In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.
Abstract

This study was directed towards the development of rigorous, systematic, computer-assisted statistical procedures for the interpretation of quantitative and qualitative data commonly encountered in practical exploration-oriented surveys. A suite of data analysis tools were developed to evaluate the quality of geochemical data sets, to investigate the value and utilization of categorical field data, and to recognize and rank anomalous samples. Data obtained from regional stream sediment surveys as undertaken by the British Columbia Ministry of Energy, Mines and Petroleum Resources in southern British Columbia were examined as a case history.

A procedure based on a statistical analysis of field-site duplicates was developed to evaluate the quality of regional geochemical silt data. The technique determines: (1) whether differences in metal concentrations between sample sites reflect a real trend related to geological and geochemical features and not merely a consequence of sampling and analytical error, and (2) absolute precision estimates at any particular accumulation across a metal's concentration range. Results for metals Zn, Cu, Ni, Co, Fe and Mn indicated that combined variability due to local and procedural error averaged less than 5% of the total error and that precision estimates at the 95th percentile concentration value averaged less than 6.0%. Results presented indicate duplicates are more in accord with splits of individual samples (analytical duplicates) rather than separate field-site
duplicates. This type of systematic approach provides a basis for interpreting geochemical trends within the survey area, while simultaneously allowing evaluation of the method of sampling and laboratory analysis.

A procedure utilizing Duncan's Multiple Range Test examined the relationships between metal concentrations and class-interval and categorical observations of the drainage catchment, sample site and sediment sample. Results show that, many field observations can be systematically related to metal content of drainage sediments. Some elements are more susceptible than others to environmental factors and some factors influence few or many elements. For example, in sediments derived from granites there are significant relationships between bank type and concentration of 8 elements (Zn, Cu, Ni, Pb, Co, Fe, Mn and Hg). In contrast, the texture of these sediments, using estimates of fines contents as an index, did not significantly affect the concentration of any of the elements studied. In general, results indicate that groups of environmental factors acting collectively are more important than any single factor in determining background metal contents of drainage sediments.

A procedure utilizing both a graphical and multiple regression approach was developed to identify and characterize anomalous samples. The procedure determines multivariate models based on background metal values which are used to describe very general geochemical relations of no interest for prospecting purposes.
These models are then applied to sample subsets selected on the basis of factor/s known to strongly influence geochemical results. Individual samples are characterized after comparisons with relevant determined threshold levels and background multielement models. One hundred and fifteen anomalous samples for zinc from seven provenance groups draining 1259 sample sites were identified and characterized by this procedure. Forty three of these samples had zinc concentrations greater than its calculated provenance threshold, while 72 of these anomalous samples were identified solely because their individual metal associations were significantly different than their provenance multivariate background model. The method provides a means to reduce the effects of background variations while simultaneously identifying and characterizing anomalous samples.

The data analysis tools described here allow extraction of useful information from regional geochemical data, and as a result provide and effective means of defining problems of geological interest that warrant further investigation.
Table of Contents

Abstract .............................................................. ii
List of Tables .......................................................... vi
List of Figures .......................................................... vii
Acknowledgement ....................................................... viii

Chapter I
INTRODUCTION ......................................................... 1

Chapter II
STATISTICAL EVALUATION OF DUPLICATE SAMPLES, REGIONAL
SEDIMENT SURVEYS (MAP-AREA 92H, 92I AND 92J) ................. 7

Chapter III
STATISTICAL EVALUATION OF THE SIGNIFICANCE OF CATEGORICAL
FIELD PARAMETERS IN INTERPRETATION OF REGIONAL GEOCHEMICAL
SEDIMENT DATA (MAP-AREA 82F) ..................................... 30

Chapter IV
RECOGNITION AND RANKING PROCEDURE FOR MULTI-ELEMENT
REGIONAL STREAM SEDIMENT SURVEYS (MAP-AREA 82F) ............ 63

Chapter V
CONCLUSIONS .......................................................... 88

BIBLIOGRAPHY .......................................................... 92
List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Analysis of Variance</td>
<td>18</td>
</tr>
<tr>
<td>2.2</td>
<td>Test for systematic bias</td>
<td>18</td>
</tr>
<tr>
<td>2.3</td>
<td>Analysis of Variance Results for 132 duplicate pairs of stream sediments</td>
<td>22</td>
</tr>
<tr>
<td>2.4</td>
<td>Relative precision estimates, Thompson-Howarth method</td>
<td>23</td>
</tr>
<tr>
<td>3.1</td>
<td>Field-recorded categorical variables</td>
<td>39</td>
</tr>
<tr>
<td>3.2</td>
<td>Spatial density of stream sediment samples</td>
<td>39</td>
</tr>
<tr>
<td>3.3</td>
<td>5% level Duncan's Multiple Range Test of three unequally replicated means and calculation of D-ratio</td>
<td>45</td>
</tr>
<tr>
<td>3.4</td>
<td>D-Ratio calculated for categorical field variables from sediments draining granitic provenances</td>
<td>55</td>
</tr>
<tr>
<td>3.5</td>
<td>D-Ratio calculated for categorical field variables from sediments draining quartzitic provenances</td>
<td>55</td>
</tr>
<tr>
<td>3.6</td>
<td>Specific categories of individual categorical field variables with highest means, Granitic provenance</td>
<td>57</td>
</tr>
<tr>
<td>3.7</td>
<td>Specific categories of individual categorical field variables with highest means, Quartzitic provenance</td>
<td>57</td>
</tr>
<tr>
<td>3.8</td>
<td>Specific categories of individual categorical field variables with lowest means, Granitic provenance</td>
<td>58</td>
</tr>
<tr>
<td>3.9</td>
<td>Specific categories of individual categorical field variables with lowest means, Quartzitic provenance</td>
<td>58</td>
</tr>
<tr>
<td>3.10</td>
<td>Nonparametric correlation matrix, Granitic provenance</td>
<td>59</td>
</tr>
<tr>
<td>4.1</td>
<td>Examples of multivariate regression background models for zinc</td>
<td>77</td>
</tr>
<tr>
<td>4.2</td>
<td>Summary statistics for multivariate background zinc models</td>
<td>78</td>
</tr>
<tr>
<td>4.3</td>
<td>Part of a table listing anomalous samples in order of decreased ranking code</td>
<td>80</td>
</tr>
</tbody>
</table>
List of Figures

1.1 Location map of study area 82F, 92H, 92I and 92J ......4
2.1 Location map of study area 92H, 92I and 92J ..........13
2.2 Absolute differences of paired data versus corresponding mean values of pairs, copper in stream sediments ......24
2.3 Linear model of average error as a function of concentration, copper in stream sediments ..........25
3.1 Location map of study area 82F .........................35
3.2 Simplified geology of the study area 82F ............37
3.3 Duncan's Multiple Range Test for the influence of content of stream sediments associated with a granitic provenance .........................46
3.4 Duncan's Multiple Range Test for the influence of sediment colour of stream sediments associated with a Granitic provenance .........................47
4.1 Location map of study area 82F .........................68
4.2 Sample ranking in relation to fields on a plot of observed value versus value calculated from a multivariate model ........................................74
4.3 Plot of an area of anomalous samples from study area 82F ........................................81
4.4 Observed versus calculated zinc values for provenance group Argillite; calculated values based on a model determined from all samples with background zinc values ........................................83
4.5 Observed versus calculated zinc values for provenance group Argillite; calculated values based on a model determined from those samples with zinc background values that are also not anomalous in any other element ........................................84
Acknowledgement

Special thanks are extended to Drs. A.J. Sinclair and W.K. Fletcher for providing guidance and encouragement throughout this study. The author would like to thank the British Columbia Ministry of Energy, Mines and Petroleum Resources for providing financial support. I am especially grateful to Asger Bentzen for his valuable assistance with the computer work. A special mention goes to Kaarina Talvila for her personal encouragement.
I. INTRODUCTION
The completion of large multi-element geochemical surveys leads to the accumulation of enormous amounts of data. It is probable that only the most obvious features of the data can be recognized by visual and manual interpretation of such complex and voluminous data and that subtle though significant information has remained undetected. Consequently, a premium is placed upon computer processing by mathematical and statistical procedures that provide a more detailed interpretation of geochemical data, specifically, those techniques which can evaluate quality of geochemical data sets, identify regional trends in the data, and more importantly, distinguish variations related to bedrock mineralization from secondary environment influences. However, it must be stressed that results from these techniques must be reviewed critically in terms of their geochemical implications before acceptance and incorporation into an interpretation.

The studies described here are concerned with the development of computer-based statistical procedures for extracting useful information from regional geochemical data, and as a result, provide an effective means of defining problems of geological interest that warrant further investigation. More formally, this study was directed towards the development of rigorous, systematic, computer-assisted, interpretative
techniques for the evaluation both of quantitative and categorical data of the types commonly encountered in practical exploration-oriented surveys. A suite of data analysis tools with appropriate presentation techniques were developed to:

(a) evaluate the quality of geochemical survey data
(b) investigate the value and utilization of categorical field data
(c) recognize and rank anomalous samples

Data obtained from regional geochemical surveys undertaken by the British Columbia Ministry of Energy, Mines and Petroleum Resources (1981) and the Department of Energy, Mines and Resources, Geological Survey of Canada were (1977) were examined. These data consist of analyses of stream sediments and water samples collected at an average density of one per 17 square kilometers over NTS mapsheets 82F, 92H, 92I and 92J (Figure 1.1). Twelve hundred samples were collected on average per mapsheet. Semi-quantitative and qualitative class interval and categorical observations on characteristics of the drainage catchement, sample site and sediment sample were also recorded.

Samples were field dried and the minus 80 mesh (177 microns) fraction was retained for subsequent analyses. The samples were analyzed for Zn, Cu, Pb, Ni, Co, Ag, Mn, Fe, Mo, As and Sb by atomic absorption. Tungsten was determined colorimetrically. Field site duplicates were reported to be taken at a density of one per block of 20 samples and could not
Fig. 1.1 Location map of study area 82F, 92H, 92I and 92J, and index map of regional geochemical surveys available in British Columbia (McMillan, 1982)
be distinguished from other silt samples by the contracted commercial laboratory. These data are available publicly on magnetic tape from the British Columbia Ministry of Energy, Mines and Petroleum Resources, Geological Branch, Mineral Resources Division, Victoria. Due to the voluminous nature of this type of data, computer usage is essential if the data are to be compiled in a timely and efficient manner. As a result, each procedure investigated was computer-based, and incorporated a variety of relatively simple, commonly used univariate and multivariate techniques.

This thesis is organized in the form of a series of papers. Three interpretative procedures are examined and each one forms a subsequent chapter in this thesis. Individual procedures have been reviewed previously and preliminary versions have been published (Matysek et al, 1981; Matysek et al, 1982; Matysek et al, 1983; and Matysek and Sinclair, 1984). A brief description of each follows.

An improved method of examining quality of regional geochemical survey data is presented in Chapter Two. The method is based on field-site duplicates and incorporates a bias test, followed by an analysis of variance technique and the Thompson and Howarth (1976) approach to quantifying precision. This type of systematic procedure provides a basis for interpreting geochemical trends within the sampling area, while simultaneously allowing evaluation of the method of sampling and laboratory analysis.
A procedure utilizing Duncan's multiple range test to evaluate the significance of commonly collected categorical field observations is the subject of Chapter Three. The procedure examines the relationships between metal concentrations and categorical observations of the drainage catchment, sample site and sediment sample.

Chapter Four presents a multivariate interpretative technique that combines both a graphical and statistical approach to the identification, characterization and significance of anomalous samples. The procedure involves developing multivariate models based on background metal values and applying them to sample subsets selected on the basis of factors known to strongly influence geochemical results.

General conclusions are summarized in Chapter Five.
II. STATISTICAL EVALUATION OF DUPLICATE SAMPLES, REGIONAL SEDIMENT SURVEYS (MAP-AREA 92H, 92I AND 92J), BRITISH COLUMBIA
ABSTRACT

A systematic computer-assisted method of evaluating the quality of geochemical survey data based on field site duplicates is described. The procedure incorporates a bias test, an analysis of variance technique and the Thompson and Howarth (1976) approach to quantifying precision. The technique combines both a graphical and statistical approach to determine whether differences in metal concentrations between sample sites reflect a real trend related to geological and geochemical features and not merely a consequence of sampling and analytical error, and estimation of absolute precision as a function of metal concentration.

Duplicate data obtained from regional sediment surveys undertaken by the British Columbia Ministry of Energy, Mines and Petroleum Resources in southern British Columbia were evaluated by the proposed method. Analysis of 132 pairs of field-site duplicate silt samples indicated that: (1) Errors for metals Zn, Cu, Ni, Co, Mn and Fe resulting from sampling, sample preparation and analysis were small (<5% within-site variance component) compared to the regional geochemical variation and, (2) Precision estimates calculated by the Thompson and Howath method for the six metals examined averaged less than 6.5% at the 90th concentration value percentile, 6.0% at the 95th percentile and 5.5% at the 99th percentile.
Results presented here are more in accord with splits of individual samples (analytical duplicates) rather than analyses of two separate field samples. If this is the case, the loss of information on the variability associated with the sampling procedure and local environment seriously hampers the validity of these duplicate data in the recognition of significant geochemical trends. Since there appears to be some question over the type of duplicate analyzed, every effort should be made to ensure that field site duplicates are collected in future surveys. This type of systematic procedure utilizing field-site duplicates provides a basis for interpreting geochemical trends within the sampling area, while simultaneously allowing evaluation of the adequacy of the method of sampling and laboratory analysis.
INTRODUCTION

The ability to discriminate real trends related to geological and geochemical causes from those that result from spurious factors such as sampling and analytical errors is of considerable importance in the success of geochemical data interpretation. An estimate of reproducibility (precision) allows us to quantify the amount of variation due to sampling and analysis, and it is an integral part of the evaluation of geochemical data. Quality control procedures including duplicate samples should be conducted prior to carrying out any detailed data interpretation. Utilization of paired duplicates provide information on precision and aid in assessing:

1. Metal variability within and between sample sites.
2. Analytical precision over the range of values encountered.
3. Definition of realistic detection limits.
4. Sample homogeneity by comparing precision obtained on standards versus precision for paired duplicates at approximately the same concentration.

As part of our continuing study of rapid thorough evaluation procedures for multielement stream sediment data (for example Matysek et al., 1981, 1982 and 1983) we have designed a systematic, computer-oriented method of evaluating the quality of geochemical data based on field-site duplicates incorporating a bias test, an analysis of variance technique and the Thompson and Howarth (1976) approach to quantifying precision.

This detailed procedure is applicable to the type and
quality of data incorporated in various regional programs undertaken by the British Columbia Ministry of Energy, Mines and Petroleum Resources but can be adapted easily for other data types.

GENERAL METHODOLOGY

In brief, our general approach to evaluation of the quality of geochemical data sets involves the following steps:

1. Extraction of at least 50 independent, field-site duplicate pairs from a geochemical data set for use in subsequent data analysis.

2. Determination of the degree of systematic bias between duplicate pairs, based upon the number of cases in which the first observation is greater than the duplicate.

3. Evaluation of the duplicated geochemical data set in terms of metal variability at the regional (between-sample site) and at the local, sampling and analytical levels (within-sample site) by a two factor analysis of variance technique.

4. Quantification of the within-sample site variability by estimating the precision utilizing the Thompson and Howarth (1976) method.

DATA

Data obtained from three recent regional geochemical surveys undertaken by the British Columbia Ministry of Mines, Energy and Petroleum Resources (1981) were evaluated by the proposed method. The base data consist of analyses of stream sediments and water samples collected at an average density of one sample per 17.3 square kilometers over NTS mapsheet areas
92H, 92I and 92J (Figure 2.1). Silt samples were field dried and the minus 80 mesh (177 microns) fraction retained for subsequent analyses.

The samples were analyzed for zinc, copper, lead, nickel, cobalt, silver, manganese, iron, molybdenum, arsenic and antimony by atomic absorption. Tungsten was determined colorimetrically, uranium in water samples was determined by a fluorimetric method and fluoride in stream waters was determined using a specific ion electrode.

Field site duplicates were also collected to provide information on analytical precision over a range of concentrations and to give some impression of sample representivity or geologic variation. Altogether, 132 pairs of duplicate field site samples were extracted from the data set. Individual duplicate samples were recorded as first of field site duplicate and second of field site duplicate in the publically available general information guide describing the data base (British Columbia Ministry of Energy, Mines and Petroleum Resources, 1981). Duplicates were reported to have been collected at a density of one per block of 20 samples. Duplicate pairs within individual blocks were analyzed in batches and could not be distinguished from other silt samples by the analyst. These data are available publically on magnetic tape from the British Columbia Ministry of Energy, Mines and Petroleum Resources, Geological Branch, Mineral Resources Division, Victoria.

Sampling and analytical error associated with zinc, copper,
Fig. 2.1 Location map of study area 92H, 92I, 92J, and index map of regional geochemical surveys available in British Columbia (McMillan, 1982).
nickel, cobalt, manganese and iron were investigated by this study. This particular suite of metals was selected on the following basis:

1. Their distributions approximated the Gaussian curve.

2. Majority of reported values were above their published detection limits.

ASSUMPTIONS

The assumption that analytical and sampling variation closely follows the Gaussian distribution underlies most of this discussion, and it is necessary to examine the assumption and to be aware of situations where it may not apply. The mathematical conception of the Gaussian curve is that it results from the combination of a large number of small independent errors. In practice, these small errors can be regarded as resulting from variations in manipulative operations such as sampling, sub-sampling, weighing, dissolution, etc. Despite the fact that some of these contributions may not be small, they are usually independent and the resulting distribution is not usually distinguishable from a random sample drawn from a Gaussian population.

Non-Gaussian forms are apparent in practice when two particular conditions are met: one or more contributing factors is no longer relatively small (i.e., it makes a substantial contribution to the total variation); and its particular frequency distribution is non-Gaussian. The uncritical use of
statistics calculated on metals characterized by non-Gaussian forms is likely to lead to erroneous conclusions, especially when the variability of measurement is high relative to the actual metal abundance. Non-Gaussian metal distributions are recognized when the following conditions arise:

1. Metal abundances are dependent on rare grains
2. Precision of the method of analysis is poor
3. Concentration levels are near the detection limit
4. Data contains outliers

BIAS TEST

A characteristic of numerical measurement is inconsistency in repeated measurements of the same quantity. Two types of error contribute to the unreliability of a measurement: random errors arising from the variations inherent in any sampling or measurement process and non-random errors causing systematic negative or positive deviations from the true results.

A bias test was utilized to assess the degree of systematic non-random error between the determined metal contents of the first sample and its duplicate. The test is based upon the number of instances in which the first observation is greater than the duplicate. If there is no bias this number (m) should be close to one half the total number of pairs (n), and the frequency distribution of possible results should correspond with successive terms of the binomial expansion of:
\[
\left[ \left( \frac{1}{2} \right) + \left( \frac{1}{2} \right) \right]^n
\]

The probability of obtaining a particular deviation from \( n/2 \) or greater can be calculated exactly. If the number of duplicate pairs exceeds 50, its frequency distribution approximates that of the normal distribution, having mean and standard deviation as \( n/2 \) and \( n\sqrt{2} \) respectively. The observed incidence \((m)\) is converted into the standardized normal deviate. The probability of obtaining a result of \((m)\) or greater can be obtained from the usual tables for the areas in the tails of the normal distribution.

ANALYSIS OF VARIANCE

The second stage in our procedure is to evaluate the geochemical data set in terms of variability at regional (between-sample sites) and at local levels (within-sample site). Only if a significant proportion of the data variability is at the regional level can one be confident that differences in metal concentrations between-sample sites reflect a real trend related to geological and geochemical features and not merely a consequence of sampling or analytical error.

Significance of metal variability within-sample sites due to sampling and analytical errors versus dispersion between-sample sites can be determined by standard analysis of variance techniques. For our purposes, the significance of the various sources of variation can be determined from logarithmic values.
of duplicate samples using a two factor analysis of variance. Theory and assumptions inherent in this method are described by Krumbein and Graybill (1965) and Koch and Link (1971). The formal design for the analysis of variance is given in Table 2.1.

The significance of metal variability between-sample sites (geochemical variation) and metal variability within-sample sites (sampling and analytical errors) is determined from the F-statistic. The value of F which is significant can be obtained from statistical tables and is a function of the number of sites at which duplicates are collected, the number of duplicates collected at each site and the significance level selected for the investigation. Relative variance components have also been calculated as described by Garrett (1969, 1973) and correspond to the average percentage of variability explained by each source at a sample site.

THOMPSON AND HOWARTH PRECISION METHOD

The final stage in our evaluation of the quality of geochemical data sets is to quantify the amount of variation due to sampling and analytical error. This variation can be expressed in terms of precision which, in geochemical practice, is specified as the percent relative variation at the two standard deviation (95%) confidence level:

\[ P_c = \frac{2S_c}{c} \times 100\% \]
Table 2.1 Analysis of Variance

<table>
<thead>
<tr>
<th>SOURCE OF VARIATION</th>
<th>DEGREE OF FREEDOM</th>
<th>SUM OF SQUARES</th>
<th>MEAN SQUARE</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Error (temping, sub-sampling, analysis)</td>
<td>$i - 1$</td>
<td>$m \sum (\bar{z}_i - \bar{z})^2$</td>
<td>$S_i^2 = \frac{m}{i - 1} \sum (\bar{z}_i - \bar{z})^2$</td>
<td>$F_i = S_i^2 / S_S^2$</td>
</tr>
<tr>
<td>Geochemical variation</td>
<td>$m - 1$</td>
<td>$\sum \sum (z_{ij} - \bar{z}_i)^2$</td>
<td>$S_i^2 = \frac{1}{m - 1} \sum (\bar{z}_i)$</td>
<td>$F_i = S_i^2 / S_S^2$</td>
</tr>
<tr>
<td>Residual</td>
<td>$V - 11(m - 1)$</td>
<td>$\sum \sum (z_{ij} - \bar{z}_i - \bar{z}_j)^2$</td>
<td>$S_i^2 = \frac{1}{V - 11(m - 1)} \sum \sum (\bar{z}_{ij} - \bar{z}_i - \bar{z}_j)^2$</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>$N - 1$</td>
<td>$\sum \sum (z_{ij} - \bar{z})^2$</td>
<td>$S^2 = \frac{1}{N - 1} \sum \sum (\bar{z}_{ij} - \bar{z})^2$</td>
<td></td>
</tr>
</tbody>
</table>

$z_{ij}$ = result for the $ith$ replicate from the $jth$ site; $i = 1, 2, ..., ith$ result $V$ is usually 21;

$j = 1, 2, ..., mth$ result; $N = lm$ = total number of results; $\bar{z}_i$ = mean of $ith$ replicate group;

$\bar{z}_j$ = mean of $jth$ site group; and $\bar{z}$ = overall mean.

Table 2.2 Test for systematic bias

<table>
<thead>
<tr>
<th>METAL</th>
<th>NO DIFFERENCES</th>
<th>POSITIVE DIFFERENCES</th>
<th>NEGATIVE DIFFERENCES</th>
<th>PROBABILITY OF OBTAINING NO. OF POSITIVE DIFFERENCES per cent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nickel</td>
<td>43 (32.6)</td>
<td>46</td>
<td>43</td>
<td>36</td>
</tr>
<tr>
<td>Arsenic</td>
<td>49 (37.1)</td>
<td>44</td>
<td>39</td>
<td>30</td>
</tr>
<tr>
<td>Cobalt</td>
<td>56 (42.4)</td>
<td>42</td>
<td>34</td>
<td>22</td>
</tr>
<tr>
<td>Copper</td>
<td>31 (28.5)</td>
<td>55</td>
<td>46</td>
<td>19</td>
</tr>
<tr>
<td>Manganese</td>
<td>14 (10.6)</td>
<td>64</td>
<td>54</td>
<td>17</td>
</tr>
<tr>
<td>Iron</td>
<td>31 (23.5)</td>
<td>64</td>
<td>37</td>
<td>&lt;1</td>
</tr>
<tr>
<td>Zinc</td>
<td>23 (17.4)</td>
<td>68</td>
<td>41</td>
<td>&lt;1</td>
</tr>
</tbody>
</table>
where \((Pc)\) is the precision in percent at concentration \((c)\) and 
\((Sc)\) is an estimate of the analytical standard deviation \((\sigma_c)\) 
at concentration \((c)\).

Application of analysis of variance techniques can only 
determine an average precision value for a range of 
concentrations. In actual fact, it has been shown (Thompson and 
Howarth, 1976) that where there is a wide range of 
concentrations in a set of samples both the absolute and 
relative errors in analytical determinations can vary across the 
range. To deal with this complexity, alternative ways of 
estimating precision using randomly selected duplicates, have 
been considered in detail by Thompson and Howarth 

Briefly, their preferred method involves dividing 50 or 
more duplicate samples into groups with narrow concentration 
ranges, and employing the median of absolute differences between 
pairs of duplicate analyses \(\left( x_i, x_i' \right)\) as an estimator 
of the standard deviation \(\sigma_c\). The group mean value of 
all the mean average values \(\left( \frac{x_i + x_i'}{2} \right)\) is 
used as an estimator of the average concentration. If this 
procedure is repeated for a number of successive narrow 
concentration ranges a set of corresponding mean concentration 
and standard deviation estimates is obtained. The relationship 
between them can be found by simple linear regression. From the 
expression:

\[
Sc = S_o + Kc \quad [2]
\]
through substitution of [1] above, Pc is given by:

\[ Pc = 200 \left( \frac{S_0}{c + K} \right) \]  \[ \text{(3)} \]

where \((S_0)\) is the standard deviation at zero concentration and \((K)\) is a constant. This linear function has been determined in many practical cases to be a satisfactory model for the expression of the variation.

The following rapid procedure is suggested for estimation of precision from a minimum of 50 pairs of duplicate samples (Thompson and Howarth, 1976):

1. From the duplicate analyses obtain a list of the means and absolute differences.

2. Arrange the list in increasing order of concentration means.

3. From the first 11 results obtain the mean concentration and median difference from that group.

4. Repeat step (3) for each successive group of 11 results, ignoring any remainder less than 11.

5. Calculate or obtain graphically, the linear regression of the median differences on the means and multiply the intercept and coefficient by 1.048 (factor derived from the properties of the half-normal distribution) to obtain \((S_0)\) and \((K)\) respectively.

In order to assess the significance of the precision parameters obtained from the linear regression, the calculated slope and intercept were evaluated individually by a t-test. Significance of the resulting regression was also determined by an analysis of variance.
RESULTS

Table 2.2 illustrates the results obtained from the bias test. From this table we observe:

1. Incidence of positive differences in metal content between duplicate pairs is greater than the number of negative differences for each of the metals analyzed.

2. Metals nickel, cobalt, copper and manganese exhibit statistically significant, but minor bias.

3. A significant bias for zinc and iron is indicated by the extremely low probability ( < 1% ) of obtaining such a high proportion of positive differences in metal contents between duplicate pairs.

4. The percentage of duplicate pairs exhibiting identical metal concentrations is about 25 percent.

These results suggest that for the majority of metals, no major systematic biases are present and application of techniques such as analysis of variance and Howarth and Thompson precision procedure should provide meaningful results.

The results of the analysis of variance are presented in Table 2.3 As expected, the between-stream variability is decidedly higher than the within-stream dispersion for all of the metals studied. This feature is encouraging because the purpose of these regional surveys is to define a regional trend related to geological and geochemical phenomena; the greater the variability in metal concentrations between-sample sites the greater the ease of defining possible trends.

Results obtained by other investigations (Bolviken and Sinding-Larsen, 1973 and Chork, 1977 using similar type data
Table 2.3 Analysis of Variance Results  
For 132 duplicate pairs of stream sediments

<table>
<thead>
<tr>
<th>METAL</th>
<th>GRAND MEAN (GEOMETRIC)</th>
<th>BETWEEN SITES</th>
<th>WITHIN SITES</th>
<th>VARIANCE COMPONENTS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$F_1$</td>
<td>$F_2$</td>
<td>BETWEEN SITES</td>
</tr>
<tr>
<td>Zinc</td>
<td>50</td>
<td>156.98</td>
<td>8.22</td>
<td>89</td>
</tr>
<tr>
<td>Copper</td>
<td>26</td>
<td>115.45</td>
<td>1.61 ns</td>
<td>97</td>
</tr>
<tr>
<td>Nickel</td>
<td>15</td>
<td>59.51</td>
<td>2.81 ns</td>
<td>91</td>
</tr>
<tr>
<td>Cobalt</td>
<td>8</td>
<td>40.70</td>
<td>0.20 ns</td>
<td>99</td>
</tr>
<tr>
<td>Manganese</td>
<td>292</td>
<td>276.88</td>
<td>0.26 ns</td>
<td>99</td>
</tr>
<tr>
<td>Iron</td>
<td>1.68</td>
<td>313.30</td>
<td>13.56</td>
<td>91</td>
</tr>
</tbody>
</table>

The test depends on the following null hypothesis:
(a) Between sample sites—mean metal contents are equal at each site. Degrees of freedom are $(m - 1, m - 1)$, where $m$ = number of sites = 132.
(b) Within sample sites—mean of samples x1, is equal to mean of duplicates x2. Degrees of freedom are $(1, m - 1)$.

All values are in ppm except for iron which is in per cent.

Probability ranges for accepting the null hypothesis:

- $0.01 > P$; $0.05 > P > 0.01$; $ns = P > 0.05$

Table 2.4 Relative precision estimates,  
Thompson-Howarth method.  
British Columbia regional survey data (92H, 92I and 92J).

<table>
<thead>
<tr>
<th>METAL</th>
<th>DETECTION LIMIT</th>
<th>MAXIMUM VALUE</th>
<th>SLOPE</th>
<th>INTERCEPT</th>
<th>REGRESSION F-VALUE</th>
<th>PRECISION ESTIMATES AT SELECTED PERCENTILES</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>90th</td>
</tr>
<tr>
<td>Zinc</td>
<td>2</td>
<td>210</td>
<td>0.027</td>
<td>0.892 ns</td>
<td>7.19</td>
<td>7.0</td>
</tr>
<tr>
<td>Copper</td>
<td>2</td>
<td>720</td>
<td>0.018</td>
<td>0.945</td>
<td>16.60</td>
<td>6.7</td>
</tr>
<tr>
<td>Nickel</td>
<td>2</td>
<td>1300</td>
<td>0.012</td>
<td>0.629</td>
<td>55.44</td>
<td>4.0</td>
</tr>
<tr>
<td>Cobalt</td>
<td>2</td>
<td>96</td>
<td>0.029</td>
<td>0.212 ns</td>
<td>3.14</td>
<td>8.0</td>
</tr>
<tr>
<td>Manganese</td>
<td>5</td>
<td>3700</td>
<td>0.036</td>
<td>2.785</td>
<td>40.72</td>
<td>7.8</td>
</tr>
<tr>
<td>Iron</td>
<td>0.02</td>
<td>5.70</td>
<td>0.011</td>
<td>0.051</td>
<td>2.57</td>
<td>5.0</td>
</tr>
</tbody>
</table>

All values are in ppm except iron which is in per cent.

Probability ranges for accepting $t$-tests on the slope and intercept and $F$-test on the regression:

- $0.01 > P$; $0.05 > P > 0.01$; $0.10 > P > 0.05$; $ns = P > 0.10$
and tests found that the variation within-sample sites was about 10-25% and the variation between-sample sites was about 75-90%. Such results are typical for low density stream sediment surveys and contrast markedly with result obtained from this study (Table 2.3). The B.C. Duplicate data exhibit surprisingly lower within-sample site variance, averaging less than 4% of the total variability. The significance of these anomalous results are elaborated upon in a later section.

Results obtained from the Thompson and Howarth method are presented in Table 2.4 and illustrated graphically for copper on Figures 2.2 and 2.3. Although, the regression of the median of absolute differences on the concentration means was only based on 12 singular points, analysis of variance on the regression proved significant at the 99% confidence interval for all metals except cobalt. A regression plot of copper illustrated in Figure 2.3 shows conclusively that the simple linear regression more than adequately accounts for the relationship between the median of absolute differences and mean concentrations; thus it provides an excellent indicator of precision over the concentration range. It would appear that the slope is the major influence on the precision estimate of a given metal, and second, its relative magnitude quantifies the precision estimate.

Examination of Table 2.4 reveals that nickel possesses the relatively lowest slope (0.012) whereas manganese has the highest (0.036). The intercept for most of the elements is not significantly different than zero, and the magnitude of the
Figure 2.2 Absolute differences of paired data versus corresponding mean values of pairs, copper in stream sediments, British Columbia Regional Geochemical Survey Data, (92H, 92I and 92J).
Figure 2.3 Linear model of average error as a function of concentration, copper in stream sediments. British Columbia Regional Geochemical Survey Data (92H, 92J and 92I).
slopes is extremely small, therefore the precision is high.

For example, Figures 2.2 and 2.3 illustrate for the metal copper (1) the plot of absolute differences between duplicate pairs against means of duplicate pairs, with the concentration range divided into equal frequency intervals; and (2) a plot of the regression of the median absolute difference against mean concentration of copper. To determine the precision as an absolute value obtain graphically the absolute difference corresponding to a selected concentration value on the regression line and multiply by 2. Thus, for copper at the:

- 50th percentile (26 ppm) absolute precision = 2.8 ppm
- 95th percentile (59 ppm) absolute precision = 3.9 ppm
- 99th percentile (80 ppm) absolute precision = 4.7 ppm

DISCUSSION OF RESULTS

In general, all data collected in stream sediment surveys contain errors that are acquired through sampling, laboratory analysis and data handling. Taking the existence of sampling errors into account a precision of 10 - 15% at the 95% confidence level is generally regarded as acceptable for laboratory variability in most exploration programmes (Fletcher, 1981). Studies tailored to the evaluation of error in stream sediment surveys such as Plant (1971), Howarth and Lowenstein (1971), Bolviken and Sinding-Larsen (1973), Plant et al. (1975) and Chork (1977) generally concluded:
1. Variable bias and variable precision introduced by secondary environment effects obscure the primary regional geochemical variation. The factors involved are complex and related to several variables in the primary and secondary environment of areas investigated.

2. Metal dispersion within sample sites depend on such factors such as:
(a) concentration of the metal under investigation
(b) concentration of other metals (eg. iron and manganese)
(c) homogeneity of sediment composition
(d) catchment size and drainage density of sampling site

3. The combined variability due to local variation and procedural error ranged from 10 - 25% of the total error.

4. Sampling errors tended to exceed analytical errors when precise analytical techniques such as atomic absorption spectrometry are used.

Precision estimates calculated for this study are significantly lower than those obtained from the other investigations cited. It seems highly unlikely that the magnitude of variability associated with the local environment, sampling and analytical techniques are adequately accounted for, especially, if we consider that:

1. A high percentage of individual duplicate pairs (Table 2.1, column 2) are characterized by identical metal contents occurring over the entire concentration range for all elements studied.

2. Relative variance components of within-sample site averaged less than five percent.

3. Precision estimates calculated by the Thompson-Howarth method for six different metals at different concentration levels averaged less than 6.5% at the 90th percentile concentration value, less than 6.0% at the 95th percentile and less than 5.5% at the 99th percentile.
Results presented are more in accord with splits of an individual sample (analytical duplicate) rather than analyses of two separate field samples. If this is the case, the loss of information on the variability associated with the sampling procedure and local environment may seriously hamper the interpretation of significant geochemical trends. Furthermore, the whole reason for these regional surveys is to facilitate provincial comparison, and if the data are not of similar precision for a particular element and consistent relative to some baseline level in terms of accuracy, such comparisons will be both difficult and unreliable. Since there appears to be some question over the type of duplicate analyzed, every effort should be made to ensure that proper field-site duplicates are collected in future surveys.

CONCLUSIONS

To make the most efficient use of geochemical analyses their reliability must be known. Quality control throughout sample collection, sample preparation and analysis should be regarded as an essential aspect of any geochemical prospecting method. The procedure outlined here combines both a graphical and statistical approach to the estimation of precision. The technique allows quantification of:

1. Significant systematic biases;

2. Relative within-sample site and between-sample site variance components;

3. Absolute precision at any particular
accumulation across its concentration range.

As a result, the method provides salient information for decisions on threshold values, the significance of anomalies and contour intervals.

This type of systematic procedure utilizing field duplicates provides a basis for interpreting geochemical trends within the study area, while simultaneously allowing evaluation of the adequacy of the method of sampling and laboratory analysis. This information can also be used to design more efficient sampling schemes to improve contrast in geochemical results.

ACKNOWLEDGEMENTS

The author would like to thank the British Columbia Ministry of Energy, Mines and Petroleum Resources for providing financial support.
III. STATISTICAL EVALUATION OF THE SIGNIFICANCE OF CATEGORICAL
FIELD PARAMETERS IN INTERPRETATION OF REGIONAL GEOCHEMICAL
SEDIMENT DATA (MAP-AREA 82F), BRITISH COLUMBIA
ABSTRACT

In an attempt to study the value and utilization of categorical data collected during regional stream sediment surveys, we have taken data from the Canadian Uranium Reconnaissance Programme in S.E. British Columbia. After initial classification of the data into six subsets on the basis of catchment geology, probability plots were constructed for each of 11 elements (Zn, Cu, Pb, Ni, Co, Fe, Mn, Mo, W, Hg and U) and used to select thresholds to reject anomalous samples. The remaining background populations then were successively subdivided into groups according to their classification with respect to four sediment characteristics (abundance of fines, sand, organic matter and sediment colour) and six environmental parameters (physiography, water flow rate, stream class, drainage pattern, bank type and contamination). After calculation of log means for each group and a pooled common standard error, differences between group means were tested for significance using Duncan's multiple range test. The relative degrees of confidence in the significance of difference between categorical means for any single field parameter were determined using a ratio method.

Results of Duncan's multiple range test, show that, many field observations can be related systematically to metal content of drainage sediments. Some elements are more susceptible than others to environmental factors and some factors influence few or many elements. For example, in
sediments derived from granites there are significant relationships between bank type and concentration of 8 elements (Zn, Cu, Ni, Pb, Co, Fe, Mn and Hg). In contrast, the texture of these sediments, using estimates of fines content as an index, did not significantly affect the concentration of any of the elements studied. In general, results indicate that groups of environmental factors acting collectively are more important than any single factor in determining background metal content of drainage sediments.
INTRODUCTION

During collection of drainage sediments for regional geochemical surveys, it has become customary to record class-interval and categorical observations on characteristics of the drainage catchment, sample site and sediment sample. However, both the subjective character, and qualitative and interdependent nature of these categorical observations has made it difficult to make quantitative evaluations of their significance with respect to elemental concentrations. Consequently, apart from classifying data in terms of geology and screening anomalous samples (Sinclair and Fletcher, 1980; Brummer et al., 1978), utilization of categorical information is often so limited or non-existent that the effort and cost involved in its acquisition become difficult to justify. This led Meyer et al., (1979) to suggest that serious consideration be given to reducing the number of observations to those of proven significance.

In an attempt to study the value and utilization of categorical field data collected during regional stream sediment surveys, data from the Joint Federal-Provincial Uranium Reconnaissance Program in south-eastern British Columbia (Geological Survey of Canada Open File Report OF514) have been evaluated (Matysek et al., 1981). Here we describe a systematic, computer-assisted method of investigating the significance of a non-parametric data utilizing Duncan's multiple range test (Duncan, 1955, 1957).
The more extreme ranges of the small populations of anomalous values complicate the probability density functions of geochemical variables and confuse the potentially different relation between categorical variables on the one hand and metal content on the other.

Our philosophy has been to extract recognizable anomalous information from the data sets we are studying in order to better understand controls on the great bulk of background analytical data normally acquired during large-scale regional stream sediment surveys.

DESCRIPTION OF STUDY AREA

The study area (Fig. 3.1), including approximately 14,500 km² of the Columbia Mountains in south-eastern British Columbia, is bounded by the 49th and 50th parallels of north latitude and by 116th and 118th meridians of west longitude (NTS map-sheet 82F). Topographically, the region is characterized by northerly trending mountain ranges rising in altitude from wooded hills (1800m to 2100m asl) in the south to rugged mountainous terrains (2400m to 2800m asl) in the north.

Prevailing westerly winds result in an average annual precipitation of about 750 mm and there is a well developed active drainage system over the entire region. Major rivers include the Columbia, Kootenay, Slocan, Moyie and South Mary. The mountain ranges, particularly to the west of Kootenay lake
Figure 3.1 Location map of study area 82F, and index map of regional geochemical surveys available in British Columbia (McMillan, 1982).
are dissected by the major east-west or northwest-southeast flowing tributaries of the main river systems.

The area is underlain by severely folded and faulted sedimentary and volcanic rocks, ranging from Purcell (late Precambrian) to Cretaceous age, intruded by Mesozoic plutons (Fig. 3.2). These plutons, which include granodiorites, granites, syenites and diorites, occupy much of the terrain in the Selkirk Mountains west of Kootenay Lake and enclose large gneissic roof pendants. East of Kootenay Lake the principal lithologies are Proterozoic quartzites of the Creston and Aldridge Formations. Less extensive, but nonetheless significant rock units, include slates, argillites and andesitic volcanic rocks all of which occur at several levels in the stratigraphic succession, (Rice (1941); Little (1960)).

Over much of the region, to an elevation of 2400m, bedrock units are covered by a thin mantle of glacial till with reworked fluvio-glacial deposits in the valleys. Above 2400m till is replaced by talus and felsenmeer.

The region has a long mineral exploration history with four important producing mining camps; Rossland (Au-Cu), Slocan (Ag-Pb-Zn), Ainsworth (Zn-Pb-Ag) and Sheep Creek (Au) (Fig. 3.2). The Slocan and Ainsworth camps alone account for 350 vein deposits found in the area (Sinclair, 1979).

METHODOLOGY

The data base for the study comprises field observations
Figure 3.2 Simplified geology of the study area, modified from Geological Survey of Canada Open File Report 514.
(Table 3.1) and analytical results (minus 80-mesh, Zn, Cu, Pb, Ni, Co, Mn, Fe, Mo, W, Hg, U) for drainage sediments collected at 1313 sites as part of the Joint Federal-Provincial Uranium Reconnaissance Program. Details of sampling and analytical procedures are to be found in Geological Survey of Canada Open File Report OF514. These data are available publically on magnetic tape from the British Columbia Ministry of Energy, Mines and Petroleum Resources, Geological Branch, Mineral Resources Division, Victoria, and the Federal Department of Energy, Mines and Resources, Ottawa.

For the purposes of this study, geology has been simplified by combining similar lithologies, irrespective of their age, into the six rocktypes - granite, quartzite, slate, argillite, andesite and gneiss - that characterize the catchments of the sediment samples (Fig. 3.2). Spatial density and number of samples for these units are summarized in Table 3.2.

Evaluation of the field observations then included the following steps:

1. Sorting of data into provenance groups, i.e., predominant rock type in drainage basin upstream from the sample.

2. Evaluation of simple statistics and probability graphs for each element.

3. Threshold selection using the method of Sinclair (1976) to isolate anomalous samples from background samples.

4. Subdivision of background populations into groups according to their classification with respect to sediment characteristics and environmental variables.

5. Calculation of log means and pooled standard error for each group: testing of
Table 3.1

Field-recorded categorical variables, regional stream-sediment survey, Geological Survey of Canada OF514 (NGR 25-1977)

<table>
<thead>
<tr>
<th>Sample variables (4):</th>
<th>Categories</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Fines Content</td>
<td>Absent, Minor (0—33%), Medium (34—66%), Major (&gt; 67%)</td>
</tr>
<tr>
<td>(2) Sand Content</td>
<td>Absent, Minor (0—33%), Medium (34—66%), Major (&gt; 67%)</td>
</tr>
<tr>
<td>(3) Organic Content</td>
<td>Absent, Minor (0—33%), Medium (34—66%), Major (&gt; 67%)</td>
</tr>
<tr>
<td>(4) Sediment Colour</td>
<td>Red, White, Black, Yellow, Green</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Environmental variables (6):</th>
<th>Categories</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Banktype</td>
<td>Alluvial, Colluvial, Glacial Till, Other</td>
</tr>
<tr>
<td>(2) Waterflow Rate</td>
<td>Slow, Moderate, Fast</td>
</tr>
<tr>
<td>(3) Stream Class</td>
<td>Primary, Secondary, Tertiary, Quaternary</td>
</tr>
<tr>
<td>(4) Physiography</td>
<td>Hilly, Mountainous Mature, Mountainous Youthful</td>
</tr>
<tr>
<td>(5) Drain Pattern</td>
<td>Poor, Dendritic, Herringbone, Other</td>
</tr>
<tr>
<td>(6) Contamination</td>
<td>None, Possible, Probable, Definite</td>
</tr>
</tbody>
</table>

Table 3.2

Spatial density of stream-sediment samples: Nelson map-area

<table>
<thead>
<tr>
<th>Lithology</th>
<th>Map unit</th>
<th>Area (km²)</th>
<th>No. of samples</th>
<th>Samples/km²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Granite</td>
<td>20</td>
<td>5 825</td>
<td>485</td>
<td>12.0</td>
</tr>
<tr>
<td>Quartzite</td>
<td>1</td>
<td>3 775</td>
<td>318</td>
<td>11.9</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>430</td>
<td>32</td>
<td>13.4</td>
</tr>
<tr>
<td></td>
<td>Combined</td>
<td>4 205</td>
<td>350</td>
<td>12.0</td>
</tr>
<tr>
<td>Slate</td>
<td>6</td>
<td>585</td>
<td>61</td>
<td>9.6</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>420</td>
<td>51</td>
<td>8.2</td>
</tr>
<tr>
<td></td>
<td>14</td>
<td>390</td>
<td>33</td>
<td>11.8</td>
</tr>
<tr>
<td></td>
<td>Combined</td>
<td>1 395</td>
<td>145</td>
<td>9.6</td>
</tr>
<tr>
<td>Argillite</td>
<td>2</td>
<td>630</td>
<td>58</td>
<td>10.9</td>
</tr>
<tr>
<td></td>
<td>17</td>
<td>420</td>
<td>31</td>
<td>13.5</td>
</tr>
<tr>
<td></td>
<td>Combined</td>
<td>1 050</td>
<td>89</td>
<td>11.8</td>
</tr>
<tr>
<td>Andesite</td>
<td>5</td>
<td>40</td>
<td>5</td>
<td>8.0</td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>50</td>
<td>6</td>
<td>8.3</td>
</tr>
<tr>
<td></td>
<td>18</td>
<td>495</td>
<td>62</td>
<td>8.0</td>
</tr>
<tr>
<td></td>
<td>Combined</td>
<td>585</td>
<td>73</td>
<td>8.0</td>
</tr>
<tr>
<td>Gneiss</td>
<td>13</td>
<td>660</td>
<td>73</td>
<td>9.0</td>
</tr>
<tr>
<td>Totals</td>
<td></td>
<td>13 720</td>
<td>1 215</td>
<td>11.3</td>
</tr>
</tbody>
</table>
difference between means using Duncan's Multiple Range Test.

6. Determination of relative degree of confidence and/or consistency in the significance of the differences between group means using a ratio method.

7. Results of (5) visually presented in Venn diagram form.

This procedure has been computerized to allow rapid systematic evaluation of field parameters, and utilizes programs by Le (1970) and the senior author.

Analysis of histograms and probability plots indicates that individual metal frequency distributions for each rock type are generally both lognormal and bimodal in form. Data were therefore partitioned using log-probability plots into low (probably background) and high (probably anomalous) populations (Sinclair, 1976). Anomalous samples were then rejected leaving only background samples for classification into groups based on the field observations. These were then further subdivided into groups according to their classification with respect to four sediment characteristics (fines, sand, organic content and colour) and six environmental variables (physiography, waterflow rate, stream class, bank types, drainage pattern and contamination) as described in Table 3.1.

Log means and a pooled standard error from an analysis of variance on these categorical means for each field variable were calculated. The significance of differences between the means of all groups were then tested with Duncan's multiple range test (Duncan, 1955, 1957). This test determines if differences among
means of several groups are significant at a given level of confidence. The test has been applied previously to regional geochemical data by Miesch (1976) and Doyle and Fletcher (1979).

The test assumes that the means \( x_1, x_2, \ldots, x \) are drawn independently from ‘n’ normal populations having true means \( m_1, m_2, \ldots, m \) respectively. The conventional statistic for the comparison of two means is:

\[
t = \frac{(\bar{x}_1 - \bar{x}_2)}{S \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}
\]

(1)

where \( \bar{x}_1 \) and \( \bar{x}_2 \) are the estimates of the two means from the populations with equal variance, estimated by \( S^2 \bar{x} \) and are based on \( n_1 \) and \( n_2 \) independent measurements of variables \( x_1 \) and \( x_2 \), respectively.

For a test at the 95 per cent confidence level \( (\alpha=0.05) \) the computed \( t \) is compared with the tabled values of Student's \( t \) distribution \( (1-\alpha/2, v) \) where \( v \) is the degrees of freedom on which the estimate of \( S \) is based (equal to \( n_1 + n_2 - 2 \)). If \( n_1=n_2=n \), and the critical value of \( T(1-\alpha/2) \) is used for \( t \), equation 1 then becomes:

\[
X_1 - x_2 = Sr = 2 \sqrt{tSp}
\]

(2)

\( Sr \) is the shortest significant studentized range for the two means, and \( Sp \) is the pooled standard error common to \( x_1 \) and \( x_2 \). The quantity \( 2 \sqrt{tSp} \) is referred to as the shortest significant studentized range range by Duncan (1955).

For \( \alpha=0.05 \) and say \( v=60 \), the significant studentized range is 2.83. Thus, differences of at least 2.83 times the standard
error of the means are required in order to declare the means as significantly different.

Duncan (1957) extended this method by calculating significant studentized ranges (zp) when more than two means (p>2) are compared and variances are not equal. The means are first ordered by magnitude then if two adjacent means are being compared, p=2; if the two means compared are separated by one mean, p=3, and so on. The significant studentized range increases with an increase in p. For example where p=4 with v=60 the value is 3.08, compared with 2.83 where p=2 (a=0.05). If all the means are equally replicated, the significant studentized ranges (tsp) are multiplied by the pooled standard error, and the means are ranked and divided into homogeneous groups. More commonly, if the means are unequally replicated, shortest significant studentized ranges are calculated in the same manner. These must be compared with an adjusted difference (AD) between the p means which takes into consideration the number of observations each mean is based upon (n₁ and n₂ respectively):

\[ AD = (x₁ - x₂) * \sqrt{2n₁n₂/(n₁+n₂)} \]  

Any subset of p means is homogeneous if the largest adjusted difference in the subset fails to exceed the critical value Sr. Two means not in the same homogeneous subset are significantly different.

The relative degree of confidence in the significance of differences between categorical means for any single field
parameter can be determined by comparing the adjusted difference between the highest categorical mean \( (x_h) \) and the lowest categorical mean \( (x_l) \) with their shortest significant studentized range \( (S_r) \).

\[
D = \frac{(x_h - x_l) \sqrt{(2n_1 * n_2)/(n_1 + n_2)}}{S_r}
\]

A ratio less than 1 indicates that the differences between means are just significant at the 95\% confidence level. Greater confidence in the significance of the difference between means is reflected by an increased ratio. Comparison of the D-ratio thus provides a guide to the relative strength or consistency of the influence of the various field parameters on trace element composition.

For example, considering zinc in sediments associated with granites, 43 samples (from a total of 485) are rejected as anomalous after interpretation of the zinc probability graph. If sediment colour is the field observation of interest, the remaining 442 sediments can be subdivided into five (red, white, yellow, black, and grey) groups, and log means and an analysis of variance on the means calculated. However, the only colours recorded with reasonable frequency \( (n>10) \) are red \( (n=311) \), white \( (n=105) \) and black \( (n=26) \) with corresponding means of 63, 49 and 70 ppm. Table 3 illustrates the steps involved in application of Duncan's multiple range test and D-ratio method to these three categorical means.

In this case the test establishes that concentrations of zinc in black and red sediments are indistinguishable but those
in white sediments are significantly lower. The results can be presented as in Table 3.3 and below, where any two means not appearing within the same parentheses are significantly different, or as Venn diagrams in which overlapping parts of circles indicate groups whose means are not significantly different (Fig. 3.3 and 3.4).

RESULTS

The significance of differences of means among categorical field observations were tested and presented as Venn diagrams as shown in Figs. 3.3 and 3.4. Although, the procedure was tested on six different drainage basin lithologies, only results from sediments associated with streams draining granite and quartzite terrains will be presented in detail.

Bulk composition - Fines content

Categories: 2=minor(0-33%) 3=medium(34-66%) 4=major(>67%)  
Granite: (no significant differences)  
Quartzite: Zn,Cu,Ni,Co,Fe,Mn (23)(31)  

Variation in the content of fines in sediments derived from granites has no apparent influence on any of the eleven metal concentrations studied, whereas for quartzites concentrations of zinc, copper, nickel, cobalt, iron and manganese are significantly affected. However, there is a considerable degree
Table 3.3

5% level Duncan's Multiple Range Test of three unequally replicated means and calculation of D-ratio

(1) Ranked means and replication numbers

<table>
<thead>
<tr>
<th>Sediment colour</th>
<th>White</th>
<th>Red</th>
<th>Black</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log mean</td>
<td>1.69</td>
<td>1.80</td>
<td>1.845</td>
</tr>
<tr>
<td>Geometric mean</td>
<td>49</td>
<td>63</td>
<td>70</td>
</tr>
<tr>
<td>Number</td>
<td>(105)</td>
<td>(311)</td>
<td>(26)</td>
</tr>
</tbody>
</table>

(2) Analysis of variance

<table>
<thead>
<tr>
<th>Source</th>
<th>Degrees of Freedom</th>
<th>Mean Square</th>
<th>F-ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between means</td>
<td>2</td>
<td>0.50781</td>
<td>13.815</td>
</tr>
<tr>
<td>Error</td>
<td>439</td>
<td>0.03676</td>
<td></td>
</tr>
</tbody>
</table>

Pooled standard error \( (S_p) = \sqrt{0.03676} = 0.1917 \), F-crit = 4.567

(3) Critical values

\[ S_r = S_p z_p \]

\[ p: \quad (2) \quad (3) \]

\[ z_p: \quad 2.7853 \quad 2.9290 \]

\[ S_r: \quad 0.5340 \quad 0.5616 \]

(4) Calculated adjusted differences \( (AD) \) between \( p \) means

Black-White = \( (1.845 - 1.69) \sqrt{2(26)(105)/(26 + 105)} = 1.000 \)

Black-Red = \( (1.845 - 1.80) \sqrt{2(26)(311)/(26 + 311)} = 0.312 \)

Red-White = \( (1.80 - 1.69) \sqrt{2(105)(311)/(105 + 311)} = 1.378 \)

(5) Test sequence: comparing \( S_r \) to calculated adjusted difference

Black-White = 1.000 > \( S_r = (0.5616) \) means significantly different

Black-Red = 0.312 < \( S_r = (0.5340) \) means not significantly different

Red-White = 1.378 > \( S_r = (0.5340) \) means significantly different

(6) Result

(White) (Red, Black)

Any two means not appearing together within the same parentheses are significantly different. Any two means appearing together within the same parenthesis are not different.

(7) D-ratio: ratio comparing the adjusted difference \( (AD) \) between the highest and lowest means with their critical value \( S_r \)

\[ D = 1.000/0.5616 = 1.78 \]
Figure 3.3 Duncan's multiple range test for the influence of organic content of stream sediments associated with a granitic provenance. Common or overlapping circles indicate that group means are not significantly different at the 0.05 confidence level.

**VARIABLES**
1 = ABSENT
2 = MINOR (0-34%)
3 = MEDIUM (34-67%)

ALL ELEMENTS ARE IN ppm. EXCEPT: Hg ppb. Fe %
<table>
<thead>
<tr>
<th>Element</th>
<th>Category</th>
<th>Group 1</th>
<th>Group 2</th>
<th>Group 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zn</td>
<td>Pb</td>
<td>Cu</td>
<td>Zn</td>
<td>Pb</td>
</tr>
<tr>
<td>X-</td>
<td>X-</td>
<td>X-</td>
<td>N-</td>
<td>N-</td>
</tr>
<tr>
<td>50</td>
<td>9</td>
<td>8</td>
<td>105</td>
<td>102</td>
</tr>
<tr>
<td>64</td>
<td>12</td>
<td>311</td>
<td>311</td>
<td>302</td>
</tr>
<tr>
<td>70</td>
<td>18</td>
<td>26</td>
<td>26</td>
<td>15</td>
</tr>
<tr>
<td>Ni</td>
<td>Co</td>
<td>Hg</td>
<td>Ni</td>
<td>Co</td>
</tr>
<tr>
<td>X-</td>
<td>X-</td>
<td>X-</td>
<td>N-</td>
<td>N-</td>
</tr>
<tr>
<td>8</td>
<td>5</td>
<td>14</td>
<td>336</td>
<td>336</td>
</tr>
<tr>
<td>8</td>
<td>5</td>
<td>21</td>
<td>109</td>
<td>23</td>
</tr>
<tr>
<td>13</td>
<td>5</td>
<td>31</td>
<td>23</td>
<td>22</td>
</tr>
<tr>
<td>Fe</td>
<td>Mn</td>
<td>U</td>
<td>Fe</td>
<td>Mn</td>
</tr>
<tr>
<td>X-</td>
<td>X-</td>
<td>X-</td>
<td>N-</td>
<td>N-</td>
</tr>
<tr>
<td>1.4</td>
<td>317</td>
<td>9</td>
<td>105</td>
<td>109</td>
</tr>
<tr>
<td>1.5</td>
<td>339</td>
<td>13</td>
<td>320</td>
<td>320</td>
</tr>
<tr>
<td>1.5</td>
<td>382</td>
<td>14</td>
<td>28</td>
<td>28</td>
</tr>
<tr>
<td>Mo</td>
<td>W</td>
<td>VARIABLES</td>
<td></td>
<td></td>
</tr>
<tr>
<td>X-</td>
<td>X-</td>
<td>VARIABLES</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>1 - RED</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>2 - WHITE</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>3 - BLACK</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N-</td>
<td>N-</td>
<td>ALL ELEMENTS ARE IN ppm.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>96</td>
<td>27</td>
<td>EXCEPT: Hg ppb. Fe%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>96</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>303</td>
<td>305</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 3.4 Duncan's multiple range test for the influence of sediment colour of stream sediments associated with a granitic provenance. Common or overlapping circles indicate that group means are not significantly different at the 0.05 confidence level.
of overlap between categories (i.e., no single categorical group forms an independent statistically significant population) and although relatively higher concentrations of metals are found in sediments containing major amounts of fines, concentrations are not significantly different to those in sediments lacking fines.

**Bulk composition - Sand content**

Categories: 1=absent 2=minor(0-33%) 3=medium(34-66%) 4=major(>67%)

    Mo(342)(421), Hg(324)(41), U(34)(421)

Quartzite: Zn(234)(341)

Texture, as reflected by sand content significantly affects concentrations of seven elements in sediments derived from granites but only zinc for quartzites. As with fines, there is considerable overlap between categories and the greatest concentrations of zinc, lead, manganese, molybdenum and mercury appear to occur in sediments containing either major quantities of sand (category 4) or no sand at all (category 1).

**Bulk composition - Organic content** (Fig. 3)

Categories: 1=absent 2=minor(0-33%) 3=medium(34-66%)


Quartzite: Pb(13)(32) Hg(13)(32)

Variation in the content of organic matter significantly influences zinc, lead, mercury and molybdenum concentrations in
sediments derived from granitic terrane, and lead and mercury in those derived from quartzites. In general, although sediments from both rock types display increasing metal contents with increase in organic matter, those derived from granites have more statistically significant independent categorical groups (Fig. 3.3). In all cases, lowest metal concentrations are found in sediments judged to contain no organic matter with this category forming a statistically independent group for lead, mercury and zinc for sediments derived from granites.

**Sediment colour** (Fig. 3.4)

Categories: 1=red 2=white 3=black

Granite: Zn(2)(13), Cu(21)(3), Pb(2)(1)(3), Ni(12)(3), Mn(23)(31), U(23)(31), Hg(2)(1)(3)

Quartzite: Zn(23)(31), Pb(23)(31), W(13)(32), Hg(21)(3)

Variation in sediment colour is related significantly to concentrations of six elements in granitic catchments and four elements in quartzitic catchments. Zinc, lead and mercury concentrations are significantly affected in both groups and the general trend throughout is for relatively high metal content to be associated with red or black sediments.

For granites, greater average concentrations of copper, lead, nickel, mercury and zinc are associated with black sediments and except for zinc these form statistically significant independent groups. In the case of mercury there is approximately a twofold difference between average concentrations in white (14) and black (31) sediments.
In contrast, red sediments contain somewhat higher average concentrations of manganese and uranium.

**Banktype**

Categories: 1=alluvial 2=colluvial 3=glacial till

Granite: Zn(1)(32), Cu(1)(23), Pb(1)(32), Ni(12)(23), Co(1)(23), Fe(1)(32), Mn(1)(2)(3), Hg(1)(23)

Quartzite: U(3)(21), Mo(31)(2), Hg(13)(32)

Variation in bank type significantly affects concentrations of eight elements associated with predominantly granitic catchment basins and three elements in samples derived from quartzites. Mercury is the only element influenced in both rocktypes.

For granitic sediments, relatively high average concentrations of zinc, copper, lead, nickel, cobalt, iron, manganese and mercury are associated with glacial and colluvial banktypes, and for all of these elements, except nickel, significantly lower average concentrations are associated with alluvial banks. In contrast, for quartzitic sediments, relatively higher average concentrations of uranium, molybdenum and mercury are associated with colluvial banktypes and relatively lower contents with samples obtained from glacially derived banktypes. However, only uranium forms a statistically independent group for the latter category.

**Waterflow rate**
Categories: 1=slow 2=moderate 3=fast

Granite: Zn(3)(2)(1), Pb(3)(2)(1), Ni(2)(3)(1), Cu(32)(1), Hg(32)(1)

Quartzite: Cu(23)(1), Ni(23)(1), Mn(21)(13), Hg(23)(31)

Variation in stream flow significantly relates to metal contents for five elements in granitic terranes and for four elements at sample sites with a quartzite provenance. Both groups exhibit similar trends with sediments associated with the slower flow rates having the highest metal contents. This is most apparent with granitic sediments in which zinc, lead, nickel, copper and mercury all have their greatest concentrations, forming a statistically significant different category, in slow flowing streams. Nickel and copper exhibit a similar phenomenon in sediments derived from quartzites whereas for granites the lowest and statistically independent concentrations of zinc, lead, copper and mercury are associated with fast flowing streams.

Stream class

Categories: 1=secondary 2=tertiary 3=quaternary

Granite: Zn(2)(31), Cu(32)(1), Pb(2)(31), Ni(32)(1), Co(23)(1), Hg(32)(21)

Quartzite: Zn(2)(3), Cu(2)(3), Pb(2)(3), Ni(2)(3)

Co(2)(3), Mn(2)(3)

Differences in stream class are significantly related to concentrations of zinc, copper, lead, nickel and cobalt in sediments from both granites and quartzites. In addition
mercury and manganese concentrations are affected in granite and quartzite derived sediments, respectively. In all cases for quartzites, together with zinc and lead in sediments derived from granites, the lowest metal concentrations are found in samples from tertiary streams. In sediments on granites the secondary drainage class (category 1) forms a significantly distinctive category characterized by above average copper, nickel and cobalt concentrations.

**Drainage pattern**

Categories: 1=poor 2=dendritic 3=herringbone

Granite: Zn(13)(32), Cu(1)(32), Pb(31)(2), Co(13)(32),
Mn(13)(32), U(12)(3), Hg(31)(12)

Quartzite: Cu(3)(21) U(12)(3)

Variation in drainage pattern type are significantly related to concentrations of seven and two elements from granitic and quartzitic terrains, respectively. Only uranium is influenced for both rock types. Although there is considerable overlap between drainage pattern categories and metal concentrations, there is a general tendency for relatively high metal concentrations to occur in sediments of streams characterized by either dendritic or herringbone drainage patterns.

**Contamination**

Categories: 1=absent 2=possible 3=probable 4=definite
Granite: U(432)(321)
Quartzite: Mn(32)(241) Hg(324)(41)

Of the ten field parameters examined, contamination appears to influence the least number of metals. However, it should be noted, that anomalous concentrations of the metals resulting from contamination have for the most part been eliminated from the data set. For example, for granitic sediments 24, 25 and 34 samples have been removed from the test as a result of their anomalous content of zinc, copper and lead, respectively.

Physiography

Categories: 1=hilly  2=mountainous mature  3=mountainous youthful
Granite: Zn(3)(12), Pb(3)(12), Mn(3)(12), U(31)(12), Hg(13)(32)
Quartzite: Zn(23)(1), Cu(3)(21), Ni(32)(1), Fe(32)(1)

Variation in physiographic classification of stream catchments are significantly related to five elements for sample sites on granites and two on quartzites. In both cases, higher metal concentrations are associated with sediments obtained from hilly and mountainous mature terrains. Concentrations of zinc, nickel and iron form a significantly unique category of above average concentrations in catchments classified as hilly. In contrast, on granites a significantly distinctive category with below average concentrations of zinc, lead and manganese is associated with the mountainous young physiographic classification.
From the foregoing it is apparent that despite their qualitative, subjective character field observations can be related to variations in the trace element content of sediments associated with a single rock unit. These findings are summarized conveniently in Tables 3.4 and 3.5, for granites and quartzites respectively, where it is apparent which parameters influence the most elements or which elements are most susceptible to particular environmental factors. As an example, for streams draining granites (Table 3.4) we note that the bank type significantly influences 8 of 11 elements, whereas the content of fines apparently influences none of the elements. Conversely, the susceptibility of an element for its concentration to influence to the field parameters decreases from zinc, lead and mercury abundances of which are affected significantly by eight factors, to those elements whose abundances are affected by fewer recorded factors, eg, copper, manganese and uranium (5 factors), nickel (4 factors), cobalt (3 factors), molybdenum (2 factors), iron and tungsten (1 factor). Furthermore, comparisons of D ratios indicate, for example, that of the eight parameters significantly affecting concentrations of mercury, sediment colour and organic content have the most consistent or strongest influence. These results provide an objective criteria for assessing susceptibility of trace element dispersion to environmental factors and determining which are most relevant to geochemical exploration programs.

DISCUSSION
Table 3.4

*D*-Ratio calculated for categorical field variables. Granitic provenance, with group means shown to be significantly different by Duncan's Multiple Range Test (95% confidence)

<table>
<thead>
<tr>
<th></th>
<th>Zn</th>
<th>Pb</th>
<th>Hg</th>
<th>Cu</th>
<th>Ni</th>
<th>Co</th>
<th>Fe</th>
<th>Mn</th>
<th>U</th>
<th>W</th>
<th>Mo</th>
<th>Tot</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fines content</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sand content</td>
<td></td>
<td>1.79</td>
<td>2.37</td>
<td>1.28</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Organic content</td>
<td></td>
<td>1.63</td>
<td>2.43</td>
<td>2.32</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sediment colour</td>
<td>1.78</td>
<td>1.79</td>
<td>2.42</td>
<td>1.51</td>
<td>1.34</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Water flow rate</td>
<td></td>
<td>1.99</td>
<td>1.94</td>
<td>1.99</td>
<td>1.68</td>
<td>1.74</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physiography</td>
<td></td>
<td>2.59</td>
<td>2.80</td>
<td>1.46</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sediment colour</td>
<td>1.98</td>
<td>1.98</td>
<td>1.84</td>
<td>1.35</td>
<td>1.26</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Water flow rate</td>
<td></td>
<td>1.99</td>
<td>1.94</td>
<td>1.99</td>
<td>1.68</td>
<td>1.74</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physiography</td>
<td></td>
<td>2.59</td>
<td>2.80</td>
<td>1.46</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stream class</td>
<td></td>
<td>1.07</td>
<td>1.16</td>
<td>1.98</td>
<td>1.84</td>
<td>1.35</td>
<td>1.26</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Drain pattern</td>
<td></td>
<td>1.47</td>
<td>2.57</td>
<td>1.98</td>
<td>1.81</td>
<td>1.15</td>
<td>1.14</td>
<td>1.26</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bank type</td>
<td></td>
<td>1.34</td>
<td>1.95</td>
<td>2.06</td>
<td>1.62</td>
<td>1.73</td>
<td>1.48</td>
<td>1.42</td>
<td>3.23</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contamination</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Totals</td>
<td>(8)</td>
<td>(8)</td>
<td>(8)</td>
<td>(5)</td>
<td>(4)</td>
<td>(3)</td>
<td>(1)</td>
<td>(5)</td>
<td>(5)</td>
<td>(1)</td>
<td>(2)</td>
<td>(50)</td>
</tr>
</tbody>
</table>

Dashes indicate that significant differences among groups were not identified by Duncan's Multiple Range Test.

Table 3.5

*D*-Ratio calculated for categorical field variables. Quartzitic provenance, with group means shown to be significantly different by Duncan's Multiple Range Test (95% confidence)

<table>
<thead>
<tr>
<th></th>
<th>Zn</th>
<th>Pb</th>
<th>Hg</th>
<th>Cu</th>
<th>Ni</th>
<th>Co</th>
<th>Fe</th>
<th>Mn</th>
<th>U</th>
<th>W</th>
<th>Mo</th>
<th>Tot</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fines content</td>
<td>1.28</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sand content</td>
<td></td>
<td>1.14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Organic content</td>
<td></td>
<td>1.56</td>
<td>1.33</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sediment colour</td>
<td>1.18</td>
<td>1.22</td>
<td>1.29</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Water flow rate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.37</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physiography</td>
<td></td>
<td>1.41</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.91</td>
<td>1.29</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stream class</td>
<td></td>
<td>1.22</td>
<td>1.57</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.59</td>
<td>1.83</td>
<td>2.59</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Drain pattern</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.80</td>
<td>1.72</td>
<td></td>
</tr>
<tr>
<td>Bank type</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.44</td>
<td>1.26</td>
<td></td>
</tr>
<tr>
<td>Contamination</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.44</td>
<td>1.26</td>
<td></td>
</tr>
<tr>
<td>Totals</td>
<td>(5)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
<td>(4)</td>
<td>(2)</td>
<td>(2)</td>
<td>(4)</td>
<td>(3)</td>
<td>(1)</td>
<td>(1)</td>
<td>(34)</td>
</tr>
</tbody>
</table>

Dashes indicate that significant differences among groups were not identified by Duncan's Multiple Range Test.
Results clearly indicate that, although there are differences both between rock units and elements, the various categorical and non-parametric field observations can be related to significant variations in trace metal concentrations. Furthermore, it is encouraging to note that, despite their subjective character, results are reasonably consistent with well known controls on the behaviour of the trace elements. For example, association of enhanced concentrations of zinc, lead and mercury with organic-rich sediments on granites is consistent with scavenging of these elements by organic matter. Similarly sediment colour exhibits a consistent relationship with above average metal concentrations of zinc, copper, lead, nickel, manganese, uranium and mercury on granites being associated with strongly coloured red or black sediments - presumably a result of scavenging by iron sesquioxides and organic matter, respectively.

Consistency between the results and anticipated geochemical trends is further emphasized in Tables 3.6 and 3.7 (3.8 and 3.9) in which those factors favoring above (or below) average concentrations of the metals are summarized. Thus, on granites, it is apparent that relatively high organic content, black, fine grained sediment and a slow water flow rate all favor above average metal contents. In contrast, low organic content in relatively coarse sediments from fast flowing streams are likely to be characterized by low metal concentrations. Physiographic features, stream class and drainage pattern are also implicated as important factors. However, it is not so clear what their
Table 3.6

Specific categories of individual categorical field variables with highest means, Granitic provenance

<table>
<thead>
<tr>
<th></th>
<th>Zn</th>
<th>Pb</th>
<th>Hg</th>
<th>Cu</th>
<th>Ni</th>
<th>Mn</th>
<th>U</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fines content</td>
<td>(Major)</td>
<td>(Major)</td>
<td>(Major)</td>
<td>(Major)</td>
<td>(Major)</td>
<td>(Major)</td>
<td>(Major)</td>
</tr>
<tr>
<td>Sand content</td>
<td>Absent</td>
<td>Absent</td>
<td>Absent</td>
<td>(Absent)</td>
<td>(Major)</td>
<td>Absent</td>
<td>Absent</td>
</tr>
<tr>
<td>Organic content</td>
<td>Medium</td>
<td>Medium</td>
<td>Medium</td>
<td>(Medium)</td>
<td>(Medium)</td>
<td>(Medium)</td>
<td>(Medium)</td>
</tr>
<tr>
<td>Sediment colour</td>
<td>Black</td>
<td>Black</td>
<td>Black</td>
<td>Black</td>
<td>Black</td>
<td>Red</td>
<td>Red</td>
</tr>
<tr>
<td>Water flow rate</td>
<td>Slow</td>
<td>Slow</td>
<td>Slow</td>
<td>Slow</td>
<td>Slow</td>
<td>(Slow)</td>
<td>(Moderate)</td>
</tr>
<tr>
<td>Drain pattern</td>
<td>Dendritic</td>
<td>Dendritic</td>
<td>Dendritic</td>
<td>Dendritic</td>
<td>(Herringbone)</td>
<td>Dendritic</td>
<td>Herringbone</td>
</tr>
<tr>
<td>Bank type</td>
<td>Colluvial</td>
<td>Colluvial</td>
<td>Glacial</td>
<td>Glacial</td>
<td>Glacial</td>
<td>Colluvial</td>
<td>Alluvial</td>
</tr>
</tbody>
</table>

Unbracketed categories are significantly different (95% confidence level) from at least one other category for each variable. Categories within brackets were not found to be significantly different from other categories.

Table 3.7

Specific categories of individual categorical field variables with highest means, Quartzitic provenance

<table>
<thead>
<tr>
<th></th>
<th>Zn</th>
<th>Pb</th>
<th>Hg</th>
<th>Cu</th>
<th>Ni</th>
<th>Mn</th>
<th>U</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fines content</td>
<td>Minor</td>
<td>(Minor)</td>
<td>(Minor)</td>
<td>Minor</td>
<td>Minor</td>
<td>(Major)</td>
<td>(Minor)</td>
</tr>
<tr>
<td>Sand content</td>
<td>Absent</td>
<td>(Absent)</td>
<td>(Major)</td>
<td>(Absent)</td>
<td>(Major)</td>
<td>(Major)</td>
<td>(Minor)</td>
</tr>
<tr>
<td>Organic content</td>
<td>(Medium)</td>
<td>Minor</td>
<td>Minor</td>
<td>(Medium)</td>
<td>(Absent)</td>
<td>(Minor)</td>
<td>(Minor)</td>
</tr>
<tr>
<td>Sediment colour</td>
<td>Red</td>
<td>Red</td>
<td>Red</td>
<td>(Red)</td>
<td>(Red)</td>
<td>(Fast)</td>
<td>(Fast)</td>
</tr>
<tr>
<td>Water flow rate</td>
<td>(Slow)</td>
<td>(Slow)</td>
<td>Slow</td>
<td>Slow</td>
<td>Slow</td>
<td>(Fast)</td>
<td>(Fast)</td>
</tr>
<tr>
<td>Physiography</td>
<td>Hilly</td>
<td>(Hilly)</td>
<td>Mt. Mature</td>
<td>Hilly</td>
<td>Hilly</td>
<td>(Mt. Youth)</td>
<td>Mt. Youth</td>
</tr>
<tr>
<td>Stream class</td>
<td>Quaternary</td>
<td>Quaternary</td>
<td>Quaternary</td>
<td>Quaternary</td>
<td>Quaternary</td>
<td>Quaternary</td>
<td>(Tertiary)</td>
</tr>
<tr>
<td>Drain pattern</td>
<td>(Poor)</td>
<td>(Poor)</td>
<td>(Herringbone)</td>
<td>Poor</td>
<td>(Poor)</td>
<td>Poor</td>
<td>Poor</td>
</tr>
<tr>
<td>Bank type</td>
<td>(Alluvial)</td>
<td>(Alluvial)</td>
<td>(Colluvial)</td>
<td>(Colluvial)</td>
<td>(Alluvial)</td>
<td>(Alluvial)</td>
<td>Alluvial</td>
</tr>
</tbody>
</table>

Unbracketed categories are significantly different (95% confidence level) from at least one other category for each variable. Categories within brackets were not found to be significantly different from other categories.
### Table 3.8

Specific categories of individual categorical field variables with lowest means, Granitic provenance

<table>
<thead>
<tr>
<th></th>
<th>Zn</th>
<th>Pb</th>
<th>Hg</th>
<th>Cu</th>
<th>Ni</th>
<th>Mn</th>
<th>U</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fines content</td>
<td>(Minor)</td>
<td>(Minor)</td>
<td>(Medium)</td>
<td>(Minor)</td>
<td>(Minor)</td>
<td>(Minor)</td>
<td>(Minor)</td>
</tr>
<tr>
<td>Sand content</td>
<td>Medium</td>
<td>Medium</td>
<td>Medium</td>
<td>(Minor)</td>
<td>(Minor)</td>
<td>Medium</td>
<td>Medium</td>
</tr>
<tr>
<td>Organic content</td>
<td>Absent</td>
<td>Absent</td>
<td>Absent</td>
<td>Absent</td>
<td>(Absent)</td>
<td>(Absent)</td>
<td>(Absent)</td>
</tr>
<tr>
<td>Sediment colour</td>
<td>White</td>
<td>White</td>
<td>White</td>
<td>White</td>
<td>Red</td>
<td>White</td>
<td>White</td>
</tr>
<tr>
<td>Water flow rate</td>
<td>Fast</td>
<td>Fast</td>
<td>Fast</td>
<td>Moderate</td>
<td>Moderate</td>
<td>(Fast)</td>
<td>(Fast)</td>
</tr>
<tr>
<td>Physiography</td>
<td>Mt. Youth</td>
<td>Mt. Youth</td>
<td>Mt. Youth</td>
<td>Mt. Youth</td>
<td>Mt. Youth</td>
<td>Mt. Youth</td>
<td>Hilly</td>
</tr>
<tr>
<td>Stream class</td>
<td>Tertiary</td>
<td>Tertiary</td>
<td>Tertiary</td>
<td>Quaternary</td>
<td>Quaternary</td>
<td>Tertiary</td>
<td>(Tertiary)</td>
</tr>
<tr>
<td>Drain pattern</td>
<td>Poor</td>
<td>Poor</td>
<td>Poor</td>
<td>Herringbone (Poor)</td>
<td>Poor</td>
<td>Poor</td>
<td>Poor</td>
</tr>
<tr>
<td>Bank type</td>
<td>Alluvial</td>
<td>Alluvial</td>
<td>Alluvial</td>
<td>Alluvial</td>
<td>Alluvial</td>
<td>Alluvial</td>
<td>(Glacial)</td>
</tr>
</tbody>
</table>

Unbracketed categories are significantly different (95% confidence level) from at least one other category for each variable. Categories within brackets were not found to be significantly different from other categories.

### Table 3.9

Specific categories of individual categorical field variables with lowest means, Quartzitic provenance

<table>
<thead>
<tr>
<th></th>
<th>Zn</th>
<th>Pb</th>
<th>Hg</th>
<th>Cu</th>
<th>Ni</th>
<th>Mn</th>
<th>U</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fines content</td>
<td>Medium</td>
<td>(Major)</td>
<td>(Major)</td>
<td>Medium</td>
<td>Medium</td>
<td>Medium</td>
<td>(Medium)</td>
</tr>
<tr>
<td>Sand content</td>
<td>Minor</td>
<td>(Medium)</td>
<td>(Medium)</td>
<td>(Minor)</td>
<td>(Minor)</td>
<td>(Medium)</td>
<td>(Minor)</td>
</tr>
<tr>
<td>Organic content</td>
<td>(Absent)</td>
<td>Absent</td>
<td>Absent</td>
<td>Absent</td>
<td>(Absent)</td>
<td>(Medium)</td>
<td>(Medium)</td>
</tr>
<tr>
<td>Sediment colour</td>
<td>White</td>
<td>White</td>
<td>White</td>
<td>(Black)</td>
<td>(White)</td>
<td>(White)</td>
<td>(White)</td>
</tr>
<tr>
<td>Water flow rate</td>
<td>(Moderate)</td>
<td>(Moderate)</td>
<td>Moderate</td>
<td>Moderate</td>
<td>Moderate</td>
<td>Moderate</td>
<td>(Moderate)</td>
</tr>
<tr>
<td>Physiography</td>
<td>Mt. Youth</td>
<td>Mt. Mature</td>
<td>(Hilly)</td>
<td>(Mt. Mature)</td>
<td>Mt. Youth</td>
<td>(Mt. Mature)</td>
<td>(Hilly)</td>
</tr>
<tr>
<td>Stream class</td>
<td>Tertiary</td>
<td>Tertiary</td>
<td>(Tertiary)</td>
<td>Tertiary</td>
<td>Tertiary</td>
<td>Tertiary</td>
<td>(Quaternary)</td>
</tr>
<tr>
<td>Drain pattern</td>
<td>(Herringbone)</td>
<td>(Herringbone)</td>
<td>(Poor)</td>
<td>Dendritic</td>
<td>(Herringbone)</td>
<td>(Herringbone)</td>
<td>Poor</td>
</tr>
<tr>
<td>Bank type</td>
<td>(Glacial)</td>
<td>(Colluvial)</td>
<td>(Alluvial)</td>
<td>(Alluvial)</td>
<td>(Colluvial)</td>
<td>(Glacial)</td>
<td>Glacial</td>
</tr>
</tbody>
</table>

Unbracketed categories are significantly different (95% confidence level) from at least one other category for each variable. Categories within brackets were not found to be significantly different from other categories.
Table 3.10

Nonparametric correlation matrix, Granitic provenance (italicized $r$-values are significant at 95% confidence level, $N = 485$)

<table>
<thead>
<tr>
<th></th>
<th>Fines content</th>
<th>Sand content</th>
<th>Organic content</th>
<th>Water flow rate</th>
<th>Physiography</th>
<th>Stream class</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fines content</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Sand content</td>
<td>$-0.79$</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Organic content</td>
<td>$-0.09$</td>
<td>$-0.95$</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Water flow rate</td>
<td>$-0.05$</td>
<td>$0.22$</td>
<td>$-0.41$</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Physiography</td>
<td>$0.24$</td>
<td>$0.19$</td>
<td>$0.11$</td>
<td>$0.11$</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Stream class</td>
<td>$0.47$</td>
<td>$-0.29$</td>
<td>$0.13$</td>
<td>$0.11$</td>
<td>$0.08$</td>
<td>—</td>
</tr>
</tbody>
</table>
role is.

It is apparent that composition of a particular sample will reflect the interactions and relative strengths of many factors which, by reinforcing or counteracting each other, influence metal content associated with a particular bedrock unit. However, many of the environmental factors are interdependent; for example, a non-parametric correlation (Table 3.10) clearly indicates the interrelationships between stream class, fines, sand and organic content, and waterflow rate for sediments derived from granites. Consequently, groups of factors, acting collectively, become more important than any single factor in both characterizing the sediment and sample site, and in determining its metal content.

CONCLUSIONS

As far as we are aware, this is the first systematic attempt to evaluate the significance of field observations in relation to background variations in metal contents of drainage sediments. We conclude:

(1) Despite their subjective character field observations can related to significant variations in metal content of drainage sediments associated with a single rock unit.

(2) Variations in metal content of drainage sediments observed for most categorical field parameters are consis with well known controls on the behaviour of the trace
elements.

(3) Results indicate that groups of environmental factors, acting collectively are more important than any single factor characterizing the sediment and sample site, and in determining metal content.

(4) Results provide an objective criteria for assessing susceptibility of trace element dispersion to environment factors and deciding which factor(s) warrant recording in future surveys.

(5) The procedure presented, enables the significance of field observations to be systematically related to variations in background metal content of sediments and provides a guide to the relative strength or consistency of the influence of the various field parameters on trac element composition.

(6) The procedure also provides a means of characterizing those surficial environments favouring abnormally high or low metal contents.

ACKNOWLEDGEMENTS

The authors would like to thank the British Columbia
Ministry of Energy, Mines and Petroleum Resources for providing financial support. We are especially grateful to A. Bentzen for his valuable assistance with the computer work.
IV. RECOGNITION AND RANKING PROCEDURE FOR MULTI-ELEMENT REGIONAL STREAM SEDIMENT SURVEYS (MAP-AREA 82F), BRITISH COLUMBIA
ABSTRACT

A systematic, computer-oriented method of recognizing and ranking anomalous samples from geochemical data sets is described. The technique employs a multivariate approach (multiple regression) to identify and reduce the effects of background variations by re-expressing background multi-element data in metal associations (background regression models). These models are subsequently applied to sample subsets selected on factor/s (environmental controls) known to strongly influence geochemical results. Briefly the method involves:

1. Sorting of data into environmental control groups (usually rocktype)

2. Threshold selection using the method of Sinclair (1976) to isolate anomalous samples from background samples.

3. Selection of one or more elements to serve as focus of the study.

4. Backward stepwise regression of each environmental group to develop background models for selected metal in terms of other elements.

5. Ranking individual samples in terms of the regression model and threshold.

Multielement data obtained from a regional stream sediment survey undertaken by the British Columbia Ministry of Energy, Mines and Petroleum Resources in S.E. British Columbia (map-
area 82F) were evaluated. Samples were grouped initially as to dominant rocktype in the provenance region. Background multivariate models for zinc in terms of other log-transformed elements (Cu, Pb, Ni, Co, Fe, Mn, Hg, Mo and W) were obtained for each of the seven selected provenance groups. Individual samples were subsequently characterized and ranked after comparison with relevant threshold levels and background multi-element models. From a survey draining 1259 sample sites, 115 samples anomalous for zinc were identified and characterized. 43 of these samples contained zinc concentrations greater than its determined threshold value, while 72 of these anomalous samples were only identified because their individual metal associations were significantly different than their relevant provenance background model.

The method provides a means to reduce the effects of background variations while simultaneously identifying and characterizing anomalous samples. The technique promotes recovery of geological information and is applicable where geochemical data are complexly interrelated and the geological features within the survey area can be characterized by diagnostic metal associations.
INTRODUCTION

Exploration geochemistry utilizes geochemical patterns, that is natural variation, of elements, as a guide to the presence or absence of ore mineralization. Traditionally, interpretation of the bulk of exploration geochemical data has been based on the evaluation of single element distributions. This approach, however, frequently proves inadequate since metal distributions are often characterized by multiple populations of which presence ore or potential ore is only one of a host of genetic factors that may play a part in the development of the overall geochemical pattern. Consequently, a premium is placed upon interpretative techniques which can identify and reduce the influence of local geochemical patterns and secondly sort out all samples of background character. A number of previous investigations (Rose, Dalberg and Keith, 1970; Rose and Suhr, 1971; Closs and Nichol, 1975; and Austria and Chork, 1976) have indicated the potential of various interpretational procedures as aids in identifying background and anomalous populations in geochemical data. These studies commonly employed a multivariate approach to identify and reduce background variations by re-expressing the entire multi-element geochemical data set in metal associations (multi-element background models) which characterized general geologic, geochemical and environmental processes. Unfortunately, many of these studies included obviously anomalous samples into their regression analyses, tending to compromise the validity of their multi-
element background models. In order to overcome this limitation, a systematic computer-oriented method of recognizing and ranking anomalous samples was developed. Our detailed procedure utilizes the type and quality of data incorporated in various regional programs undertaken by the British Columbia Ministry of Energy, Mines and Petroleum Resources (Fig. 4.1) but can be adapted easily for data for other programs.

Regional multi-element stream sediment surveys of the type carried out in British Columbia under terms of the Uranium Reconnaissance Program contain coded information on the principal rock unit forming the provenance region of each sample. Rock-type coding for this purpose is never perfect: some basins may be miscoded, perhaps because of the scale of geological base maps available. In any case it is apparent that some apparently anomalous metal concentrations arise from incorrect assignment of the dominant rock-type or from mixing of sediments derived from several rock types. Consequently, the following procedure for determining multi-element background models is intended to be applied to individual samples so that its anomalous nature can be assessed for a variety of lithologies hosting the drainage basin.

GENERAL METHODOLOGY

In brief, our general approach to the recognition and ranking of anomalous samples involves the following steps:
Figure 44  Location map of study area 82F, and index map of regional geochemical surveys available in British Columbia (McMillan, 1982).
1. Sorting of data into provenance groups, that is, predominant rock type in drainage basin above the sample site

2. Evaluation of simple statistics and probability graphs for each element in each provenance group.

3. Threshold selection using the method of Sinclair (1976) to isolate anomalous samples from background samples.

4. Selection of one or more elements to serve as the focus of the study (for example, zinc).

5. Backward stepwise regression of each provenance group to develop background models for zinc in terms of other elements.

6. Ranking individual samples in terms of (a) their contamination code and (b) the regression model and threshold.

7. Output of sample information in a manner convenient for practical use in follow-up examination.

SORTING INTO PROVENANCE GROUPS

The basic assumption here is that for most samples the abundant rock type in the provenance area will exert a dominant control on the chemical character of derivative stream sediments. As a result data for each provenance group was dealt with separately. Means and standard deviations of all raw and log-transformed metal abundances provide insight into levels of abundance, dispersion, and general aspect of population densities (histogram). Correlation coefficients indicate metal associations of geological importance (for example, Sinclair and Tessari, 1980). If only background values are considered, these
associations commonly reflect differences in background environments and are not related to anomalous samples.

THRESHOLD SELECTION

Separation of background and anomalous samples is essential to our method because it leads directly to statistical models for background metal abundances. Samples with unusual values in any metal will tend to have an unfortunately strong influence on correlations and other multivariate statistics. Consequently, the method of threshold recognition is important. We have adopted the probability graph approach of Sinclair (1976) because this procedure is systematic and has been shown by numerous examples to provide effective thresholds for many types of geochemical data.

ELEMENT SELECTION

We must decide which element or elements are of direct concern to our search problem. Are we interested in silver-lead-zinc, copper-molybdenum, tungsten-uranium, or others? Of course, we may want to investigate many associations of the sort listed, but in our approach each association would be dealt with separately. Within a particular metal association it may not be necessary to deal thoroughly with all elements because some may be redundant, others may not show adequate geochemical contrast,
and still others may present limitations resulting from analytical problems. In our case we will use zinc data as a basis for evaluating regional silt samples in terms of silver-lead-zinc associations typical of our study area (map-area 82F).

MULTIVARIATE MODELLING OF BACKGROUND VALUES

Multiple regression has been shown by many to be an effective method of demonstrating empirical relationships between a particular element (dependent variable) and a group of other elements (independent variables). In many cases a high proportion of the variability of the dependent variable is explained in terms of variations in the independent variables (Sinclair and Fletcher, 1980). Where such methods are applied to background samples only, the abundance of a dependent variable (for example, zinc) can be expressed as a linear combination of the abundances (or logarithms of abundances) of many other elements to provide a multivariate background model. It should however, be emphasized that in order for the independent variables or predictor elements to be efficient predictors of the expected background levels of the dependent variable (element), the initial data should preferably be chosen only to describe the normal range of geochemical background variation. If numerous anomalous samples are included the regression line will be pulled away from the trend of the background samples into a compromise position, and some anomalous samples will not be recognizable in terms of their
deviations from this trend.

We have experimented with two approaches to the selection of samples used to establish a multiple regression model. In our first attempts sample selection was based on the dependent variable for a single provenance group with only those values below the threshold (based on probability graphs) being selected. In a later refinement we edited the data base for a single provenance group by omitting samples that were also obviously anomalous with respect to any of the independent variables.

The specific method we use for multivariate background modelling is backward, stepwise regression which starts with all independent variables in the data base and sequentially drops those that make no statistically significant contribution to explaining the variability of the dependent variable. Eventually a point is reached where all remaining variables are statistically significant (at the 0.05 level, for example) and an equation is obtained of the form:

$$\log (Zn) = B_0 + B_1 \log(x_3) + B_2 \log(x_4) + B_3 \log(x_5)$$

where B's are constants and x's are abundances of metal i.

RECOGNITION AND RANKING OF ANOMALOUS SAMPLES

For each sample we determine a series of ranks from 0 to 3 by comparing the observed value of the dependent variable with the values calculated by each of the provenance group
multivariate models. Significance of the rank numbers is shown on Figure 4.2. We then calculate a 4-digit ranking code for each sample where the first digit is the number of rock types for which rank 3 was obtained, the second digit is the number of rock types for which rank 2 was obtained, and so on. If there are seven rock types all with very high zinc values (rank 3) the ranking code would be 7000; in another case rank might be (3) for two rock types, (2) for three rock types, (1) for two rock types, and (0) for one rock type to give a ranking code of 2331.

The main advantage of this procedure is as a refinement in the selection of anomalous values relating to the probability graph procedure and the assigning of relative priorities to anomalous samples. Values above \( t_1 \) (Figure 4.2) are recognized as being anomalous without the aid of multiple regression. In addition, however, values below \( t_1 \) that depart substantially from the expectation according to a multiple regression model (1 and 2 on Figure 4.2) are also out of the ordinary and warrant examination. In particular, we are interested in those values below \( t_1 \) that are much higher than the corresponding calculated values. Such samples are anomalous in one element, relative to a linear combination of other elements. On Figure 4.2 the suggestion is made graphically that samples are anomalous if observed values are more than two standard errors greater than values calculated according to the multiple regression model.

OUTPUT PROCEDURES
Figure 4.2 Sample ranking in relation to fields on a plot of observed value versus a value calculated from a multivariate model.
We have designed an output system by which samples can be ordered in terms of decreasing priority for follow-up exploration. All anomalous samples recognized by the foregoing procedures are ranked according to the estimated likelihood of sample contamination from such factors as known mines, man-made metallic features, or fertilizer, on a scale of 0 to 3. Our first rank of anomalous samples is based on this coded parameter, zero contamination being of most interest. Within this group we code a sample for each background model as 3, 2, 1, or 0 as described previously and a 4-digit ranking code is used to list samples within each contamination group in order of decreasing ranking code. Locations for each sample are listed as is the observed abundance of the dependent variable and the sample number. These items are arranged in a manner as to promote efficiency of evaluation of each sample. In addition, we use plot locations of anomalous samples with their identification number and ranking code.

CASE HISTORY (MAP AREA 82F)

The data base for the study comprises field observations and analytical results (minus 80-mesh, Zn, Cu, Pb, Ni, Co, Mn, Fe, Mo, Hg and U) for drainage sediments collected at 1313 sites in S.E. British Columbia (Figure 4.1) as part of the Joint Federal-Provincial Uranium Reconnaissance Program. Details of sampling and analytical procedures are to be found in Open File 514 of the Department of Energy, Mines and Resources, Canada and

Samples were grouped initially on the basis of coding as to dominant rock type in the provenance region. Data for each element in each provenance group were examined as a probability graph and a threshold selected separating two populations (presumably anomalous and background) using the method of Sinclair (1976). We chose to examine zinc as the dependent variable described here because of the association silver-lead-zinc in known vein deposits in the area. Background multivariate models for zinc in terms of other elements were obtained for each of the seven provenance groups for which we have adequate samples. Three of these models are summarized in Table 4.1 to illustrate the type of results obtained. Statistics for all seven provenance group models for zinc are given in Table 4.2 to illustrate the statistical quality of the background models.

All samples coded in one the seven provenance groups for which we could calculate background models were treated by each of the background equations separately. The calculated zinc background according to a given model was then compared with expectations for that model so that for each background model a sample received a ranking from 0 to 3 inclusive (compare Figure 4.2). In our case each sample was ranked seven times, once for
Table 4.1 Examples of multivariate regression background models for zinc, Map area 82F.

**GRANITE**

\[
\log(Zn) = 0.4726 + 0.0713 \log(Cu) + 0.2420 \log(Pb) + 0.0529 \log(Ni) \\
+ 0.3994 \log(Mn) + 0.4189 \log(Fe) - 0.2334 \log(Co) \\
R^2 = 0.62 \\
S_e = 0.1277 \\
n = 393
\]

**QUARTZITE**

\[
\log(Zn) = 1.1020 + 0.2721 \log(Pb) + 0.1316 \log(Ni) + 0.4891 \log(Fe) \\
+ 0.1412 \log(Mo) + 0.0399 \log(Hg) \\
R^2 = 0.74 \\
S_e = 0.0915 \\
n = 287
\]

**SCHIST**

\[
\log(Zn) = 0.8392 + 0.4100 \log(Pb) + 0.2244 \log(Ni) + 0.5603 \log(Fe) \\
+ 0.2412 \log(W) \\
R^2 = 0.76 \\
S_e = 0.1008 \\
n = 27
\]
<table>
<thead>
<tr>
<th>PROVENANCE GROUP</th>
<th>GRNT</th>
<th>QRTZ</th>
<th>SLTE</th>
<th>ANDS</th>
<th>ARGL</th>
<th>GNSS</th>
<th>SCST</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>393</td>
<td>287</td>
<td>100</td>
<td>.57</td>
<td>56</td>
<td>53</td>
<td>27</td>
</tr>
<tr>
<td>R</td>
<td>.79</td>
<td>.86</td>
<td>.84</td>
<td>.84</td>
<td>.92</td>
<td>.85</td>
<td>.87</td>
</tr>
<tr>
<td>$R^2$</td>
<td>.62</td>
<td>.74</td>
<td>.70</td>
<td>.70</td>
<td>.85</td>
<td>.71</td>
<td>.76</td>
</tr>
<tr>
<td>$S_e$</td>
<td>.1277</td>
<td>.0915</td>
<td>.1184</td>
<td>.1031</td>
<td>.0812</td>
<td>.0940</td>
<td>.1008</td>
</tr>
</tbody>
</table>

Table 4.2 Summary statistics for multivariate background zinc models for seven provenance groups, Map area 82F.
each model. These rankings were accumulated into a single ranking code. Samples recognized as anomalous or potentially anomalous were divided into three contamination classes with priority decreasing as certainty of contamination increases. For each contamination category samples are ranked according to decreasing numeric value of the ranking code. An example is shown in Table 4.3, where a small part of the O-contamination category is listed.

From a total of 1259 samples, this procedure produced 115 anomalous samples in the O-contamination category. Forty three of these samples had zinc concentrations greater than its simple threshold, calculated for its observed provenance group, while the remaining 72 samples were identified solely on the basis of having significantly different metal associations than their particular provenance multivariate background model. Our procedure is to list these samples in tabular form in Table 4.3 and to produce computer-drawn plots of anomalous sample locations as illustrated on Figure 4.3.

In addition to ranking information, original raw data, and coordinates, the output table contains a simple consecutive numeric identifier used for clarity on the map output and permitting easy combined use of the tabulated data and the output map. The output map is of particular use because it identifies the most obvious anomalous samples (for example 7000) from those that might escape detection (for example, 0520). The scale of the location plot should be identical to geological base maps of the area so the two can studied together without
Table 4.3 Part of a table listing anomalous samples in order of decreased ranking code

<table>
<thead>
<tr>
<th>UTME</th>
<th>UTMN</th>
<th>GNSS</th>
<th>ORTZ</th>
<th>GRNT</th>
<th>SLTE</th>
<th>ANDS</th>
<th>SCST</th>
<th>ARGL</th>
<th>TOTALS (3210)</th>
<th>ZN (PPM)</th>
<th>ROCK TYPE</th>
<th>CONTAM</th>
<th>URPS (82F)</th>
<th>ID #</th>
</tr>
</thead>
<tbody>
<tr>
<td>54</td>
<td>558659</td>
<td>5532787</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>3400</td>
<td>181</td>
<td>GRNT</td>
<td>0</td>
<td>775229</td>
<td></td>
</tr>
<tr>
<td>55</td>
<td>538165</td>
<td>5479004</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>2302</td>
<td>GRNT</td>
<td>0</td>
<td>773116</td>
<td></td>
</tr>
<tr>
<td>56</td>
<td>548189</td>
<td>5536835</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>2302</td>
<td>ORTZ</td>
<td>0</td>
<td>775248</td>
<td></td>
</tr>
<tr>
<td>57</td>
<td>518892</td>
<td>5514798</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>2203</td>
<td>ORTZ</td>
<td>0</td>
<td>775391</td>
<td></td>
</tr>
<tr>
<td>58</td>
<td>525483</td>
<td>5477274</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>2203</td>
<td>ORTZ</td>
<td>0</td>
<td>779153</td>
<td></td>
</tr>
<tr>
<td>59</td>
<td>444220</td>
<td>5512900</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2005</td>
<td>GRNT</td>
<td>0</td>
<td>771294</td>
<td></td>
</tr>
<tr>
<td>60</td>
<td>483724</td>
<td>5474162</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>601</td>
<td>123</td>
<td>ANDS</td>
<td>0</td>
<td>779063</td>
<td></td>
</tr>
<tr>
<td>61</td>
<td>489795</td>
<td>5497195</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>601</td>
<td>89</td>
<td>SLTE</td>
<td>0</td>
<td>777215</td>
<td></td>
</tr>
<tr>
<td>62</td>
<td>447145</td>
<td>5447184</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>601</td>
<td>79</td>
<td>GNSS</td>
<td>0</td>
<td>775011</td>
<td></td>
</tr>
<tr>
<td>63</td>
<td>511528</td>
<td>5479314</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>601</td>
<td>77</td>
<td>GRNT</td>
<td>0</td>
<td>779204</td>
<td></td>
</tr>
<tr>
<td>64</td>
<td>479680</td>
<td>5464101</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>502</td>
<td>123</td>
<td>ARGL</td>
<td>0</td>
<td>771091</td>
<td></td>
</tr>
<tr>
<td>65</td>
<td>501640</td>
<td>5430728</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>502</td>
<td>69</td>
<td>ANDS</td>
<td>0</td>
<td>777078</td>
<td></td>
</tr>
<tr>
<td>66</td>
<td>510994</td>
<td>5485659</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>403</td>
<td>ORTZ</td>
<td>0</td>
<td>773211</td>
<td></td>
</tr>
<tr>
<td>67</td>
<td>485042</td>
<td>5493023</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>403</td>
<td>127</td>
<td>GRNT</td>
<td>0</td>
<td>777234</td>
<td></td>
</tr>
<tr>
<td>68</td>
<td>466702</td>
<td>5510212</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>403</td>
<td>123</td>
<td>GNSS</td>
<td>0</td>
<td>771328</td>
<td></td>
</tr>
<tr>
<td>69</td>
<td>512650</td>
<td>5440330</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>403</td>
<td>46</td>
<td>ORTZ</td>
<td>0</td>
<td>775264</td>
<td></td>
</tr>
<tr>
<td>70</td>
<td>568389</td>
<td>5428516</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>403</td>
<td>47</td>
<td>ORTZ</td>
<td>0</td>
<td>773096</td>
<td></td>
</tr>
<tr>
<td>71</td>
<td>537227</td>
<td>5485766</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>304</td>
<td>141</td>
<td>ORTZ</td>
<td>0</td>
<td>775385</td>
<td></td>
</tr>
<tr>
<td>72</td>
<td>511127</td>
<td>5487241</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>304</td>
<td>121</td>
<td>ORTZ</td>
<td>0</td>
<td>773214</td>
<td></td>
</tr>
</tbody>
</table>
Figure 4.3 Plot of an area of anomalous samples from map-area 82F
ambiguity. We tested sensitivity of the regression procedure for determining a multivariate background for zinc by establishing such models based on two training sets: (1) all samples indicated as having background zinc values, and (2) the same data set minus any samples that appeared to be anomalous in any element other than zinc. Tables 4.1 and 4.2 are based entirely on the second training set. Figures 4.4 and 4.5 illustrate the contrasting results obtained in background definition. It is clear that the 'cleaner' data set (number 2 previously) leads to a better multiple regression relationship, that is, with less scatter of calculated and observed values. The problem with using the second training set is that more work is required to set it up and more samples will be included in the anomalous category.

DISCUSSION

The methodology described here would appear to have a wide range of applications to geochemical data evaluation, perhaps with minor modifications to suit particular data sets. For example, many geochemical surveys may not record the likelihood that a sample is contaminated, and this level of ranking might have to be omitted. The precise limits to the coding regions illustrated on Figure 4.2 can be changed to suit a particular bias to anomaly selection, resulting in a slightly different listing of anomalous samples.

One of the serious problems is the question of initial
Figure 4.4 Observed versus calculated zinc values for provenance group 'ARGL' (argillite), map-area 82F; calculated values based on a model determined from all samples with background zinc values.
Figure 4.5 Observed versus calculated zinc values for provenance group 'ARGL' (argillite), map-area 82F; calculated values based on a model determined from those samples with zinc background values that also are not anomalous in any other element (that is, a 'cleaner' subset of the data used in Figure 4.4).
grouping of data on the basis of dominant rock type that underlies the drainage basin of each sample, a classification which is fundamental to our procedures. A substantial amount of effort is required to code this rock-type information even if the data are available. If rock type has not been coded it may be necessary to use some less satisfactory method of approximation of background geology for each sample. In some environments, of course, some other parameter may be more useful than rock type for grouping data.

CONCLUSIONS

A method of anomaly selection and ranking for multi-element regional stream sediment data has been described. The procedure offers the following advantages:

(1) The method is rigorous in making use of established statistical method for treating geochemical data such as a probability graph analysis and backward stepwise regression.

(2) The procedure is computer based and is rapid and thorough.

(3) The methodology ensures that some anomalous values which are not obvious (that is, are not higher than a simple threshold) will be recognized.

(4) A novel ranking procedure is described that
assigns relative priorities to samples for further investigation. Details of the ranking procedure are subjective but a system for ranking codes clearly describes the manner in which a sample is anomalous.

(5) Because samples are tested against every rock type, the procedure incorporates an evaluation as to whether other rock types might be contributing to the provenance area of a particular sample. Possible additional rock types are identified and can be compared with available geological maps.

(6) The method provides a means to reduce the effects of background variations, while simultaneously identifying and characterizing anomalous samples.

Consideration should be given to its implementation in all geochemical surveys where the data are complexly interrelated and the geological features within the survey area can be characterized by diagnostic metal associations.
ACKNOWLEDGEMENTS

This project has been funded by the British Columbia Ministry of Energy, Mines and Petroleum Resources.
V. CONCLUSIONS
A suite of systematic, computer-assisted, statistical and graphical procedures have been developed and applied to data from published government stream sediment surveys in southeastern British Columbia and the geochemical interpretation of the data analysis had been presented.

Results show that statistical analysis of field-site duplicate samples provide a rapid an effective means of interpreting geochemical trends, while simultaneously allowing evaluation of the adequacy of sampling and laboratory analysis. Government silt data appear to be of very high quality and are characterized by low precision estimates (6%) at the 95th percentile concentration value for metals Zn, Cu, Ni, Co, Fe and Mn, and a combined variability due to local and procedural error of less than 5% of the total error. Metal concentrations between drainage sample sites reflect real trends related to geological and geochemical features. Results presented are more in accord with splits of an individual sample (analytical duplicate) rather than analyses of two separate field samples.

Relationships between background metal concentrations and field observations of the drainage catchment, sample site and sediment sample commonly recorded in geochemical surveys were examined systematically for the first time. We conclude:

1. Despite their subjective character field observations can be related systematically to
significant variations in metal content of drainage sediments associated with a single rock unit.

2. Variations in metal content of drainage sediments observed for most categorical field parameters are consistent with well known controls on the behaviour of the trace elements.

3. Results indicate that groups of environmental factors, acting collectively are more important than any single factor characterizing the sediment and sample site, and in determining metal content.

4. Results provide an objective criteria for assessing susceptibility of trace element dispersion to environmental factors and deciding which factor(s) warrant recording in future surveys.

A method of anomaly selection and ranking for multi-element regional stream sediment data has been described. The procedure allows the geochemical character of individual silt samples draining various provenance regions to be screened against relevant threshold levels and background multi-element models, so as to insure that some anomalous samples which are not obvious (that is, are not higher than a simple threshold) are recognized. Because samples are tested against every rock type the procedure incorporates an evaluation as to whether other rock types might be contributing to the provenance area of a particular sample. Possible additional rock types are identified and can be compared with available geological maps. The method provides a means to reduce the effects of background variations while simultaneously identifying and characterizing anomalous samples. Consideration should be given to its implementation in all geochemical surveys where the data are complexly interrelated and the geological features within the
search area can be characterized by diagnostic metal associations.

In conclusion, it is believed that the adoption of such procedures as outlined here will provide rapid and promote maximum recovery of geochemical information from geochemical data sets.
BIBLIOGRAPHY


Garrett, R.G., 1969. The determination of sampling and
analytical errors in exploration geochemistry. Econ. Geol., 64: 568-569.


Meyer, W.T., Theobald, P.K. and Bloom, H., 1979. Stream sediment geochemistry in Geophysics and Geochemistry in the


