Mathematical Geology, Vol. 24, No. 2, 1992

Estimating a Favorability Equation For the Integration of Geodata and Selection of Mineral Exploration Targets[1]

Guocheng Pan[2] and D. P. Harris[3]

The notion of a favorability function for delineation of exploration targets has attracted attention among geologists and geomathematicians over the last decade, as indicated by the number of publications on this subject. Traditional estimation methods for a favorability equation carry several ambiguities in the meaning of the estimate. In order to avoid these problems, a special type of geological variable, referred to as the target variable appears in a favorability equation. Explanatory variables are usually physical and chemical descriptors of geologic objects, while target variables are usually available only in best explored regions. A favorability function should be defined as a linear combination of the explanatory variables, while the meaning of the function should be in terms of the target variables. Two objective methods, canonical correlation and weighted canonical correlation, are proposed in this paper. The estimation of a favorability equation by these methods is predicted upon a criterion that maximizes the correlation of the estimate of the favorability function and the target variables. Both methods are demonstrated on a case study of epithermal gold-silver vein deposits in the 2° Walker Lake quadrangle of Nevada and California. Targets for mineral exploration of gold-silver deposits were identified on the basis of the favorability functions by means of optimum discretization.

KEY WORDS: exploration target, favorability equation, weighted canonical correlation, gold-silver deposit, resource assessment.

INTRODUCTION

Multiple regression models have been widely used in geodata analysis, either separately or jointly with other statistical methods (Harris, 1965; Agterberg and Gabilio, 1969; Jones 1972; McCammon, 1973; Ethridge and Davies, 1973; Mark and Church, 1977; Davis, 1986). Let Y be a random dependent variable and $X_1, X_2, \ldots, X_m$ be m independent variables. The basic linear regression model is

[1] Received 20 June 1990; accepted 5 June 1991.

[2] NERCO Exploration Company, 8100 NE Parkway Drive, Vancouver, Washington 98668.
[3] Department of Mining and Geological Engineering, University of Arizona, Tucson, Arizona 85721.

0882-8121/92/0200-0177$06.50/1

$$Y = \beta_0 + \beta_1 X_1 + \ldots + \beta_m X_m + \mu$$

where μ is a random error satisfying certain statistical conditions. The requirement of a dependent variable is necessary in some cases when prediction of a measure is of primary interest, but it is not required in many other cases, for example, when the estimation of favorability of a region for mineral exploration is a major concern. In such cases, one may be more interested in the following model

$$F = a_1 X_1 + a_2 X_2 + \ldots + a_m X_m \tag{1}$$

where F is called *favorability function* in terms of random variables $X_1, \ldots, X_m$, and F is meaningful with respect to certain stated objectives. In mineral exploration, F would characterize a particular type of mineralization given a collection of observations. F is sometimes regularized to the interval $[-1, 1]$ with 1 indicating the region most favorable to the mineralization of interest, and -1 representing most unfavorable.

A slightly generalized form of Eq. (1) is

$$F = X\mathbf{W}\mathbf{a} = a_1 w_1 X_1 + a_2 w_2 X_2 + \ldots + a_m w_m X_m \tag{2}$$

where $\mathbf{W} = \text{diag}(w_1, w_2, \ldots, w_m)$ with w_j being an *a priori* selected weight for X_j, $X = (X_1, X_2, \ldots, X_m)$ and $\mathbf{a} = (a_1, a_2, \ldots, a_m)^T$. *A priori* information which should be independent of the information contained in the input data is not always available. An appropriate description of such information requires a thorough understanding of the geological features and their intrinsic relations to the mineralization processes.

Several methods have been developed for the estimation of a favorability equation for the selection of exploration targets. Characteristic analysis (McCammon et al., 1983; Pan and Wang, 1987) involves the estimation of Eq. (1) by the principal components method. Reddy and Koch (1988a,b) estimated Eq. (1) by a subjective method based on the weighting factors from control sample units. Each of these weighting factors was determined using *a priori* information by recording the presence of transformed digitized variables in the control cells as a fraction of all the control cells. The most recent work was given by Luo (1990), who proposed a method for estimating Eq. (1) without defining a control area.

As these methods present possible ambiguities in the meaning of the favorability function, Harris and Pan (1987) and Pan (1989) proposed a method to improve the estimate, called weighted and targeted multivariate criterion. This paper examines some alternative objective methods of estimating Eqs. (1) or (2) by seeking a means of enhancing important information about the mineralization. Two such methods, *canonical correlation* and *weighted canonical correlation*, are proposed. A major goal of these methods is to explicitly identify and incorporate a set of special geological features ($Z_1, \ldots, Z_p$) that provide firm evidence about the mineralization of interest. This reduces ambiguities in the favorability function estimated by traditional methods, and hopefully captures much of the useful information. These methods are demonstrated by a case study of exploration target areas for epithermal gold–silver deposits in the Walker Lake quadrangle of Nevada and California. The regions being potentially favorable to epithermal gold–silver mineral deposits are identified on the basis of the optimum favorability functions by means of optimum discretization.

BRIEF REVIEW

Characteristic analysis (McCammon et al., 1983; Pan and Wang, 1987) estimates Eq. (1) by means of a principal component method. Let **S** be the similarity matrix of $X = (X_1, X_2, \ldots, X_m)$. The estimate of Eq. (1) by characteristic analysis is given by

$$\hat{F} = X\mathbf{e} \tag{3}$$

where **e** is the unit eigenvector associated with the largest eigenvalue of **S**. The magnitude of the coefficient, e_j, measures the importance of the jth variable, X_j, to function F. The characteristic analysis assumes that geological characteristics are expressed as ternary or binary variables. One appealing feature of the method is that it can quantitatively integrate multiple forms and kinds of data, such as various modes of occurrence as well as qualitative geological variables which are often readily accessible but difficult to quantify.

The principal component with the largest eigenvalue explains the largest portion of the total variability contained in a data structure; but unless the variables are purposefully chosen, they may not necessarily contain any meaningful information about a particular type of mineralization. However, characteristic analysis is useful for the integration of mineral exploration data due to its particular coding design. For example, when geological characteristics are coded in ternary form, a variable having a value of positive one at a location means that its presence is favorable to the mineralization at that location, while a variable taking a negative one implies that the presence of this variable is unfavorable. Accordingly, the implications of F in Eq. (3) are partially implied *a priori* by the data coding requirements. Determination of which values of a variable would represent favorable or unfavorable state (1 or -1) is still subjective.

The estimation of the favorability function given by Luo (1990) is essentially a constrained least squares analysis. The criterion of his approach is to minimize the variance Var $(F) = E[F - \mu_F]^2$ with constraint $\Sigma_{j=1}^{m} a_j \mu_j = \mu_F$, where $\mu_F = E(F)$ and $\mu_j = E(X_j)$. The estimate for Eq. (1) is given by

$$\hat{F} = X\mathbf{C}_{XX}^{-1}\boldsymbol{\mu} \tag{4}$$

where $\mathbf{\mu} = (\mu_1, \ldots, \mu_m)^T$ and $\mathbf{C}$ is the covariance matrix. This approach is subject to possible ambiguities in the meaning of the estimate of F, for the "optimization" is solely predicated upon the internal statistical structure (covariance matrix) of X. In other words, the function F estimated above is useful to characterize a typical linear combination with the minimized variance of the favorability function, but not necessarily meaningful as being a "favorability indicator" for new mineral discoveries, unless the variables are appropriately pre-selected.

Luo's method was developed without using a control area. However, it is necessary to select control areas when the study region is large, because of the non-homogeneities in the degree of mineral exploration and variability of geological settings as some sub-regions might be better explored than others. A choice between using or not using control areas seems to be equivalent to the choice between preventing losses of useful information and avoiding difficulties in selecting control areas. Without defining control areas, one has to ignore those variables that are useful but available only in a small part of the region.

The weighted and targeted multivariate criterion proposed by Harris and Pan (1987) and Pan (1989) may be considered as a variation of the principal components method. Enhancements of useful information about mineral resources are sought by maximizing the correlation between geological and preselected target variables which characterize the mineralization of interest. The techniques developed in this paper are predicated upon a similar idea, i.e., seeking an "external" enhancement.

EXPLANATORY AND TARGET VARIABLES

In regional mineral resources assessment, it is usually possible to identify two types of geological features. The first type, referred to as *explanatory variables*, consists of the geological measurements, $X_1, X_2, \ldots, X_m$, which are primary to Eq. (1). They are usually chemical and physical descriptors of geologic objects, which are generally observable throughout the entire study region by conventional sensing technologies. Stream sediment geochemical measurements, gravity observations, and intensity of magnetic flux are a few examples. The second type, referred to as *target variables*, includes mineral resource descriptors, critical genetic factors, or those directly related to these descriptors. These features are usually known only in the best explored parts of the study region and are unevaluated in the rest of the region. Their precision is highly dependent upon exploration levels. They usually characterize the major objective of a study and are important criteria for the selection of control areas. Table 1 contains an example of the two types of geological variable identified for epithermal gold–silver vein deposits in the Walker Lake quadrangle of Nevada and California. The variables in the table are explained in a later section.

Table 1. The Explanatory and Target Variables for Epithermal Gold–Silver Deposits in the Walker Lake Quadrangle

	Variable name	Geological implication
Explanatory variables set	$X1$	Synthesized geochemical fields
	$X2$	Synthesized structural fields
	$X3$	Band-pass gravity fields
	$X4$	Band-pass magnetic fields
	$X5$	Ratio of geophysical fields
	$X6$	Correlation of geophysical fields
	$X7$	Host rocks of Au–Ag deposits
	$X8$	Estimate of depth to intrusives
Target variables set	$Z1$	Hydrothermal alterations
	$Z2$	Tertiary intrusives
	$Z3$	Au–Ag mines or prospects

Since the two sets of variables are not equal in terms of their contributions to the prediction of mineral targets or resources, their treatment in a lump would mask and even result in losses of important information. Ideally, these two different types of data should be identified, separated, and then correlated. Equation (1) contains only explanatory variables, while the meaning of the favorability function is characterized by the information carried by target variables. The optimum estimation of favorability Eq. (1) is obtained by maximizing the correlation between the explanatory and target variables.

CANONICAL CORRELATION METHOD

Let $X = (X_1, X_2, \ldots, X_m)$ and $Z = (Z_1, Z_2, \ldots, Z_p)$ be two vectors of m explanatory and p target variables, respectively. Consider the linear combination

$$G = Z\mathbf{b} = b_1 Z_1 + \ldots + b_p Z_p$$

with $\mathbf{b} = (b_1, \ldots, b_p)^T$. Then, the coefficients $\mathbf{a}$ in Eq. (1) are selected such that the correlation of F and G is maximized for all possible real values of $\mathbf{b}$. Clearly, this is a canonical correlation problem. Since the standard form of canonical correlation analysis for two sets of variables is well-known (see Mardia et al., 1979; Seber, 1984; Johnson and Wichern, 1988), only the major results are summarized below.

Let $\mathbf{C}_{XX} = \text{Cov}\,(X^T, X)$, $\mathbf{C}_{ZZ} = \text{Cov}\,(Z^T, Z)$, $\mathbf{C}_{XZ} = \text{Cov}\,(X^T, Z)$ and $\mathbf{C}_{ZX} = \mathbf{C}_{XZ}^T$. The objective is now to find $\mathbf{a}$ and $\mathbf{b}$ such that $\text{Corr}\,(F, G) = \mathbf{a}^T\mathbf{C}_{XZ}\mathbf{b}/\sqrt{(\mathbf{a}^T\mathbf{C}_{XX}\mathbf{a})(\mathbf{b}^T\mathbf{C}_{ZZ}\mathbf{b})}$ is maximized; this leads to the following main result

$$\mathbf{C}_{XX}^{-1/2}\mathbf{C}_{XZ}\mathbf{C}_{ZZ}^{-1}\mathbf{C}_{ZX}\mathbf{C}_{XX}^{-1/2}\,\mathbf{d}_j = \lambda_j\,\mathbf{d}_j \tag{5}$$

$$\mathbf{C}_{ZZ}^{-1/2}\mathbf{C}_{ZX}\mathbf{C}_{XX}^{-1}\mathbf{C}_{XZ}\mathbf{C}_{ZZ}^{-1/2}\,\mathbf{e}_j = \lambda_j\mathbf{e}_j \tag{6}$$

The best solution for Eq. (1) is then given by

$$\hat{F} = X\,\mathbf{C}_{XX}^{-1/2}\,\mathbf{d}_1 \tag{7}$$

where $\mathbf{d}_1$ is the unit eigenvector associated with the largest eigenvalue λ_1. Alternatively, the favorability function F may be defined as G predicted through regressing the first canonical variable $\hat{G} = Z\mathbf{C}_{ZZ}^{-1/2}\,\mathbf{e}_1$ of the target variables on $\hat{F}$, the first canonical variable of m explanatory variables, by least squares method. Interpretation of the estimate may be aided by computing the correlation coefficients between $\hat{F}$ and each of explanatory and target variables (Johnson and Wichern, 1988).

Derivation of Eq. (7) considers only the first pair of canonical variables in Eqs. (5) and (6). It may be useful to ask whether the estimate in Eq. (7) is significantly correlated with the first canonical variate of the target variables. Several large sample tests are available (Seber, 1984; Fujikoshi, 1977), but each of them is established on normality assumption. One of the most commonly used tests is the test that $p - q$ $(q \leq p)$ of the smallest correlations are zero, that is, H_{0q}: $\rho_{q+1} = \rho_{q+2} = \ldots = \rho_p = 0$ $(\rho_q > 0)$. When H_{0q} is true, the canonical variables $(F_j, G_j)(j = q + 1, \ldots, p)$ have no predictive value for comparing X and Z, so that the relationship between $X_1, \ldots, X_m$ and $Z_1, \ldots, Z_p$ can be summarized by means of the first q canonical variables. Since we are concerned only with the first pair of canonical variables, set $q = 0$ and then test H_0: $\rho_1 = \ldots = \rho_p = 0$. For this special case, the likelihood ratio test for testing H_0 leads to the statistic (Bartlett, 1947; Seber, 1984; Davis, 1986)

$$\phi = (1 - \rho_1^2)(1 - \rho_2^2)\ldots(1 - \rho_p^2) \tag{8}$$

and $-2n \log \phi$ is asymptotically χ_v^2 with $v = mp$. Unfortunately, the statistic ϕ is not robust to departures from normality, particularly to the presence of outliers or long-tailed distributions.

A WEIGHTED CANONICAL CORRELATION METHOD

In the canonical correlation analysis, correlation between the two sets $X_1, \ldots, X_m$ and $Z_1, \ldots, Z_p$ has been isolated in the pairs of canonical variables. By design, the coefficient vector $\mathbf{a}$ in Eq. (1) is selected in accordance with the maximization of the correlation between F and G, where G is a particular linear combination of $Z_1, \ldots, Z_p$. This criterion, however, does not necessarily provide estimates that account for the covariances between target variables, $Z_1, \ldots, Z_p$. The canonical variable G characterizes a certain direction in the target variable space, but not necessarily the direction with the maximum variability.

A simple numerical example may well illustrate this point (Johnson and Wichern, 1988). Let $X = (X_1, X_2)$ and $Z = (Z_1, Z_2)$ and their covariance matrix be

$$\mathbf{C} = \begin{pmatrix} \mathbf{C}_{XX} & \mathbf{C}_{XZ} \\ \mathbf{C}_{ZX} & \mathbf{C}_{ZZ} \end{pmatrix} = \begin{bmatrix} 100 & 0 & 0 & 0 \\ 0 & 1 & 0.95 & 0 \\ 0 & 0.95 & 1 & 0 \\ 0 & 0 & 0 & 100 \end{bmatrix}$$

It can be readily calculated that $\hat{F} = X_2$ and $\hat{G} = Z_1$, which have correlation of 0.95. However, $\hat{G}$ provides a very poor summary of the variability in the target variable set. Most of the variability in this set is in Z_2, which is uncorrelated with $\hat{G}$ and $\hat{F}$. Thus, the variability in the target variables is not satisfactorily captured by the estimated favorability function $\hat{F}$.

To characterize the largest variability in target variables, the coefficient vector $\mathbf{b}$ in canonical variate G may be pre-determined by the principal components method based upon solely the covariance matrix of target variables ($\mathbf{C}_{ZZ}$). Furthermore, the variability in a large data set may not be sufficiently represented by only one principal component, but may be approximated by several components. Therefore, instead of using only a single principal component, typically, q components ($q \leq p$) are simultaneously considered when $\mathbf{a}$ is estimated. To be more general, Eq. (2) will be considered in the following development:

Let Y_j be the jth principal component of the covariance matrix $\mathbf{C}_{ZZ}$, i.e.,

$$Y_j = Z\mathbf{b}_j = b_{1j}Z_1 + b_{2j}Z_2 + \ldots + b_{pj}Z_p, \qquad j = 1, 2, \ldots, p$$

Var $(Y_j) = \lambda_j$ (the jth largest eigenvalue of $\mathbf{C}_{ZZ}$) and Cov $(Y_j, Y_k) = 0$ $(j \neq k)$. Suppose that the first q $(\leq p)$ principal components account for a sufficiently large proportion of the total variability in p target variables. Let $\mathbf{B} = (\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_q)$ be the $p \times q$ matrix containing q eigenvectors associated with the first q largest eigenvalues of $\mathbf{C}_{ZZ}$. Denote $Y = (Y_1 \ldots, Y_q) = Z\mathbf{B}$.

Consider the criterion to determining coefficient vector $\mathbf{a}$ by maximizing the sum of the squared covariance between the favorability function $F = X\mathbf{Wa}$ and each of the q principal components, Y_j. Let

$$\phi^2(\mathbf{a}) = \sum_{j=1}^{q} \text{Cov}^2\,(Y_j, F) \tag{9}$$

Cov $(Y_j, F) =$ Cov $(Z\mathbf{b}_j, X\mathbf{Wa}) = \mathbf{b}_j^T\mathbf{C}_{ZX}\mathbf{Wa}$. Thus, Eq. (9) is rewritten as

$$\phi^2(\mathbf{a}) = \sum_{j=1}^{q} \mathbf{a}^T\mathbf{W}\mathbf{C}_{XZ}\mathbf{b}_j\mathbf{b}_j^T\mathbf{C}_{ZX}\mathbf{Wa} = \mathbf{a}^T\mathbf{Ha} \tag{10}$$

where $\mathbf{H} = \mathbf{W}\mathbf{C}_{XZ}\mathbf{B}\mathbf{B}^T\mathbf{C}_{ZX}\mathbf{W}$. The variance of F is given by

$$\alpha^2(\mathbf{a}) = \text{Var}(F) = \mathbf{a}^T\mathbf{G}\mathbf{a} \tag{11}$$

where $\mathbf{G} = \mathbf{W}\mathbf{C}_{XX}\mathbf{W}$. Using Eqs. (10) and (11) and a Lagrange multiplier, construct an objective function

$$L(\mathbf{a}, \mu) = \mathbf{a}^T\mathbf{H}\mathbf{a} - \mu(\mathbf{a}^T\mathbf{G}\mathbf{a} - 1)$$

The coefficients $\mathbf{a}$ are chosen such that L is maximized. Taking the first-order partial derivatives of function L with respect to $\mathbf{a}$ and μ, and setting them to zero, we have

$$\mathbf{H}\mathbf{a} = \mu\mathbf{G}\mathbf{a}, \quad \mathbf{a}^T\mathbf{G}\mathbf{a} = 1 \tag{12}$$

If $\mathbf{G}$ is positive definite, then $\mathbf{G} = \mathbf{G}^{1/2}\mathbf{G}^{1/2}$. Thus, Eq. (12) is rewritten as

$$\mathbf{H}_g\mathbf{e} = \mu\mathbf{e} \tag{13}$$

where $\mathbf{H}_g = \mathbf{G}^{-1/2}\mathbf{H}\mathbf{G}^{-1/2}$ and $\mathbf{e} = \mathbf{G}^{1/2}\mathbf{a}$. Hence, the estimate of Eq. (2) is given by

$$\hat{F} = X\mathbf{W}\mathbf{G}^{-1/2}\mathbf{e} \tag{14}$$

where $\mathbf{e}$ is the unit eigenvector associated with the largest eigenvalue of $\mathbf{H}_g$.

The interpretation of the estimates can be aided by computing the correlation coefficients between the estimated favorability function and explanatory and target variables as well as the principal components of the target variables. The covariances between F and X are

$$\text{Cov}(X, F) = \text{Cov}(X, X\mathbf{W}\mathbf{a}) = \mathbf{C}_{XX}\mathbf{W}\mathbf{a}$$

Given the best estimate of $\mathbf{a}$ (meaning that $\text{Var}(F) = \mathbf{a}^T\mathbf{G}\mathbf{a} = 1$), the correlation coefficient vector between F and X is

$$\text{Corr}(X, F) = \mathbf{C}^{-1/2}\mathbf{C}_{XX}\mathbf{W}\mathbf{a} \tag{15}$$

where $\mathbf{C} = \text{diag}(\sigma_1^2, \ldots, \sigma_m^2)$ with $\sigma_j^2 = \text{Var}(X_j)$. The correlation coefficient between F and X_j is

$$\text{Corr}(X_j, F) = \sum_{k=1}^{m} w_k a_k \sigma_k \rho_{jk}$$

where $\rho_{jk} = \text{Corr}(X_j, X_k)$. Thus, the correlation between the favorability function F and the jth explanatory variable is a weighted sum of the correlation coefficient between X_j and each of the explanatory variables. A larger coefficient a_k or weight w_k would suggest a bigger contribution from the kth explanatory variable to the correlation between F and X_j. This implies that interdependencies of the explanatory variables are accounted for in the quantification of F.

The covariance between F and Z is

$$\text{Cov}(Z, F) = \text{Cov}(Z, X\mathbf{W}\mathbf{a}) = \mathbf{C}_{ZX}\mathbf{W}\mathbf{a} \tag{16}$$

Given the optimum estimate of **a**, their correlations can be computed by

$$\text{Corr}\,(Z, F) = \Gamma^{-1/2}\mathbf{C}_{ZX}\mathbf{Wa} \tag{17}$$

where $\Gamma = \text{diag}\,(\gamma_1^2, \ldots, \gamma_q^2)$ with $\gamma_k^2 = \text{Var}\,(Z_k)$. Furthermore, the correlation coefficient between F and Z_j is

$$\text{Corr}\,(Z_j, F) = \sum_{k=1}^{m} w_k a_k \gamma_k \alpha_{jk}$$

where $\alpha_{jk} = \text{Corr}\,(Z_j, Z_k)$. This shows that larger weights a_k and w_k determine a larger contribution of the kth target variable to the correlation between the favorability function and the jth target variable.

Similarly, the covariance between F and Y is

$$\text{Cov}\,(Y, F) = \text{Cov}\,(\mathbf{Z}\mathbf{B}, X\mathbf{Wa}) = \mathbf{B}^T\mathbf{C}_{ZX}\mathbf{Wa}$$

$$\|\text{Cov}\,(Y, F)\|^2 = \mathbf{a}^T\mathbf{W}\mathbf{C}_{XZ}\mathbf{B}\mathbf{B}^T\mathbf{C}_{ZX}\mathbf{Wa} = \mathbf{a}^T\mathbf{Ha} = \mu_{\max}$$

where $\mu_{\max}$ is the largest eigenvalue of matrix **H** given the optimum estimate of **a**. The correlation coefficient vector between F and Y is given by

$$\text{Corr}\,(Y, F) = \mathbf{\Lambda}^{-1/2}\mathbf{B}^T\mathbf{C}_{ZX}\mathbf{Wa} \tag{18}$$

where $\mathbf{\Lambda} = \text{diag}\,(\lambda_1, \ldots, \lambda_q)$ with $\lambda_j = \text{Var}\,(Y_j)$.

Furthermore, given the best estimate of **a**, the ϕ^2 in Eq. (9) is rewritten as

$$\phi^2 = \sum_{j=1}^{q} \text{Cov}^2\,(Y_j, F) = \sum_{j=1}^{q} \lambda_j \rho_{jF}^2 = \mu_{\max}$$

where $\rho_{jF} = \text{Corr}\,(Y_j, F)$. Thus, maximization of ϕ^2 with constraint $\mathbf{a}^T\mathbf{Ga} = 1$ is equivalent to the maximization of the weighted sum of the squared correlation coefficients between the favorability function and each of the q principal components of target variables. The weights λ_j distinguish different contributions for different components. The maximization gives more consideration to those principal components characterizing the directions with larger variabilities in the target variable space and suppresses others with smaller variabilities.

Now consider the sample estimate for Eq. (2). Suppose that both sets of explanatory and target variables are observed on a control sample of size n. Let $\mathbf{X} = (\mathbf{x}_1, \ldots, \mathbf{x}_m)$ be an $n \times m$ data matrix for explanatory variables and $\mathbf{Z} = (\mathbf{z}_1, \ldots, \mathbf{z}_p)$ is an $n \times p$ data matrix for target variables. On this sample, let $\mathbf{f} = (f_1, \ldots, f_n)^T$ be the vector of realizations of F given the set of observations on the m explanatory variables. Thus, $\mathbf{f} = \mathbf{XWa}$.

Furthermore, the first q sample principal components of the p target variables are denoted by $\mathbf{Y} = (\mathbf{y}_1, \ldots, \mathbf{y}_q) = \mathbf{ZU}$, where $\mathbf{U} = (\mathbf{u}_1, \ldots, \mathbf{u}_q)$ with $\mathbf{u}_j$ being the unit eigenvector associated with the jth largest eigenvalue of the sample covariance matrix between p target variables. Define

$$\phi_n^2(\mathbf{a}) = \sum_{j=1}^{q} (\mathbf{y}_j^T\mathbf{Qf})^2 = \mathbf{a}^T\mathbf{Sa}$$

where $\mathbf{Q} = \mathbf{I} - (1/n)\mathbf{J}$ with $\mathbf{J} = (1)_{n\times n}$ and $\mathbf{S} = \mathbf{WX}^T\mathbf{QZUU}^T\mathbf{Z}^T\mathbf{QXW}$. The sample variance of F is $\alpha_n^2(\mathbf{a}) = \mathbf{a}^T\mathbf{Da}$ with $\mathbf{D} = \mathbf{WX}^T\mathbf{QXW}$. The coefficient vector $\mathbf{a}$ is determined by maximizing ϕ_n^2 subject to the constraint $\alpha_n^2 = 1$, leading to the following

$$\mathbf{Sa} = \mu\mathbf{Da} \tag{19}$$

Hence, the estimate of Eq. (2) based on this sample is given by

$$\hat{\mathbf{f}} = \mathbf{XWD}^{-1/2}\mathbf{e} \tag{20}$$

where $\mathbf{e}$ is the unit eigenvector associated with the largest eigenvalue of matrix $\mathbf{D}^{-1/2}\mathbf{SD}^{-1/2}$. Similarly, the sample correlations between F and X, Z, or Y can be computed on the basis of the sample estimates.

A CASE STUDY

The Walker Lake quadrangle, which comprises the area between 38° and 39° North latitude and 118° and 120° West longitude and includes parts of the states of California and Nevada (Fig. 1), is located at the western edge of the Great Basin. Epithermal gold–silver vein deposits of the Walker Lake quadrangle serve as a case study for the demonstration of the methods proposed above.

General Features of the Walker Lake Quadrangle

John et al. (1989), who examined timing and stratigraphic distribution in the Paradise Range, pointed out that the general age range of the neighboring regions, including the Walker Lake quadrangle, is very similar to the ages of

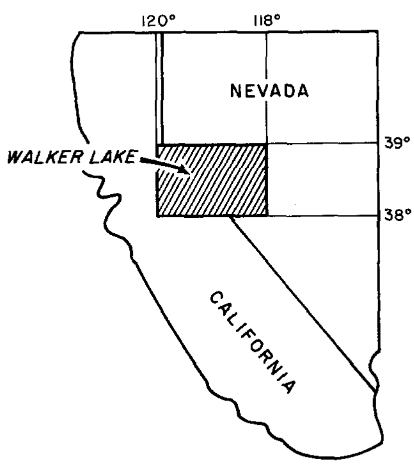

Fig. 1. The index map of the Walker Lake quadrangle, California and Nevada.

other well-studied sequences of late Tertiary rocks. In these adjacent areas, the extrusion is dominantly andesitic and thyodacitic intermediate lavas formed about 20–19 Ma. The geochemical samples from the Yerington district to the north of the Walker Lake quadrangle showed that the intermediate lavas are mainly calc-alkaline to calcic. Recent studies indicate that faulting, titling, dike orientations, and formation of late Tertiary sedimentary basins extended across several periods of Late Cenozoic in many parts of the Great Basin. Earlier extensions tends to have closely spaced normal faults and dominantly intermediate composition magmatism (John, 1986; John et al., 1989). There is a strong northwest-trending structural pattern in the northeastern part of the Walker Lake quadrangle. A prominent northwest-trending lineament occurs in the southern part of the region, along which a number of precious metal deposits were localized.

Silberman et al. (1976) (also see Mckee and Noble, 1986) noted that the Walker Lane is a major tectonic feature in the Walker Lake quadrangle and contains a large number of late Tertiary-age precious (gold and silver) metal deposits. Stewart (1988) divided the Walker Lane into eight geological blocks according to the tectonic characteristics. Three major blocks are entirely or partly included in the Walker Lake quadrangle. The Walker Lake block contains the most important precious metal deposits and provides critical information about the relationships of volcanism, pre-Basin and Range extensional faulting, and precious metal mineralization.

Most silicic and alunitic alteration is located along inferred high-angle faults within areas where low-angle faults are absent. In contrast, little hydrothermal alteration is associated with low-angle faults. In the Gabbs Valley Range, only throughgoing high-angle faults are mineralized, whereas detachment faults and listric normal faults are unmineralized (John et al., 1989). These relations suggest that circulation of hydrothermal fluids, which resulted in hydrothermal alteration and local precious metal mineralization, were restricted to areas of deep, possibly recurrent, fracturing. During the Cenozoic, igneous activity, largely extrusive, produced varied assemblages of volcanic rocks, including bimodal basalt-rhyolite sequences. In this region, intermediate lavas are the most important host rocks for late Tertiary-age volcanic-hosted precious metal deposits.

A number of types of mineralization have been found in the Walker Lake quadrangle, including epithermal precious vein deposits, porphyry and skarn deposits, and placer deposits. More than seven different metals are found in the mineralization: gold, silver, copper, lead, zinc, tungsten, and iron. Mineralization of different types appear to be zoned. Major porphyry and skarn copper deposits are distributed chiefly in the northern part of the Walker Lake, whereas the bulk of epithermal gold–silver vein deposits are found mainly in the southern and western parts of the region. In addition to these two typical belts of mineralization, a transitional zone, where mineralization of different types cooccur, lies between the two belts, mainly containing silver (gold)–lead and copper–gold (silver) (Harris and Pan, 1991). Strong hydrothermal alterations, chiefly argillization and silicification, are spatially associated with the occurrence of gold–silver vein deposits.

Data

The map folio for the 2° Walker Lake quadrangle published by the U.S. Geological Survey, supplemented by various larger scale U.S. Geological Survey maps, served as the sources for data on mineral occurrences, gold–silver deposits and mining districts, and lithology. A map of hydrothermal alteration, produced for the Walker Lake quadrangle through field evaluation of a limonite map compiled from digitally processed images from Landsat Satellite Multispectral Scanner (Rowan and Purdy, 1984), was the source of alteration information.

The digital data of faults in the Walker Lake quadrangle prepared by the U.S. Geological Survey were used for structural analysis. Geochemical data (Chaffee et al., 1980) on the Walker Lake quadrangle consisted of 1116 stream-sediment samples analyzed for Fe, Mg, Ca, Ti, Mn, As, Au, B, Ba, Bi, Cd, Co, Cr, Cu, La, Mo, Nb, Ni, Pb, Sr, V, N, Y, Zn, and Th using a six-step semiquantitative emission spectrographic method; the elements with low detection levels, Zn, Sb, Cd, Bi, and Au, also were analyzed by atomic absorption spectrometric analysis; these data were obtained from a magnetic tape through NTIS. Gravity data consisting of 3447 irregularly distributed measurements of isostatic gravity residuals (Plouff, 1982) were obtained from the U.S. Geological Survey. Several sets of digital magnetic data were collected to provide coverage of the Walker Lake quadrangle (Kucks and Hildenbrand, 1987; USGS, 1979a,b, 1981, 1982).

Processing of these data sets (Harris and Pan, 1987, 1988; Pan, 1989; Pan and Harris, 1991) produced the following eight synthesized explanatory variables:

X_1: This is the geochemical field synthesized from 14 elements (Au, Ag, Cu, Pb, Zn, Fe, Ca, Sb, Zr, V, Bi, Mo, Be, and B) sampled from drainage basins and adjusted for mobility (Pan and Harris, 1990a). The synthesized geochemical scores were produced by a weighted linear combination of the 14 elements. The scores were then filtered for noise.

X_2: This is a high-pass structural field synthesized from ten fault descriptors: numbers and lengths of major and minor faults, fault orientations, numbers of intersections of major and minor faults, etc., within a moving data window. The synthesized structural scores were generated by a weighted linear combination of these descriptors. These scores were then filtered in the frequency domain to remove noise (random error). A final filtering to remove low frequencies produced the high-pass structural field.

X_3: This is the band-pass gravity field derived from original isostatic residuals through three major steps. In the first step, the high-pass data were generated by filtering out low frequencies. Second, the high-pass field was correlated with the filtered geochemical field in the frequency domain. Based upon this analysis, a range of frequencies for which coherencies between the gravity and geochemical data are maximized was determined. Finally, the field within the selected frequency band was selected by filtering out the data associated with other frequencies.

X_4: This variable is the band-pass magnetic field, derived by analysis similar to that applied to gravity data. Densities of several magnetic data sets were adjusted, and then merged into a uniform large set. After reduction to the pole and the filtering out of low frequencies, coherency of the high-pass magnetic and geochemical fields was analyzed. Based on this analysis, a coherent frequency band was selected and the band-pass magnetic field was obtained.

X_5: This variable represents the ratio of magnetic susceptibility contrast to density contrast estimated by a Poisson moving window, based upon high-pass gravity and high-pass magnetic fields. The derivation of the estimates of density and susceptibility contrasts is based upon the concept of pseudo-geophysical fields. Using the Poisson relation, pseudo-gravity fields were generated from the observed magnetic fields. Then, regression analysis was applied to data sets within moving windows to obtain estimates of the ratio of magnetic susceptibility contrast to density contrast.

X_6: This is the value of the statistical correlation of high-pass gravity and high-pass magnetic fields within a data window. It was estimated in the same way as X_5, based upon the Poisson moving window.

X_7: This is the area (km^2) of the cell that consists of host rocks for epithermal gold–silver mineral deposits. The major types of the host rocks include Tertiary volcanic rocks of dacite, rhyolite, and andesite compositions.

X_8: This is the estimate of the depth (in km) from ground to the top surface of a Tertiary intrusive obtained by the inverse analysis of high pass gravity anomalies. The inverse analysis was performed in two-dimensional space by cutting the entire region into a number of parallel traverses along each of the north–south and east–west directions. Then, the final estimates were obtained by averaging the separate estimates along the two directions.

The following three special features are selected as target variables, because they are important descriptors of epithermal gold–silver deposits; the occurrence of any of them at a location assures the existence of a heat source, which is a critical genetic factor for epithermal gold–silver mineralization.

Z_1: This is the number of epithermal gold–silver mineral occurrences (deposits or prospects) within a cell.

Z_2: This is the area (km^2) of Tertiary intrusives that outcrop within a cell.

Z_3: This is the area (km^2) of the cell with hydrothermal alteration (mainly argillaceous and silicious).

The entire Walker Lake quadrangle was subdivided into a 55 $\times$ 55 grid and all 11 variables described above are evaluated on the centers of the cells. To implement analysis, an area containing 300 grid points in the best explored regions, including the Aurora, Bodie, and Masonic mining districts (see Fig. 2) was selected to serve as a control for estimating Eq. (1).

Estimation by the Canonical Correlation Method

Table 2 contains the correlation matrix among the 11 variables, showing that in the control samples, the three target variables are highly intercorrelated. This is not unexpected because all of the target variables relate to a heat source, a critical genetic factor for the hydrothermal deposits. Among the eight explanatory variables, geochemical field and host rock are most highly correlated with the three target variables, meaning that these two explanatory variables should reflect critical variations in the target variables. Among explanatory variables, the two comprehensive geophysical measures (ratio and correlation between magnetization and density contrast) are highly correlated, indicating similarity in both gravity and magnetic characteristics of geologic bodies. In addition, band pass gravity and magnetic fields are also moderately correlated with these measures.

The result of canonical correlation analysis shows that the correlation (0.89)

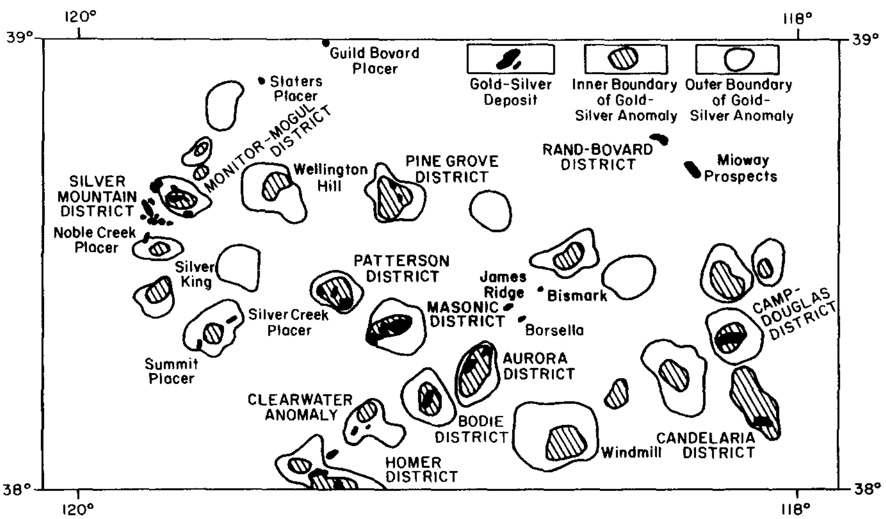

Fig. 2. Distribution of mining districts in the Walker Lake quadrangle (from Pan and Harris, 1990b).

Table 2. Correlation Matrix Between Eight Exploratory and Three Target Variables

	*Z*1	*Z*2	*Z*3	*X*1	*X*2	*X*3	*X*4	*X*5	*X*6	*X*7	*X*8
*Z*1	1.000										
*Z*2	0.505	1.000									
*Z*3	0.637	0.588	1.000								
*X*1	0.347	0.495	0.392	1.000							
*X*2	0.001	−0.037	0.025	−0.255	1.000						
*X*3	0.072	0.053	0.134	0.222	0.040	1.000					
*X*4	0.008	−0.005	0.073	0.069	0.001	0.368	1.000				
*X*5	0.009	0.046	0.024	−0.028	0.055	0.311	0.322	1.000			
*X*6	0.096	0.232	0.022	0.071	0.133	0.251	0.188	0.605	1.000		
*X*7	0.411	0.787	0.491	0.318	0.008	0.044	0.043	0.054	0.202	1.000	
*X*8	0.008	−0.273	−0.080	−0.034	−0.040	0.171	0.024	−0.132	−0.126	−0.275	1.000

between the first pair of canonical variables is highly significant at a level of one percent according to Bartlett's chi-square test. The favorability Eq. (1) is estimated by treating eight explanatory variables and three target variables as the first and second data sets and is given by

$$\hat{f}_i = 0.354x_{i1} + 0.029x_{i2} - 0.030x_{i3} - 0.058x_{i4} - 0.025x_{i5} + 0.105x_{i6} + 0.791x_{i7} - 0.068x_{i8}, \qquad (21)$$

where $\hat{f}_i$ is the estimate of favorability function, F, at location (grid) i. The coefficients for the first canonical variable of the three target variables are 0.029, 0.956, and 0.038, respectively, suggesting that Tertiary intrusives are a core target variable represented by Eq. (21). The values of the first canonical variable for target variables were also computed for the control area. The two sets of canonical variate values are plotted in Fig. 3. The favorability values computed from Eq. (21) for the entire Walker Lake quadrangle were contoured (Fig. 4). Comparing with Fig. 2, most of the known mining districts are associated with the highest favorability values. For example, several areas with outstanding favorability values are consistent with the major mining districts at Candelaria, Pine Grove, Patterson, etc. The mining district at Camp-Douglas is associated with moderate magnitude of favorability values. Several other regions having moderately high favorability values correspond to locations of mineral prospects.

For cross-validation, each of the three target variables was transformed into a binary variable with two possible responses: 1 for presence, i.e., having a value greater than zero and 0 for absence, i.e., having a zero value. Fig. 5 shows the spatial distribution of the logical combination of the three transformed variables. The observed target variables are primarily lined up along northeast and northwest zones.

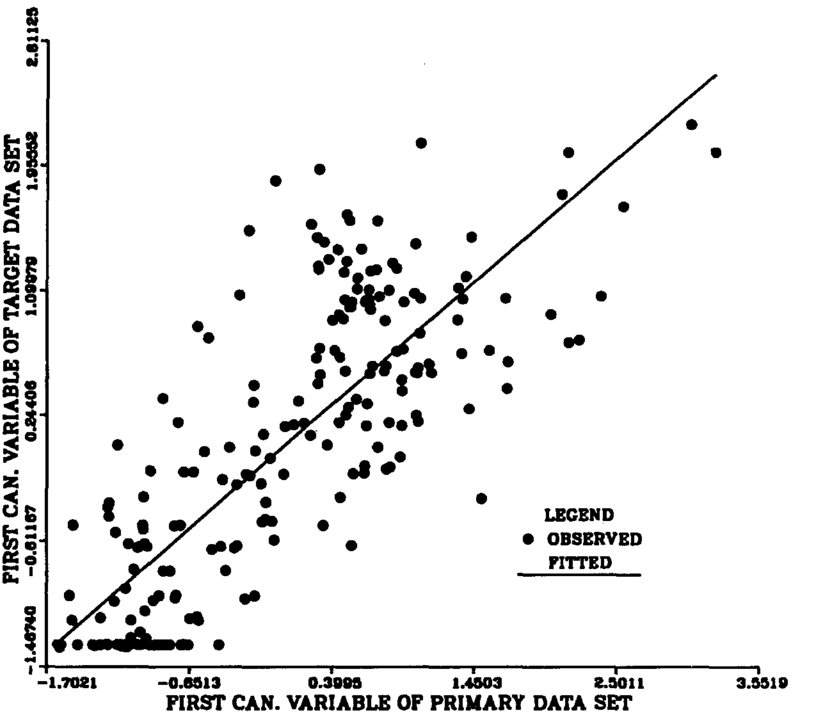

Fig. 3. The scatterplot of the values for the first pair of canonical variables obtained from the canonical correlation analysis.

Using an optimum discretization method, called partial rank correlation, proposed by Pan and Harris (1990b), the favorability values are transformed into a ternary form. Two optimum cutoffs for this conversion are 7.5 and 13.8. Fig. 6 shows the distribution of the transformed favorability values. A favorability value greater than 13.8 is highlighted by a solid rectangle, a value between 7.5 and 13.8 by a cross symbol, and a value less than 7.5 is left blank. The distribution of the discretized favorability values is broadly correlated with that of discretized target variables in Fig. 5. The patterns of the largest favorability values are roughly consistent with those of mineral occurrences except for some regions in the north and the southwest corners.

The correlation coefficients of the estimated favorability function (Eq. 21) with the explanatory and target variables were computed (Johnson and Wichern, 1988) and listed in Table 3. $\hat{F}$ is most strongly correlated with X_7 (the host rock) and X_1 (geochemical field), moderately correlated with X_6 (a combined geophysical field), and negatively correlated with X_8 (the estimate of depth to in-

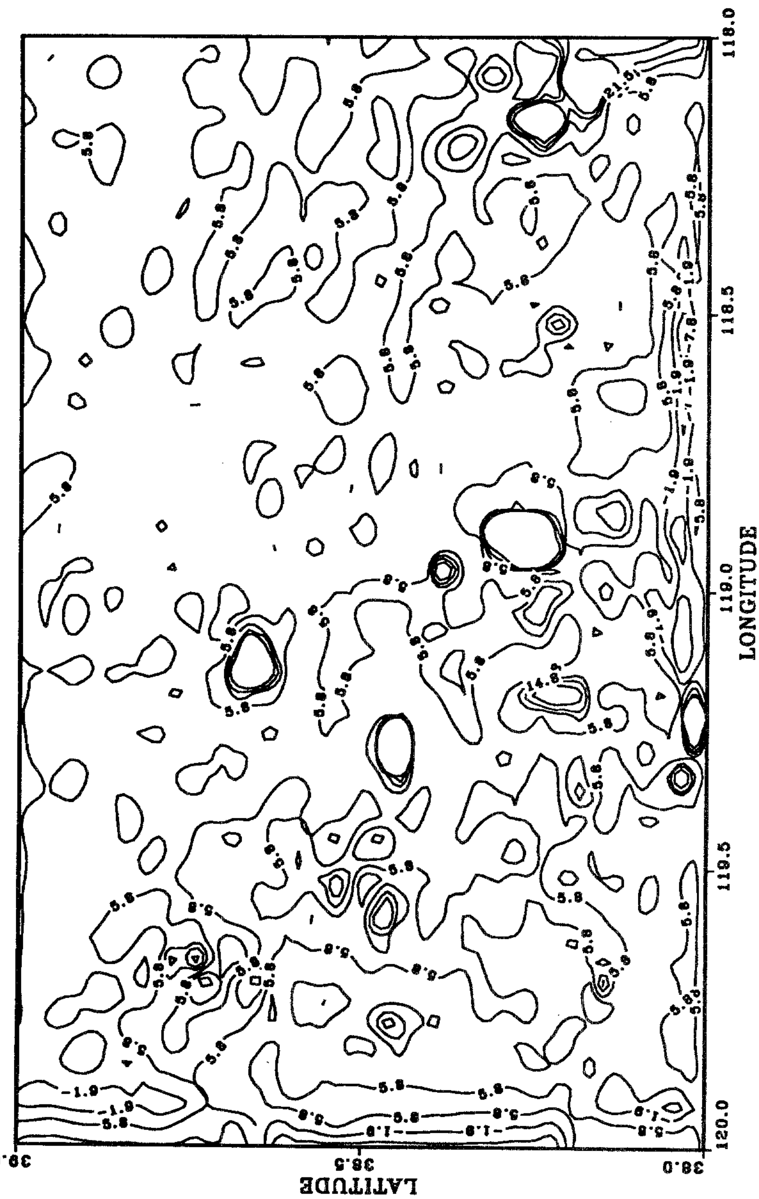

Fig. 4. The favorability values for epithermal gold–silver deposits by the canonical correlation method.

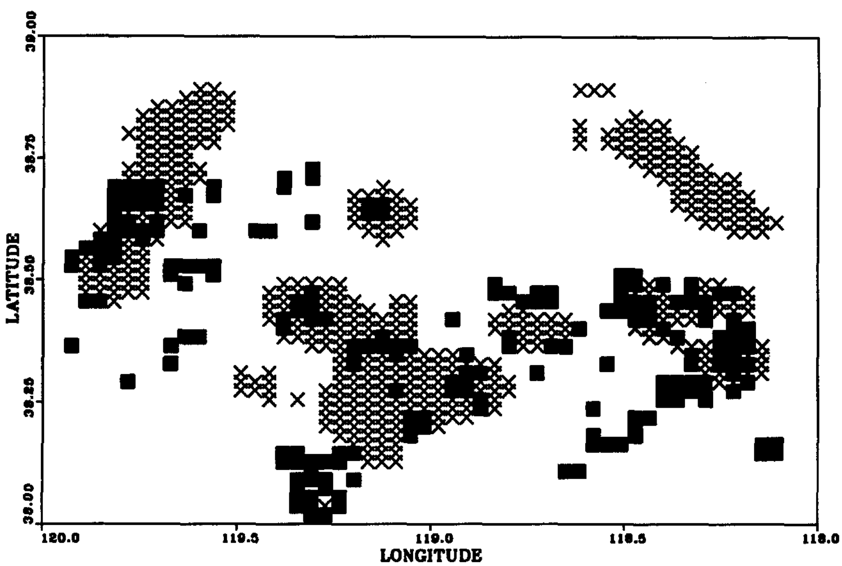

Fig. 5. The distribution of the discretized target variables (a solid rectangle represents a cell with at least one mineral occurrence observed and a cross for a cell with occurrence of Tertiary intrusives or hydrothermal alterations but no mineral occurrences).

trusives). All three target variables are highly correlated with the estimated favorability function, indicating that $\hat{F}$ captured the major information contained in the target variables. The strongest information captured by $\hat{F}$, however, is the variability in Z_2 (Tertiary intrusive), as indicated by the largest correlation coefficient.

Estimation by Weighted Canonical Correlation Method

In order to apply the weighted canonical correlation method, the principal components method was employed to convert the three target variables into orthogonal components. Table 4 summarizes the basic results of this analysis. The first component loadings indicate that all three target variables are almost equally important, although mineral occurrence has a slightly higher weight. The first component accounts for more than 70% of the total variability in the data, while the first two components account more than 88% of the total variations. All three components ($q = 3$) are used in the subsequent analysis. The principal component scores at each sample location were computed based on the loadings in Table 4.

The estimate of Eq. (1) by the weighted canonical correlation method is

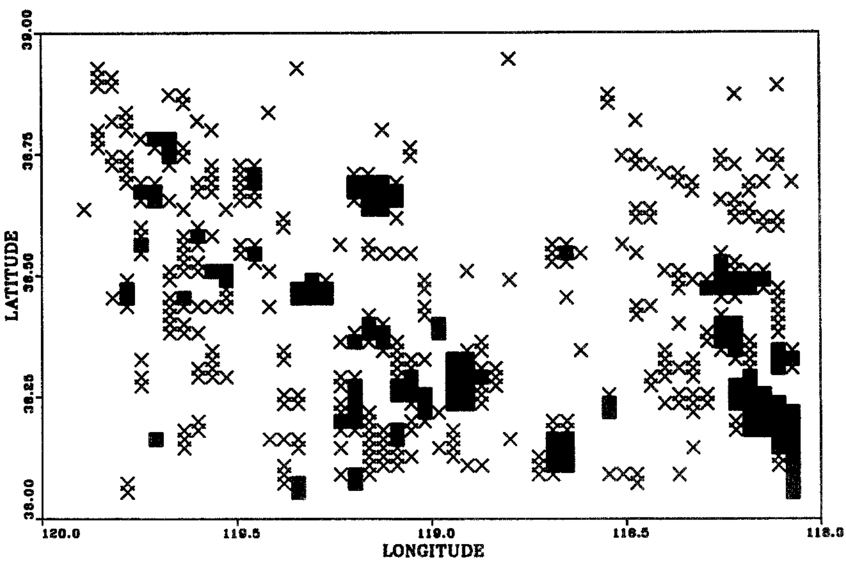

Fig. 6. The distribution of the discretized favorability values estimated by canonical correlation method (a solid rectangle represents a cell with favorability value greater than 13.8, an X for a cell with the value in the interval 7.5–13.8, and the cells with values less than 7.5 are left blank).

Table 3. Correlation Coefficients Between the F Estimated by Canonical Correlation Method with X and Z

Corr. F and X				Corr. F and Z	
$X1$	0.5958	$X5$	0.0534	$Z1$	0.4532
$X2$	−0.0406	$X6$	0.2676	$Z2$	0.8347
$X3$	0.0702	$X7$	0.9353	$Z3$	0.5249
$X4$	−0.0006	$X8$	−0.3137		

$$\hat{f}_i = 0.542x_{i1} + 0.113x_{i2} - 0.0083x_{i3} - 0.047x_{i4} + 0.037x_{i5} \\ -0.0035x_{i6} + x_{i7} + 0.027x_{i8}. \qquad (22)$$

According to the eigenvalues, this estimate accounts over 95% of the total variability related to the covariances between explanatory variables and three principal components of the target variables.

The estimated favorability value at each location was evaluated from Eq. (22). Fig. 7 is the contoured map of these values in the entire Walker Lake quadrangle, showing that all known mining districts, except for part of the Silver Mountain, are associated with highest favorability values. The general fea-

Table 4. Principal Components Analysis of the Three Target Variables

	Eigenvalue	Proportion	Cumulative
$Y1$	2.15523	0.718409	0.71841
$Y2$	0.49956	0.166521	0.88493
$Y3$	0.34521	0.115071	1.00000
		Eigenvectors	
	Y_1	Y_2	Y_3
$Z1$	0.574469	−0.610119	0.545655
$Z2$	0.556673	0.779942	0.286016
$Z3$	0.600084	−0.139444	−0.787690

tures of this map are very similar to those in Fig. 4, but some known mining districts, like Camp-Douglas, are better outlined by this method. Using the same optimum discretization approach as used in the canonical correlation analysis, two optimum thresholds, 9.5 and 14.1, were obtained. Based upon these cutoffs, the favorability values were discretized into a ternary variable, which is shown in Fig. 8.

The cells having no known mines but having the favorability values greater than 13.8 in canonical correlation analysis and greater than 14.1 in weighted canonical correlation analysis are separated and shown in Fig. 9 (the cells located at pre-Tertiary regions are not included). These regions may be considered as potential exploration targets for new discoveries of epithermal gold–silver deposits.

Several correlation vectors computed from Eqs. (16), (17), and (18) are collected in Table 5. As expected from Eq. (22), the estimated favorability function is most strongly correlated with host rock (X_7) and geochemical field (X_1), meaning that these two explanatory variables are most critical in determining the favorability of a sample unit with respect to the occurrence of the target variables. X_6 and X_8 are both moderately correlated with the estimated favorability function. Furthermore, the estimated favorability function is highly correlated with all three target variables, meaning that the estimate strongly reflects the heat source characterized by the target variables. The correlation of $\hat{F}$ with the Tertiary intrusive (Z_2) is particularly high, being consistent with the canonical correlation estimate.

More interestingly, $\hat{F}$ is more strongly correlated with Y_1, the first principal component of the target variables, moderately correlated with the second principal component, the least correlated with the last component. Table 4 shows

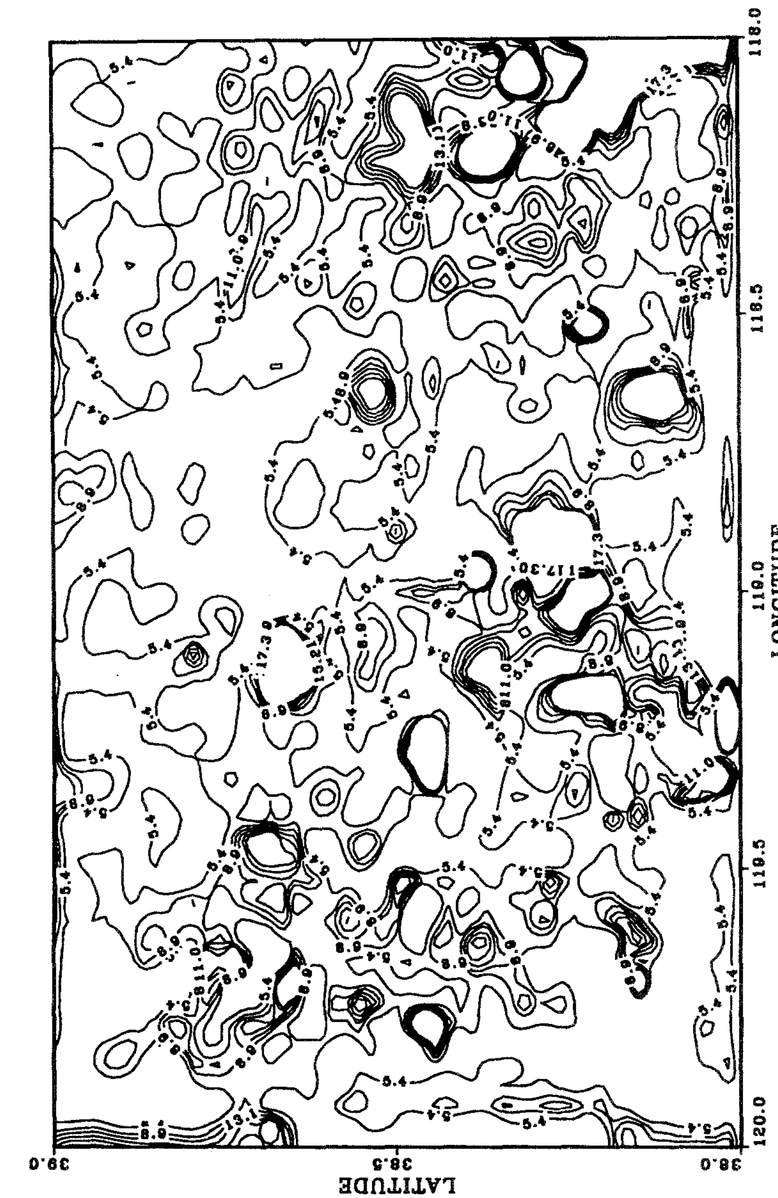

Fig. 7. The favorability values for epithermal gold–silver deposits estimated by the weighted canonical correlation method.

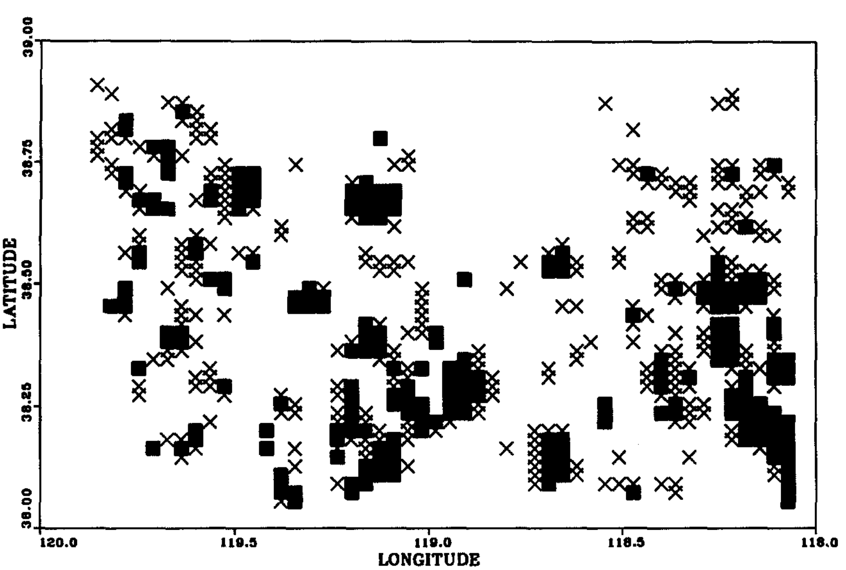

Fig. 8. The distribution of the discretized favorability values estimated by weighted canonical correlation method (a solid rectangle represents a cell with the value greater than 14.1, an X for a cell with the value in the interval 9.5–14.1, and other cells with the values less than 9.5 are left blank).

that the first principal component is determined by the coordinate vector (0.5745, 0.5567, 0.6001). The estimate $\hat{F}$ in Eq. (21) by the canonical correlation method characterizes its maximum correlation with a linear combination of the target variables with the coordinate vector (0.029, 0.956, 0.048). Clearly, the two coordinate vectors characterize two very different directions in the target variable space. The coordinate vector for the first principal component represents the direction with the largest variability. Therefore, the estimate by the weighted canonical correlation method captures the information related to the maximum variability in the three target variables, while the estimate by the canonical correlation method extracts the information primarily carried by the Tertiary intrusive.

CONCLUDING REMARKS

The notion of a favorability equation, which was introduced in the last decade, is useful in the integration of multivariate geological data into mineral resources assessment. Unlike regression analysis, estimating a favorability equation does not require a dependent variable to characterize some measure to be fitted. This notion is particularly useful when the relative magnitude of an

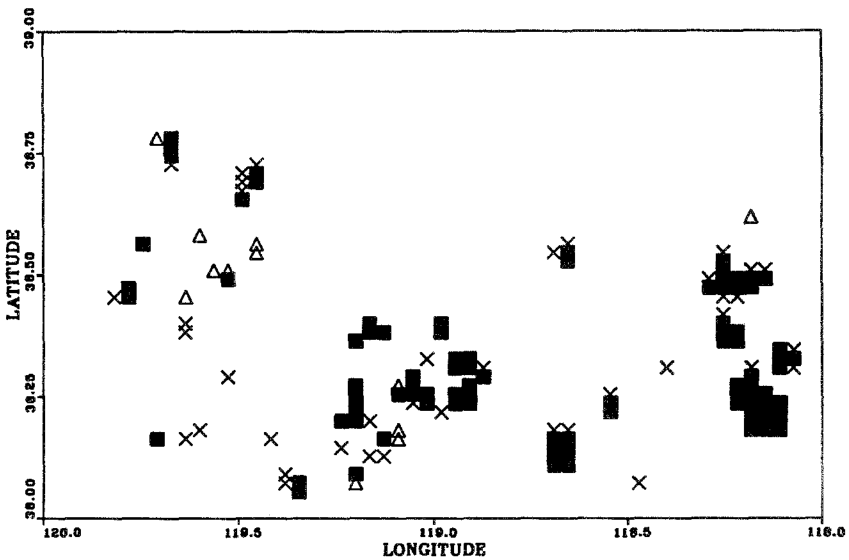

Fig. 9. The distribution of the potential target areas for epithermal gold–silver deposits (a triangle represents a cell selected as a target by the canonical correlation method, an X represents a cell selected by the weighted canonical correlation method, and a solid rectangle represents a cell selected by both methods).

Table 5. Correlation Coefficients of the F Estimated by Weighted Canonical Correlation Method with Variables X, Z, and Y

Corr. F and X				Corr. F and Z		Corr. F and Y	
$X1$	0.6492	$X5$	0.0456	$Z1$	0.4716	$Y1$	0.9244
$X2$	−0.0135	$X6$	0.2046	$Z2$	0.8235	$Y2$	0.2775
$X3$	0.1250	$X7$	0.9178	$Z3$	0.5517	$Y3$	0.0582
$X4$	0.0331	$X8$	−0.2189				

objective feature is of primary interest. In mineral exploration, for instance, delineation of drilling targets often involves the separation of anomalies from background based upon relative measures.

A primary goal of this paper is to define the meaning of a favorability function in terms of a set of target variables, other than explanatory variables themselves, and to avoid possible ambiguities in the estimate of favorability function by traditional methods. The ordinary principal components method tends to lead to an estimate of F which characterizes the direction with the maximum variability in explanatory variable space, while the methods proposed here give an estimate that maximizes the correlation between explanatory and target variables. The introduction of target variables is a key improvement in that regard. A target variable is usually identified according to a prespecified objective. The favorability analysis for mineral potential should maximally utilize the information carried by target variables on critical controlling genetic factors of ore deposit, spatial existential evidence of mineral descriptors, as well as other geological events related to mineral formation processes.

Although the case study is satisfactory, there are still some important and fundamental issues that remain to be further explored. For instance, there is a need for establishing a theoretical structure that allows one to make a rigorous statistical inference about the estimate of the favorability function and to derive interval estimates for an established confidence level.

ACKNOWLEDGMENTS

This research was partly supported by a grant (No. 14-08-0001-G1399) from the U.S. Geological Survey. Grateful acknowledgment is made of this support and for the data provided by various U.S.G.S. offices and personnel. Special acknowledgment is made of the assistance provided by Dr. D. Menzie. We also wish to acknowledge Dr. George Koch for his careful review of the manuscript, which substantially improved the paper. Finally, the authors are appreciative of the assistance of Tim Heiner in preparation of the manuscript.

REFERENCES

Agterberg, F. P., and Gabilio, P., 1969, Two-Stage Least-squares Model for the Relationship Between Mappable Geological Variables: Math, Geol., v. 1, p. 137–153.

Bartlett, M. S., 1947, Multivariate Analysis: J. R. Stat. Soc. (Suppl.), v. 9, p. 176–190.

Chaffee, M. A., Hill, R. H., Speckman, W. S., and Sutley, 1980, Preliminary Data Set Containing Geochemical Analyses of Rock, Stream Sediment, and Nonmagnetic Heavy Mineral Concentrate, Walker Lake 2 Degree Quadrangle, California and Nevada: U.S.G.S. Open-File Rept 80-881, 235 p.

Davis, J. C., 1986, Statistics and Data Analysis in Geology, 2nd ed.: John Wiley & Sons, New York, 646 p.

Ethridge, F. G., and Davies, D. K., 1973, Grouped Regression Analysis-A Sedimentologic Example: Math. Geol, v. 5, p. 377–388.

Fujikoshi, Y., 1977, Asymptotic Expansions of the Distributions of Some Multivariate Tests, *in* P. R. Krishnaiah (Ed.), Multivariate Analysis, Vol. iv: p. 55–71.

Harris, D. P., 1965, An Application of Multivariate Statistical Analysis to Mineral Exploration: Ph.D. Dissertation, the Pennsylvania State University, Univ. Park, Pennsylvania, 261 p.

Harris, D. P., and Pan, G. C., 1987, An Investigation of Quantification Methods and Multivariate Relations Designed Explicitly to Support the Estimation of Mineral Resources—Intrinsic Samples: Report on Research Sponsored by U.S. Geological Survey Grant No. 14-08-0001-G1399, 200 p.

Harris, D. P., and Pan, G. C., 1990, Intrinsic Sample Methodology: *in* G. Gaal and D. F. Merrlam

(Eds.), Computer Applications in Resource Exploration: Prediction and Assessment for Petroleum, Metals and Nonmetals, Vol. 6, of Computers and Geology Series: Pergamon Press, p. 53–74.

Harris, D. P., and Pan, G. C., 1990, Updated Concepts of Intrinsic Samples and Methodology for Simultaneous Estimation of Discovered Resources and Endowments: Report on Research Sponsored by U.S. Geological Survey, 300 p.

Harris, D. P., and Pan, G. C., 1991, Consistent Geologic Areas for Epithermal Gold-Silver Deposits in the Walker Lake Quadrangle of Nevada and California: Delineated by Quantitative Methods: Econ. Geol., v. 86, p. 142–165.

John, D. A., 1986, Geology of Tertiary Volcanic Rocks in the Paradise Range and Northern Pactolus Hills, Nye County, Nevada [abs]: U.S. Geol. Survey Open File Rept. 86-467, p. 13.

John, D. A., Thomason, R. E., and McKee, E. H., 1989, Geology and K-Ar Geochronology of the Paradise Peak Mine and Relationship of Pre-Basin and Range Extension to Early Miocene Precious Metal Mineralization in West-Central Nevada: Econ. Geol., v. 84, p. 631–649.

Johnson, R. A., and Wichern, D. W., 1988, Applied Multivariate Statistical Analysis: Prentice Hall, 607 p.

Jones, T. A., 1972, Multiple Regression with Correlated Independent Variables: Math. Geol., v. 4, p. 203–218.

Kucks, R. P., and Hildenbrand, T. G., 1987, Description of Magnetic Tape Containing Nevada State Magnetic Anomaly Data: U.S. Dept. of the Interior, EROS Data Center, U.S.G.S., Denver, Colorado.

Luo, J., 1990, Statistical Mineral Prediction Without Defining a Training Area: Math. Geol., v. 22, p. 253–260.

Mardia, K. V., Kent, J. T., and Bibby, J. M., 1979, *Multivariate Analysis*: Academic Press, London, 521 p.

Mark, D. M., and Church, M., 1977, On the Misuse of Regression in Earth Science: Math. Geol., v. 9, p. 63–75.

McCammon, R. B., 1973, Nonlinear Regression for Dependent Variables: Math. Geol., v. 5, p. 365–375.

McCammon, R. B., Botbol, J. M., Sinding-Larsen, R., and Bowen, R. W., 1983, Characteristic Analysis-1981: Final Program and a Possible Discovery: Math. Geol., v. 15, p. 59–83.

McKee, E. H., and Noble, D. C., 1986, Tectonic and Magmatic Development of the Great Basin of Western United States During late Cenozoic Time: Modern Geol., v. 10, p. 39–49.

Pan, G. C., 1985, Mineral Resources Appraisal on Pegmatitic Ni-Ta Deposits in Fujian Province of China: M.S. Thesis, Changchun College of Geology, 141 p.

Pan, G. C., 1989, Concepts and Methods of Multivariate Information Synthesis for Mineral Resources Estimation: Ph.D. Dissertation, University of Arizona, Tucson, 302 p.

Pan, G. C., and Harris, D. P., 1990a, Quantitative Analysis of Anomalous Sources and Geochemical Signatures in the Walker Lake Quadrangle of Nevada and California: J. Geochem. Explor., v. 38, p. 299–321.

Pan, G. C., and Harris, D. P., 1990b, Three Nonparametric Techniques for Optimum Discretization of Quantitative Geological Measurements: Math. Geol., v. 22, p. 699–722.

Pan, G. C., and Harris, D. P., 1991, Geology-Exploration Endowment Models for Simultaneous Estimation of Discoverable Mineral Resources and Endowment: Math. Geol., v. 23, p. 507–540.

Pan, G. C., and Wang, Y., 1987, Weighted Characteristic Analysis and Its Application to Quantitative Estimation of Pegmatitic Nb-Ta Mineral Resources: Geol. Prospect., v. 23, p. 34–42.

Plouff, D., 1982, Bouguer Gravity Anomaly and Isostatic Residual Gravity Maps of the Walker Lake 1^0 by 2^2 Quadrangle, California and Nevada: U.S.G.S. OF Rept. 169.

Reddy, R. K. T., and Koch, Jr., G. S., 1988a, A Generalized Model for Evaluating Area-Potential in a Mineral Exploration Program: Math. Geol., v. 20, p. 227–241.

Reddy, R. K. T., and Koch, Jr., G. S., 1988b, A Case Study of Silver Exploration in Parts of Idaho and Montana Using Mathematical Modeling Techniques: Econ. Geol., v. 83, p. 1008–1014.

Rowan, L. C., and Purdy, T. L., 1984, Map of the Walker Lake 1° by 2° Quadrangle, California and Nevada, Showing the Regional Distribution of Hydrothermally Altered Rocks: U.S. Geological Survey, Dept. of the Interior.

Seber, G. A. F., 1984, Multivariate Observations: John Wiley & Sons, New York, 686 p.

Silberman, M. L., Stewart, J. H., and McKee, E. H., 1976, Igneous Activities, Tectonics, and Hydrothermal Precious-Metal Mineralization in the Great Basin During Cenozoic Time: Transactions, v. 260, p. 253–263.

Stewart, J. H., 1988, Tectonics of the Walker Lane Belt, Western Great Basin-Mesozoic and Cenozoic Deformation in a Zone of Shear, *in* W. G. Ernst, (Ed.), Metamorphism and Crustal Evolution of the Western United States, Rubey, Vol. vii: Englewood Cliffs, New Jersey, Prentice Hall, p. 683–713.

USGS, 1979a, Aeromagnetic Map of the Hoover–Walker Lake, California: 1 over-size sheet, scale 1:250000, OF Rept. 79-1194.

USGS, 1979b, Aeromagnetic Map of the Mokelumne Areas, California: 1 over-size sheet, scale 1:62500, OF Rept. 79-1233.

USGS, 1981, Aeromagnetic Map of the Dardaneller Area, California: 2 over-size sheets, scale 1:62500, OF Rept. 81-0431.

USGS, 1982, Aeromagnetic Map of the Sweetwater Area, California-Nevada: 1 over-size sheet, scale 1:62500, OF Rept. 82-0066.