



O T H E R I N V E S T I G A T I O N S

MULTIVARIATE MODELS FOR RELATIVE MINERAL POTENTIAL
SLOCAN SILVER-LEAD-ZINC-GOLD CAMP

(82F)

By A.J. Sinclair
Department of Geological Sciences
University of British Columbia

ABSTRACT

One hundred and thirty-eight vein deposits forming the western part of Slocan camp are used to develop statistical models based on known ore production and average grades of production. Relative size of deposits as estimated by the biased variable 'production' (short tons of ore) is used as a relative value measure to (a) serve as a dependent variable for multiple regression models, and (b) form the basis for subdividing data into low, medium, and high-tonnage groups for discriminant analysis models. Both types of models appear useful judging by their ability to forecast known production for a target production of deposits from the eastern part of Slocan camp. Results suggest that multiple regression is more useful than is discriminant analysis in classifying deposits with respect to potential size.

INTRODUCTION

Statistical methods have been applied over a wide range of scales in attempts to develop a rigorous approach to measuring absolute or relative resource potential (Kelly and Sheriff, 1969; Sinclair and Woodsworth, 1970; Orr and Sinclair, 1971; Godwin and Sinclair, 1979). Comparable approaches are used widely in evaluating oil and gas potential (for example, Harbaugh, et al., 1977). Little effort has been expended on the application of statistical models to the evaluation of specific polymetallic mineral deposits in terms of grades of economically important metals, particularly in vein camps.

Such models have two important applications viz:

- (1) Newly located deposits can be sampled and the average grades used to estimate potential size and thereby provide an evaluation of the usefulness of additional exploration.
- (2) A statistical model may identify known deposits with low productivity that have grade characteristics of large deposits, thus indicating specific known deposits that appear to warrant more detailed examination.

Orr and Sinclair (1971) attempted to use multiple regression in Slocan City vein camp as a means of estimating relative value of a deposit. Average production grades for silver, gold, lead, and zinc were used as the independent variables and recorded production tonnage was the relative value indicator (dependent variable). Statistically significant models were described but no data existed by which the geological validity of a model could be tested. Furthermore, their models were of little practical interest because they pertained to a camp in which deposits were very small and of little general interest.

More than 200 vein deposits in Slocan camp have produced from 1 to 700,000 tons of ore for which average grades of total recorded production are more or less completely known (Orr and Sinclair, 1971). These data provide an unusually comprehensive information base for testing the validity of various statistical approaches to the estimation of 'value' of a deposit, and the potential of such models as an exploration tool.

PROCEDURE

Deposits in Slocan camp were divided into two groups for the purpose of this study. One hundred and thirty-eight deposits forming the western two-thirds of the camp were taken as a training set for the development of statistical models to determine relative 'value' of individual deposits. The eastern part of the camp contains 65, more dispersed deposits that form a target population. Average production grades and tonnages are known for these 65 deposits so real values can be compared with values calculated according to various models, and a rigorous test of the predictive capability of a model can be conducted.

RELATIVE VALUE MEASURE

Sinclair (1979) has discussed the problem of relative value measures for multi-commodity deposits. Metal content is an ideal value estimator for single commodity deposits but where several commodities contribute significantly to total value of a vein deposit, for example, size is commonly a more meaningful relative value measure. Production ore tonnage, despite obvious limitations, can be used as an adequate estimator of relative deposit size, and therefore of relative value.

We never know the true size (volume or tonnage) of a deposit because we are never sure that it has been completely worked out or had all the reserves outlined. Consequently, ore production tonnage is a biased estimator of size and of relative value that is always low relative to the true value. Furthermore, the bias may not be the same for all sizes because of selective mining, hand sorting, and so on. Nevertheless, as a relative measure, production tonnage is still adequate except for:

- (1) many pairs of deposits with identical or nearly similar production tonnages (a trivial case), and
- (2) some deposits that really are large despite the fact they have produced only small amounts of ore to date.

this latter case is not a problem, rather it is a situation that we hope exists so that such deposits can be recognized eventually by our predictive models.

In the present case, production tonnage is accepted as an adequate relative value measure. A probability graph for production tonnages shown on Figure 1 for both the training set and the target population, indicates close similarity of statistical density functions between the two areas. Two populations are present in identical proportions in each area, although the average size is somewhat lower in the target population than in the training set. The probability graph for the training set has been partitioned into two populations with mean sizes that differ by several orders of magnitude. This fact emphasizes the desirability of a numerical model that would allow distinction between high tonnage deposits of economic interests and the low tonnage population of little economic interest.

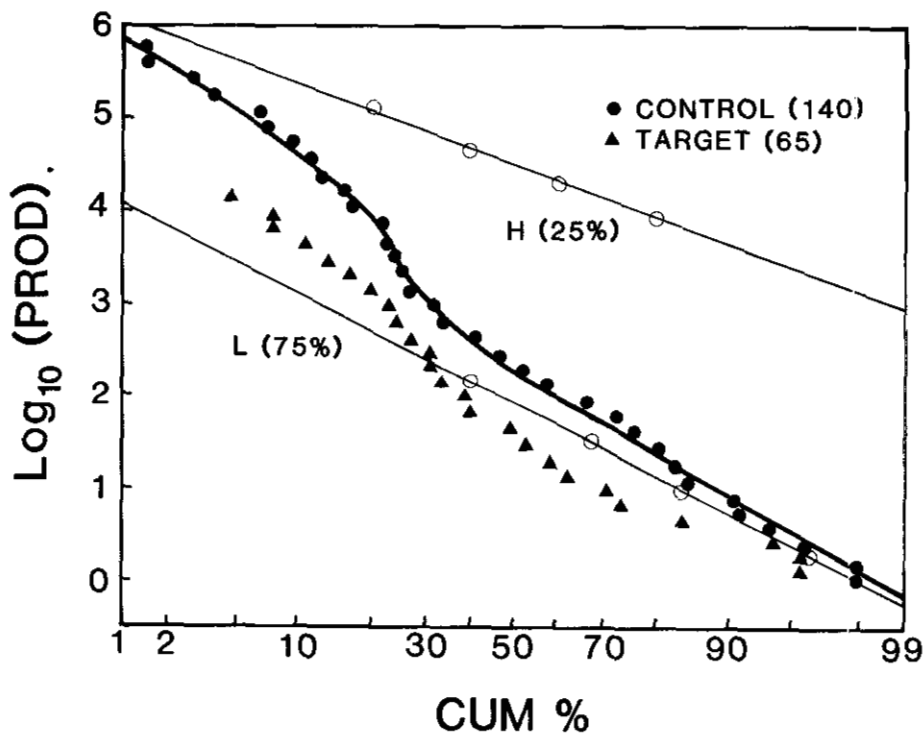


Figure 1. Probability graphs of ore production tonnage for the training set (filled circles) and the target population (filled triangles). Two straight lines are the two interpreted lognormal populations drawn through construction points (open circles) derived from the training set curve.

The principal metal contributor to value in Slocan camp is silver, which is plotted versus production tonnage on Figure 2. From this diagram it is difficult to ascertain whether two distinctive populations exist, each with its own average grade and dispersion of average grades, or whether, on average there is a continuous variation in metal content and size (production tonnage). Both possibilities will be tested, the continuous case using multiple regression and the discontinuous case using discriminant analysis.

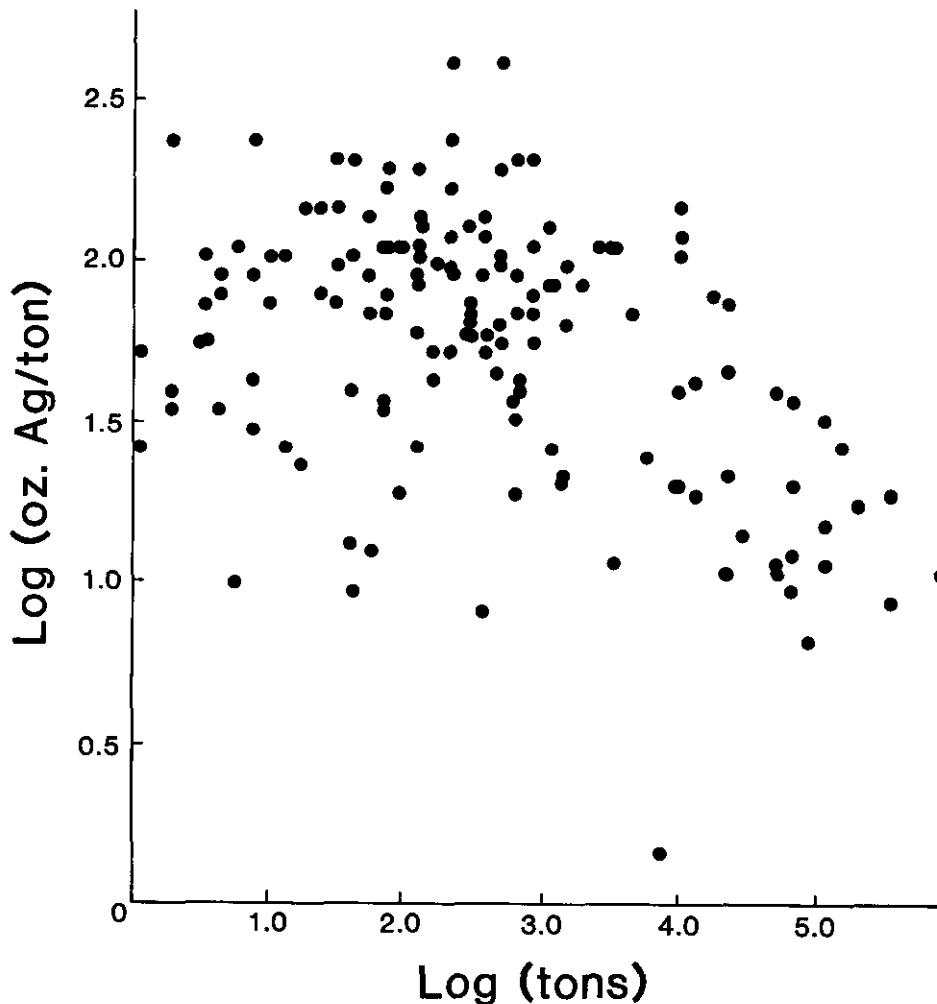


Figure 2. Scatter diagram of average Ag grade versus ore production (tonnage for training set).

MULTIPLE REGRESSION MODELS

Multiple regression models have the form

$$V = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n + e,$$

where V is a dependent variable (value measure in our case), β 's are constants, e is the standard error, and X 's are independent (geological) variables. We have examined a variety of potential models and quote two here to illustrate the variation in results. Model 1 relates ore production tonnage to the four metal grades (silver, gold, lead, zinc);

model 2 relates ore production tonnage to three metals (silver, lead, zinc), and the percentage of vein sulphides.

Model 1

$$\begin{aligned} \text{Log (tons)} &= 1.6853 - 0.4222 \log (\text{Ag}) - 0.0146 (\text{Pb}) \\ &\quad - 0.2607 \log (\text{Zn}) - 1.0429 \log (\text{Au}) \\ R^2 &= 0.6565 \\ S^2 &= 0.9770 \end{aligned}$$

Model 2 (omitting Au)

$$\begin{aligned} \text{Log (tons)} &= 3.836 - 1.1784 \log (\text{Ag}) - 0.0140 (\text{Pb}) \\ &\quad - 0.4518 \log (\text{Zn}) + 1.151 \log (\% \text{ sulphides}) \\ R^2 &= 0.3467 \\ S^2 &= 1.1489 \end{aligned}$$

In both models log transformations were used where dictated by the form of the density function. Statistically the results are significant. However, we want to know if the models have practical application over their entire range.

The five highest calculated and observed ore tonnages for the training set are shown as a scatter diagram (Figure 3) for model 1. The fit is

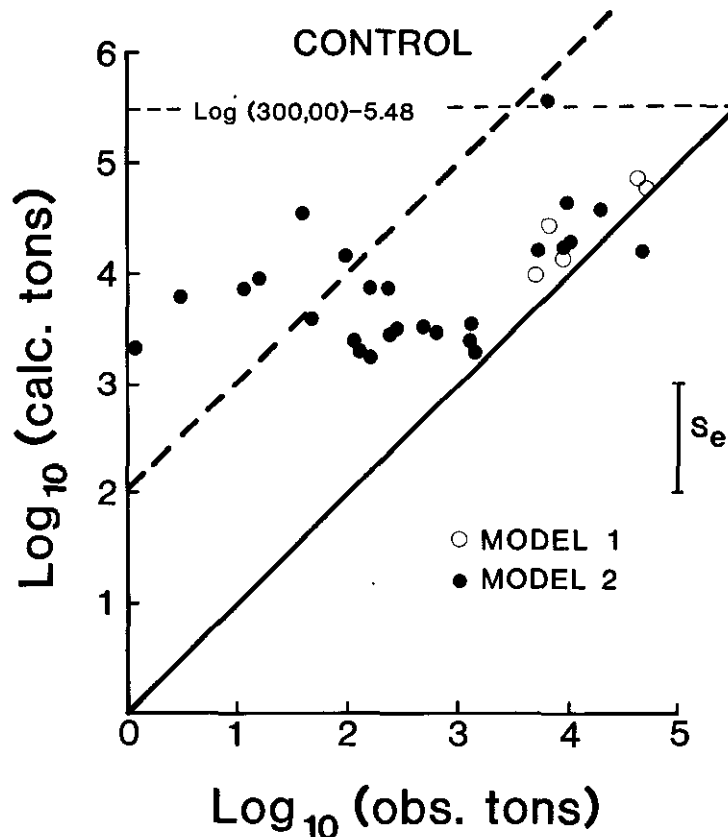


Figure 3. Observed versus calculated ore production for highest values of models 1 and 2, Slocan camp training set.

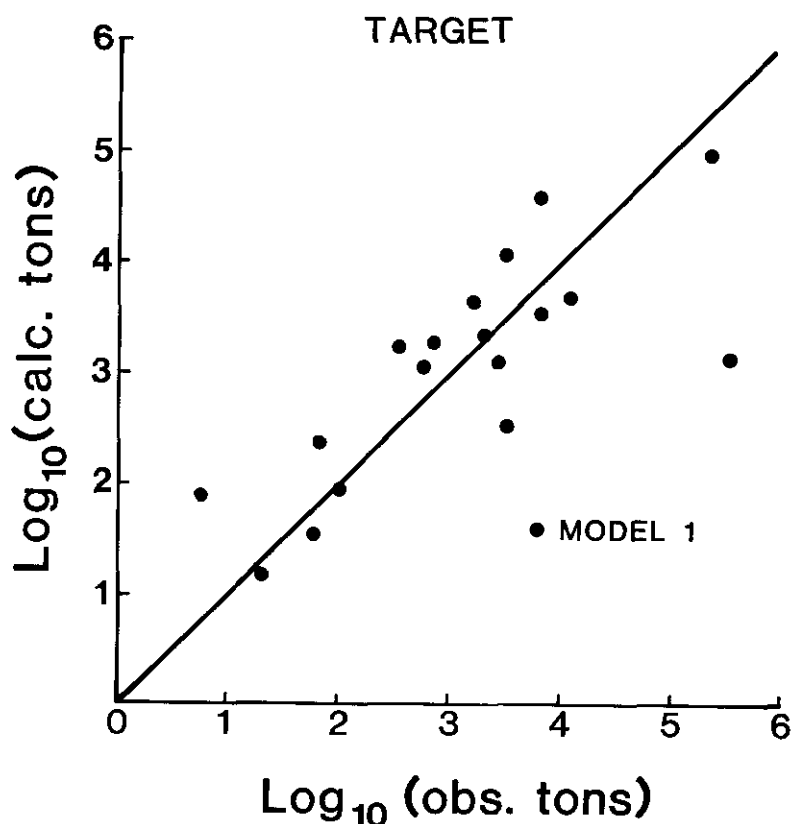


Figure 4. Observed versus calculated ore production tonnages for 19 deposits in the target area, Slocan camp.

reasonably good visually and Figure 3 provides a quantitative appreciation of the internal consistency of the model for large tonnage potential. However, a proper test of the model involves application to the target population, a set of data not used in developing the model. Such a comparison is shown on Figure 4 for the 19 deposits in the training set for which average grades were available for all silver, gold, lead, and zinc. It is apparent that the model is a relatively good forecaster! Other multiple regression models (for example, model 2) were not so successful and results will not be discussed here, although high values for model 2 are plotted on Figure 3.

The model has been shown to be an adequate forecaster of relative value (ore production tons) of individual veins in Slocan camp. Unfortunately application of the model is limited because average grades of the four metals incorporated in the model are available only for about one-third of the deposits in the camp. It was for this reason that we attempted to develop models omitting gold (for example, model 2). It is unfortunate that gold grades are not available for more of the deposits because this work has shown that gold is by far the most important single component in our models. Gold assays, along with silver, lead, and zinc, should be an integral part of evaluating any vein deposit in Slocan camp.

DISCRIMINANT ANALYSIS

Discriminant analysis is a method by which k clusters of data in n -dimensional space are separated as efficiently as possible by $(k-1)$ n -dimensional lines. These 'discrimination' lines are arranged between pairs of clusters such that the two clusters project onto the line with a minimum overlap (Figure 5, from Klovan and Billings, 1967).

Slocan data were divided into high-tonnage, medium-tonnage, and low-tonnage groups based on thresholds of 1,000 tons and 16,000 tons selected using the probability graph of Figure 1. The two groups of average production grades corresponding to high and low-tonnage groups were transformed as in the multiple regression study and linear 'discriminant'

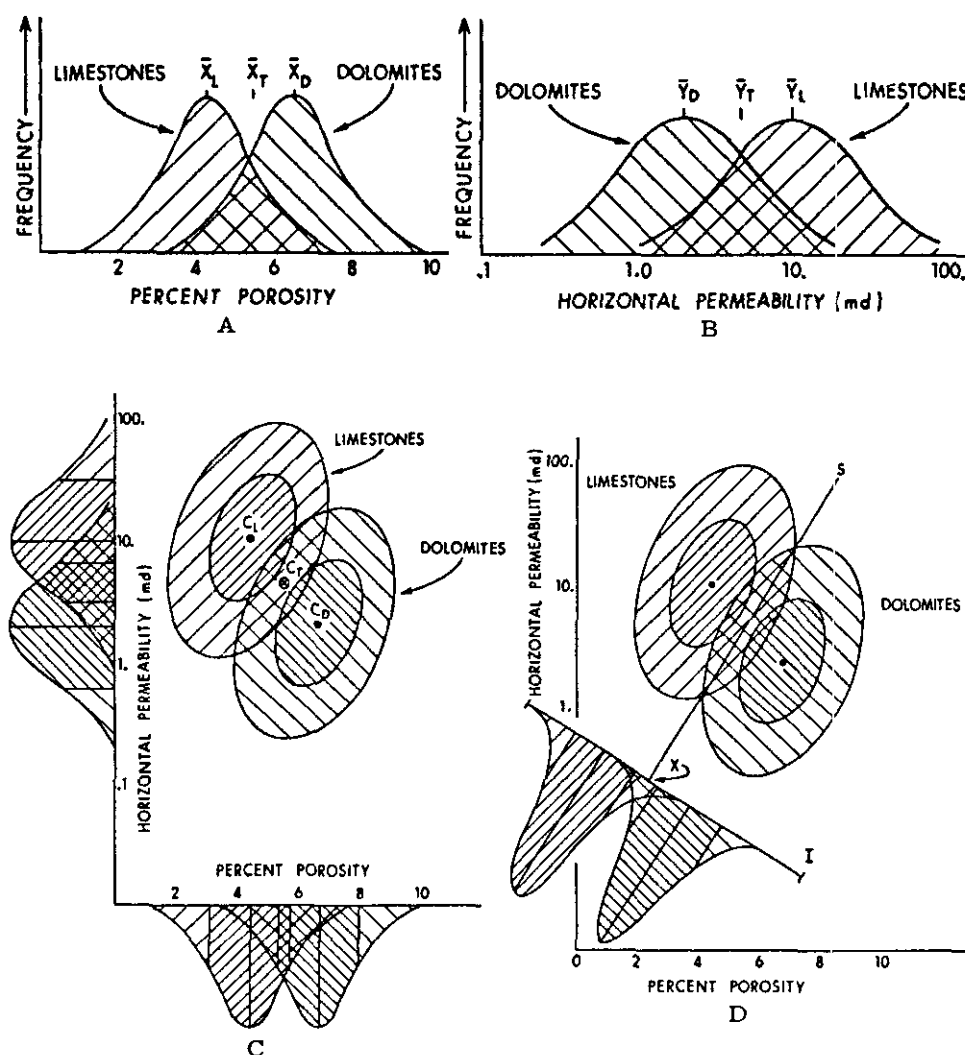


Figure 5. Conceptual model of discriminant function analysis as a more efficient means of separating clusters in two-dimensional space than either of the variables alone; from Klovan and Billings (1967).

functions were calculated between adjoining data clusters. These functions were then used to classify deposits into high-tonnage and low-tonnage categories. Results for both the training set and the target population showed very high proportions of apparently correct classifications (that is, >80 per cent correct) when the discriminant classifications were compared with the corresponding known ore production tonnages.

The misclassifications of particular interest are those thought to be small based on low ore production tonnages, but which appear to be high tonnage according to the discriminant classification. Such deposits are listed in Table 1 for the training set. Of course, these results must be interpreted in light of the validity of the model. Consequently, the model has been applied to the same subset of 19 deposits that formed the basis of a test of the multiple regression model. Results are summarized in Table 2 where it is apparent that more ambiguity exists than is apparent in the multiple regression model. Consequently, results of the discriminant analysis must be viewed cautiously and further validation is desirable.

TABLE 1

Deposits with low known production classed as having high-tonnage potential, by discriminant function analysis of training set deposits, Slocan camp.

Deposit Name	Production	Probability*
Silver Ridge	395	0.99
Grey Copper	66	0.86
Freddy Lee	817	0.91
Deadman	616	0.73
Leadsmith-Noonday	383	0.72
Echo and Graphic	896	0.69

*The probability that a deposit falls within the high-tonnage category according to a discriminant model based on average grades of Ag, Pb, Zn, and Au.

TABLE 2

Low-tonnage (<1,000 tons) and medium-tonnage (1,000-1,600 tons) deposits in target production classed as having high-tonnage potential (>16,000 tons) by discriminant model, Slocan camp

Deposit Name	Recorded Production (short tons)	Probability*
Doherty	6,097	1.00
Black Fox	1,577	1.00
Highland Surprise	3,027	0.96
Jackson	6,314	0.98
Gibson-Daybreak	676	0.92
Flint	351	0.85
Wellington	1,961	0.74

*Probability that the deposit falls in the high-tonnage category according to the discriminant model.

CONCLUSIONS

Statistical modelling of relative value of vein deposits appears an attainable goal providing data of adequate quality are available. The study reported here is only a small part of much more extensive research into exploration modelling. The principal application of statistical models is in (1) identifying deposits with high tonnage potential which, to date, have produced only small tonnages, and (2) evaluating newly found deposits in the camp to which a model applies. Multiple regression appears a more appropriate approach to exploration modelling in Slocan camp than does discriminant analysis. Gold assays should form an integral part in the evolution of all deposits in the Slocan camp.

ACKNOWLEDGMENTS

Funds for this work were obtained from Prism Resources Ltd., the British Columbia Ministry of Energy, Mines and Petroleum Resources, and the Geological Survey of Canada. Technical assistance has been provided at various times by J.F.W. Orr (deceased) and A. Bentzen.

REFERENCES

- Godwin, C.I. and Sinclair, A.J. (1979): Application of Multiple Regression Analysis to Drill Target Selection, Casino Porphyry Copper-Molybdenum Deposit, Yukon Territory, C.I.M., Transactions, Vol. 88, pp. B93-B106.
- Harbaugh, J.W., Doveton, J.H., and Davis, J.C. (1977): Probability Methods in Oil Exploration, John Wiley and Sons, New York, 269 pp.
- Kelly, A.M. and Sheriff, W.J. (1969): A Statistical Examination of the Metallic Mineral Resources of British Columbia, Symposium on Decision-making in Mineral Exploration, II, Extension Dept., University of British Columbia, pp. 221-243.
- Klován, J.E. and Billings, G.K. (1967): Classification of Geological Samples by Discriminant-function Analysis, Bull., Cnd. Pet. Soc., Vol. 15, No. 3, pp. 313-330.
- Orr, J.F.W. and Sinclair, A.J. (1971): A Computer-processible File for Mineral Deposits in the Slocan and Slocan City Areas of British Columbia, Western Miner, Vol. 44, pp. 22-34.
- Sinclair, A.J. (1974): Probability Graphs of Ore Tonnages in Mining Camps - A Guide to Exploration, C.I.M., Bull., Vol. 67, pp. 71-75.
- (1979): Preliminary Evaluation of Summary Production Statistics and Location Data for Vein Deposits, Slocan, Ainsworth and Slocan City Camps, Southern British Columbia, Geol. Surv., Canada, Current Research, Part B, Paper 79-1B, pp. 173-178.
- Sinclair, A.J. and Woodsworth, G.J. (1970): Multiple Regression as a Method of Estimating Exploration Potential in an Area near Terrace, B.C., Econ. Geol., Vol. 68, pp. 998-1003.