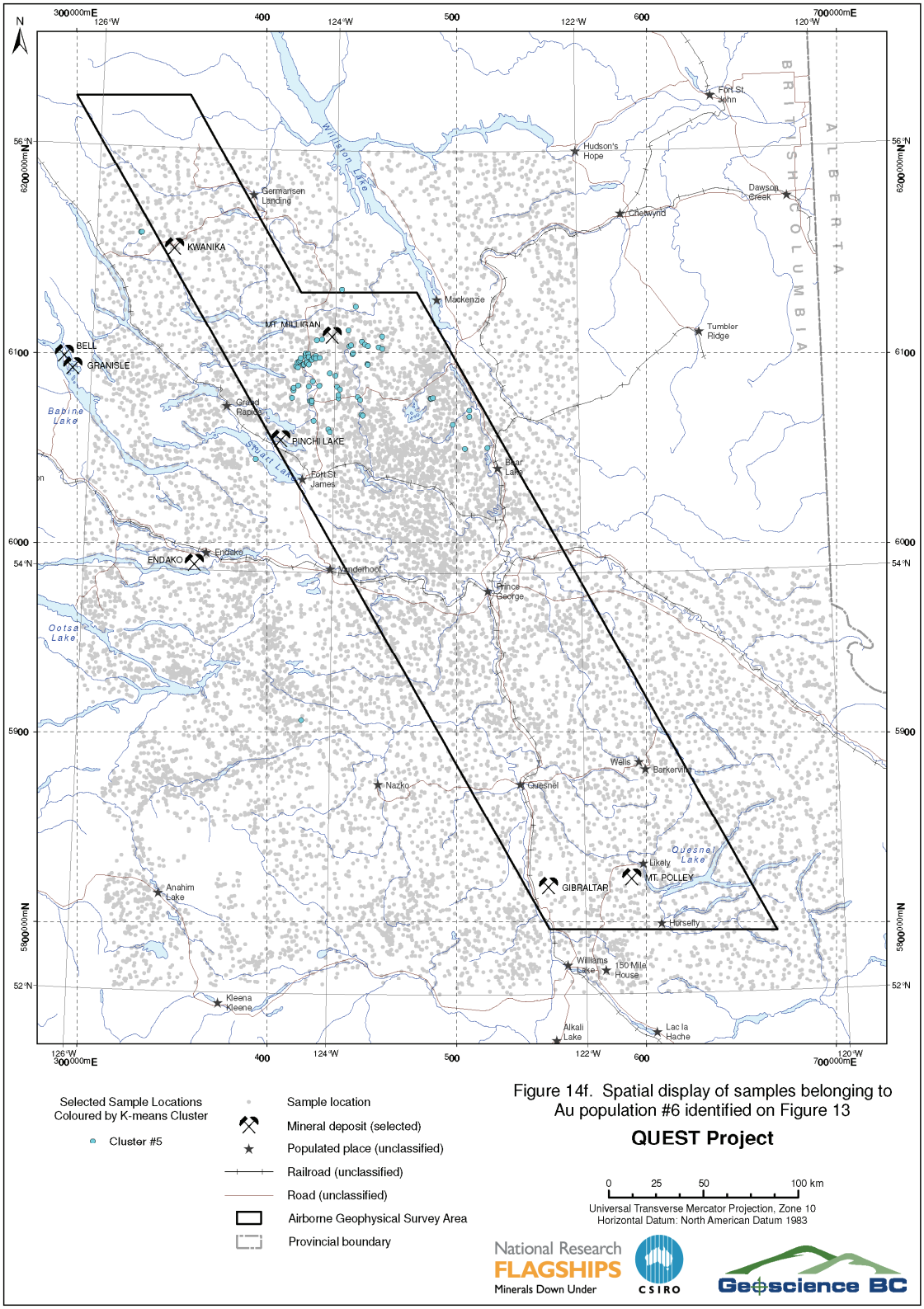


Figure 14f. Spatial displays of samples belonging to Au population #6 that is identified on Figure 13.

4.4.2 Top Thirty Anomalous Au Samples based on Samples from the Selected Au Node Populations and their QER values.

Using all samples from the six Au populations identified in the previous section, we have produced a top-thirty listing of anomalous samples based on an ordering of their QER values in descending order.

Table 2 lists these "Top 30" samples and Figure 15 shows the spatial location of these samples. (An ASCII file containing these Top 30 Au Samples -Top30_Au _Samples.csv- has been included in the digital data that accompany this report.)

Table 3.Listing of the Top 30 Anomalous Au Samples from the Six Selected Au Populations based on their QER values.

Easting_UTM8	Northing_UTM8	somx	somy	BMU	Cluster	q-error	Au_lg
354441	6129712	24	28	833	11	12.592	1.04
302058	5943340	25	28	868	11	12.078	2.86
360296	5969597	26	28	903	12	10.395	2.67
364100	6133308	25	29	869	11	10.297	0.66
357180	6131223	7	22	232	8	10.266	0.74
304415	5964378	22	10	745	3	9.43	2.25
393422	5848457	22	10	745	3	8.545	1.81
661375	5801519	24	12	817	3	8.299	1.08
659954	5802338	24	12	817	3	8.266	1.11
359512	6176091	28	22	967	8	8.16	0.99
617498	5763142	7	20	230	8	7.768	0.68
611216	5863404	23	27	797	11	7.465	1.66
382178	6171692	23	9	779	3	7.235	0.77
360288	6168459	28	21	966	8	7.209	1.33
333883	6163745	10	16	331	5	7.205	0.64
678509	5796361	24	26	831	11	7.205	1.13
334436	6163818	10	16	331	5	7.168	0.62
361597	5969640	26	28	903	12	6.889	1.97
361597	5969640	26	28	903	12	6.889	1.97
585189	5858284	23	26	796	11	6.815	0.6
606531	5826764	24	28	833	11	6.811	1.31
606531	5826764	24	28	833	11	6.811	1.31
658896	5803178	24	11	816	3	6.717	1
597048	5881409	23	27	797	11	6.611	1.68
607504	5827947	24	27	832	11	6.528	1.39
666578	5795503	24	26	831	11	6.497	0.72
470468	6118890	26	27	902	11	6.451	1.75
472068	5799246	22	11	746	3	6.231	1.52
365422	6165491	6	31	206	2	6.166	0.36
610556	5863795	23	27	797	11	6.015	1.41

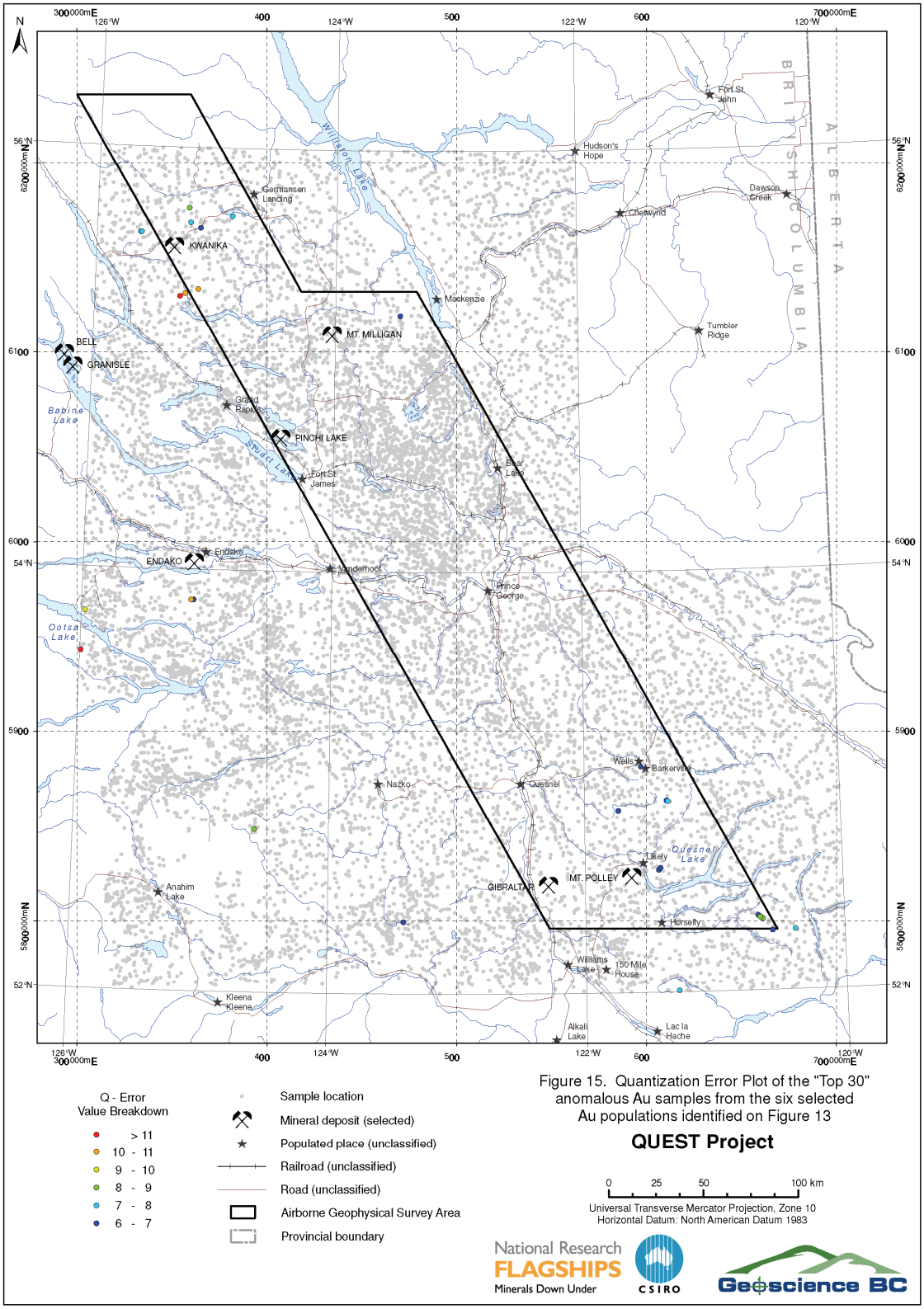


Figure 15. The spatial locations of the Top 30 Anomalous Au Samples from the Six Selected Au populations based on their QER values. Coloured according to QER magnitude (red - high).

4.4.3 Analysis of the Cu Component Plot Nodes

Figure 16 shows the Cu component plot as extracted from Figure 3 with nodes containing elevated Cu values in hotter-temperature colours. The four populations representing the highest node values are highlighted on Figure 16 (b) and the actual nodes selected are on Figure 16 (c).

The spatial locations of the samples belonging to the nodes highlighted on Figure 16 are displayed on Figure 17, coloured by their K-means cluster colour.

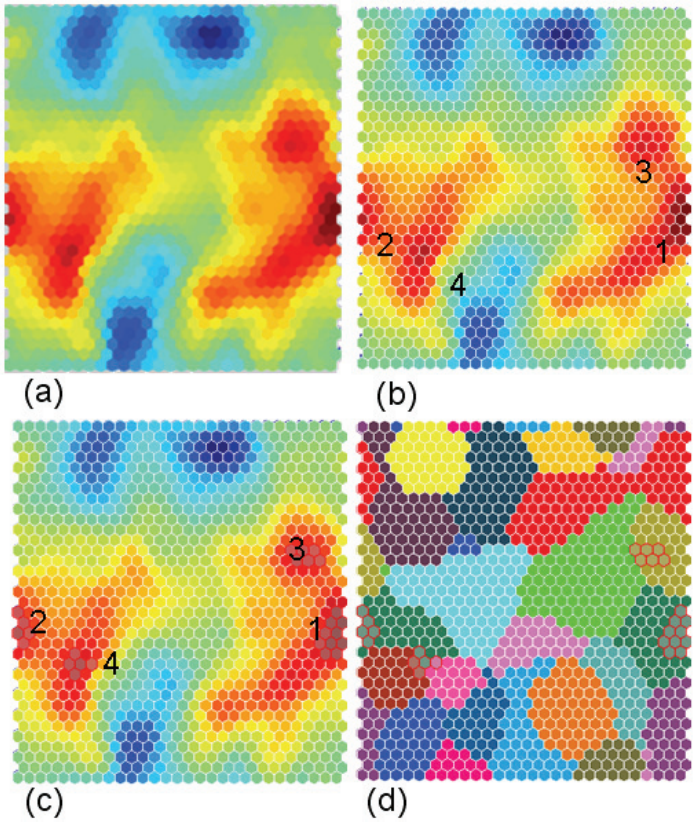


Figure 16. **(a) The Cu component plot as shown on Figure 3. (b) Four Cu populations were identified and ranked #1-#4, with #1 being the highest ranked. (c) Samples belonging to the selected nodes, identified with red outlines, were output to an ASCII file. (d) The selected nodes superimposed over the K-means clustering of the "map".**

(An ASCII file containing information on the samples plotted on Figure 17 - "Cu Nodes QUEST_out_data.csv"- has been included in the digital data that accompany this report).

4.4.4 Top Sixty Anomalous Cu Samples based on Samples from the Selected Cu Node Populations and their QER values.

Using samples from the Cu populations selected elevated nodes from the Cu component plot, we have produced a top sixty listing of anomalous samples based on an ordering of their QER values in descending order (Table 4). Figure 18 shows the spatial locations of these samples coloured according to QER values. An ASCII file containing these samples ("Top 60 Cu Nodes QUEST_ out_data.csv") is included with the digital data that accompany this report.

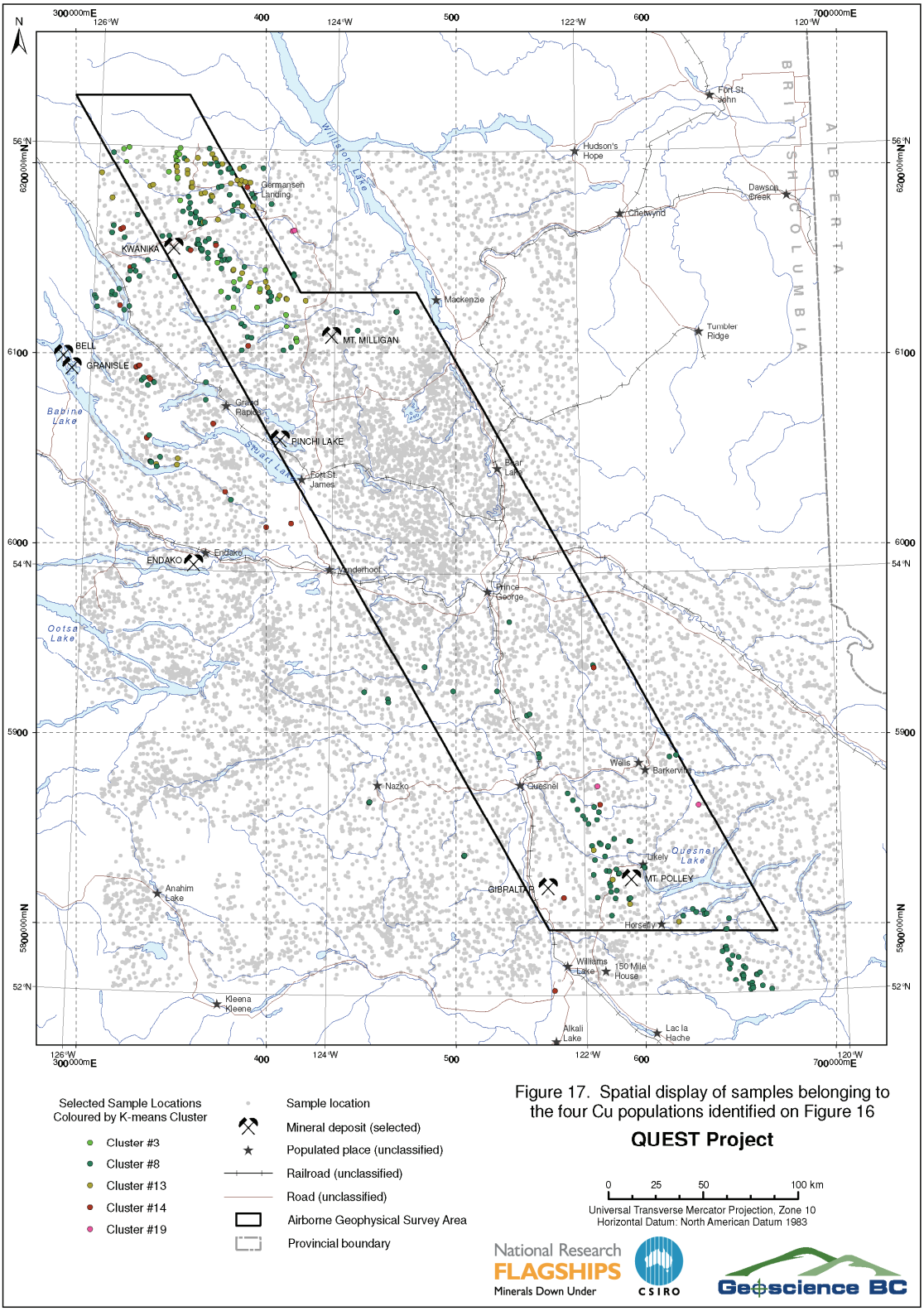


Figure 17. Spatial Locations of samples belonging to the Cu populations identified on Figure 16. Coloured according to K-means cluster.

Table 4:Listing of the Top 60 Anomalous Cu Samples from the Four Selected Cu Populations based on their QER values.

Easting_UTM83	Northing_UTM83	somx	somy	BMU	Cluster	q-error	Cu_lg
591111	5842836	6	24	199	8	12.811	2.44
574109	5853952	1	21	21	8	10.443	2.47
591697	5842595	6	24	199	8	9.407	2.09
643116	5780206	28	23	968	8	9.018	2.1
583135	5843852	6	24	199	8	8.988	2.25
352040	6140450	6	24	199	8	8.604	2.38
360194	6178850	29	21	1001	8	8.296	2.14
360386	6154120	6	25	200	14	8.286	1.94
359512	6176091	28	22	967	8	8.16	2.31
323713	6165228	6	25	200	14	7.395	1.37
360288	6168459	28	21	966	8	7.209	1.76
319998	6205873	29	21	1001	8	7.186	2.14
319998	6205873	29	21	1001	8	7.186	2.14
324930	6165769	6	25	200	14	7.048	1.41
331825	6092627	6	25	200	14	6.989	1.85
331662	6092428	6	25	200	14	6.957	1.83
358445	6201435	26	14	889	13	6.639	2.16
337880	6148888	28	23	968	8	6.574	1.53
349178	6161734	6	24	199	8	6.574	1.63
627625	5862313	7	24	234	19	6.342	1.71
357073	6202616	29	23	1003	8	6.319	2.41
552085	5763696	6	25	200	14	6.283	1.84
413978	6164056	7	24	234	19	6.08	1.36
353060	6197817	25	14	854	13	5.91	2.28
353060	6197817	25	14	854	13	5.91	2.28
353060	6197817	25	14	854	13	5.91	2.28
352724	6203364	24	14	819	3	5.906	2.1
353300	6195266	25	14	854	13	5.889	2.17
376151	6194858	29	19	999	8	5.838	2.03
360451	6173109	29	21	1001	8	5.827	1.84
353311	6205610	24	14	819	3	5.711	2.07
572346	5934043	5	24	164	14	5.68	1.44
414993	6164340	7	24	234	19	5.669	1.54
352790	6198193	25	14	854	13	5.63	2.19
388141	6171978	6	24	199	8	5.604	1.77
644572	5778579	28	23	968	8	5.541	1.81
325628	6160804	28	21	966	8	5.498	1.49
390183	6187022	6	25	200	14	5.497	1.44
382320	6174752	29	19	999	8	5.474	1.7
528023	5958220	28	23	968	8	5.467	1.62
340665	6188325	25	14	854	13	5.466	1.31
365129	6168159	28	22	967	8	5.436	1.89
571623	5840174	28	23	968	8	5.426	1.57
353496	6201029	24	14	819	3	5.425	1.82
571901	5935311	29	19	999	8	5.41	1.56
353763	6204797	24	14	819	3	5.313	2.02
333217	6093208	6	25	200	14	5.311	1.67
331470	6143176	29	20	1000	8	5.292	1.49
359221	6154809	28	23	968	8	5.241	1.79
335756	6086868	28	21	966	8	5.188	1.69
630091	5806424	29	20	1000	8	5.184	1.89
355018	6208033	24	14	819	3	5.156	1.99
333478	6093274	6	25	200	14	5.155	1.59
326711	6200778	24	14	819	3	5.142	1.54
591004	5812423	1	21	21	8	5.134	1.93
353141	6041292	26	14	889	13	5.113	1.86
334646	6085634	28	21	966	8	5.11	1.61
334646	6085634	28	21	966	8	5.11	1.61
334646	6085634	28	21	966	8	5.11	1.61
334646	6085634	28	21	966	8	5.11	1.61

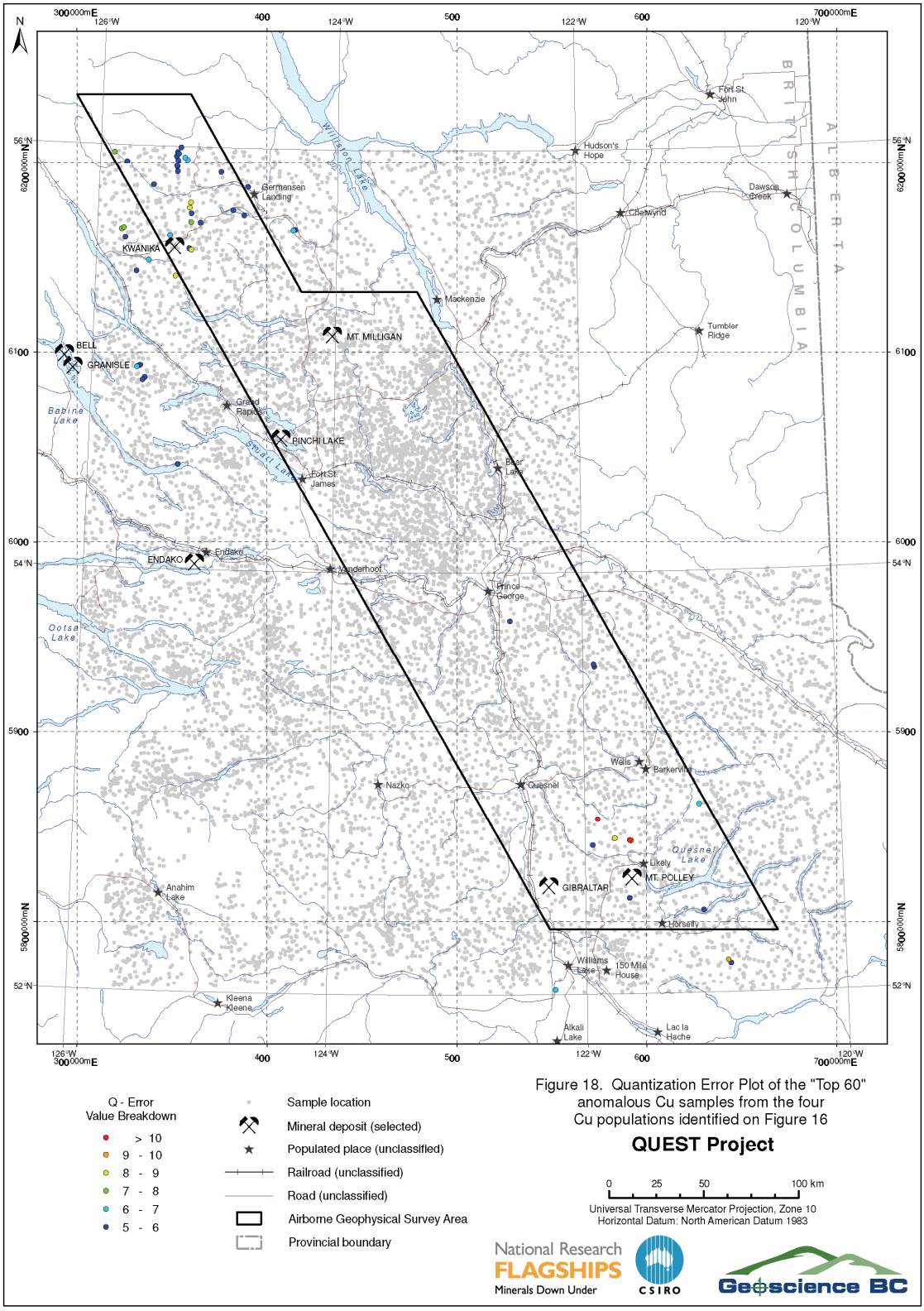


Figure 18. Spatial locations of the Top 60 Anomalous Cu Sample populations based on their QER values. Coloured according to QER magnitude (red - high magnitude).

4.4.5 Analysis of the Mo Component Plot Nodes

Figure 19 shows the Mo component plot extracted from Figure 3 with nodes containing elevated Mo values in hotter-temperature colours. The three populations representing the highest node values are highlighted on Figure 19 (b) and the actual nodes selected are on Figure 19 (c). Figure 19 (d) shows the selected nodes over the K-means clusters on the self-organized map.

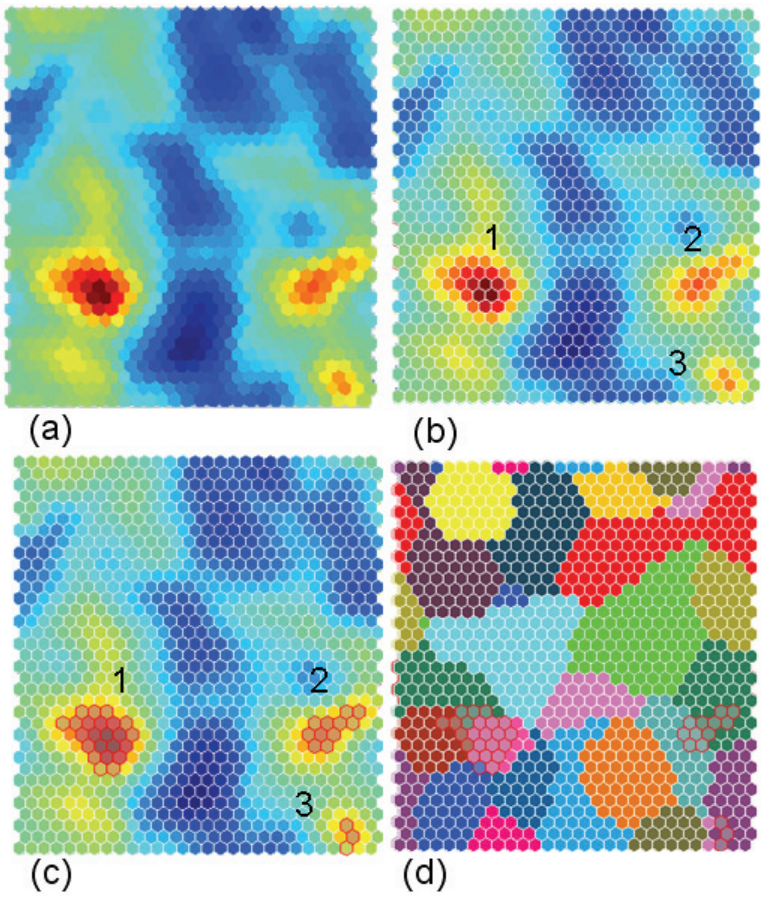


Figure 19. **(a) The Mo Component Plot as shown on Figure 3; (b) Three Mo populations were identified and ranked #1-#3, with #1 being the highest ranked; (c) Samples belonging to the selected nodes, identified with red outlines, were output to an ASCII file; (d) The selected nodes superimposed over the K-means clustering of the "map".**

The spatial locations of the samples belonging to the nodes highlighted on Figure 19 are displayed on Figure 20, coloured by their K-means cluster colour.

(An ASCII file containing information on the samples plotted on Figure 20 - "Mo Nodes QUEST_out_data.csv"- has been included in the digital data that accompany this report).

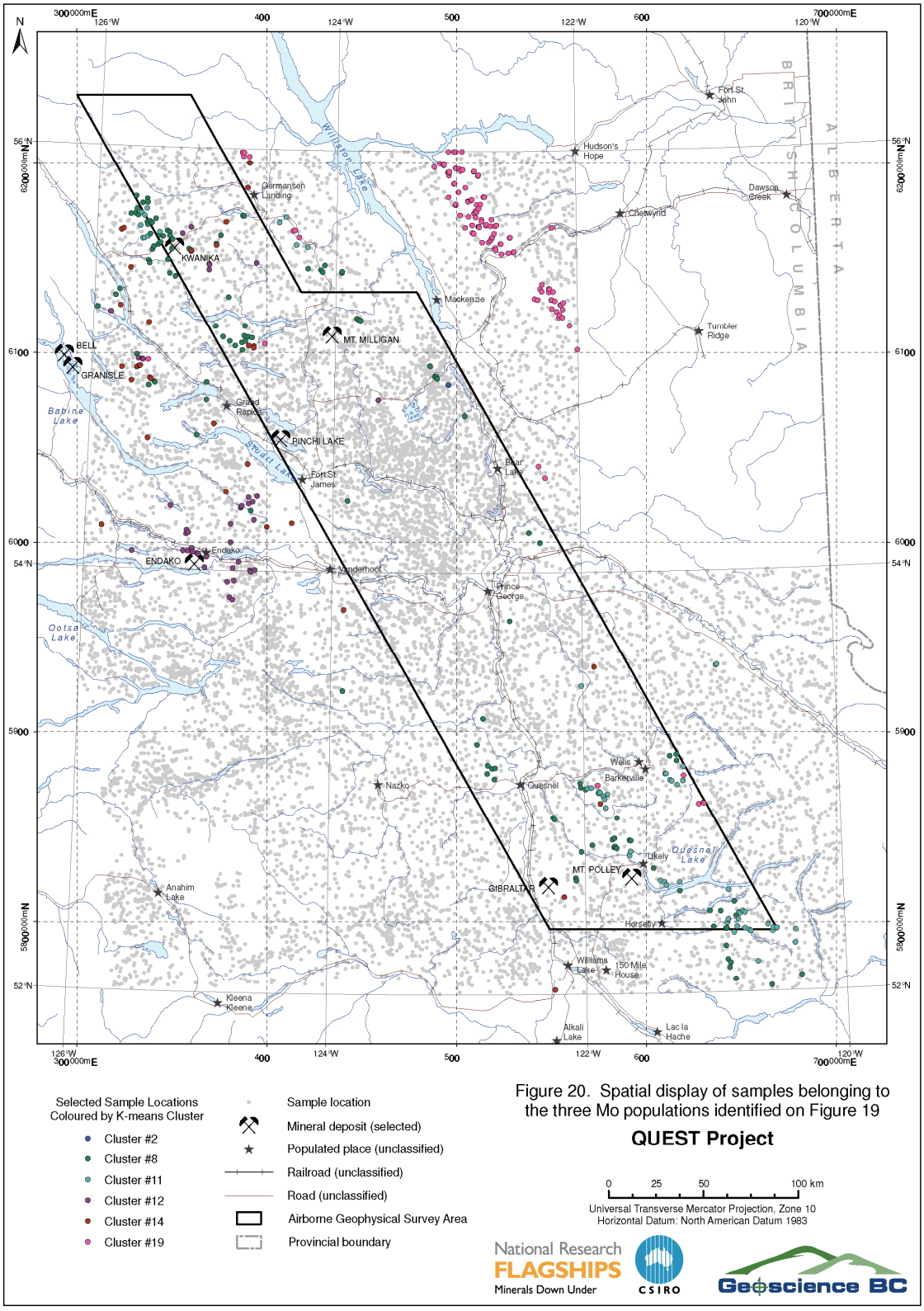


Figure 20. Spatial Locations of samples belonging to the Mo populations identified on Figure 19. Coloured according to K-means cluster.

4.4.6 Top Sixty Anomalous Mo Samples based on Samples from the Selected Cu Node Populations and their QER values

Using samples belonging to the elevated Mo nodes as shown on the Mo component plot, we have produced a top sixty listing of anomalous samples based on an ordering of their QER values (Table 5). Figure 21 shows the spatial locations of these samples coloured according to QER values. An ASCII file containing these samples ("Top 60 Mo Nodes QUEST_ out_data.csv") is included with the digital data that accompany this report.

5. DISCUSSION ON THE INTERPRETATION OF THE SOM RESULTS

5.1 General

In this study, SOM has been used to analyse stream and lake sediment geochemical concentration data to highlight anomalism indicative of mineralization. During the SOM procedure, input samples with similar characteristics are assigned to particular BMU nodes that are arranged on the self-organized map so that adjacent BMU nodes on the SOM "map" also have a degree of similarity. We used a K-means approach to group the BMU nodes into similar clusters. Hence there is a hierarchy of "similarity" (and dissimilarity) produced.

"Anomalism" can occur at various levels in terms of the self-organized map. At the lowest level, individual samples with particular characteristics (high QER values and/or elevated or reduced elemental values) can be "outliers" to the population of samples assigned to a particular node. At a higher level, samples that belong to a particular node, or group of nodes as defined by a cluster, may represent "anomalous" samples, compared with all other samples in the global population being examined.

The following sections are non-exhaustive, but are intended to alert the reader to the possible analysis approaches available as a consequence of the SOM analysis. Also, because many of the SOM outputs and enhancements can be used in conjunction with each other, there is overlap between the approaches mentioned.

Table 5: Listing of the Top 60 Anomalous Mo Samples from the Three Selected Mo Populations based on their QER values on Figure 19.

Easting_UTM83	Northing_UTM83	somx	somy	BMU	Cluster	q-error	Mo_lg
591111	5842836	6	24	199	8	12.811	1.57
337687	6096453	7	25	235	19	11.13	1.74
391056	6203402	7	25	235	19	11.016	0.68
391056	6203402	7	25	235	19	11.016	0.68
391056	6203402	7	25	235	19	11.016	0.68
391056	6203402	7	25	235	19	11.016	0.68
363392	5994111	27	34	944	12	10.02	2.08
591697	5842595	6	24	199	8	9.407	1.11
363933	5993979	27	34	944	12	9.371	2.13
387998	6203566	7	25	235	19	9.182	1.01
643116	5780206	28	23	968	8	9.018	2.11
583135	5843852	6	24	199	8	8.988	1.44
352040	6140450	6	24	199	8	8.604	0.63
338117	6116205	6	26	201	14	8.416	0.95
367278	5986795	27	34	944	12	8.338	2
360386	6154120	6	25	200	14	8.286	1.64
503230	6196283	8	26	271	19	8.266	1.34
584327	5838595	7	23	233	8	7.656	1.29
334708	6096709	27	33	943	12	7.599	1.07
334829	6096524	27	33	943	12	7.46	1.05
323713	6165228	6	25	200	14	7.395	0.7
360892	5994934	27	34	944	12	7.356	1.72
388322	6205648	7	25	235	19	7.232	0.93
678509	5796361	24	26	831	11	7.205	0.4
368234	5992319	27	34	944	12	7.115	1.39
324930	6165769	6	25	200	14	7.048	0.47
331825	6092627	6	25	200	14	6.989	1.14
331662	6092428	6	25	200	14	6.957	1.13
398985	6104697	7	27	237	19	6.944	0.69
398985	6104697	7	27	237	19	6.944	0.69
619551	5877321	8	24	269	19	6.756	0.97
391098	6200267	6	26	201	14	6.688	0.49
439921	5921332	27	24	934	8	6.686	0.45
387393	6205739	7	27	237	19	6.619	0.79
337880	6148888	28	23	968	8	6.574	0.55
349178	6161734	6	24	199	8	6.574	0.55
448511	6117147	7	23	233	8	6.572	0.58
391220	6200052	6	26	201	14	6.527	0.45
666578	5795503	24	26	831	11	6.497	1.18
603827	5828126	7	23	233	8	6.425	0.55
491395	6199691	8	26	271	19	6.418	0.91
332584	6097082	27	33	943	12	6.396	1.06
627625	5862313	7	24	234	19	6.342	0.81
552085	5763696	6	25	200	14	6.283	0.54
331636	6099276	26	25	900	8	6.235	1.07
335930	6172531	25	25	865	11	6.232	0.95
355761	6162915	27	33	943	12	6.215	0.8
510031	6171081	8	26	271	19	6.109	0.93
510031	6171081	8	26	271	19	6.109	0.93
510031	6171081	8	26	271	19	6.109	0.93
510031	6171081	8	26	271	19	6.109	0.93
358379	5995475	27	34	944	12	6.095	1.61
413978	6164056	7	24	234	19	6.08	0.62
440634	5964253	4	24	129	14	6.024	0.65
546506	6034102	9	24	304	19	5.995	0.82
636386	5935258	24	25	830	11	5.988	0.85
636915	5935730	24	25	830	11	5.973	0.84
630082	5862943	8	24	269	19	5.964	0.79
318069	6122150	27	34	944	12	5.936	1.1
421898	6144197	27	23	933	8	5.93	0.63

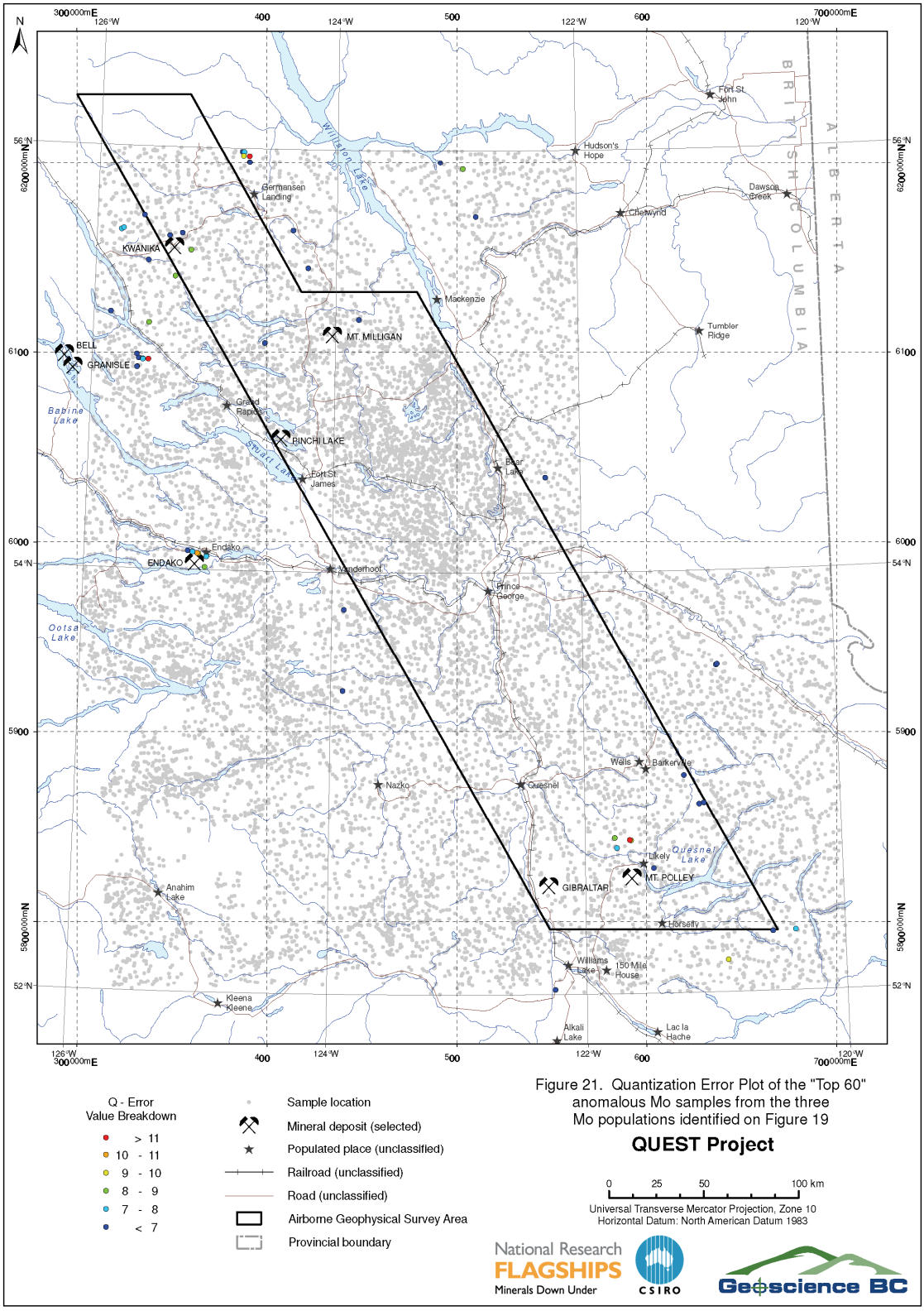


Figure 21. Spatial locations of the Top 60 Anomalous Mo Sample populations based on their QER values. Coloured according to QER magnitude (red -high).

5.1.1 Quantization Errors

To recount, the quantization error (QER) is a measure of how much a sample is different (i.e., an outlier) compared to the SOM node vector that represents it on the self-organized map. It is important to understand, however, that a sample's variation in characteristics compared to that of the node vector can be caused by either an increase or decrease in the abundance of any element or combination thereof. Hence, care should be used in the interpretation of samples with high QER values. It is recommended that samples with high QER values should be assessed in terms of other supporting evidence (elevated values in elements of interest and the spatial context of nearby samples' QER and concentration values).

Figure 5 is a spatial map showing the locations of samples coloured-coded by the size of their QER (red-high; blue-low). Particular areas on that map (in both the north-west and south-east) show higher concentrations of elevated QER values and these potentially correspond to the presence of a variety of sediment source materials. Sediment trains with elevated QERs should be of interest as these occurrences could represent an anomaly train and hence should be targeted for subsequent follow-up.

The QER values can be used as a parameter to assist in the rating of samples. In Section 4.4, we used the QER values to refine the ranking of samples that were selected on the basis of being assigned to BMU nodes with particular elevated elemental values.

5.1.2 SiroSOM-derived K-means Clustering

The K-means clustering of the nodes on the self-organized map (Figure 2) is an attempt to see patterns and perhaps even process within the larger data set. It is possible that a cluster represents sediment samples from a particular source rock that has been affected by similar weathering processes and or transport history; whereas an adjacent cluster may represent a similar lithology, affected by both mineralizing and weathering effects.

Plotting the spatial location of samples coloured by the cluster assignment of the nodes to which those samples belong, (Figure 4) should, to some degree, reflect the local upstream geology and sediment transport environment. Hence there could be a relationship of the domains identified on this spatial map to bedrock geology. However, unlike the Barnett and Williams (2009) study, which specifically intended to infer bedrock geology from the sediment elemental concentrations (see their Figure 12), Figure 4 should be more similar to a surface geology map.

5.1.3 "Cluster-Normalized" Element Anomaly Maps

The reasoning behind the "cluster-normalized" element plots lies in the assumption that each SOM-derived, K-means cluster represents a population of samples with particular elemental distribution characteristics that are related to both surface processes and geology. By normalizing each sample's elemental concentration on a per assigned cluster basis, we have attempted statistically to produce a "level playing field" in terms of geochemical baselines across the whole of the data set, to allow comparison across variations in background elemental responses.

The maps and data produced by this process are products that readily can be assessed for evaluation purposes, via either hardcopy (printing the digital maps) or by plotting the ASCII files in a spatial information system.

5.1.4 Component Plots

The simplest way to find samples belonging to populations with above-average values for an element is to look for hotter-temperature coloured nodes on the component plots (Figure 3) and examine the samples that are assigned to those nodes. Note that, as a consequence of the SOM analysis, samples are assigned to a node on the basis of that sample's total response. For example, it is possible for a sample with a low Au value to be assigned to a population on a node with a much higher average Au concentration; this is considered an advantage of the SOM method as subtle anomalies can be identified.

Case studies involving the selection of elevated nodes on component plots for Au, Cu and Mo are presented in Sections 4.4.1, 4.4.2 and 4.4.3, respectively. The Au and Mo component plots show multiple "bulls-eye style" regions on the self-organized map (Figures 13(a) and 19(a)), which can be interpreted as unique or compositional variations of sub-populations within the larger data set. In comparison, the Cu component plot (Figure 16(a)) essentially shows a broad elevated population, but this elevated zone has even higher sub-populations within it (remember that the map wraps from top-to-bottom, and side-to-side).

The component plots of the elements, Au, Cu and Mo, appear to be related. The second highest Au population (#2 on Figure 13) coincides with an elevated "ridge" on the Cu component (Figure 16) and an area between Mo populations #2 and #3 (Figure 19). The highest ranked Cu nodes (population #1 on Figure 16) corresponds to Au population #3 (Figure 13) and the side of Mo population #2 (Figure19). The highest ranked Mo nodes (population #1 on Figure 19) coincides with Cu population #4 (Figure 16), which is also positioned between Au populations #4 and #5 (Figure 13). These relationships suggest that, at least for the elemental populations mentioned, there are various genetic relationships between them.

For convenience, the samples belonging to BMU nodes on the self-organized map are often displayed or plotted (spatially or otherwise) in the colour of the K-means cluster assignment for the node in question

5.1.5 Cross Plots of BMU Node Component Values

Cross-plots of BMU node values are another approach we have used in this study to highlight nodes and clusters that may be important for exploration purposes. Because the BMU nodes are a condensed representation of the data set, cross-plots of BMU node values are typically more easily interpreted than if the raw input data are plotted. Such plots can be even more meaningful if the element selections are driven by prior knowledge regarding the geochemical attributes of possible mineralizing systems that could be represented in the data set.

The Au vs Ag BMU node cross-plot (Figure 7) indicates there are two elevated Au trends in these data: one with a lower, and the other with a higher set of Ag values. The spatial distribution of samples belonging to the nodes within the box on Figure 7 is shown on Figure 8, and this indicates there are distinct zones and belts where samples belonging to these trends

occur. It would be of interest to determine whether there is other geological information available to validate these zones spatially. While most of the nodes highlighted belong to Clusters #11, #12 and #3, the single node belonging to Cluster #4 (red), which is included in the low Ag trend, appears to be worthy of further investigation. Figure 8 indicates that there are at least three (red) samples associated with the Cluster#4 node.

The Cu vs Ni cross-plot (Figure 9) indicates that there is an association between higher Ni values and Cu, and a lesser association between higher Cu values and Ni. The spatial locations of samples belonging to the nodes outlined within Figure 9 are plotted on Figure 10. Samples belonging to these high Ni (Cluster#13 - olive-green) and high Cu (Cluster #8 - teal-blue) nodes are shown on Figure 10 and these also form a weak north-west to south-east zone across the area. The more localized spatial grouping of samples associated with the elevated Ni BMU nodes should also be investigated.

The Cu vs Mo cross-plot (Figure 11) again shows the highest BMU nodes associated with Cu to be associated with Cluster #8 (teal-blue). The BMU nodes associated with the highest Mo values belong to Cluster #19 (pink). The spatial plots of samples belonging to the nodes with the highest Mo values shows what could be a lithologically or structurally controlled (NW-SE) belt in the north-central area and three well defined spatial groupings of samples in the north-western corner of the map area. In the south-east corner of the map there is another group of samples that should be followed up.

6. RECOMMENDATIONS

i) This SOM analysis complements, rather than replaces, more traditional methods of assessing stream and lake sediment geochemical survey data. It is recommended that the results of both methods need to be assessed to ensure confidence and maximize discovery opportunities.

ii) To be as objective as possible, this study has been undertaken without reference to known mineral occurrences, nor detailed knowledge of the study area. However, now that the processing has finished, the locations and distributions of samples highlighted by this study need to be assessed in the context of known mineral occurrences and with regards to the structural evolution and metallogenesis of the Quesnellia Terrane.

iii) In this report we have demonstrated how the results of this SOM analysis can be used for Au, Cu and Mo (Ni ?) exploration. If there is interest in other elements, the reader is recommended to use a similar strategy. To assist, maps and digital ASCII data have been provided.

ACKNOWLEDGEMENTS

This study has been built on the shoulders of many. Our analysis would not be possible without the dedication and foresight of those agencies and geologists who were responsible for collecting, analysing and cataloguing the geochemical stream, and lake sediment data sets. The pivotal work of Colin Barnett and Peter Williams is acknowledged as their levelled and imputed grids were the foundations for this study; and despite our efforts, we were not able to find fault with their grids nor of the data contained within. Many thanks go to Peter Kowalczyk for mentoring this study, reviewing this report and his overall assistance. Fion Ma (Geoscience BC) is acknowledged for her GIS skills, and thanked for the manner in which she transformed our SiroSOM outputs into quality maps. Geoscience BC is thanked for supporting this project and for allowing Fion to assist in map production. Christa Sluggett is thanked for her professionalism and enthusiasm in her role as our prime Geoscience BC contact. Last but not least, the authors would like to thank Cameron Huddlestone-Holmes and Binzhong Zhou from CSIRO Exploration and Mining for reviewing and editing this manuscript.

REFERENCES

Barnett, C.T., and Williams, P.M., 2009: Using Geochemistry and Neural Networks to Map Geology under Glacial Cover. Final Report for Geoscience Project 2008-003. Geoscience BC Report 2009-003.

Briqueu, L., Gottlib-Zeh, S, Ramadan, M., and Brulhet, J, 2002: Traitement des disgraphies a l'aide d'un reseau de neurons du type <<carte auto-organisatrice>>: application a l'etude lithogique de la couche silteuse de Marcoule (Gard France). C.R. Geoscience 334 (2002) 31-337.

Davies, D., and Bouldin, D., 1979: A Cluster Separation Measure, IEEE Trans. Pattern Anal. Mach. Intell. 1(2) 224-227.

Fraser, S.J., and Dickson, B.L., 2005: Ordered vector quantization for the integrated analysis of geochemical and geoscientific data sets. "IGES 2005 – 22nd International Geochemical Exploration Symposium incorporating the 1st International Applied Geochemistry Symposium. Program and Abstracts, 19-23 September, 2005 Perth, Western Australia pp 52-53.

Fraser, S.J., and Dickson, B.L., 2007: A New Method for Data Integration and Integrated Data Interpretation: Self-Organising Maps. In "Proceedings of Exploration 07: Fifth Decennial International Conference on Mineral Exploration" edited by B. Milkereit, 2007, p. 907-910.

Fraser, S.J., and Hodgkinson, J.H., 2008: Data Mining Geoscientific Data Using Self Organizing Maps: A New Approach to Geological Data Analysis. (CSIRO Exploration and Mining Report P2008/1839). In, Proceedings of the 44th Brazilian Geological Congress; October 21-27, 2008, Curitiba, Brazil. pp 679.

Jackaman, W., 2006: Regional drainage sediment and water geochemical data, Anahim Lake and Nechako River, central British Columbia (NTS 93C & 93F); Geoscience BC, Report 2006-4, 463 p.

Jackaman, W., 2007: Geoscience BC Mountain Pine Beetle Data Repository Version 1.0. July, 2007. Geoscience BC Report 2007-9.

Jackaman, W., 2008: QUEST Project sample reanalysis; Geoscience BC, Report 2008-3, 4 p.

Jackaman, W., 2008: Regional Lake Sediment and Water Geochemical Data Northern Fraser Basin, Central British Columbia (Parts of NTS 93G, H, J, K, N & O). Geoscience BC Report 2008-5.

Jackaman, W., 2008: Regional stream sediment and water geochemical data, Pine Pass (NTS 093O), British Columbia; Geoscience BC, Report 2008-7, 262 p.

Jackaman, W. and Balfour, J.S., 2008: QUEST Project geochemistry: field surveys and data reanalysis, central British Columbia (parts of NTS 093A, B, G, H, J, K, N, O); in Geoscience BC Summary of Activities 2007, Geoscience BC, Report 2008-1, p. 7–10.

Kaski, S, 1997: Data exploration using self-organizing maps; Acta Polytechnica Scandinavica, Mathematics, Computing and Management in Engineering Series No. 82, Espoo 1997, 57 pp. Published by the Finnish Academy of Technology.

Kohonen, T, 2001: Self-Organizing Maps. Third Extended Edition, Springer Series in Information Sciences, Vol. 30, Springer, Berlin, Heidelberg, New York, 2001.

Lett, R.E.W., 2005: Regional Geochemical Survey Database on CD, BC Ministry of Energy, Mines and Petroleum Resources, Geofile 2005-17.

Lett, R.E.W., and Bluemel, B., 2006: Re-analysis of regional geochemical survey stream sediment samples from the McLeod Lake area (NTS map sheet 093J); BC Ministry of Energy, Mines and Petroleum Resources, Geofile 2006-09, 220 p.

Mikula, P.A., Fraser, S.J., and Lee, M.F., 2008: Self Organizing Map Analysis of Seismicity Associated with Mining at Mount Charlotte Mine. (CSIRO Exploration and Mining Report P2008/1843). In Proceedings of the First Southern Hemisphere International Rock Mechanics Symposium (SHRIMS) Editors Potvin, Y., Carter, J., Dyskin, A., and Jeffery, R..; September 16-19, 2008, Perth, Australia. Volume 1, pp 653-666.

Penn, S.B., 2002: Using self-organizing maps, histograms, and standard deviation to detect anomalies in hyperspectral imagery data. In: Proceedings of the Fifth International Airborne Remote Sensing Conference. 22–24, Miami, Florida.

Penn, S.B., 2004: Using self-organizing maps to visualize high-dimensional data. Computers & Geosciences, 31(5), 531-544.

Scheffe, H., 1953: A Method for Judging All Contrasts in the Analysis of Variance: Biometrika, v. 40, no. 1-2, p. 87-104.

Selim, S.Z., and Ismail, M.A., 1984: K-Means-Type Algorithms - a Generalized Convergence Theorem and Characterization of Local Optimality: IEEE Transactions on Pattern Analysis and Machine Intelligence, v. 6, no. 1, p. 81-87.

Sliwa, R., Fraser, S.J., and Dickson, B.L., 2003: Application of self-organizing maps to the recognition of specific lithologies from borehole geophysics. In Proceedings of the 35th Sydney Basin Symposium on "Advances in the study of the Sydney Basin" Editors: Hutton, AC, Jones, BG, Carr, PF, Ackermann, B and Switzer, AD, September 29-30, 2003, University of Wollongong, Pages:105-113.

Strecker, U., and Uden, R., 2002: Data mining of poststack seismic attribute volumes using Kohonen self-organizing maps. The Leading Edge, October 2002, pp1032-1036.

Vesanto J., 2000: Using SOM in Data Mining. Licentiate's thesis at the Helsinki University of Technology, Booklet, April, 2000.

Xu, R., and Wunsch, D., 2005: Survey of clustering algorithms: IEEE Transactions on Neural Networks, v. 16, no. 3, p. 645-678.

Zhou, B., Fraser, S., Borsaru, M., Aizawa, T., Sliwa, R., and Hashimoto, T., 2005: New approaches for rock strength estimation from geophysical logs. In Beeston, J.W., (Editor): Bowen Basin Symposium 2005 — The future for coal — Fuel for thought. pp151-164; Geological Society of Australia Inc. Coal Geology Group and the Bowen Basin Geologists Group, Yeppoon, October 2005, 1–6.

APPENDIX 1: Descriptions of the Self-Organized Map Co-Ordinate & BMU Numbering Systems

Introduction:

There is often a requirement to refer to or describe the locations of particular "best matching unit" (BMU) nodes on a self-organized map (SOM). In particular the SOM ASCII output files (described in Appendix 2) refer to a SOM X and a SOM Y co-ordinate system and a BMU number. It is critical to understand these references as patterns of interest on the map may include BMUs that occur on adjacent rows and columns and hence will have non-sequential BMU numbers.

The Self-Organized Map Co-ordinate System:

The origin (1, 1) of the SOM co-ordinate system is a location upper-left on the map, with the SOM Y value increasing downwards, and SOM X increasing to the right. Figure 1 displays this system for the "honey-comb" of best matching unit nodes that are shown as an overlay to the U-matrix resulting from the analysis in this report.

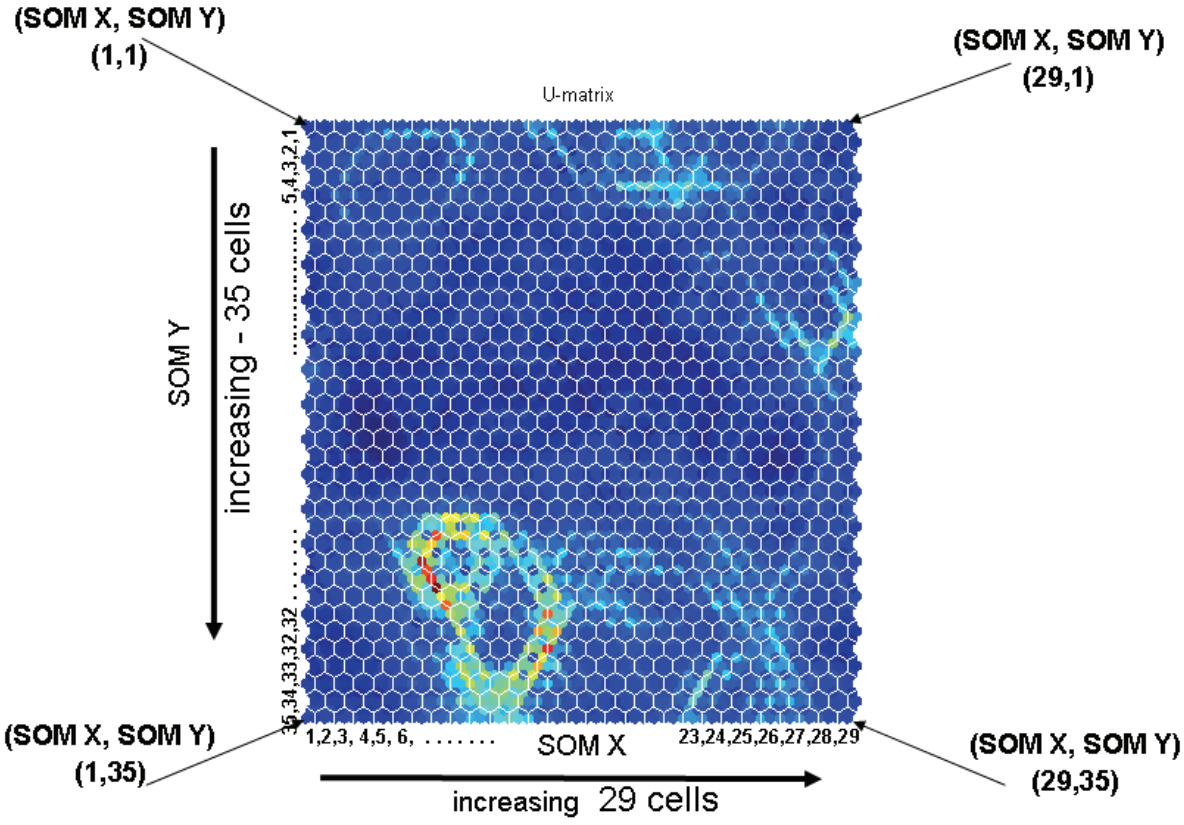


Figure 1: The Self-Organized Map Co-ordinate System for the 35x29 map from this report.

The Self-Organized Map BMU Node Numbering Convention:

The top-left hand BMU node (or cell) is referred to as #1 and the convention is that the BMU numbering then increases down the first column till the last BMU of that column is reached. The top-most BMU of the next column to the right of the first is then assigned the next cell number and the pattern is repeated till all BMUs are numbered. Figure 2 displays this system for the "honey-comb" of best matching unit nodes that are shown as an overlay to the U-matrix resulting from the analysis in this report.

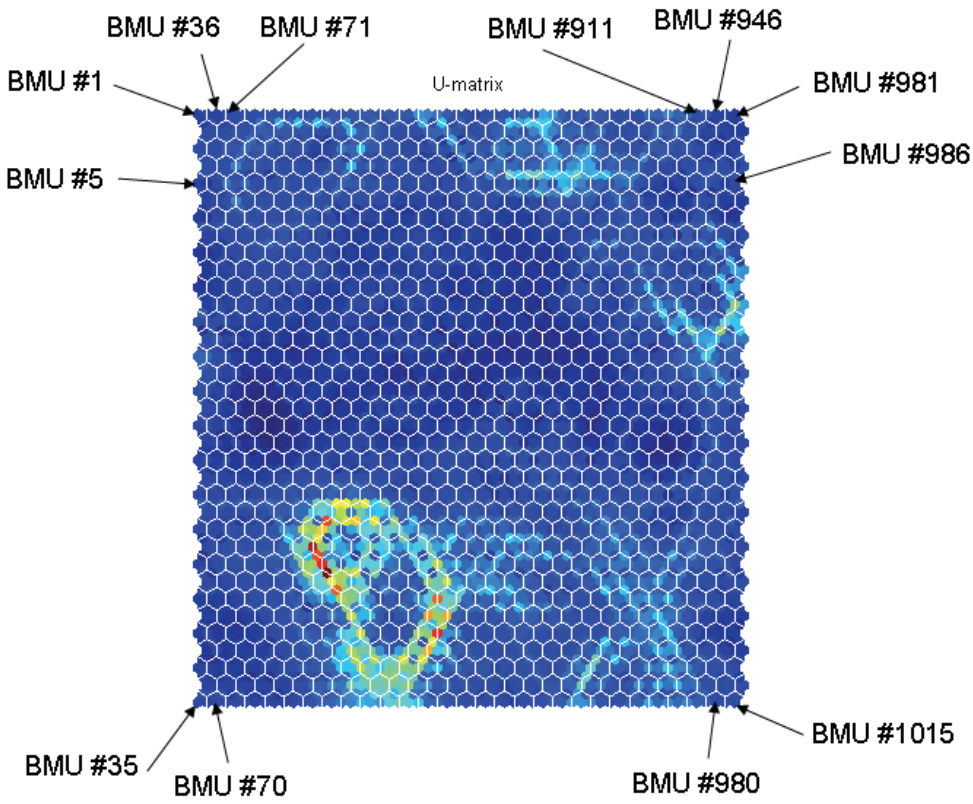


Figure 2: The Self-Organized Map Numbering System for the 35x29 map from this report.

APPENDIX 2: Descriptions of the SiroSOM ASCII Output Files

Introduction:

There are two standard ASCII output files from a SiroSOM analysis. The "*_out_map" file contains all the information that is required to recreate the self-organized map display and the component plot displays. The "*_out_data" file is essentially the input data but in addition each sample has been tagged with extra SOM-derived attributes for that sample. The two files can be used to visualize self-organized map or to further analyse or visualize the spatial distribution of samples that have been tagged with SOM-derived attributes.

The "*_out_map" SiroSOM Output File:

This file essentially describes the self-organized map itself. Columns refer to particular attributes of each BMU (also called a node or cell) and the BMU values are arranged in rows. The variable headers in the "*_out_map" output file for this study (QUEST_out_map.csv) are as follows:

> SOMx, SOMy, cluster, BMU, Ag_lg, Al_lg, As_lg, Au_lg, Ba_lg, Bi_lg, Br_lg, Ca_lg, Cd_lg, Ce_lg, Co_lg, Cr_lg, Cs_lg, Cu_lg, Fe_lg, Ga_lg, Hf_lg, Hg_lg, K_lg, La_lg, Lu_lg, Mg_lg, Mn_lg, Mo_lg, Na_lg, Ni_lg, P_lg, Pb_lg, Rb_lg, S_lg, Sb_lg, Sc_lg, Se_lg, Sm_lg, Sr_lg, Tb_lg, Th_lg, Ti_lg, Tl_lg, U_lg, V_lg, Zn_lg, red, green, blue, U-MAT.
>
> **"SOMx"** and **"SOMy"** are the SOM co-ordinates of the BMU (node/cell) in question.
> **"Cluster"** is the K-means cluster number assigned to the BMU in question.
> **"BMU"** refers to the BMU node number.
> **"Ag_lg" …… "Zn_lg"** are the SOM-determined elemental attributes of the BMU in question (in log10). The measurement units are given in Table 1 of the report proper.
> The **"red", "green"** and **"blue"** variables are the 8-bit colour components used to display a BMU on the map - typically the "cluster-colour".
> **"U-MAT"** is the average "distance" a BMU is from its neighbours.

The "*_out_data" SiroSOM Output File:

As indicated, this file contains the input sample data with additional SiroSOM tags added to each sample. Variable information is contained in columns and sample information in rows. The variable headers in the "*_out_data" output file for this study (QUEST_out_data.csv) are as follows:

> Easting, Northing, somx, somy, BMU, Cluster, q-error, UTME_83, UTMN_83, Ag_lg, Al_lg, As_lg, Au_lg, Ba_lg, Bi_lg, Br_lg, Ca_lg, Cd_lg, Ce_lg, Co_lg, Cr_lg, Cs_lg, Cu_lg, Fe_lg, Ga_lg, Hf_lg, Hg_lg, K_lg, La_lg, Lu_lg, Mg_lg, Mn_lg, Mo_lg, Na_lg, Ni_lg, P_lg, Pb_lg, Rb_lg, S_lg, Sb_lg, Sc_lg, Se_lg, Sm_lg, Sr_lg, Tb_lg, Th_lg, Ti_lg, Tl_lg, U_lg, V_lg, Zn_lg, L_MID (excluded)
>
> **"Easting"** and "**Northing"** refer to SiroSOM assigned geo-spatial location of the sample.

The **"somx"** and **"somy"** refer to the SOM_X and SOM_Y co-ordinated s of the BMU the sample has been assigned to.
"BMU" is the BMU number that a sample has been assigned to.
"Cluster" is the cluster number assigned to that sample after the K-means analysis of the BMUs.
The **"q-error"** is the quantization error assigned to a sample.
"UTME_83" and **"UTMN_83**" are the input geo-spatial sample locations.
"Ag_lg ……..Zn_lg" are the (log10) input sample values for each sample. The measurement units are given in Table 1 of the report proper.
"L_MID (excluded)" refers to a "label" attribute that was carried through the analysis but excluded from the actual analysis.

These data can be used as input to 3rd party data analysis packages or to display the SiroSOM results spatially in geographic information or display systems.

APPENDIX 3: Listing of Digital Data to Accompany this Report

Report Digital ASCII Data

Section 3

QUEST_out_data.csv
QUEST_out_map.csv

Section 4.2

QUEST_out_data_Cluster_Normalized.csv

Section 4.3.1

Au vs Ag Box_QUEST_out_data.csv

Section 4.3.2

Cu vs Ni CuBox QUEST_out_data.csv
Cu vs Ni NiBox QUEST_out_data.csv

Section 4.3.3

Cu vs Mo Box QUEST_out_data.csv

Section 4.4.1

AU_1_6_ QUEST_out_data.csv
AU_1_ QUEST_out_data.csv
AU_2_ QUEST_out_data.csv
AU_3_ QUEST_out_data.csv
AU_4_ QUEST_out_data.csv
AU_5_ QUEST_out_data.csv
AU_6_ QUEST_out_data.csv
Top30_ Au _Samples.csv

Section 4.4.2

Cu Nodes QUEST_out_data.csv
Top 60 Cu Samples QUEST_out_data.csv

Section 4.4.3

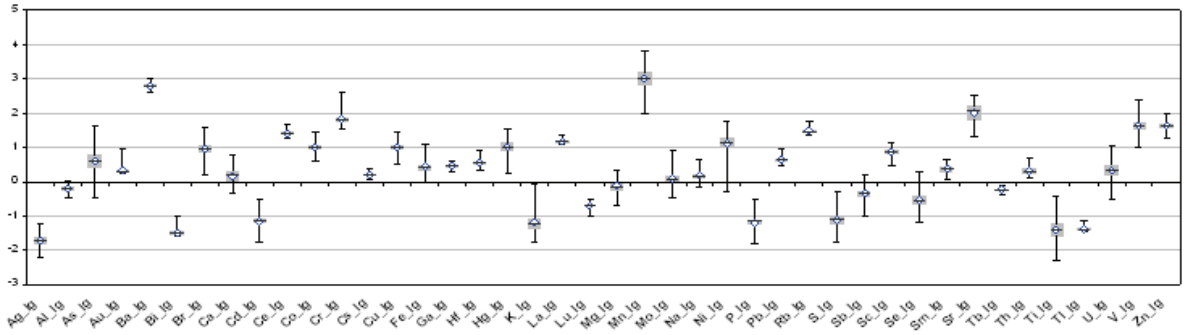
Mo Nodes QUEST_out_data.csv
Top 60 Mo Nodes QUEST_out_data.csv

Report Maps

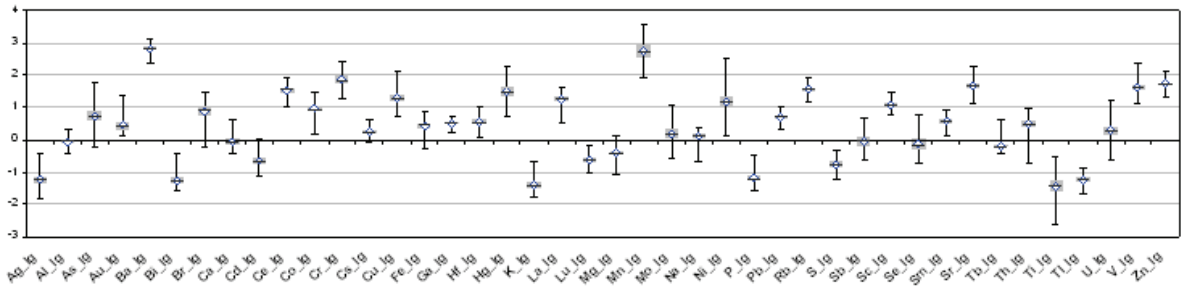
Quest_CSIRO - Fig 1 Location.pdf
Quest_CSIRO - Fig 4_Kmeans.pdf
Quest_CSIRO - Fig 5_QError.pdf
Quest_CSIRO - Fig 6_GoldAnom.pdf
Quest_CSIRO - Fig 8_AuVSAg.pdf
Quest_CSIRO - Fig 10_CuVSNi.pdf
Quest_CSIRO - Fig 12_CuVSMo.pdf
Quest_CSIRO - Fig 14a_Au_1.pdf
Quest_CSIRO - Fig 14b_Au_2.pdf
Quest_CSIRO - Fig 14c_Au_3.pdf
Quest_CSIRO - Fig 14d_Au_4.pdf
Quest_CSIRO - Fig 14e_Au_5.pdf
Quest_CSIRO - Fig 14f_Au_6.pdf
Quest_CSIRO - Fig 15_Top30Au.pdf
Quest_CSIRO - Fig 17_Cu.pdf
Quest_CSIRO - Fig 18_Top60Cu.pdf
Quest_CSIRO - Fig 20_Mo.pdf
Quest_CSIRO - Fig 21_Top60Mo.pdf

Geoscience BC QUEST Project Maps - Geochemistry - CSIRO SOM Analysis	
(available for download from the Geoscience BC web: http://www.geosciencebc.com)	
Field Sites Gridded and Coloured by K-means Clustering of their SOM-Derived Nodes	
GBCMAP 2009-14- 01	CSIRO SOM ANALYSIS: Sample Sites Coloured by SiroSOM-Derived K-means Clustering
Field Sites Coloured by the Magnitude of their Quantization-Error	
GBCMAP 2009-14- 02	CSIRO SOM ANALYSIS: Quantization Error - Global Anomalism of Geochemical Samples
Field Sites Coloured by the Magnitude of their "Cluster Normalized" value for each Element	
GBCMAP 2009-14- 03	CSIRO SOM ANALYSIS: "Cluster-Normalized " Silver Anomalism
GBCMAP 2009-14- 04	CSIRO SOM ANALYSIS: "Cluster-Normalized " Aluminum Anomalism
GBCMAP 2009-14- 05	CSIRO SOM ANALYSIS: "Cluster-Normalized " Arsenic Anomalism
GBCMAP 2009-14- 06	CSIRO SOM ANALYSIS: "Cluster-Normalized " Gold Anomalism
GBCMAP 2009-14- 07	CSIRO SOM ANALYSIS: "Cluster-Normalized " Barium Anomalism
GBCMAP 2009-14- 08	CSIRO SOM ANALYSIS: "Cluster-Normalized " Bismuth Anomalism
GBCMAP 2009-14- 09	CSIRO SOM ANALYSIS: "Cluster-Normalized " Bromine Anomalism
GBCMAP 2009-14- 10	CSIRO SOM ANALYSIS: "Cluster-Normalized " Calcium Anomalism
GBCMAP 2009-14- 11	CSIRO SOM ANALYSIS: "Cluster-Normalized " Cadmium Anomalism
GBCMAP 2009-14- 12	CSIRO SOM ANALYSIS: "Cluster-Normalized " Cerium Anomalism
GBCMAP 2009-14- 13	CSIRO SOM ANALYSIS: "Cluster-Normalized " Cobalt Anomalism
GBCMAP 2009-14- 14	CSIRO SOM ANALYSIS: "Cluster-Normalized " Chromium Anomalism
GBCMAP 2009-14- 15	CSIRO SOM ANALYSIS: "Cluster-Normalized " Cesium Anomalism
GBCMAP 2009-14- 16	CSIRO SOM ANALYSIS: "Cluster-Normalized " Copper Anomalism
GBCMAP 2009-14- 17	CSIRO SOM ANALYSIS: "Cluster-Normalized " Iron Anomalism
GBCMAP 2009-14- 18	CSIRO SOM ANALYSIS: "Cluster-Normalized " Gallium Anomalism
GBCMAP 2009-14- 19	CSIRO SOM ANALYSIS: "Cluster-Normalized " Hafnium Anomalism
GBCMAP 2009-14- 20	CSIRO SOM ANALYSIS: "Cluster-Normalized " Mercury Anomalism
GBCMAP 2009-14- 21	CSIRO SOM ANALYSIS: "Cluster-Normalized " Potassium Anomalism
GBCMAP 2009-14- 22	CSIRO SOM ANALYSIS: "Cluster-Normalized " Lanthanum Anomalism
GBCMAP 2009-14- 23	CSIRO SOM ANALYSIS: "Cluster-Normalized " Lutetium Anomalism
GBCMAP 2009-14- 24	CSIRO SOM ANALYSIS: "Cluster-Normalized " Magnesium Anomalism
GBCMAP 2009-14- 25	CSIRO SOM ANALYSIS: "Cluster-Normalized " Manganese Anomalism
GBCMAP 2009-14- 26	CSIRO SOM ANALYSIS: "Cluster-Normalized " Molybdenum Anomalism
GBCMAP 2009-14- 27	CSIRO SOM ANALYSIS: "Cluster-Normalized " Sodium Anomalism
GBCMAP 2009-14- 28	CSIRO SOM ANALYSIS: "Cluster-Normalized " Nickel Anomalism
GBCMAP 2009-14- 29	CSIRO SOM ANALYSIS: "Cluster-Normalized " Phosphorous Anomalism
GBCMAP 2009-14- 30	CSIRO SOM ANALYSIS: "Cluster-Normalized " Lead Anomalism
GBCMAP 2009-14- 31	CSIRO SOM ANALYSIS: "Cluster-Normalized " Rubidium Anomalism
GBCMAP 2009-14- 32	CSIRO SOM ANALYSIS: "Cluster-Normalized " Sulphur Anomalism
GBCMAP 2009-14- 33	CSIRO SOM ANALYSIS: "Cluster-Normalized " Antimony Anomalism
GBCMAP 2009-14- 34	CSIRO SOM ANALYSIS: "Cluster-Normalized " Scandium Anomalism
GBCMAP 2009-14- 35	CSIRO SOM ANALYSIS: "Cluster-Normalized " Selenium Anomalism
GBCMAP 2009-14- 36	CSIRO SOM ANALYSIS: "Cluster-Normalized " Samarium Anomalism
GBCMAP 2009-14- 37	CSIRO SOM ANALYSIS: "Cluster-Normalized " Strontium Anomalism
GBCMAP 2009-14- 38	CSIRO SOM ANALYSIS: "Cluster-Normalized " Terbium Anomalism
GBCMAP 2009-14- 39	CSIRO SOM ANALYSIS: "Cluster-Normalized " Thorium Anomalism
GBCMAP 2009-14- 40	CSIRO SOM ANALYSIS: "Cluster-Normalized " Titanium Anomalism
GBCMAP 2009-14- 41	CSIRO SOM ANALYSIS: "Cluster-Normalized " Thallium Anomalism
GBCMAP 2009-14- 42	CSIRO SOM ANALYSIS: "Cluster-Normalized " Uranium Anomalism
GBCMAP 2009-14- 43	CSIRO SOM ANALYSIS: "Cluster-Normalized " Vanadium Anomalism
GBCMAP 2009-14- 44	CSIRO SOM ANALYSIS: "Cluster-Normalized " Zinc Anomalism

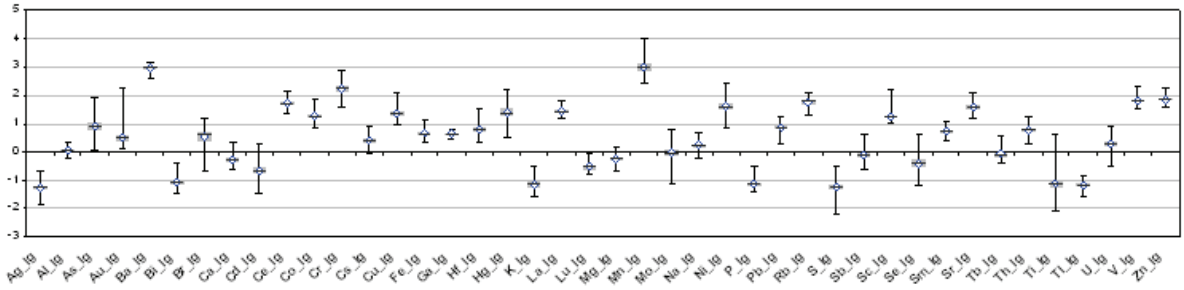
APPENDIX 4: Box and Whisker Plots of the 20 K-Means Classes Based on the Nodes of the Self Organized Map



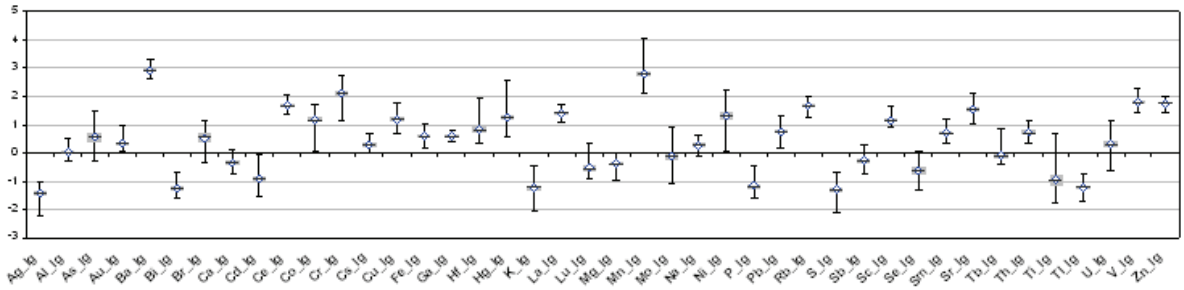
Cluster #1



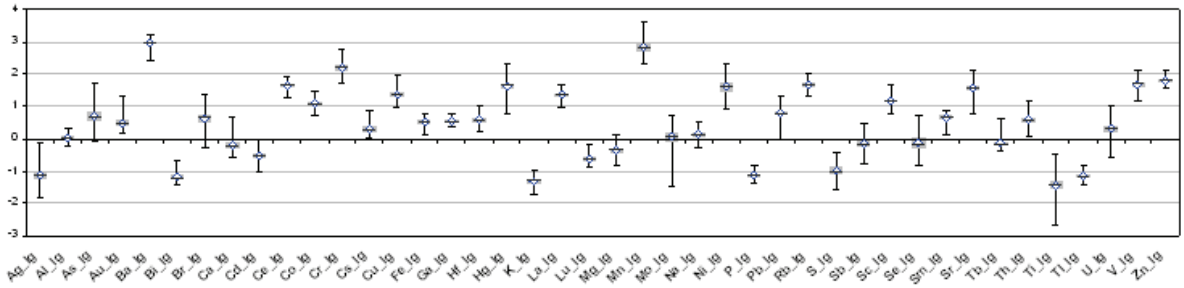
Cluster #2



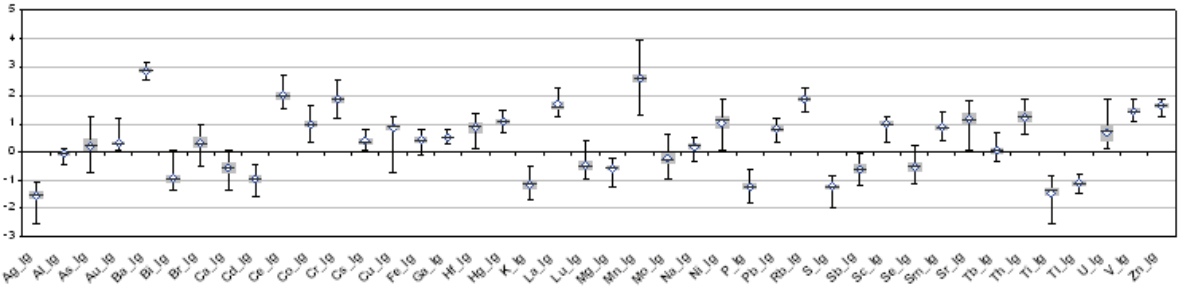
Cluster #3



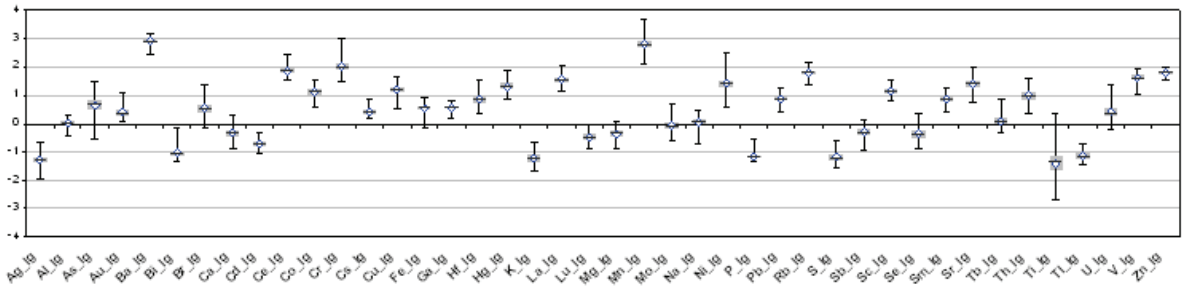
Cluster #4



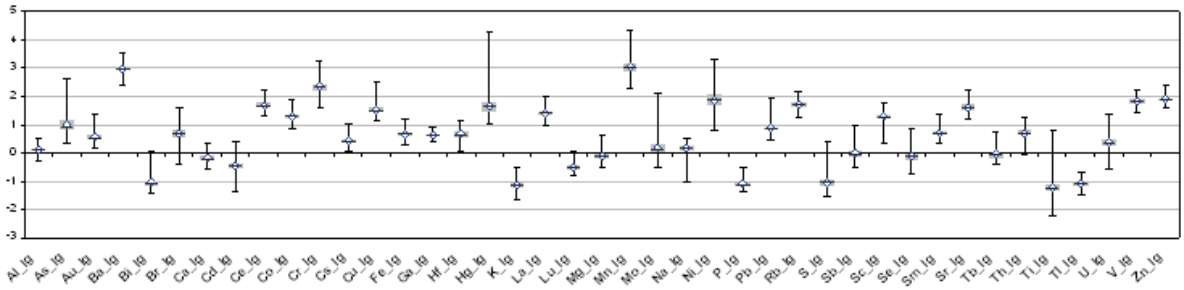
Cluster #5



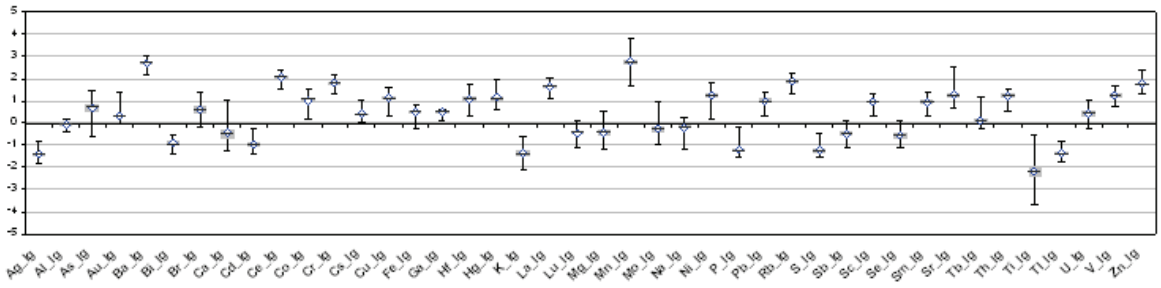
Cluster #6



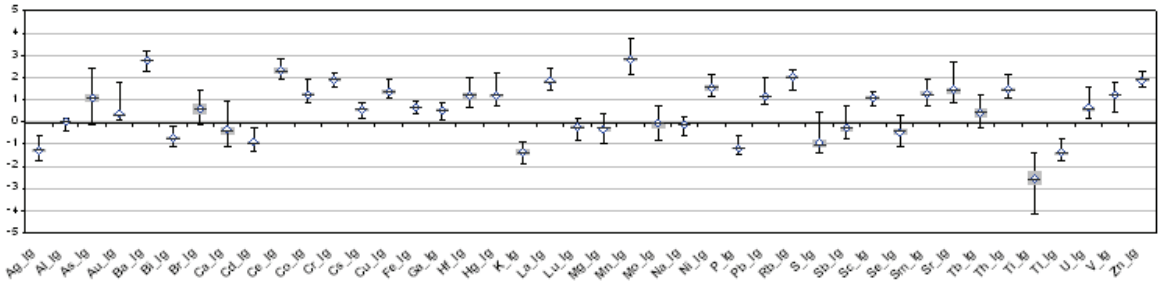
Cluster #7



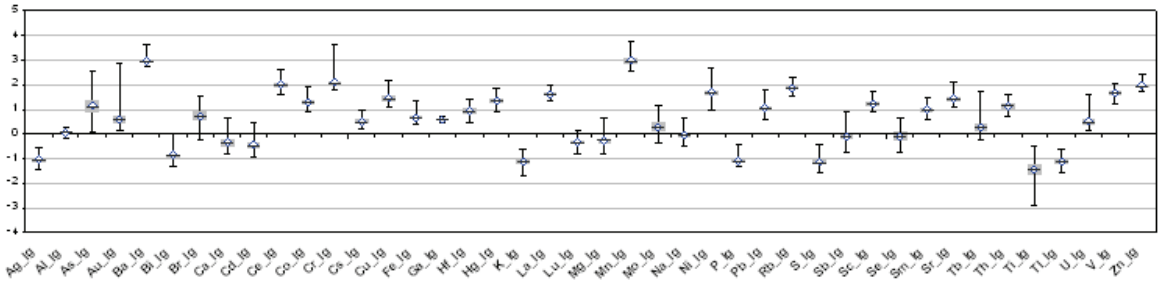
Cluster #8



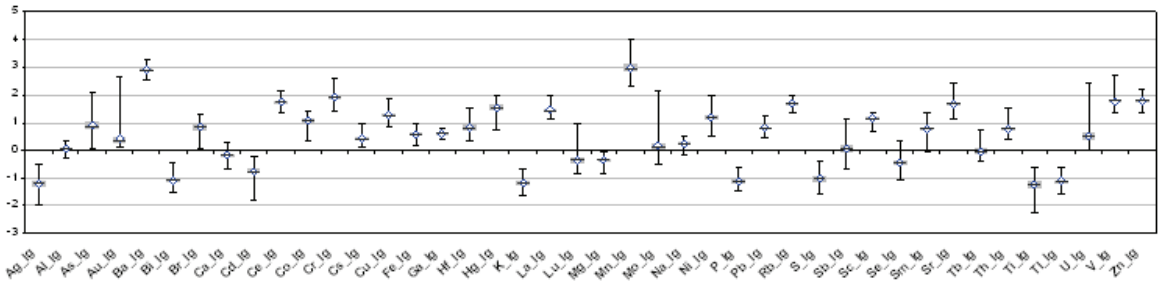
Cluster #9



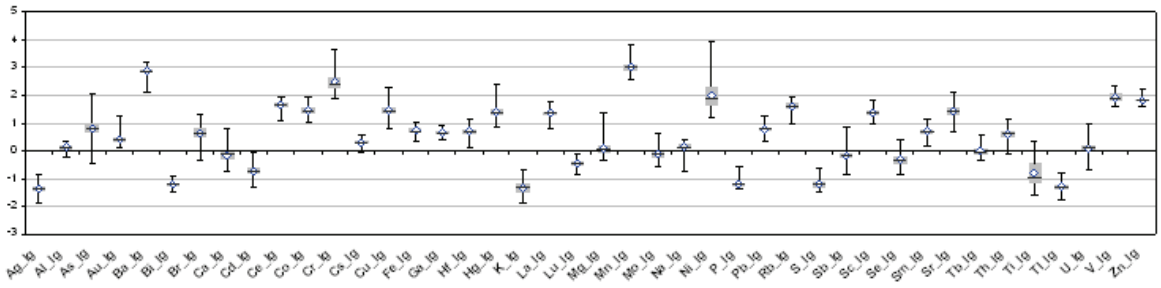
Cluster #10



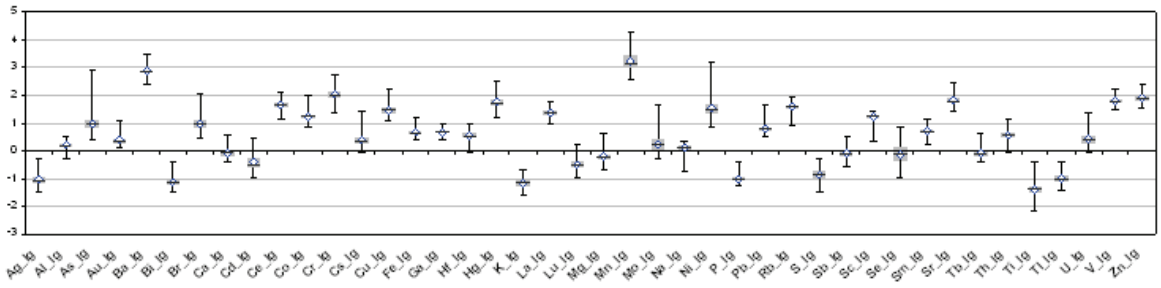
Cluster #11



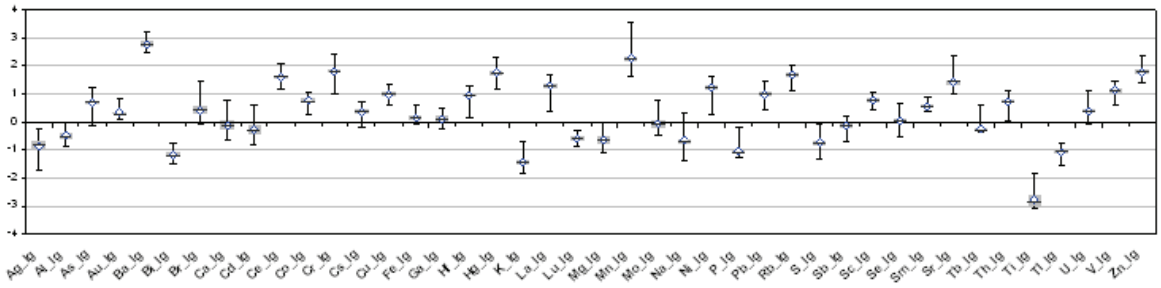
Cluster #12



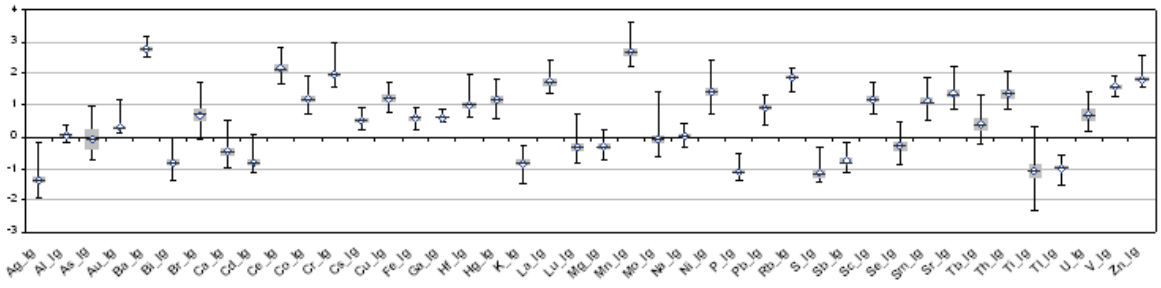
Cluster #13



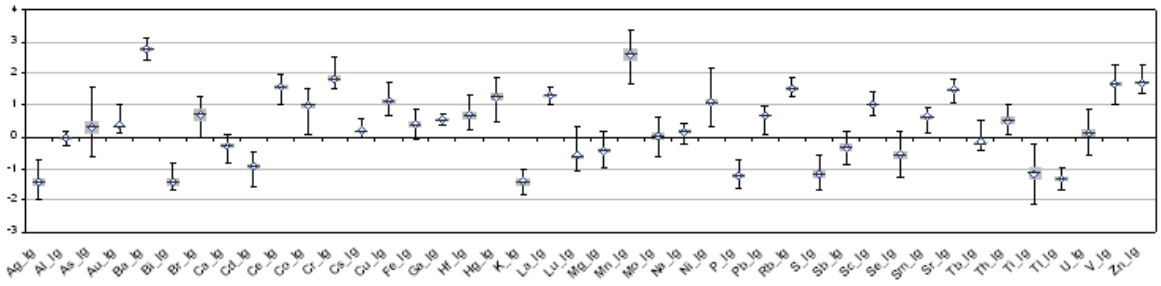
Cluster #14



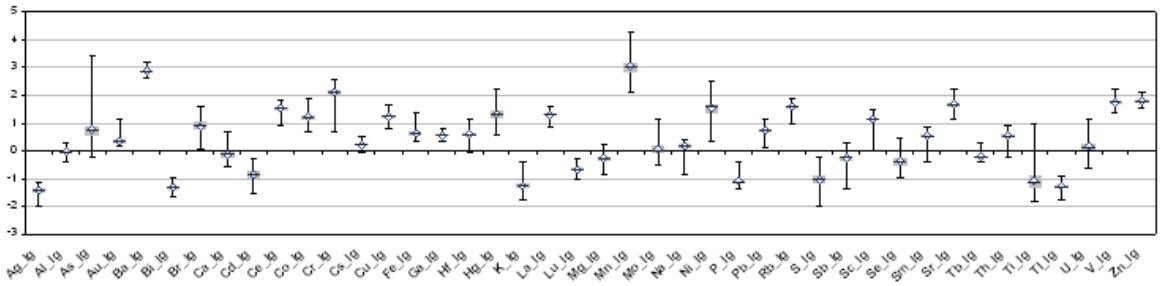
Cluster #15



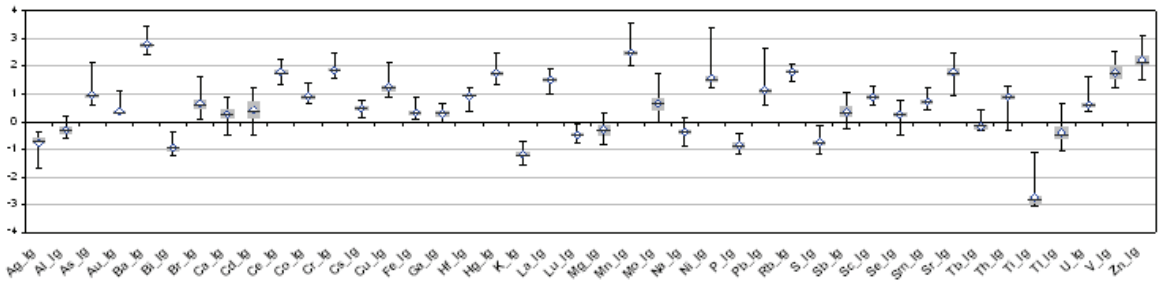
Cluster #16



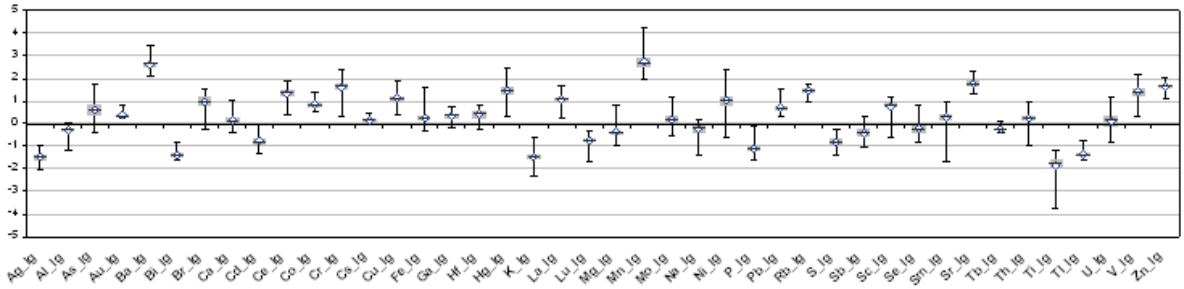
Cluster #17



Cluster #18



Cluster #19



Cluster #20

Contact Us
Phone: 1300 363 400
+61 3 9545 2176
Email: enquiries@csiro.au
Web: www.csiro.au

Your CSIRO

Australia is founding its future on science and innovation. Its national science agency, CSIRO, is a powerhouse of ideas, technologies and skills for building prosperity, growth, health and sustainability. It serves governments, industries, business and communities across the nation.