Natural Resources Research, Vol. 8, No. 3, 1999

Integration of Sparse Geologic Information in Gold Targeting Using Logistic Regression Analysis in the Hutti–Maski Schist Belt, Raichur, Karnataka, India—A Case Study

Nihar Ranjan Sahoo[1,3,4] and Hari Shankar Pandalai[2]

Received 3 January 1999; accepted 25 June 1999

Logistic regression has been used in the study to integrate indicator patterns for estimation of the probability of occurrence of gold deposits in a part of the auriferous Archaean Hutti–Maski schist belt. Data used consist of categorical and continuous variables obtained from a coded lineament map and geochemical anomaly maps of the pathfinder elements of gold in soil and groundwater. Main effects and interactions of the variables studied were used in formulating the logistic regression model. Regression models using lineament-proximity data, combined with soil and groundwater geochemical anomalies were tested on parts of the schist belt with data not used in estimation of model parameters. Predicted probabilities greater than 0.9 identified known deposit locations in the area.

KEY WORDS: Mineral exploration; multielement geochemical data-integration; logistic regression; gold targeting; greenstone terrains.

INTRODUCTION

The problem of poor sample representativity and poor analytical reproducibility of gold in geochemical samples (Harris, 1982; Clifton and others, 1969; Nichol, Closs, and Lavin, 1989; Boyle, 1987) requires the application of multielement geochemical studies for targeting gold deposits (Boyle, 1979, 1987). This also permits broad sample spacing without the risk of missing mineralization. Several successful applications of multielement geochemistry in locating gold mineralization are available from literature. In this study, geochemical dispersion halos of As, Sb, Hg, and Bi in soil and As and Sb in groundwater have been used, along with a careful selection of geological variables to compute probabilities of occurrences of gold deposits by data integration using logistic regression.

Prediction of mineral endowment in virgin areas and computation of probability or favorability measures of mineral occurrence of a specific type in a selected area are problems that require use of various statistical models. Quantitative approaches to regional mineral potential mapping can be grouped into two categories: data-driven (objective) and knowledge-driven (subjective) approaches(Bonham-Carter, 1994; Reddy, Bonham-Carter, and Galley, 1992).

In data-driven approaches, map patterns used as evidence of mineralization are combined numerically using various statistical models to compute a probability or favorability measure of deposit occurrence. Multiple linear regression (Agterberg, 1974; Chung and Agterberg, 1980; Harris, 1984), weighted and targeted

[1] Department Of Earth Sciences, Indian Institute Of Technology, Bombay, Powai, Mumbai 400 076, India. (e-mail: nrsahoo@geos.iitb.ernet.in)

[2] Department Of Earth Sciences, Indian Institute Of Technology, Bombay, Powai, Mumbai, India 400 076.

[3] Present address: Geoinformatics, Applied Technology Group, Tata Infotech Limited, SDF5, SEEPZ, Mumbai, India 400096. (e-mail: nihar@darkstar.atg.tatainfotech.com)

[4] To whom all correspondence should be addressed.

1520-7439/99/0900-0193$16.00/1 © 1999 International Association for Mathematical Geology

multivariate criterion (Harris and Pan, 1987), logistic regression analysis (Chung, 1977; Agterberg, 1992; Bonham-Carter and Chung, 1989; Harris and Pan, 1991), cannonical correlation analysis (Pan and Harris, 1992), factor analysis (Harris and Pan, 1990, 1991), modified component analysis (Pan and Harris, 1992; McCammon, 1992), characteristic analysis (Botbol and others, 1978), discriminant analysis (Agterberg, 1974; Harris, 1966), and cluster analysis (Briggs and Press, 1977) are some of the multivariate statistical techniques, that have been used for such prognostication. Important reviews on mineral targeting includes Singer and Mosier (1981), Reddy, Bonham-Carter, and Galley (1992), Harris (1984), Bonham-Carter (1994), Agterberg (1989), and Pan, Harris, and Heiner (1992).

Integration of data for computation of probabilities or favorability measures have used the weights of evidence modeling (Bonham-Carter, Agterberg, and Wright, 1988, 1989; Agterberg, 1989; Reddy, Bonham-Carter, and Galley, 1992; Reddy and Bonham-Carter, 1991; Agterberg and others, 1993), Indicator favorability theory (Pan, 1993) and evidential belief functions (Moon, 1990; An, Moon, and Bonham-Carter, 1994a, 1994b). Weights of evidence modeling, which is a technique for integration of binary map patterns, has become popular because of the relative ease of implementation. Evidence of mineralization is combined from several predictive maps, using a formulation of Bayes rule, as was done earlier in FINDER (Singer, 1985; Singer and Kouda, 1988) in a direct use of Bayes rule. Calculation of prior probabilities for hypotheses and evidences require a well-explored training area of similar type as the area being explored. The major disadvantage of the method is that of testing and assessing the conditional independence of the predictor maps given the deposit map in known areas. Extended weights of evidence model (Pan, 1996) has been proposed for use with categorical explanatory variables for obtaining pseudometal estimates, which provide a quantitative method for delineation of exploration targets.

Knowledge-driven approaches usually use forward-chaining expert systems in which the method of propagation of the favorability measure through the inference network may include the Bayesian updating and fuzzy logic for computation of posterior values of favorability given evidence(s). Expert systems such as Prospector (Duda and others, 1978), Prospector II (McCammon, 1989), Prospector III (McCammon, 1990), and MAPS (Katz, 1988, 1991) have been used for estimating undiscovered mineral potential on a regional scale. One of the more recent advances is the use of artificial neural networks for assessment of mineral potential zones (Singer and Kouda, 1996). Other methods which have been used include the favorability model of McCammon (1992), the descriptive model of Cox and Singer (1986), and artificial intelligence (Campbell, Hollister,and Duda, 1982).

Because the data used in this study consists of categorical and continuous variables, logistic regression was selected for the purpose of predicting probabilities of mineral occurrences. The main effects and interactions of the variables have been modeled by a systematic study.

GEOLOGICAL OVERVIEW

The Archaean Hutti–Maski greenstone belt is a narrow, elongated, hook shaped, NNW-trending belt, consisting of predominant metavolcanic rocks and subordinate metasediments. The volcanics are represented by pillowed tholeiitic metabasalts, minor rhyolites, acid tuffs, and pyroclastic rocks. Metasediments include banded ferrugenous chert, quartzite, carbonaceous phyllite, and garnetiferous mica–schist (Roy, 1979, 1991). These rocks occur as thin impersistent bands in metabasalts (Giritharan and Rajamani, 1998; Srikantia, 1995). The rocks of greenstone association are surrounded by multiple phases of intrusive diapiric granitoids (Fig. 1).

Vesicular metabasalt is the host rock for the auriferous lodes. The auriferous reefs are mylonitized, brecciated, and most of them are localized along the major shear zone in the western segment of the greenstone belt (Fig. 1). The geometry and orientation of the auriferous lodes are affected by the structural fabric of the shear zone. Gold occurs in native state with quartz, as well as in association with sulfides in the alteration zones.

PHYSIOGRAPHY

The Hutti–Maski schist belt is a region of low rainfall with small catchment areas, leading to non-availability of stream sediments. Although weathering is deep at places, the soil profile in general, is poorly developed. The soil profile may consist partly or wholly of transported overburden. This type of terrain, which is particularly widespread in tropical Precambrian greenstone terrains, suggests that groundwater and weathered bedrock are likely to be the most suitable media for detecting dispersion halos related to

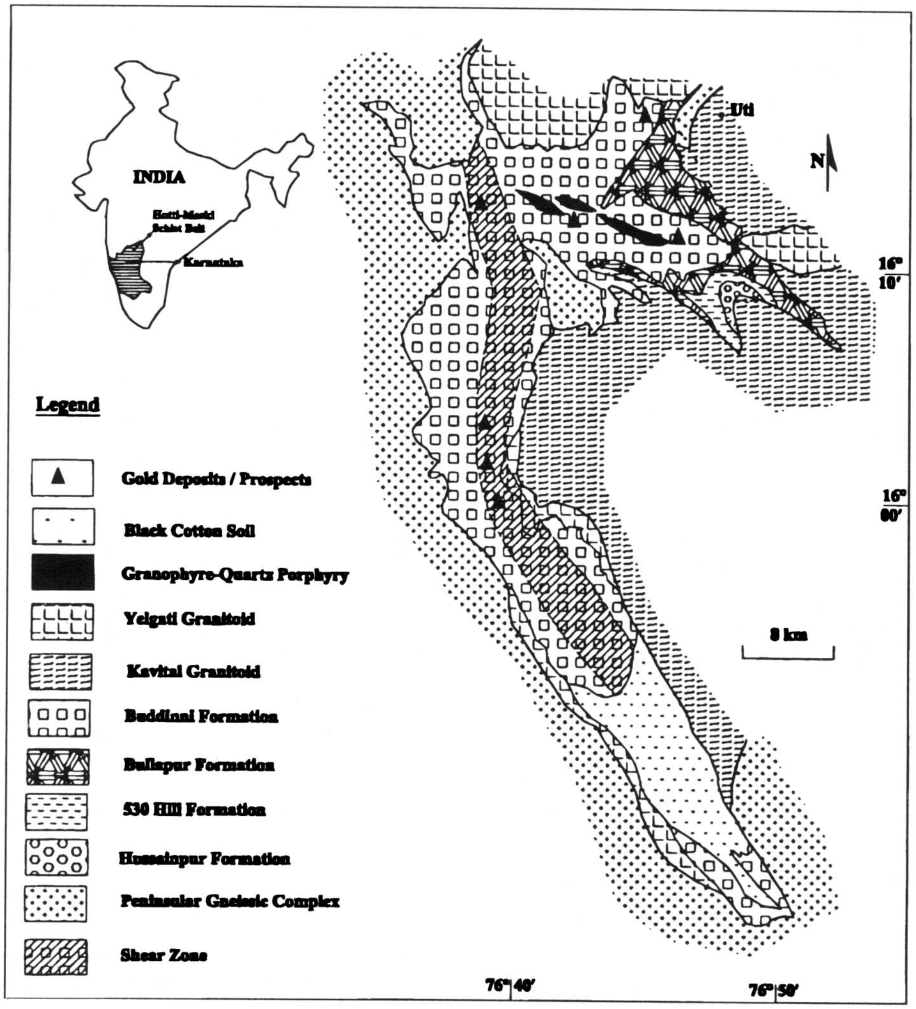

Figure 1. Geological map of Hutti–Maski schist belt (after Srikantia, 1995).

mineralization (Butt and Smith, 1980). This study thus was carried out on soil and groundwater samples.

DATA GENERATION

Minus 80 mesh-sieved fraction of the soil samples, collected from just above the C horizon were digested for total analyses of As, Sb, Bi, and Hg by hydride generation–induced-coupled plasma–optical emission spectroscopy (HG–ICP–OES) technique (Trafford, 1986). International USGS soil standards (GXR-2 and GXR-6) were used for calibration.

Nonacidified samples of groundwater were analyzed for alkalinity, chloride, and sulfate. Alkalinity and chloride were estimated by titration methods, and sulfate was estimated photometrically (Greenberg, Clesceri, and Eaton, 1992). The acidified portions of the groundwater samples were preconcentrated and analyzed for As and Sb by HG–ICP–OES.

Selenium, Ag, and Mo also were analyzed, but were below detection limits in both types of samples from mineralized areas. Antimony, Bi, and Hg were detected only in soil samples located close to known gold mineralized areas. Bismuth and Hg were not detected in any of the groundwater samples. The details of sampling and chemical-analytical technique is given in Sahoo (1998).

SECONDARY DISPERSION

The arsenic values were interpolated with an inverse distance weighting algorithm. The contour diagram of the As values shows anomalous values over the mineralized areas. The anomalous areas show good spatial association with the known gold mineralized areas (Fig. 2). A regional background of 40 ppm and threshold of 150 ppm were defined for As, on the basis of the known mineralized areas. The elements Hg, Sb, and Bi were detected only in areas indicated anomalous on the basis of the As anomalies. The values of these elements were interpolated with inverse distance weighting algorithm. Figure 3 shows the extent of the anomalies of these elements. It is seen that Hg anomalies are wider than those of Sb and Bi. High values of Sb and Bi are restricted to areas almost directly above mineralization.

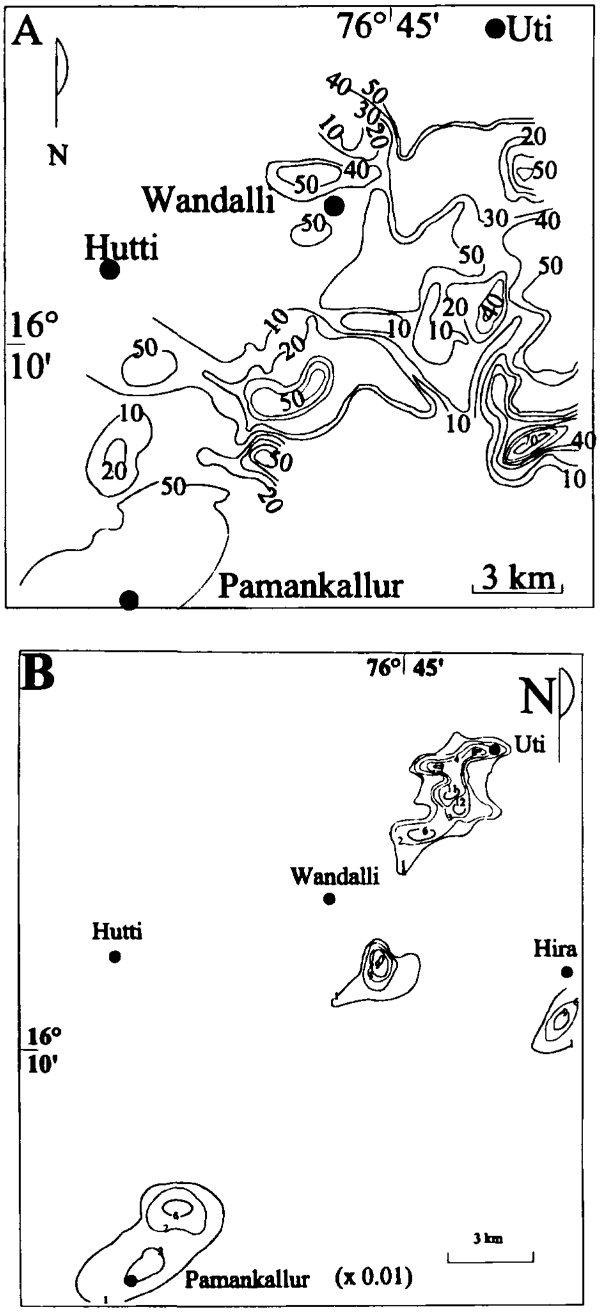

Figure 2. Contours (in ppm) of As in soil. Only contours below 100 ppm are shown in A. Contours above 100 ppm are shown in B. Contours in B identify As anomalies around known mineralization. A and B represent same area.

Hydrogeochemical contours of alkalinity, sulfate, chloride, As, and Sb were prepared with an inverse distance weighting algorithm (Fig. 4). Anomalies of alkalinity, sulfate, and chloride had wider spread whereas those of As and Sb were narrower and located almost directly over the mineralized areas. Values in groundwater, of alkalinity > 2000, sulfate > 50, and chloride >20 ppm delineate the mineralized areas.

Anomalies of alkalinity, sulfate, chloride, As, and Sb in groundwater and those of As, Sb, Hg, and Bi in soil delineate the known mineralized areas. Mercury alone can not identify the gold mineralized areas. Shear zones and granite–metabasalt contacts were determined to be major structural controlling features localizing auriferous lodes.

Preliminary tests on the data revealed that (1) multielement data of soil were not conditionally independent of each other given deposit occurrence and (2) water chemistry variables also were not conditionally independent of each other, given mineral occurrence. The conditional independence of water and soil data was not tested as the sampling sites were different. The weights of evidence method thus was not suitable for handling the data for the purpose of mineral targeting. Logistic regression was selected for the purpose of computation of probability of mineral occurrence, because it is possible to model binary (dichotomous) target variable using continuous and discrete categorical or ordinal data. The main effects and interaction terms of the explanatory variables has been studied and used for modeling.

LOGISTIC REGRESSION

The logistic regression model falls in the general category of log-linear models. The theory of logistic regression analysis is discussed in Hosmer and Lemeshow (1989), Cox and Snell (1989), and Freeman (1987). The following discussion on the theory of logistic regression analysis is developed after the treatments given in these references.

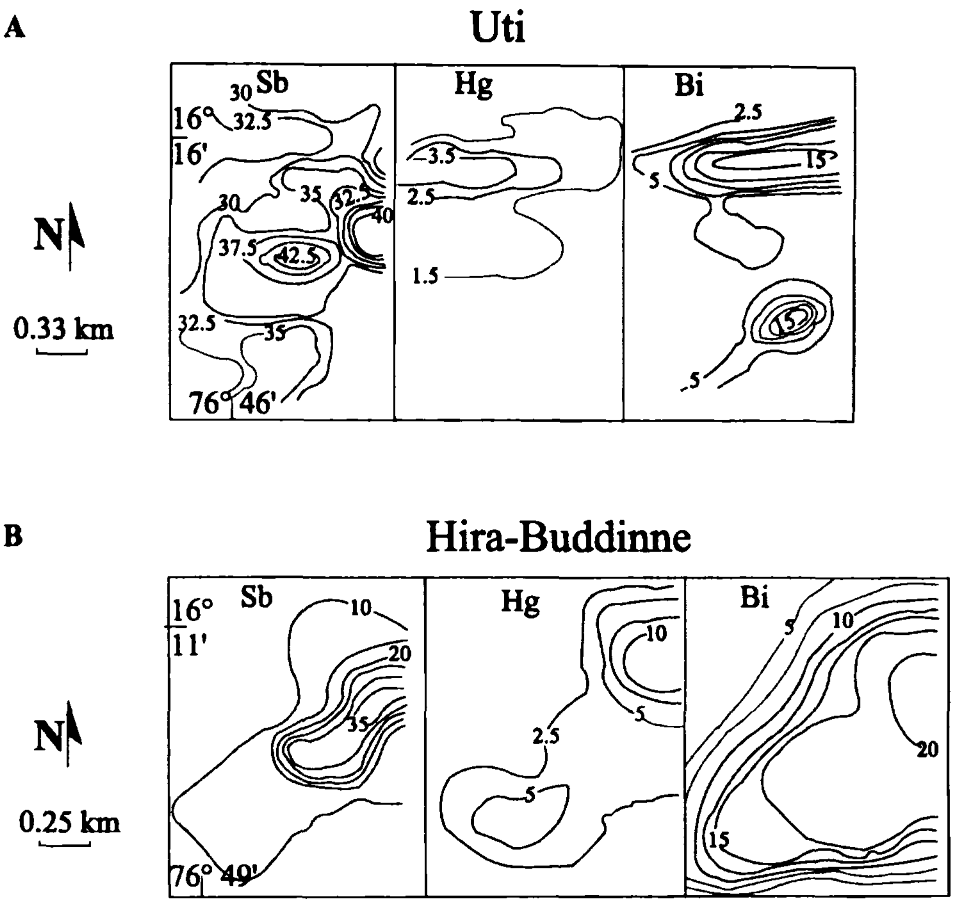

Figure 3. Contours (in ppm) of Sb, Bi, and Hg in soil from Uti block, A, and Hira-buddinni block B. Contours show characteristic anomalies in both blocks, which are known to be mineralized.

A dichotomous response variable Y (say) of an experimental unit takes one of two possible values, denoted for convenience by 0 and 1 (for example $Y = 1$, if the response is a favorable event, otherwise $Y = 0$). Let X be an explanatory variable and $p = Pr(Y = 1|x)$ be the response probability to be modeled. A logistic regression model can be written in the form $Y = E(Y|x) + \varepsilon$, where, ε assumes one of the two possible values. If $Y = 1$ then, $\epsilon = 1 - p$ with probability p, and if $Y = 0$ then $\varepsilon = -p$ with probability $(1 - p)$. Thus ε has a distribution with mean zero and variance equal to $p(1 - p)$.

In general, when several explanatory variables are used, the logistic regression model has the form

$$\text{logit}(p) = \log(p/(1 - p)) = \beta_0 + \boldsymbol{\beta}'\mathbf{X}$$

where, $\mathbf{X}$ is the vector of explanatory variables, β_0 is the intercept parameter, and $\boldsymbol{\beta}'$ is the vector of slope coefficients.

FITTING A LOGISTIC REGRESSION MODEL

Let, (x_i, y_i), $i = 1, 2, 3 \ldots n$, be a pair of observations on the ith object of the dichotomous response variable (y_i) and an independent variable (x_i), among n such pairs of observations. Fitting a logistic regression model of the form $p(x) = e^{\beta_0+\beta_1 x}/(1 + e^{\beta_0+\beta_1 x})$ requires estimation of β_0 and β_1. These values are selected such that the sum of squared deviations of the observed values y from the predicted values $(\hat{y})$, based on the model are minimized.

If Y is dichotomous variable, the expression for $p(x)$ in the given equation, provides $Pr(Y = 1| x)$. The quantity $(1 - p(x))$ is defined as $Pr(Y = 0|x)$. Thus for those pairs (x_i, y_i), where $y_i = 1$, the contribution to the likelihood function is $p(x_i)$, and for those pairs with $y_i = 0$, the contribution to likelihood function is $(1 - p(x_i))$, where $p(x_i)$ denotes the value of $p(x)$,

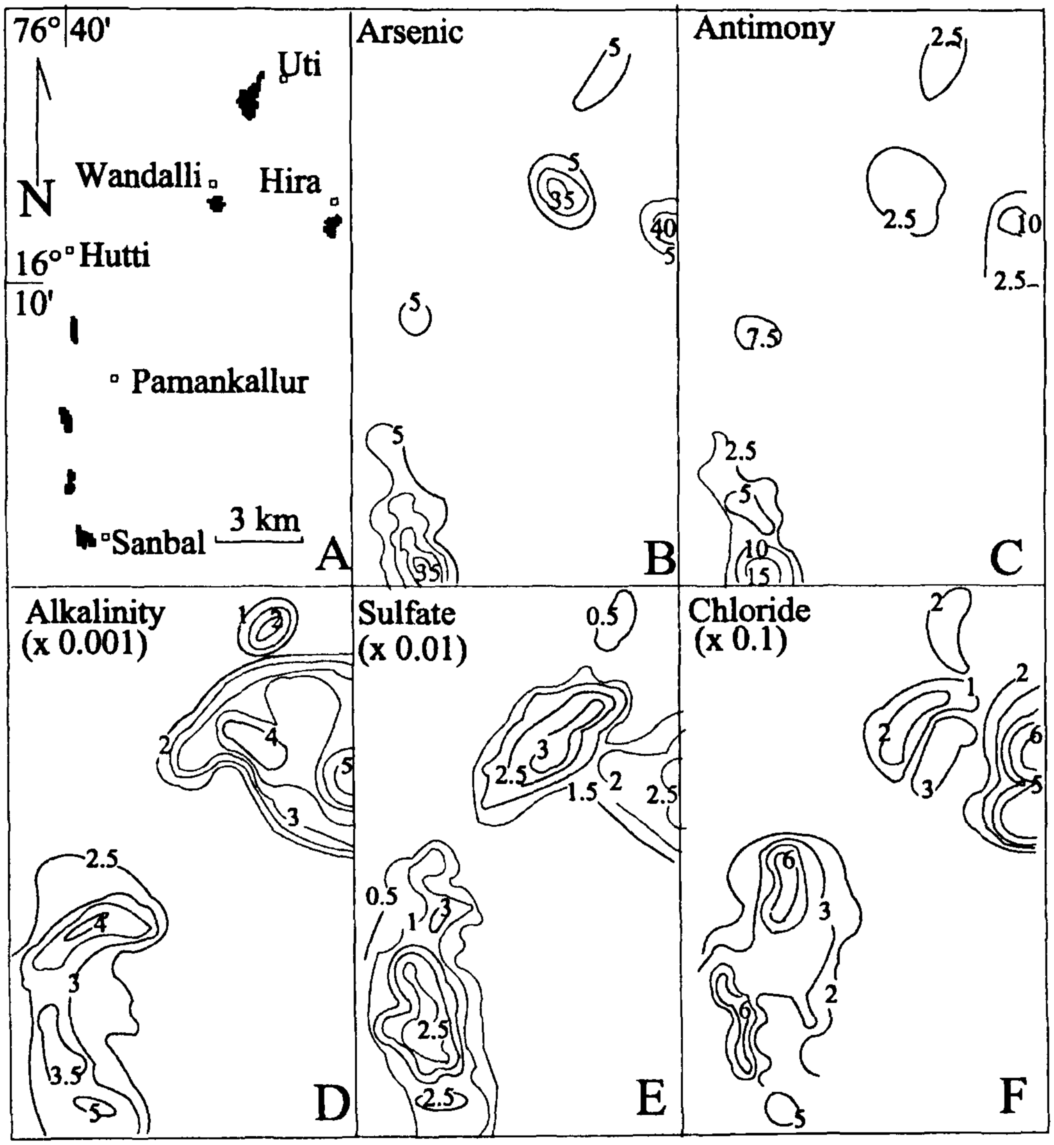

Figure 4. Contours in ppm for alkalinity, sulfate, and chloride and in ppb for Sb and As. Location of known mineralized areas are shown in A. Hydrogeochemical anomalies of As, Sb, alkalinity, sulfate, and chloride delineate gold-mineralized areas.

computed at x_i. The contribution to the likelihood function is obtained (Hosmer and Lemeshow, 1989) by:

$$\zeta(\beta) = p(x_i)^{y_i} (1 - p(x_i)^{1-y_i})$$

Because the observations are assumed to be independent, the likelihood function is obtained as $1(\beta) = \sum_{i=1}^{n} = \xi(X_i)$

Maximum likelihood estimation gives the estimates that maximize the likelihood function

$$L(\beta) = \ln(1(\beta))$$

$$= \sum_{i=1}^{n} y_i \ln(p(x_i) + (1 - y_i) \ln(1 - p(x_i))$$

To determine the values of β_0 and β_1, that maximize

$l(\beta)$, the given expression is differentiated with respect to β_0 and β_1 and setting the resulting expressions equal to 0. The equations are

$$\sum_{i=1}^{n} (y_i - p(x_i)) = 0$$

$$\sum_{i=1}^{n} x_i(y_i - p(x_i)) = 0$$

Because these equations are nonlinear in β_0 and β_1, iteration methods similar to Newton–Raphson and iteratively reweighted least-squares (IRLS) algorithm are used to solve for β's. The solution β's are termed estimated coefficients.

A multivariate logistic regression model may be expressed by the equation

$$p(\mathbf{x}) = e^{g(\mathbf{x})}/(1 + e^{g(\mathbf{x})})$$

where, $g(\mathbf{x}) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 \cdots + \beta_k x_k$ and $x_1, x_2, x_3, \ldots x_k$ are observations on k-independent variables.

This requires estimation of the vector $\hat{\boldsymbol{\beta}} = (\hat{\boldsymbol{\beta}}_0, \hat{\beta}_1, \hat{\beta}_2 \ldots \hat{\beta}_k)$. There are $(k + 1)$ likelihood equations which are obtained by differentiating the likelihood function with respect to $(k + 1)$ coefficients.

Estimators of variances and covariances are obtained from the matrix of second partial derivatives of the loglikelihood function as described. When interpreting statistically adjusted log odds ratio and odds ratio, the effectiveness of the adjustment is dependent entirely on the adequacy of the assumptions of model linearity and constancy of slope. Departure from these indicates the possibility of interaction and confounding of variables. When an explanatory variable is associated with the other covariates as well as the response variable, the relationship between the response and this explanatory variable is said to be confounded. When interaction is present, the association between the response and the covariates depends, in some way, on the levels of the covariates.

Criteria for Assessing Model Fit

The four criteria generally used for assessing the fit of the model are Wald's statistic, likelihood ratio test statistic (G), Akaike Information Criterion (AIC), and the Schartz Criterion (SC).

Wald's statistic is given by

$$W_i = (\hat{\beta}_i/\mathrm{SE}(\hat{\boldsymbol{\beta}}_i))^2$$

where, $\hat{\boldsymbol{\beta}}_i$ is the ith coefficient and $\mathrm{SE}(\hat{\beta}_i)$ is the standard error of the estimate of $\boldsymbol{\beta}_i$

-2Loglikelihood is given by

$$2 \text{ Log } L = -2 \Sigma \log(\hat{\rho}_i)$$

where L is the likelihood ratio.

Akaiki information criterion

$$\mathrm{AIC} = -2 \log L + 2(k + s)$$

where k is the number of ordered values for the response and s is the number of explanatory variables.

Schwartz criterion

$$\mathrm{SC} = -2 \log L + (k + s) \log(N)$$

where N is the total number of observations and L, k, and s are as given above.

The Wald's statistic checks the significance of the variables in the model. When the Wald's statistic exceeds the critical value of 2, the level of significance of is approximately 0.05 ($\alpha = 0.05$). The G statistic, which is chi-squared distributed, is obtained as the difference between the -2loglikelihood of two models, one of which has an additional new variable. The computed value of the statistic is used to check the significance of incorporation of the new variable in the model. The AIC and SC statistic also are used to compare two models. Lower values of these statistics indicate a more desirable model.

Linear Predictor, Predicted Probability, and Confidence Limits

When Y can take values $1,2, \ldots k$ and $\mathbf{X}$ is vector of explanatory variables, the linear predictor η_i is given by

$$\eta_i = g(Pr(Y \leq i|x)) = \alpha_i + \beta'\mathbf{x}$$

(where $1 \leq i \leq k$) and its estimate is

$$\hat{\eta}_i = \hat{\alpha}_i + \hat{\beta}'\mathbf{x}$$

where the $\hat{\alpha}_i$ and $\hat{\beta}'$ are the MLEs of α_i and β.

The estimated standard error of η_i is $\hat{\sigma}(\hat{\eta}_i)$. The confidence interval for η_i is given by

$$\hat{\eta}_i \pm Z_{\alpha/2}\, \hat{\sigma}(\hat{\eta}_i)$$

where $Z_{\alpha/2}$ is the $100(1 - \alpha/2)$ percentile point of a standard normal distribution.

The predicted value and the confidence limit for $Pr(Y \leq i|x)$ are obtained by back-transforming the corresponding values of the linear predictor. For instance, if the logit link function is used, the predicted value of $Pr(Y \leq i|x)$ is

$$1/(1 + \exp(-\hat{\eta}_i))$$

and the lower and upper confidence limits are, respectively,

$$1/(1 - \exp(-(\hat{\eta}_i - Z_{\alpha/2}\hat{\sigma}(\hat{\eta}_i))))$$

and

$$1/(1 + \exp(-(\hat{\eta}_i + Z_{\alpha/2}\hat{\sigma}(\hat{\eta}_i))))$$

SELECTION OF VARIABLES AND MODELING STRATEGY

The current problem was to obtain a model for probability of occurrence of gold deposits for target area delineation in the area of study. This was carried out with soil–geochemical data, water–chemical data, and favorable structural-proximity (such as proximity to lineaments, granite–metabasalt contacts) and lithological data. The nonavailability of water–chemical and soil–chemical information from the same locations, required that they be used separately.

Arsenic, Sb, Hg, and Bi analyzed in soil were used along with "proximity to lineament" for one of the logistic regression models whereas As and Sb analyzed in water samples along with proximity to lineament was used in a separate logistic regression model for prediction of probabilities of mineral occurrences.

MODELING WITH WATER–CHEMICAL DATA

Modeling was carried out using As, Sb, and proximity to lineament (LIN) as explanatory variables and proximity to deposit (DP) as the response variable. The explanatory variables were restricted to a small number because it was determined that these would be sufficient for obtaining a stable model. The number of observations available also does not justify the inclusion of a large number of variables. All the observations selected were restricted to the areas underlain by metabasalt. The water–chemical data were continuous variables, whereas the LIN was a binary variable coded using a corridor width of 0.5 km around lineaments. Observations falling within this buffer zone were coded as 1, whereas others were coded to 0. The response variable, proximity to known deposit, was similarly coded with 0 and 1, with the value 1 accorded to a circular zone of radius 0.5 km around a known deposit or prospect.

MODELING METHODOLOGY

Modeling the probability of occurrence of mineral deposits using the factors described here was done in a stepwise manner. First, the main effects of each factor were tested independently. In the second step, the factors, whose main effects were significant, were incorporated into a multivariate model without any interaction and the significance of the main effects was tested. In the third step, the factors, whose main effects were significant in the second step, were retained and the interactions of all factors, including those whose main effects were not significant, are tested one by one. The results are given next.

Step 1. Univariate analysis was carried out using each independent explanatory variable to establish its statistical significance and possibility of inclusion into the model. Table 1 shows that all the variables are statistically significant at 0.05 significance level via Wald's statistic.

Step 2. Analysis of main effects of the multivariate models then was carried out using all these variables. The coefficients of the parameters and the relevant statistics are given in Table 2. The low p values of Wald's statistic for As and Sb allowed their inclusion in the model. The variable LIN was eliminated because of the relatively high p value of the Wald's statistic and a new multivariate model was fitted without this variable. The relevant statistics

Table 1. Wald's Statistic Computed for Univariate Analysis of Main Effects[a]

Variable	Parameter estimate	Standard error	Wald's statistic	p value
As	3.130	0.959	10.665	0.0011
Sb	3.022	0.847	12.736	0.0004
LIN	4.620	0.957	21.843	0.001

[a] All main effects are significant at $\alpha = 0.05$ level.

Table 2. Wald's Statistic for Multivariate Analysis of Main Effects[a]

Variable	Parameter estimate	Standard error	Wald's statistic	p value	Standardized estimate	
Intercept	−3.480	1.983	3.079	0.079		AIC = 20.63
As	2.637	1.217	4.695	0.030	2.584	SC = 28.6
Sb	1.732	1.136	2.324	0.127	1.284	−2 Log L = 12.603
LIN	1.085	0.992	0.453	0.501	0.294	

[a] Table shows insignificance of variable LIN (proximity to lineament) on basis of Wald's statistic at $\alpha = 0.05$ level. AIC and SC statistic are computed for model for comparison with model without containing LIN (as given in Table 3).

corresponding to various parameters of this model is given in Table 3. The G statistic [the difference between the loglikelihood of the multivariate model containing LIN ($-2 \log L = 12.603$) and the other not containing LIN ($-2 \log L = 13.407$))] is small as can be seen from Tables 2 and 3 confirming the nonsignificance of LIN. The AIC and SC statistics corresponding to the two models are given in Tables 2 and 3. The values of AIC and SC statistics are lower in the model not containing LIN confirming further that LIN is not a significant variable.

Step 3. In this step, all the interaction terms were checked for their significance. This analysis was carried out, by adding each interaction term with the earlier fitted model with only main effects. The significance was assessed via G statistics. The interaction terms As*LIN and As*Sb*LIN are significant statistically at an $\alpha = 0.05$ level (Table 4). The p values of the Wald's statistic for each parameter in the model is given in Table 5. Because the final model includes main effects and interactions, it causes the coefficients of all included parameters to be nonsignificant at the 5% level. However, the p values calculated for each parameter in the model is small (< 0.166) and, therefore, justifies their inclusion in the final integrated model. The final model fitted with As, Sb, and the interaction terms As*LIN and As*Sb*LIN (Table 5).

Table 4. Significance of Interaction Terms in Multivariate Model on Basis of G Statistic of Loglikelihood Ratio

Effects	−2Log-likelihood	G statistic	Degrees of freedom
Main effect	13.047		
+As*Sb	12.883	0.164	1
+As*LIN*	9.190	3.857	1
+Sb*LIN	12.561	0.972	1
+As*Sb*LIN*	8.130	4.817	1

[a] Asterisk marked interaction terms are statistically significant on basis of G statistic.

The model for predicted probabilities for gold target area delineation with water–chemical data and lineaments is thus given by

$$Y = 1/(1 + \exp(-(-6.286 + 3.487*\text{As} + 4.282*\text{Sb} + 3.098*\text{As}*\text{LIN} - 2.619*\text{As}*\text{Sb}*\text{LIN})))$$

Table 3. Wald's Statistic, Loglikelihood Ratio, AIC, and SC Statistics Computed by Eliminating LIN from the Model Used for Obtaining Values of Statistics Given in Table 2[a]

Variable	Parameter estimate	Standard error	Wald's statistic	p value	Standardized estimate	
Intercept	−2.798	1.445	3.746	0.053		AIC = 19.047
As	2.734	1.139	5.805	0.016	2.690	SC = 25.069
Sb	1.548	0.992	2.438	0.118	1.147	−2log L = 13.047

[a] AIC, SC, and −2Log L statistic are found to be lower in comparison to those of model given in Table 2. Main effects of As and Sb are significant at $\alpha = 0.05$ level.

Table 5. Estimates of Parameters of Final Model

Effects	DF	Parameter estimate	Standard error	Wald's statistic	p value	Standardized estimate	Odds ratio
Intercept	1	6.286	3.911	2.584	0.108	—	—
As	1	3.487	2.308	2.282	0.131	3.417	32.671
Sb	1	4.282	2.640	2.628	0.105	3.173	72.362
As*LIN	1	3.098	2.127	2.212	0.145	1.947	22.152
As*Sb*LIN	1	−2.619	1.89	1.921	0.166	−3.114	0.073

DISCUSSION OF RESULTS OF MODELING USING WATER-CHEMICAL DATA.

The given model shows that proximity to lineament (LIN) by itself, is not a significant factor for predicting the probability of gold occurrences. The main effects of As and Sb are significant, showing that high values of these elements are related to gold occurrence. Interaction of As, Sb however, is not significant, whereas the three-way interaction of As, Sb, and LIN is. This factor incorporates the fact that in areas close to lineaments, high values of As and Sb are important. The two-variable interaction between LIN and As shows that high values of As close to lineament are important for deposit location. The interaction of Sb with lineament also should be significant. However, because high values of Sb occur only close to lineaments, the main effect of Sb itself may account for the variation in the response variable.

The predicted probabilities are close to the observed values of the response variable (Table 6). The predicted probability map was prepared by interpolation using inverse distance weighting and then classified into three classes (Fig. 5B). A predicted probability of more than 0.9 discriminates all the mineralized areas from barren areas. This can be seen clearly by comparing Figure 5B with 5A. The latter shows the locations of all known mineralized areas (after Srikantia, 1995). The data from the area to the south of the schist belt, were not used in estimating the parameters of the model. The model was tested on this area where it correctly identified a known deposit.

MODELING WITH SOIL-CHEMICAL DATA

The data set used for modeling includes concentrations of As, Sb, Hg, and Bi in the soil medium. The variables LIN and DP were coded as described earlier. LIN and concentration of As, Sb, Hg, and Bi were used as the explanatory variables, whereas PD is the target variable.

MODELING METHODOLOGY

The modeling strategy followed was similar to that described earlier. One extra step of forward selection and backward elimination of main effects and interaction determined significant in the third step, was used for obtaining the most desirable model. The results are discussed next in stepwise manner.

Step 1. Univariate statistical analysis was carried out with all the five factors. The significance of the variables was tested with Wald's statistic. Table 7 shows that the main effects of all variables are significant.

Step 2. Analysis of the main effects of the multivariate model without interaction terms was carried out in this step. It was inferred from the Wald test (Table 8) that As, Sb, Bi, and LIN are significant statistically as the p values of their Wald statistic was low. Mercury was not significant, showing a high p value of its Wald statistic. Further, the models with and without Hg show a small difference (0.891) in their loglikelihood ratio, indicating nonsignificance via the G statistic at 0.05 level. The AIC and SC statistics also are lower in the model without Hg compared to with Hg, indicating nonsignificance of this variable (Table 9).

Step 3. This step was carried out to check for the statistical significance of interaction terms in the multivariate model. All the possible interaction terms were tested for their statistical significance. This analysis was carried out by adding each interaction term with the earlier fitted multivariate model containing only

Table 6. Predicted Probabilities (PP) Obtained Using Logistic Regression Model with Water-Chemical, and Lineament Data[a]

Samples no.	AS	SB	LIN	DP	PP	Residual
W1	6.77	3.08	1	1	0.996	0.004
W2	15.03	6.15	1	1	0.998	0.002
W3	6.43	2.35	1	1	0.993	0.007
W4	0.1	0.3	0	0	0	0
W5	0.1	0.3	0	0	0	0
W6	41.48	9.62	1	1	0.998	0.002
W7	0.1	0.51	0	0	0	0
W8	6.01	3.71	1	1	0.993	0.007
W9	0.2	4.32	0	0	0	0
W10	0.2	0.8	0	0	0	0
W11	7.07	8.45	1	1	0.995	0.005
W12	0.3	0.6	0	0	0	0
W13	0.69	0.2	0	0	0	0
W14	0.1	0.4	0	0	0	0
W15	0.1	0.3	0	0	0	0
W16	26.94	5.6	1	1	0.999	0.002
W17	0.5	0.2	0	0	0	0
W18	0.4	0.3	0	0	0	0
W19	0.4	0.1	0	0	0	0
W20	0.3	0.1	0	0	0	0
W21	0.2	0.2	0	0	0	0
W22	0.3	0.1	0	0	0	0
W23	12.78	3.5	1	1	0.999	0.001
W24	0.3	0.1	0	0	0	0
W25	1.15	2.3	1	0	0.109	−0.109
W26	0.7	0.3	0	0	0	0
W27	0.58	0.3	0	0	0	0
W28	0.1	0.2	0	0	0	0
W29	0.1	0.5	0	0	0	0
W30	0.1	0.4	0	0	0	0
W31	0.98	0.41	0	0	0	0
W32	0.87	0.22	0	0	0	0
W33	0.67	0.31	1	0	0	0
W34	0.99	0.41	1	1	0	0
W35	0.63	0.22	1	0	0	0
W36	5.43	1.11	1	1	0.992	0.008
W37	1.24	0.53	0	0	0.001	−0.001
W38	0.89	0.52	0	0	0	0
W39	1.31	4.22	0	1	0.689	0.311
W40	4.68	1.62	1	1	0.982	0.018
W41	2.79	0.32	1	0	0.206	−0.206
W42	2.38	0.91	1	0	0.317	−0.317
W43	0.43	0.22	0	0	0.001	−0.001
W44	4.23	1.43	1	1	0.967	0.033
W45	6.21	1.22	0	1	0.717	0.283
W46	5.87	2.54	1	1	0.993	0.007
W47	4.68	1.11	0	0	0.387	−0.387
W48	0.87	0.43	0	0	0.001	−0.001
W49	1.23	0.42	0	0	0.001	0.001
W50	4.78	4.99	1	1	0.987	0.013
W51	6.38	1.11	1	1	0.997	0.003
W52	3.25	0.7	1	1	0.741	0.259
W53	1.38	0.39	0	0	0.001	−0.001
W54	4.32	3.19	1	1	0.979	0.021
W55	5.64	1.92	1	1	0.993	0.007
W56	5.12	4.12	1	1	0.989	0.011

Table 6 Continued.

Samples no.	AS	SB	LIN	DP	PP	Residual
W57	7.43	2.47	1	1	0.998	0.002
W58	21.22	18.2	1	1	0.955	0.045
W59	12.33	1.21	1	1	0.999	0.001
W60	16.48	8.32	1	1	0.997	0.003
W61	19.88	6.32	1	1	0.999	0.001
W62	8.34	1.82	1	1	0.999	0.001
W63	3.79	2.14	1	1	0.956	0.047
W64	1.02	0.56	1	0	0.002	−0.002
W65	26.57	8.92	0	1	0.999*	0.001
W66	5.79	3.32	1	1	0.993*	0.007
W67	36.55	16.12	1	1	0.957*	0.043
W68	34.38	15.91	1	1	0.962*	0.038

[a] Values of explanatory variables (As, Sb, and LIN), response variable (DP), predicted probabilities, and residual are shown. Predicted probabilities in Sanbal area are marked with asterisk (*). Observations from this area were excluded during estimation of model parameters.

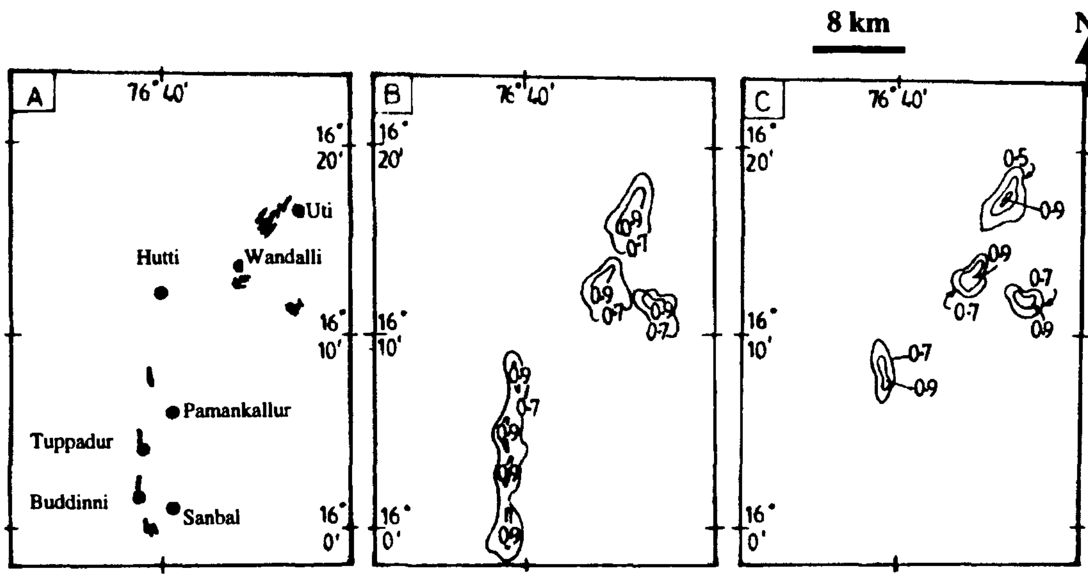

Figure 5. Predicted probability map for gold occurrences using logistic regression analysis. Location of known gold occurrences is given in A. Predicted probability map for gold occurrences using water—chemical and soil-chemical data are given in B and C, respectively.

main effects. The significance was assessed via the G statistic. From the Wald–Chi square test and the p values of the G statistics, GX, JX, NX, PX, UX, WX, XX, YX, ZX1, ZX2, ZX3, and ZX5 are statistically significant (Table 10). These interactions then were included with the main effects for a stepwise forward selection and backward elimination in the following step.

Step 4. Fitting of the model started with a test of the intercept-only model and an evaluation

Table 7. Univariate Statistical Analysis of Main Effects[a]

Effects	Parameter estimate	Standard error	Wald's statistic	p value
As	−1.288	0.321	16.160	0.0001
Sb	−1.692	0.485	11.920	0.0006
Hg	−0.837	0.347	5.821	0.0160
Bi	−1.161	0.361	10.347	0.0013
LIN	−2.590	0.548	22.310	0.0001

[a] All main effects are significant at $\alpha = 0.05$ level.

Table 8. Multivariate Statistical Analysis of Main Effects[a]

Effects	Parameter estimate	Standard error	Wald's statistic	p value	Standardized estimate	
Intercept	12.0404	3.015	12.911	0.0001		
As	−0.649	0.376	2.989	0.0438	−0.474	
Sb	−1.844	0.638	8.338	0.0039	−0.668	
Bi	−1.352	0.568	5.671	0.0173	−0.607	
Hg	−0.286	0.593	0.232	0.6299	−0.123	AIC = 66.842
LIN	−2.042	0.783	6.805	0.0091	−0.566	SC = 81.059

[a] All variables except Hg are significant at $\alpha = 0.05$ level.

Table 9. Multivariate Model with All Main Effects Except That of Hg

Effects	Parameter estimate	Standard error	Wald's statistic	p value	Standardized estimate	
Intercept	12.212	3.014	16.411	0.0001		
As	−0.676	0.372	3.303	0.0492	−0.494	
Sb	−1.858	0.641	8.398	0.0038	−0.674	
Bi	−1.412	0.560	6.357	0.017	−0.634	AIC = 65.082
LIN	−2.073	0.770	7.114	0.007	−0.575	SC = 76.632

of the likelihood ratios. Then each of the possible univariate logistic regression model was fitted and comparisons made of their respective likelihoods. The most important variable is the one with smallest p value at every step. This variable is entered into the next model. The alpha level to judge the significance of the variable was fixed at 0.15. Elimination of the insignificant variables was decided on the basis of the G statistic. The final model again was examined for the significance of the coefficients. Significance of the coefficients of the explanatory variables was checked with Wald's statistic. The finally selected main effects were Sb, Bi, and LIN. The result is given in Table 11 and the estimated coefficients of the factors and their interactions of the final model are given in Table 12.

The final model for predicting the probability for gold target area delineation using soil data and proximity to lineaments is given by

$$Y = 1 / (1 + \exp(-(15.786 - 3.542\text{Sb} - 2.023\text{Bi} - 3.869\text{LIN} - 0.773\text{As*Bi*Hg*LIN} + 1.394\text{Sb*Bi*Hg*LIN})))$$

DISCUSSION OF RESULTS OF MODELING WITH SOIL–CHEMICAL DATA

The model shows that the main effects of Sb, Bi, and LIN are significant. The interaction terms show that close to the lineaments, Bi and Hg have interaction with As and Sb separately. This indicates that high values of As and Sb close to lineaments are related to the high values of Bi and Hg giving rise to the interaction. Where Sb values are high, As is high, but because of the wide geographical spread of As anomalies in contrast to the relatively narrow geographic spread spread of Sb anomalies, wherever the As values are high, Sb values are not necessarily high. Thus two separate interaction terms are in the model. Two-way and three-way interactions are not in the model because the four-way interactions are more significant.

The predicted probabilities are given in Table 13. The predicted probability map, was prepared by contouring the predicted probabilities with the inverse distance weighting algorithm. The map was classified into four groups as shown in Figure 5C. The predicted probability values are high around locations of known deposits. This can be seen by comparing Figure 5C and A. Predicted probabilities of greater than 0.9 distin-

Table 10. Analysis of Main Effects and Their Interactions[a]

	Symbol	DF	G statistic
Main effect			55.078
+As*Sb	FX	1	55.078
+As*Bi	GX	1	47.846 *
+As*Hg	HX	1	54.971
+As*LIN	IX	1	53.332
+Sb*Bi	JX	1	52.886 *
+Sb*Hg	KX	1	54.890
+Sb*LIN	LX	1	54.690
+Bi*Hg	MX	1	54.978
+Bi*LIN	NX	1	50.683 *
+Hg*LIN	OX	1	53.385
+As*Sb*Bi	PX	1	46.180 *
+As*Bi*Hg	QX	1	54.674
+As*Hg*LIN	RX	1	54.285
+As*Sb*Hg	SX	1	54.990
+As*Sb*LIN	TX	1	54.738
+As*Bi*LIN	UX	1	52.883 *
+Sb*Bi*Hg	VX	1	54.780
+Sb*Bi*LIN	WX	1	45.766 *
+Bi*Hg*LIN	XX	1	50.465 *
+Sb*Hg*LIN	YX	1	53.045 *
+As*Sb*Bi*Hg	ZX	1	54.461
+As*Sb*Bi*LIN	ZX1	1	51.036 *
+As*Bi*Hg*LIN	ZX2	1	52.659 *
+Sb*Bi*Hg*LIN	ZX3	1	50.096 *
+Sb*Hg*LIN*As	ZX4	1	54.078
+As*Sb*Hg*BI*LIN	ZX5	1	52.496 *

[a] Asterisk marked effects are significant at $\alpha = 0.05$ level.

Table 12. Estimates of Parameters of Final Model

Effects	Estimates	Standard error	Wald's statistic	p value	Odds ratio
Intercept	15.786	3.985	15.689	0.0001	
Sb	−3.542	0.966	13.433	0.0002	0.029
Bi	−2.023	0.728	7.713	0.0055	0.132
LIN	−3.869	1.102	12.333	0.0004	0.021
As*Bi*Hg*LIN	−0.773	0.474	2.658	0.1030	0.462
Sb*Bi*Hg*LIN	1.394	0.689	4.087	0.0432	4.033

guish all known mineralized areas. The effectiveness of the model was tested using a few observations from the Wandali area, which were not used in formulating the model. The model clearly identified it as a high probability area of gold mineralization (Fig. 5*C*). Further, all areas with more than 0.9 predicted probability level had known mineral occurrences (i.e., no false positives were identified).

CONCLUSIONS

The results of logistic regression show that this method is adequate for identifying mineralized areas with the type of data used. Known deposits are identified with values of predicted probability greater than

Table 11. Stepwise Variable Selection Using Maximum Likelihood Method[a]

Step	ZX1	Sb	ZX3	Bi	LIN	ZX2	WX	As	GX	JX	HX	PX	UX	XX	YX	ZX5
0	0.0001	0.0001	0.0041	0.0005	0.0001	0.0037	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0042	0.0003	0.0039
1	0.0001	0.0116	0.0587	0.4087	0.4260	0.1630	0.0410	0.0259	0.2120	0.0460	0.1264	0.0223	0.3847	0.0450	0.9747	0.2023
2	0.0856	0.0167	0.0083	0.0964	0.0764	0.0003	0.0993	0.0778	0.0789	0.1524	0.8793	0.1304	0.1143	0.0219	0.4427	0.0428
3	0.0006	0.0084	0.0289	0.0260	0.2674	0.2117	0.4399	0.2042	0.0396	0.0414	0.8675	0.0573	0.2752	0.2355	0.2186	0.3978
4	0.0014	0.0028	0.0094	0.0362	0.0450	0.1805	0.3426	0.0925	0.6273	0.2319	0.9275	0.8225	0.3154	0.3012	0.0922	0.3349
5	0.6625*	0.0003	0.0311	0.0048	0.0001	0.0026	0.0748	0.0266	0.1252	0.3969	0.2496	0.1711	0.5166	0.7045	0.4766	0.0082
6	0.2872*	0.0002	0.0432	0.0055	0.0004	0.1030	0.1175	0.2076	0.9188	0.4292	0.2338	0.9964	0.3299	0.9878	0.3985	0.9587

where,

FX	As*Sb	OX	Hg*LIN	XX	Bi*Hg*LIN
GX	As*Bi	PX	As*Sb*Bi	YX	Sb*Hg*LIN
HX	As*Hg	QX	As*Bi*Hg	ZX	As*Sb*Bi*Hg
IX	As*LIN	RX	As*Hg*LIN	ZX1	As*Sb*Bi*LIN
JX	Sb*Bi	SX	As*Sb*Hg	ZX2	As*Bi*Hg*LIN
KX	Sb*Hg	TX	As*Sb*LIN	ZX3	Sb*Bi*Hg*LIN
LX	Sb*LIN	UX	As*Bi*LIN	ZX4	Sb*Hg*LIN*As
MX	Bi*Hg	VX	Sb*Bi*Hg	ZX5	As*Sb*Hg*BI*LIN
NX	Bi*LIN	WX	Sb*Bi*LIN		

[a] P Values of terms to right of vertical line are not included into model at that step. Terms to left vertical line are incorporated into model steps. Asterisk-marked effects eliminated from model in later step.

Table 13. Predicted Probabilities (PP) Obtained Using Logistic Regression Model with Soil-Chemical, and Lineament Data

Sample nos.	As	Sb	Hg	Bi	LIN	DP	PP	Residual
U1	12.56	15.98	1.78	1.41	0	0	0.034	−0.034
U2	17.34	30.66	1.09	1.61	0	0	0.063	−0.063
U3	58.79	32.85	1.78	1.61	0	0	0.079	−0.079
U4	114.94	40.21	1.12	1.61	1	1	0.886	0.114
U5	59.84	32.08	1.1	2.2	0	0	0.129	−0.129
U6	6.78	31.5	0.71	1.31	0	0	0.046	−0.046
U7	25.86	36.44	0.93	1.2	0	0	0.064	−0.064
U8	65.57	37.9	1.26	1.5	0	0	0.109	−0.109
U9	123.39	38.32	1.98	2.57	1	1	0.883	0.117
U10	877.32	34.36	1.85	3.73	1	1	0.971	0.029
U11	404.36	35.78	1.47	6.09	1	1	0.984	0.016
U12	90.93	28.04	2.81	5.22	1	1	0.777	0.223
U13	550.12	37.74	3.98	16.7	1	1	0.997	0.003
U14	700.98	31.72	4.06	12.5	1	1	0.998	0.002
U15	417.87	41.54	2.81	14.1	1	1	0.994	0.006
U16	136.9	32.49	1.89	14.6	1	1	0.982	0.018
U17	130.25	27.06	1.69	12.5	0	1	0.732	0.268
U18	5.32	32.99	1.89	14.6	0	1	0.882	0.108
U19	103.95	41.55	2.2	13.3	0	1	0.962	0.038
U20	167.43	36.88	2.63	5.11	1	0	0.022	−0.022
U21	474.63	29.71	1.76	2.47	1	1	0.874	0.126
U22	109.66	38.98	3.87	3.16	1	0	0.247	−0.247
U23	105.84	36.33	3.73	3.79	1	1	0.738	0.262
U24	40.61	15.65	1.24	1.76	0	0	0.007	−0.007
U25	145.83	39.52	1.73	4.22	0	1	0.237	−0.237
U26	1168.7	22.77	1.17	3.74	1	1	0.886	0.114
U27	339.08	15.65	1.24	1.76	1	1	0.721	0.279
U28	293.28	6.49	1.76	7.75	1	1	0.714	0.286
U29	95.05	25.65	0.87	4.17	0	0	0.197	−0.197
U30	90.87	6.34	0.67	4.17	0	0	0.002	−0.002
U31	91.87	35.57	0.91	4.1	1	0	0.229	−0.229
U32	231.64	33.57	1.13	4.08	1	1	0.963	0.037
U33	1258.1	28.16	2.51	4.06	1	0	0.244	−0.244
U34	261.97	28.85	2.11	4.71	1	1	0.935	0.065
U35	260.34	19.09	0.67	4.17	1	1	0.787	0.213
U36	28.38	15.37	0.73	2.76	0	0	0.017	−0.017
U37	160.47	32.87	0.78	4.28	0	0	0.383	−0.383
U38	286.38	31.51	1.34	2.93	1	0	0.199	−0.199
U39	148.24	31.94	1.76	5.86	0	0	0.315	−0.315
U40	104.07	33.85	1.08	3.79	0	0	0.250	−0.250
U41	73.83	22.78	1.89	5.81	0	0	0.239	−0.239
U42	679.05	24.71	1.03	7.26	1	1	0.971	0.029
U43	1025.1	35.25	0.37	11.3	1	1	0.991	0.009
U44	33.79	34.97	0.68	17.2	0	1	0.928	0.072
U45	41.01	22.85	0.34	4.68	0	0	0.171	−0.171
U46	52.41	20.99	0.79	3.93	0	0	0.097	−0.097
U47	14.93	16.33	0.11	4.22	0	0	0.048	−0.048
U48	52.33	27.66	0.31	4.47	0	0	0.269	−0.269
U49	38.28	24.81	0.47	5.68	0	0	0.289	−0.289
H1	593.33	32.57	10.8	21.6	1	1	0.999	0.001
H2	979.58	40.17	15.2	20	1	1	0.999	0.001
H3	67.54	32.53	7.01	21.6	1	0	0.050	−0.050
H4	294.95	32.18	4.66	24.3	1	1	0.991	0.009
H5	616.36	16.47	1.66	18.4	1	1	0.995	0.005
H6	123.67	11.29	4.32	5.2	0	0	0.020	−0.020
H7	20.37	13.47	3.37	15.2	1	0	0.255	−0.255

Table 13. Continued.

Sample nos.	As	Sb	Hg	Bi	LIN	DP	PP	Residual
H8	81.15	15.56	2.47	4.1	0	0	0.038	−0.038
H9	74.78	11.08	4.32	5.2	0	0	0.019	−0.019
H10	40.69	11.48	4.76	0.89	0	0	0.001	−0.001
H11	68.92	11.46	3.37	15.2	0	0	0.162	−0.162
H12	154.27	5.07	2.47	4.1	0	0	0.001	−0.001
H13	86.76	11	1.38	2.77	0	0	0.005	−0.005
H14	12.27	7.21	2.76	3.32	0	0	0.002	−0.002
181	2.79	5.96	1.37	4.1	0	0	0.001	−0.001
182	3.49	7.21	2.76	3.32	0	0	0.002	−0.002
183	20.63	8.37	1.47	8.01	1	0	0.336	−0.336
188	17.22	4.32	1.02	3.98	1	0	0.019	−0.019
189	43.87	4.76	1.79	3.96	0	0	0.001	−0.001
190	32.25	6.37	1.42	11.8	1	1	0.731	0.269
196	198.48	37.92	1.73	5.73	1	1	0.972	0.028
197	670.15	59.57	4.32	4.67	1	1	0.984	0.016
198	155.58	31.37	1.03	4.83	1	1	0.969	0.031
204	125.88	18.27	0.67	0.41	1	0	0.028	−0.028
205	164.39	10.23	1.21	4.12	1	0	0.348	−0.348
206	122.39	7.63	1.03	7.01	1	0	0.326	−0.326
211	116.28	6.21	1.87	6.87	1	0	0.253	−0.253
212	140.37	33.68	1.32	1.22	0	0	0.051	−0.051
213	279.99	47.42	1.74	2.31	1	1	0.951	0.049
214	108.96	29.42	2.42	7.62	1	1	0.901	0.091
215	85.91	43.41	3.56	2.78	0	1	0.711	0.289
72	79.765	12.78	3.678	12.654	0	0	0.016*	−0.016
73	3.876	9.543	4.532	11.56	0	0	0.009*	−0.009
91	20.56	12.37	2.67	6.231	0	0	0.007*	−0.007
92	438.45	11.987	3.456	4.1	1	1	0.826*	0.174
W1	657.543	12.575	3.321	11.876	0	0	0.012*	−0.012
W2	640.234	53.57	5.43	6.43	1	1	0.979*	0.021
W3	155.89	12.45	1.67	12.56	1	0	0.191*	−0.191
W4	213.67	32.675	12.567	12.45	0	1	0.843*	0.157
W5	234.78	36.78	6.78	4.89	1	1	0.907*	0.093

[a] Values of explanatory variables (As, Sb, Bi, Hg, and LIN), response variable (DP), predicted probabilities, and residual are shown. Predicted probabilities in Wandali area are marked with asterisk. Observations from this area were excluded during estimation of model parameters.

0.9 in all known situations. The geochemistry of As, Sb, Hg, and Bi in soil or As and Sb in water, in association with lineaments such as litho-contacts and shear zones, can be used in similar terrains for targeting gold-mineralized areas.

ACKNOWLEDGMENTS

The authors acknowledges the Department of Earth Sciences for providing the necessary facilities for carrying out this work. The hospitality provided by the Hutti Gold Mines Co. Ltd and the guidance given by Dr. K. K. Raju, DGM (Exploration) for carrying out field work in and around Hutti is gratefully acknowledged. The analytical work would not have been possible without the cooperation and dedication of Mr. L. S. Mombasawala, Research Engineer and the analysts Deepa and Bakul of the ICP-AES laboratory, Regional Sophisticated Instrumentation Centre, IIT, Bombay. The authors also are grateful to the USGS for providing them with GXR-2 and GXR-6 soil standards. We sincerely acknowledge Anand and Sukanta for their effort in providing the figures in a presentable form.

REFERENCES

Agterberg, F. P., 1974, Automatic contouring of geologic maps to detect target areas for mineral exploration: Math. Geology, v. 6, no. 4, p. 373–394.

Agterberg, F. P. 1989, Computer programs for mineral exploration: Science, v. 245, no. 4913, p. 76–81.

Agterberg, F. P., 1992, Combining indicator patterns in weights of evidence modelling for resource evaluation: Nonrenewable Resources, v. 1, no. 1, p. 39–50.

Agterberg, F. P., Bonham-Carter, G. F., Cheng, Q., and Wright, D. F., 1993, Weights of evidence modelling and weighted logistic regression for mineral potential mapping, *in* Davis, J. C., and Herzfield, U. C., eds, Computers in Geology - 25 Years of Progress, Oxford Univ. Press, New York, p. 13–32.

An, P., Moon, W. M., and Bonham-Carter, G. F., 1994a, Uncertainty management in integration of exploration data using the belief function: Nonrenewable Resources, v. 3, no. 1, p. 60–71.

An, P., Moon, W. M., and Bonham-Carter, G. F., 1994b, An object oriented knowledge representation structure for exploration data management: Nonrenewable Resources, v. 3, no. 2, p. 132–145.

Bonham-Carter, G. F., 1994, Geographic information systems for geoscientists: Modelling with GIS: Pergamon Press, Oxford, 398 p.

Bonham-Carter, G. F., and Chung, C. F., 1989, Integration of mineral resource data for Kasmere lake area, Northwest Manitoba, with emphasis on uranium: Computers & Geosciences, v. 15, no. 1, p. 25–45.

Bonham-Carter, G. F., Agterberg, F. P., and Wright, D. F., 1988, Integration of geological datasets for gold exploration in Nova Scotia: Photogrammetric Engineering and Remote Sensing, v. 54, no. 11, p. 1585–1592.

Bonham-Carter, G. F., Agterberg, F. P., and Wright, D. F., 1989, Weights of evidence modelling: new approach to mapping mineral potential, *in* Agterberg, F. P., and Bonham-Carter, G. F., eds, Statistical Applications in the Earth Sciences: Geol.Survey Canada Paper 89-9, p. 171–183.

Botbol, J. M., Sinding-Larsen, R., McCammon, R. B., and Gott, G. B., 1978, A regionalised multivariate approach to target selection in geochemical exploration: Econ.Geology, v. 73, no. 4, p. 534–546.

Boyle, R. W., 1979, The geochemistry of gold and its deposits: Geol. Survey Canada Bull. 280, 584 p.

Boyle, R. W., 1987, Gold: history and genesis of gold deposits: Von Nostrand, New York, 675 p.

Briggs, P. L., and Press, F., 1977, Pattern recognition applied to uranium prospecting: Nature, v. 268, no. 5616, p. 125–127.

Butt, C. R. M., and Smith, R. E., 1980, Conceptual models in exploration geochemistry: Jour. Geochem. Exploration, v. 12, no. 1, p. 89–365.

Campbell, A. N., Hollister, V. F., and Duda, R. O., 1982, Recognition of a hidden mineral deposit by an artificial intelligence program: Science, v. 217, no. 4563, p. 927–929.

Chung, C. F., 1977, An application of discriminant analysis for the evaluation of mineral potential, *in* Ramani. R. V., ed., Proc. 14th Intern. Symp. Application of Computer Methods in the Mineral Industry, (Pennsylvania State Univ.) AIME, New York, p. 299–311.

Chung, C. F., and Agterberg, F. P., 1980, Regression models estimating mineral resources from geological map data: Math. Geology, v. 12, no. 5, p. 473–488.

Clifton, H. E., Hunter, R. E., Swanson, F. T., and Phillips, R. L., 1969, Sample size and abundance of gold in selected stream sediments, Southern British Columbia, Canada: Jour. Geochem. Exploration, v. 26, no. 2, p. 203–214.

Cox, D. P., and Singer, D. A., 1986, eds., Mineral deposit models: U.S. Geol. Survey Bull. 1693, 379 p.

Cox, D. R., and Snell, E. J., 1989, The analysis of binary data: (2nd edn.) Chapman and Hall, London, 287 p.

Duda, R. O., Hart, P. E., Nilsson, N. J., and Sutherland, G. L., 1978, Semantic network representations in rule-based interference systems, *in* Waterman, D. A., and Hayes-Roth, F., eds., Pattern-directed Inference Systems: Academic Press, New York, p.203–221.

Freeman, D. H. Jr., 1987, Application of categorical data analysis: Marcel Dekker Inc., New York, 318 p.

Giritharan, T. S., and Rajamani, V., 1998. Geochemistry of metavolcanics of Hutti-Maski schist belt, South India-implications to gold metallogeny in the eastern Dharwar craton: Jour. Geol. Soc. India, v. 51, no. 5, p. 583–594.

Greenberg, A. E., Clesceri, L. S., and Eaton, A. D., ed., 1992, Standard methods for the examination of water and waste water (18th ed.): APHA-AWWA-WEF Publ., 769 p.

Harris, D. P., 1966, A probability model for mineral wealth: Trans. of Symposium Mining and Exploration, June, 1996, p. 199–216.

Harris, D. P., 1984, Mineral resources appraisal: Oxford Univ. Press, London, 445 p.

Harris, D. P., and Pan, G. C., 1987, An investigation of quantitative methods and multivariate relations designed explicitly to support the estimation of mineral resources-intrinsic samples: USGS grant 14-08-000-G1399, Research Rept., 200 p.

Harris, D. P., and Pan, G. C., 1990, Subdividing consistent geologic areas by relative exceptionalness of additional information: methods and case study. Econ. Geology, v. 85, no. 1, p. 1072–1083.

Harris, D. P., and Pan, G. C., 1991, Consistent geologic areas for epithermal gold-silver deposits in the Walker lake Quadrangle of Neveda and California: delineated by quantitative methods: Econ. Geology, v. 86, no. 1, p. 142–165.

Harris, J. F., 1982, Sampling and analytical requirement for effective use of geochemistry in exploration for gold, *in* Levinson, A. A., ed., Precious Metals in the Northern Cordillera: AEG Spec. Publ., v. 10, p.53–67.

Hosmer, D. W. Jr., and Lemeshow, S., 1989, Application of logistic regression: John Wiley & Sons, Inc., New York, 307 p.

Katz, S. S., 1988, Using the MAPS geographic information system as a geographic expert system: Proc. 5th Nat. Moss Users Workshop (New Orleans), p. 167–187.

Katz, S. S., 1991, Emulating the Prospector expert system with a raster GIS: Computers & Geosciences, v. 17, no. 7, p. 1033–1050.

McCammon, R. B., 1990, Prospector III, *in* Agterberg, F. P., and Bonham-Carter, G. F., eds., Statistical Applications in the Earth Sciences: Geol. Survey of Canada Paper 89-9, p. 395–404.

McCammon, R. B., 1989, Prospector II-The redesign of Prospector: AI systems in government (March, 27-31, 1989), Washington, D.C., p. 88–92.

McCammon, R. B., 1992, Numerical mineral deposit models, *in* Bliss, J. D., ed., New Developments in Mineral Exploration and Resource Assessment: U.S.Geol. Survey Bull. 2004, p. 6–12.

Moon, W. M., 1990, Integration of geophysical and geological data using evidential belief function: IEEE Trans. Geoscience and Remote Sensing, v. 28, p. 711–720.

Nichol, I., Closs, C. G., and Lavin, O. P., 1989, Sample representativity with reference to gold exploration, *in* Garland, G. D. ed., Proc. Exploration' 87, Ontario Geol. Survey Spec. Vol. 3, no. p. 609–624.

Pan, G., 1993, Indicator favourability theory for mineral potential mapping: Nonrenewable Resources, v. 2, no. 1, p. 7–24.

Pan, G., 1996, Extended weights of evidence modelling for the pseudo-estimation of metal grades: Nonrenewable Resources, v. 5, no. 1, p. 53–76.

Pan, G., and Harris, D. P., 1992, Estimating a favourability function for the integration of geodata and selection of mineral resources: Math. Geology, v. 24, no. 2, p. 177–202.

Pan, G., Harris, D. P., and Heiner, T., 1992, Fundamental issues in quantitative estimation of mineral resources: Nonrenewable Resources, v. 1, no. 4, p. 281–292.

Reddy, R. K. T., and Bonham-Carter G. F., 1991, A decision-tree approach to mineral potential mapping in Snow Lake area, Manitoba: Can. Jour. Remote Sensing, v. 17, no. 1, p. 191–200.

Reddy, R. K. T., Bonham-Carter, G. F., and Galley, A.G., 1992, Developing a geographic expert system for regional mapping of volcanogenic massive sulfide (VMS) deposit potential: Nonrenewable Resources, v. 2, no. 2, p. 112–124.

Roy, A., 1979, Poly-phase folding deformation in the Hutti-Maski schist Belt, Karnataka: Jour. Geol. Soc. India, v. 20, no. 12, p. 598–607.

Roy, A., 1991, The geology of gold mineralisation at Hutti in Hutti-Maski schist Belt, Karnataka, India: Indian Minerals, v. 5, no. 4, p. 229–250.

Sahoo, N. R., 1998. An analysis of secondary dispersion patterns of selected elements in soil and ground water of the auriferous Hutti-Maski schist belt, Raichur, Karnataka, India: unpub. doctoral dissertation, IIT Bombay (India), 100 p.

Singer, D. A., 1985, Preliminary version of FINDER, a Pascal program for locating mineral deposits with spatial information: U.S. Geol. Survey Open-File Rept. 85–590, 24 p.

Singer, D. A., and Kouda, R., 1988, Integrating spatial and frequency information in the search for Kuroko deposits of the Hokuroku district, Japan: Econ. Geology, v. 83, no. 1 p. 18–29.

Singer, D. A., and Kouda, R., 1996, Application of a feedforward neural network in the search for Kuroko deposits in the Hokuroku district, Japan: Math. Geology, v. 28, no. 8, p. 1017–1023.

Singer, D. A., and Mosier, D. L., 1981, A review of regional mineral resource assessment methods: Econ. Geology, v. 76, no. 5, p. 1006–1015.

Srikantia, S. V., 1995, Geology of Hutti-Maski Greenstone Belt, *in* Curtis, L. C., and Radhakrishna, B. P., eds:, Hutti Gold Mine, into the 21st Century: Geol. Soc. India (Bangalore), p. 8–27.

Trafford, J. M., 1986, Investigations into the stage of ground water containing trace levels of arsenic prior to the analysis by hydride generation combined with ICP-AES: British Geol. Survey, Tech. Rept., 86/8, p.