APPLICATION OF STATISTICAL TECHNIQUES IN THE INTERPRETATION OF STREAM SEDIMENTS DATA OF SOME RIVERS IN SRI LANKA

A PROJECT REPORT PRESENTED BY

A.PRAGALATHAN

to the Board of Study in Statistics and Computer Science of the

POSTGRADUATE INSTITUTE OF SCIENCE

in partial fulfillment of the requirement

for the award of the degree of

MASTER OF SCIENCE IN APPLIED STATISTICS

of the

UNIVERSITY OF PERADENIYA

SRI LANKA

2007

ABSTRACT

Geochemical surveys play a significant role in mineral explorations as well as for the establishment of the geological information relevant to a particular area. In the context of gem mineral exploration, stream sediments sampling and their geochemical analysis has been identified as a versatile technique.

The present study focuses on the identification of statistical signature that can be used to delineate potential gem bearing areas using stream sediment geochemical data. The stream sediment geochemical data were obtained from various previous studies. 25 known variables (elements) from samples collected from different parts of Sri Lanka were treated statistically, using multivariate techniques such as Discriminant Analysis, Principal Component Analysis and Factor Analysis. The samples were categorized based on the known gem potential and then treated statistically.

After subjecting the data into stepwise discriminant analysis, 17 elements were extracted and the discriminant function was derived using predictive discriminant analysis. In the formulation of the linear discriminant function for the high potential and low potential regions the results obtained for each variable were used. Based on the results the linear discriminant functions associated with low potential (f_l) and high potential (f_h) are as follows.

$$f_l = -835.18 + 12.79 \log Al - 65.24 \log Ca + 109.22 \log Co + 22.14 \log Cr + 111.65 \log Fe - 16.08 \log K - 33.56 \log La + 3.91 \log Mg - 1.01 \log Na - 140.26 \log P + 255.57 \log Si + 124.58 \log Sr + 22.19 \log Th - 97.24 \log Ti + 18.52 \log U + 204.57 \log Y + 50.09 \log Zr.$$

$$f_h = -774.03 + 26.64 \log Al - 68.70 \log Ca + 105.13 \log Co + 33.00 \log Cr + 96.94 \log Fe - 0.57 \log K - 37.82 \log La + 6.84 \log Mg - 3.50 \log Na - 130.15 \log P + 235.03 \log Si + 114.00 \log Sr + 24.56 \log Th - 92.06 \log Ti + 20.91 \log U + 192.10 \log Y + 52.86 \log Zr.$$

To identify a new stream sediment taken from un known potential area whether it belong to low potential or high potential group with the above attributes, both equation values would be computed and if $f_l > f_h$ then the sample belongs to low potential otherwise sample belongs to high potential.

Principle Component Analysis is necessary to group these elements into high potential and low potential areas. Data were further clustered using Factor Analysis by the varimax rotation method. The derived equations can be applied to delineate high and low gem potential regions effectively and can be applied for geochemical data obtained from a unknown terrain. However the accuracy of the equations should be verified by field studies.