We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

6,900
Open access books available

185,000
International authors and editors

200M
Downloads

154
Countries delivered to

Our authors are among the

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

BOOK CITATION INDEX
CLARIVATE ANALYTICS
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Distribution Law for Constituent Minerals and Chemical Components in Rocks and Ores

Basanta K. Sahu

Additional information is available at the end of the chapter

Abstract

Rocks and Ores comprising of multicomponent of minerals (Group I) contain chemical molecules (Group II) which further contain (Group III). Man has exploited these valuable constituents for industrial, economic, and social growth causing serious depletion of high-grade ores at surface and shallow depths. Sustainable growth now requires exploitation of low-grade ores and those occurring at depth which imply optimal and most efficient mining efforts based on spatial, temporal, and spatiotemporal distribution of these constituents. The optimal decisions resulting in profit maximization is possible by obtaining precise and accurate parameters of these distributions. Sample size required for such estimations must be at least representative elementary volume (REV) of the rocks/ores, but the data matrix is not full-rank for statistical analyses and sample mean (x; $0 < x < 1$) must be transformed to Gaussian to apply standard univariate (UND)/multivariate (MND) statistical techniques. A $\log(x/(1-x))$ or $\ln(x/(1-x))$ transform is shown to be an appropriate pre-transformation that eliminates the twin problems of full-rank, and spurious negative correlations as well as makes the distribution Gaussian for major, minor, and trace components. Mining applications using the univariate and/or Multivariate Normal Theory of pre-transformed sample mean (x) in rocks/ores is optimal for anomaly detection, drilling site selection, global reserve and grade estimations, mine planning, ore mineral liberation, blending, sustainable mine developments and maximization of profits computed on net profit value (NPV) basis to present time.

Keywords: hierarchical geochemical constituents, fractional mineral and chemical constituents (x; $0 < x < 1.0$ or 100%) in rocks and ores, logarithmic pre-transformation of sample data using $\log(x/(1-x))$, average Gaussian, statistical and geological inferences for reserves and average grades, mine life at minimum economic extractions, profit maximization through estimation of optimal cutoff grade

1. Introduction

Large-scale industrialization and urbanization with greater consumption of mineral and energy resources since the "Industrial Revolution" have greatly depleted high-grade resources available at shallow depths of crust [1–3]. Therefore, more intensive search for these finite resources is necessary at greater depths and for lower grades besides optimal mining, processing, and marketing in order to maintain sustainable development of economy and society and to have higher national growth. Mineral resources include high-grade marketable ores and/or associated low-grade ores which require blending and/or beneficiation to market the mined and/or processed products at profit [4, 5].

Rocks and ores comprise of multielement/multimineral materials containing more than three constituents (C > 3), some of which need separation and/or beneficiation to remove nonmarketable gangue minerals and waste materials [1, 6]. We are interested in static characterization (invariant in time) of the heterogeneous solids (rocks/ores) which can be achieved through sampling of adequate material (>representative elementary volume (REV) [7]) from the rock/ore, where REV means the minimum volume/weight that represents elementary volume. Thus, integration of any random variable (fractional constituent, x with 0 < x < 1) over the REV provides stable and unbiased statistics representing the characteristics of an equivalent statistically isotropic homogeneous sample. The probability density functions (PDFs) corresponding to the random fractional constituents (such as minerals, molecules, elements, or isotopes) of the rock/ore are having a binomial distribution for major and minor constituents which reduces to a Poisson distribution for trace constituents. Realistic and accurate PDFs are essential for optimal exploration, mining, mine planning, processing, blending, beneficiation, and marketing of the concerned mineral resource. In contrast to rocks and ores, mineral geochemistry involves sampling of (homogeneous) minerals which would be much smaller in volume than REV for geochemical analysis and inference [8].

Major, minor, or trace fractional constituents (x; 0 < x < 1) of rocks and ores are Gaussianized using the log(x) pre-transformation since 1954 as advocated by Ahrens [9]. However, this log (x) pre-transformation is not always valid for major constituents and not necessarily independent for these constituents; hence, a slight modification introducing multiplicative random errors and independence of transformed random constituents, such as log(x/(1 − x)), would be useful for all three levels of constituents [10, 11]. Rocks and ores usually contain 5–6 minerals (some at major and minor and a few at trace levels) and about 10–12 major elements and a large number of trace elements, which can be explained by thermodynamic principles for closed systems having 2 degrees of freedom. However, open systems have a larger number of degrees of freedom and hence a larger number of phases (minerals).

From statistical viewpoint, the main problems of geochemical analysis and inference of such data (without pre-transformation) are:

- Constituents are usually weight, volume, area, or length fractions but not numbers as needed in statistics and probability.

- Weight fractions are not necessarily equal to volume, area, or length fractions (as given in stereology textbooks) since rocks/ores contain finite pore spaces.

- Statistics obtained in 3D space are not equal to corresponding statistics on length or number basis unless the total measure tends to be infinite (but sample size is finite).

- Probability distributions of constituents are not Gaussian for statistical hypotheses tests and inferences.

Four additional and more complex problems include:

i. Constant sum or closure: This constraint on fractional constituents induces spurious negative correlations among the constituents within samples which confound statistical and geochemical inference.

ii. Data matrix has a rank (C-1) and is not full rank (C) to have a unique inverse.

iii. Fractional constituents are not independent as required for parameter (such as mean and variance of order); two are finite and hence useful for parameter estimation for mean and variance, and all higher-order cumulants are zeros and hence dropped.

iv. Parameter estimation, hypotheses tests, and risk analyses are simple and straightforward on Gaussian data.

Mineral resources are non-replaceable national assets which should be optimally utilized for sustainable economic growth, social benefits, and improvements in health and life quality without jeopardizing present ecology/environment for future growth. Mineral resources are characteristically nonrenewable and form prime assets for stable economy and quality of life [2]. Optimal extraction and marketing strategies must be evolved under dynamic conditions of high risks in estimation of reserves and grades of marketable as well as lean-grade resources for optimizing the net profits at present value under highly fluctuating pricing and marketing conditions and also resource augmentation through exploration efforts and/or technological substitutions and proper waste disposals and of learning and updating. The main aim of the mining industry is to maximize the accrued profits using the concept of net present value (NPV) while simultaneously minimizing damage to ecology and environment, and improving the social and health needs of the local community seems essential to achieve conservation of high-grade ores and sustainable development. Capital must be generated at a faster rate than deployment for new discoveries or for R&D efforts for technological substitutions, and the accumulated capital thus helps in industrial and social growths and safeguards and improves life quality of people [1, 11–14].

Mineral resources not only comprise high-grade directly marketable ores but also contain considerable lean-grade nonmarketable ores as well as waste materials of little economic value which must be suitably treated and properly disposed to protect ecology/environment. Mineral deposits, including high- and/or lean-grade resources, are generally finite nonrenewable resources which will be exhausted under constant/varying rates of extraction in some finite time, called terminal time to exhaustion, T. This terminal time can be extended by augmentation by exploration, blending, and/or beneficiation of lean ores and technological substitutions but

is still finite time. Here we consider resource augmentation through blending and/or beneficiation of lean-grade ores that cannot otherwise be marketed and create ecological hazards, and these techniques are much less risky than exploration efforts or R&D effort for substitutions. Hence, it would be optimal to extract resources with suitable treatments of blending and/or beneficiation of lean-grade ores that are inevitably associated with ore deposits. However, ore extraction rates and subsequent blending/beneficiation processes must be dynamically optimized, and the net present value of resource is maximized to provide a stable and viable industrial growth with benefits to all including mine owners, national and state governments, and public at large. Therefore, proper planning, extraction, and marketing policies would insure maximal national and social growths as well as sustainable life quality of people [2, 3, 5].

Mineral resources are considered to be national or state government properties and are usually allocated/leased to national and/or private parties for a fixed time period for extraction of ores. Ill effects of such allocations (such as many scams) can be overcome through various regulations, taxations, and royalty schemes. It is suggested that heavy penalty clauses must be included in these regulations so that parties involved would not try to break these laws. Market complexities do induce high fluctuations in future prices and costs and involve great risks in all mining operations. Risk-aversionist mine owners and mineral traders may plead for non optimal rapid depletion of valuables by nonrenewable high-grade ores. Hence, time paths generated by resource markets are not necessarily at equilibrium, and these paths are generally not the most efficient or optimal paths. Therefore, analyses of these past price and demand data must be performed with much care and foresight to provide sound decisions.

Policy measures are basic to operation of economic and market systems since they could reduce market volatility and increase accuracy of future expectations. Better forecast techniques such as ARIMA (p, d, q) for nonseasonal and SARIMA (P, D, Q) for seasonal time series data [10, 15] and widespread dissemination of these forecasts would be most useful to avert risks. Reduction of government secrecy about future intentions/policies and maintaining longer-term economic policies would yield better stability and benefits. When global price becomes low, mineral exporters should be adequately compensated for their loss and/or be provided with liberal credit facilities. If the terminal date of exhaustion, T, of a mine be known (or estimable with accuracy), then optimal extraction rate/depletion policy can be achieved for a closed economy with no import/export. Optimal rate of resource depletion depends on factors, such as discount rate, elasticity of marginal utility and substitute, and productivity of substitutes. But the generated model would still be extremely rough as accurate estimates of most of these technological parameters are not available.

Resource allocation for R&D activity to develop natural/technological substitute materials for scarce and exhaustible mineral resources is useful and very important. But, we know that finding a substitute is a random process where the date of invention is extremely uncertain (purely random) and may take as much as 20 years (lifetime of a mine). Hence, total cost of R&D may become very high, especially for small mines, and they should concentrate on small technological researches to improve mining, blending, beneficiation, marketing, etc. rather than to find technological substitutes. In any case, postponing intensive R&D activities to the last 5/6 years of a mine would be economically viable and prudent especially in a developing economy as India [14].

Although there exists extensive but finite low-grade resources which can be upgraded to marketable products by suitable blending and/or beneficiation techniques, this only extends life of mine which will finally be exhausted. The prime high-grade deposits are much less available and will be depleted in finite terminal time T (say, maximum of 70 years). Therefore, it would be prudent to estimate optimal rates of extraction, so that the net present value after discounting of the mine is maximized in the long run. Even if technological substitutes can be found in the future, the date of availability and amounts of product are random variables with unknown probability densities for any accurate estimation of their parameters [6, 16].

Optimal decisions are necessary at all stages of mining and marketing operations which are listed as follows:

a. **Mining stage:** Methodology adopted; mine plans in 3D based on the lowest mineable assay surfaces and associated risk factors; blocking of high-grade, lean-grade, and waste materials; dynamic extraction rates; horizontal and vertical extensions

b. **Blending and beneficiation:** Optimal lowest crushing and beneficiation sizes needed for blending and beneficiation [17]; identification of high-grade, lean-grade, and waste blocks for optimal mining and storing, etc.; conservation of high-grade ore blocks

c. **Waste management:** Wastes develop at mining, blending, and beneficiation stages; optimal treatment and safe disposal avoiding health hazards and damage to ecology

d. **Marketing:** Optimal classes of marketable ores by suitable mix of high-grade and lean-grade ores; future marketing of lean ores in situ or in dumps using blending and/or ore beneficiation

Thus, mining of nonrenewable resources (metallic, nonmetallic, industrial rocks, gems, etc.) are highly complex nonlinear risky dynamic processes and not yet completely understood or solved. The main problems are (i) how to allocate exhaustible resource so as to make its marginal social value equal for all uses and constant over time and (ii) how government should plan and conduct project exercises to achieve these goals. We highlight the utilization of nonmarketable lean ores by optimal blending and/or beneficiation techniques which would maximize economic growth, improve social benefits, and preserve ecology and environment (see Refs. [2, 8, 17–19] for more details).

2. Mathematical and statistical theory

2.1. Probability space and distribution for fractional constituents

Fractional constituents (x; 0 < x < 1) belong to a C-dimensional positive real space and cannot be viewed geometrically if the number of constituents, C, is greater than three which is most frequent in rocks and ores. Aitchison [7] and his coworkers [20] have proposed that fractional constituents belong to a C-dimensional Simplex. They have used rather complex pre-transforms such as log ratio (lr), centered log ratio (clr), and isometric log ratio (ilr) to Gaussianize the PDFs for statistical analyses. However, Carranza [21] has shown that these

complex pre-transformations still do not solve the rank problem (rank is C-1 instead of C) and inherent spurious negative correlations among the constituents within any sample. This is due to the fact that fractional constituents (x; 0 < x < 1) belong to interior points of the simplex but not to apexes, hyper-edges, hyperplanes, and some sub-compositional hyperplanes of this C-dimensional simplex [6, 11].

From measure theory, fractional constituents (x; 0 < x < 1) belong to the open interval (0, 1) and have a binomial/Poisson (not Gaussian) distribution excluding the non-admissible points 0 and 1 on this line. Since the fractional constituents are averaged values over the sampled REV and REV is much larger than the specific constituent, this random variable (rv) x can be considered to be a continuous rather than a discrete binomial/Poisson rv. It is absolutely necessary; otherwise, the rock/ore is not a C-dimensional simplex. In such a system, additive probability measure is not applicable, and a multiplicative probability model (odds ratio) leading to a log-Gaussian or lognormal model (as proposed by Ahrens in [9]) would be appropriate with some modification such that the log (odds ratios) of the constituents becomes independent and Gaussian (homogeneous variances). The simple modification for log (odds) as proposed in Sahu [10, 11, 16] is the following arguments:

If x is lognormal with mean (μ) and variance (σ^2), then its complement $(1 - x)$ is also a fractional constituent. Hence, it would also follow lognormal distribution using theory of linear addition (or subtraction) of Gaussian random variables to be Gaussian (that is sums of Gaussian pdfs are closed/remain Gaussian). Therefore, we obtain $\log(x/(1 - x))$ to be Gaussian, and this pre-transformation involves only one constituent (x) of the rock/ore, which makes $\log(x/(1 - x))$ transform to be independent of all other constituents in the rock/ore [4, 10, 11, 16].

This $\log(x/(1 - x))$ pre-transformation Gaussianizes the pdfs of fractional concentrations while simultaneously eliminating the other four crucial drawbacks (i)–(iv) of binomial/Poisson distributions as listed earlier. Mathematical proof based on measure-theoretic analysis is given by Le Cam [22], Le Cam and Yang [23], and Sahu [11]. The author and his many students have used the $\log(x/(1 - x))$ pre-transformation to Gaussianize fractional concentrations (x) in geochemistry and apply to exploration, estimation, modeling, and hypothesis testing for several Indian ore deposits including those of gold, iron ores, lead-zinc, copper, phosphate, etc.

However, other log ratios such as lr, central-lr, and isometric-lr, proposed by Aitchison [7, 20], are not independent since the denominator constituent includes the numerator constituent as well, and, hence, the rank of data matrix still remains (C-1) or less (and not full rank of C, as is needed for its unique inverse) (see also [16, 21]).

2.2. Optimal extraction rates

2.2.1. General considerations

An exhaustible mineral deposit may be characterized by initial reserve, S_0, at present time, t_0, and rate of extraction, R_t, (t > o), at time t (t>t_0). The reserve at time t is S_t, given by

$$S_t = S_0 - \int_0^t R_t \, dt \tag{1}$$

Thus, the definite integral of (Rt dt) from 0 to t is less than the initial reserve, So   (2)

Therefore, $dS_t/dt = -R_t$, with S_0 known and $S_t >= 0$.   (3)

Eq. 2 gives in lim(t--> 0); R_t=0. However, R_t cannot be predicted as present market price fluctuates over time and market price controls both economic extraction rates to keep profits positive and the terminal time to exhaustion, T, of mine. Assuming R_t to be constant (R) over time T, then Eq. (1) gives

$$R = S_0/T, \text{ where } S_t = 0 \text{ (i.e., mine reached exhaustion).} \tag{4}$$

In mining and many other industries, usually time is discretized to mainly years but further into quarters, months, etc. Let price at time t be p_t and rate of return/interest denoted as r. Price at time t + u,

$$p_{(t+u)} = (1 + r_t \, u \, p_t) \text{ for } r(t) > 0. \tag{5}$$

In limit u tending to 0, we have a support price of:

$$(dp_t/dt)/p_t = r_t \tag{6}$$

However, Eq. (6) does not include costs of mining, transportation, and marketing and, hence, is not useful in practice. We must maximize the net present value of property as.

$$P_t S_t = R \left(\tau \, \exp.(-r \, (\tau - t)) \, dt, \tag{7}$$

where P_t^R is spot price, p_t is initial reserve stock, and R(τ) is extraction probability subject to the condition:

$$\int_t^\infty \{R(\tau)d(\tau)\} = S_t. \tag{8}$$

If R*(τ) is the optimal (maximal) solution to Eq. 6, then.

$$(dp_t/dt)S_t + p_t(dS_t/dt) = rp_t S_t - P_t^R R_t^*). \tag{9}$$

From Eq. 7 we get $S_t = -R^*$, and hence we can write.

$$(dp_t/dt - rp_t) = (p_t - P_t^R) \, (R_t^*/S_t). \tag{10}$$

Under competitive market stock price is equal to flow price (market equilibrium),

$p_t = P_t^R$, and Eq. (10) yields $(dp_t/dt)/p_t$ = r which is the Hotelling Rule [24].

The price of mineral resource is expected to rise in the future with increasing population and consequent increase in demand as compared to the interest rate, r; the value of reserve (or stock) or maximum present value of sales is independent of the actual extraction policy provided the entire stock is exhausted over time. Thus R_t* is not time-invariant (constant) but varies with time t. Static and dynamic optimization of extraction rates, R_t*, are discussed in Sahu [2, 14], respectively.

2.3. Socially managed exhaustible resource

Let R = D(p) be the market demand decline curve for resource flow. Since dD(p)/dt is negative, we can invert to obtain p = D^{-R}= B(R) for convenience. The gross rate of consumer surplus at R is.

$\int_0^R B(dR/dt)(d(dR/dt))$. Planning board must maximize the present discounted value at an extraction rate R_t(t > = 0) in order to maximize the integral:

$$\int_0^t e^{-rt}\left\{S_0^{Rt}B\ (dR/dt)\ d(dR/dt)\right\}\ dt, \tag{11}$$

subject to the constraints $S_t = S_0 - \int_0^t R_\tau d\tau$ and R_t and S_t both greater than and equal to 0 for t > = 0.

A feasible extraction policy (rate of extraction being positive at all times) is efficient if and only if $S_0 = \int_0^\infty R_t dt$. We obtain the optimal solution as $(dp_t/dt)/p_t$ = r. As desired, this efficient optimal extraction policy satisfies the condition of $S_0 = \int_0^\infty R_t dt$.

2.4. Reserve exhaustion at infinite time horizon

Assuming demand is linearly increasing with time, so p_t = A– B.R_t, where A and B both > 0. As long as reserves are positive, market equilibrium price holds and spot price will be given by.

$$p_t = p_0 e^{rt}. \tag{12}$$

Then, we obtain R_t = A/B – [($p_0 e^{rt}$)/B], which gives R_t = 0 for p_t> = A.

Reserve will be exhausted at future time. Time of exhaustion, T, will be at A where demand falls to zero.

Initial price p_0^* is given by $p_0^* e^{rt} = A$, and then, exhaustion time, T, is obtained from the relation:

$$\int_0^T TR_t \, dt = \int_0^T \left\{ (A/B) - \left(p_0^* e^{rt}/B \right) \right\} dt = S_0. \tag{13}$$

The above analyses show that mineral resources (including lean-grade ores) are exhaustible.

2.5. Independent versus dependent sample data

Natural phenomena are very complex and may not follow simpler mathematical and probabilistic assumptions necessary for analyses, estimation, tests, and decisions needed by scientists, technologists, and managers. For example, rocks and ores in the crust are heterogeneous, anisotropic, and inelastic materials at different scales such as microscopic, mesoscopic (hand specimens), and megascopic (outcrops) levels which make the sample size (REV) vary according to the scale of heterogeneity.

Geological processes are often nonlinear which needs complex linearization pre-transformations for mathematical and statistical analyses. Data are often non-Gaussian, need Gaussian/normal prior transform for simpler univariate (scalars) or multivariate (vectors) statistical analyses, are temporally/spatially dependent, and need complex time series analyses (wavelet and/or geostatistical analyses). Because of these complexities and of space constraints, these advanced methods are omitted here, but we give a few summary tables for guiding the readers.

More details on the nonlinear, stationary, nonstationary, and seasonal time series at time-, frequency-(Fourier), and time-frequency/wavelet domains are given in Sahu [15]. Whereas nonlinear models are in use for several hundreds of years, time series models are in use for about 100 years and were popularized by Box and Jenkins in 1970 and 1984, but wavelet models are currently very popular [25] (**Tables 1** and **2**).

Statistics	Fourier transform	Wavelet transform
Variance	Frequency spectrum	Wavelet spectrum
VAR (E(X^2))	S(X) (w) = X X* = $(|X|)^2$	W(X)(s, τ) = $(|X|)^2$
Covariance	Cross-spectrum	Wavelet cross-spectrum
COV (E(XY))	S(X,Y) (ω) = XY*	W(X,Y) (s,τ) = XY*
Correlation coefficient	Coherency	Wavelet coherency
r = E(XY)/{E(X^*2) E(Y^*2)}	γ = S(XY)(ω)/Sq.rt.S(X)(ω)S(Y)(ω)	Γ = W(XY)(s,τ)/$\sqrt{W(X(s,\tau)\ W(Y)(s,\tau)}$
Coeff. of determination	Coherence	Wavelet coherence
r^2; r as above	γ^2; γ defined as above	$Γ^2$; Γ as defined above

Table 1. Equivalent statistics for statistical (independent samples), Fourier (time series) (dependent samples), and wavelet analyses (dependent samples).

Stationary parameters	Tests	Linear prediction	Linear methods	Limitations
Univariate statistical: random IID scalars as inputs				
Mean, variance, correlation coefficient	t, Chi-sq, F, r, R^2 LR	Conf. limits of the parameters	ANOVA, ANCOVA	Samples, independent
Multivariate statistical: random IID vectors as inputs				
Mean, cov−/corrln. Matrices	T^2, Chi-sq, F, R^2 LR, partial corrln.	Conf. limits of the parameters	PCA, FA, CCA, MANOVA LDF, CLASSIF, MANCOVA	Samples, independent
Univariate time series models with random IID scalars as inputs having constant lag data				
Mean vector, acv, acf, spectrum	WN, AR, MA, ARMA nonst: ARIMA	Conf. limits of the parameters	ARMA(p,q), d = 0: ARIMA(p,d > 0,q)	Other information lost
Multivariate time series models with random IID vectors as inputs having constant lag data				
Mean vector, acv, acf, spectra	WN, AR, MA, ARMA Tr. Fn. forecasts, upgrade	One- and multi-step times	Tr. Fns with delay	Heterogeneity
Univariate geostatistical models with random IID scalars as inputs having local stationarity (d = 0)				
Mean, sill, range, nugget	Model validation	Nil	Interpolation, kriging	Extreme values
Multivariate geostatistical models with random IID vectors as inputs having local stationarity (d = 0)				
Multiple spectra, coherency	Model validation	Nil	Interpolation, kriging, co-kriging	Extreme values

Table 2. Comparison of linear statistical, time series, and geostatistical models for RVs.

3. Mineral exploration

Finite natural resources including metallic and nonmetallic ores, hydrocarbons, and industrially usable earth materials have been exploited since the emergence of mankind on the earth, and hence, surface and near-surface deposits are facing exhaustion/near exhaustion. The ever-increasing demands for improving living standards have resulted in present rapid rate of exploitation of ores and hydrocarbons with corresponding critical deterioration of the environment inducing high health risks. Therefore, at present there is an acute need for intensifying exploration for new concealed and/or partially concealed ore deposits at greater depths.

Geochemical exploration is one of the cheapest and important tools for detecting such hidden targets both at regional and local scales in the initial stage and at local scale during the mine development and production stages. Optimal exploration of any region must include integration of relevant available information on geology, geochemistry, and geophysics in order to delineate homogeneous regions within heterogeneous crustal rocks, ores and mineralizations, and surfaces. Homogeneity can be achieved through clustering by nearest neighbor (NN), fractal, or inverse distance weighting methods.

At the exploration stage when sample data are sparse, the purpose is to delineate, either directly by the element or through its pathfinder(s), regional and local (positive) anomalies

which generally include many ore targets, whereas at the mine development and/or production stages, large sets of closely spaced data become available for better statistical analyses using time series, fractals, etc. to estimate regional or local background(s) for delineation of such targets. Future actions at exploration stage include recommendations of intensive exploration and a few test-drilling at high positive anomalies to make mine feasibility studies, if desired, and to suggest mining and ore beneficiation methodologies to be adopted. In production stage, the purpose of geochemical exploration would be to find additional local targets for drilling and mining (if it is ore grade). Thus geochemical exploration plays very important roles in identifying subsurface targets of ore as well as hydrocarbon deposits and their subsequent exploitation.

3.1. Univariate statistical methods

Exploitable ore deposits and hydrocarbon resources are associated with crustal rocks and rock systems at shallow depths. Valuable mineralizations are emplaced along with the formation of these rocks (syngenetic deposits) or later formed/intruded into these rocks (epigenetic deposits) through their pores and/or fractures (faults). Some deposits may be deformed, modified, or weathered by several earth processes and at several later times (tectonic phases). For mathematical and statistical purposes, ore deposits can be characterized for geochemical purposes and subsequent statistical analyses. Rocks comprise $p > 1$ number of mineral phases which sometimes include valuable ore minerals and waste/gangue minerals. Minerals can be geochemically characterized by their elemental, molecular, and isotopic constituents that form major (>10%), minor (1–10%), or trace (<1%) amounts in them on volume or weight basis.

However, the main problems for statistical analysis of geochemical data are as follows:

i.   Rank is not full rank (i.e., not p; but p-1 or less) as total sum on volume/weight basis is 1.0/100% (closure constraint) which introduces spurious negative correlations among the constituents both among and within samples. The data are not independent for usual statistical studies but are dependent [10, 11]. Besides, samples with 0 and 100% should be deleted as they are not admissible for further analyses [18].

ii. Frequency distribution of constituents is on weight/volume basis and not on number basis as needed for statistical analysis. The distributions are not Gaussian but binomial ($\log(x/(1-x))$-normal) for major and minor components and Poisson having lognormal distribution for trace components [18].

Unfortunately, binomial and Poisson distributions, being discrete, are not that easily amenable for statistical analyses and hypotheses tests as it is for Gaussian distributions. For large sample (N > 50), binomial and Poisson distributions can be Gaussianized using the central limit theorem, but this method is sampling intensive and hence very costly to be practical and useful. A better alternative as suggested by Sahu [10, 11, 15, 16, 18] is to use log-odds pretransform as Gaussian distribution of (fractional) major or minor constituents which reduces to simple log transformation of fractional values for trace constituents. These log-odds transforms also simultaneously make all the constituents independent of others in the rock/ore sample and thus eliminates the rank problem/closure constraint.

So, all univariate normal/Gaussian (UNIV) (if N > 20) and multivariate statistical analysis (MND) with full-rank models can be applied (if N > 33) as desired. There exist other pre-transformations in literature such as log ratio [7, 20], but this is still not full rank and uncorrelated as desired in statistical theory [21]. After Gaussian pre-transformation of geochemical constituents in the sample using log-odds, linear or nonlinear statistical methods of time series or spatial series or time-spatial series (which include geostatistics (= ARIMA(p,1,q)) as a special case) or fractal/multi-fractal which are nonlinear models can be used to estimate regional/local background values from which regional/local anomalies can be identified [26].

Further actions should include:

i. Intensive geochemical investigations (including a few drillings) made at positive anomalies (prospects)

ii. Feasibility for exploitation

iii. Financial decisions for optimal mining and beneficiation (of low-grade ores) and mine remediation measure [14, 17].

3.1.1 Anomaly detection using log-odds, i.e., log (x/(1 − x)); 0 < x < 1, where x is fractional assay of geochemical constituent by linear and nonlinear modeling

Exploration and initial development stages	Advanced development and production
A. Graphical univariate multivariate methods	
Background: < av. x + 2σ	Less than av. vector x + 2σ
Anomaly: not less than av. x + 2σ	Not less than av. vector x + 2σ
Breakpoint in cdf of X in probit	Breakpoint in cdf of vector X (probits)
Above 3 for pathfinder element(s)	Above 3 for pathfinder elements
B. Fractal/multifractals (nonlinear methods which have their first term as linear)	
c-N, c-L, c-A, and c-V use log–log plots	c-N, c-L, c-A, and c-V use log–log plots
Breakpoints give regional and/or local background values from which local and/or regional anomalies and targets can be identified for further exploration	
Interpolation: by inverse distance, spline	Interpolation: by inverse distance, time series signal, fractals/multifractals, RBF, SV, kriging, spline
Singularity index gives robust backgrounds that are stable which yield good targets	
Fractal analysis in conjunction with neuro-fuzzy-genetic analysis and other soft computing techniques can yield very good to excellent ore targets	

Elemental concentrations in multicomponent (p > 1) rock and/or ore mineralizations are in different proportions (major >10%; minor 1–10%; trace levels <1%) which either have binomial distributions for major/minor proportions or reduce to Poisson distribution for trace constituents. These concentrations are spatially/temporally, as well as within each sample, dependent, and

hence statistical methods dealing with independent data are not strictly applicable. Independence of various constituents within any sample can be achieved through log (odds ratio) transformation for each constituent which simultaneously achieves Gaussian probability distribution, so that univariate normal (UNIV) theory and multivariate normal distribution (MND) theory (using linear theory) become applicable. Multivariate ($p > 1$ or vectors) theory reduces to univariate theory when a number of random variables become one ($p = 1$, or scalar), so univariate model is included under multivariate model. Population parameters like mean vector (μ) and covariance matrix (D) are then estimated for each population, and appropriate statistical tests are performed to take suitable statistical and geological decisions. Ore mineralization exists in the crust and is a three-dimensional (3D) static (not time-varying) phenomenon which can be modeled as a multivariate normal system through Gaussianization of log (odds) of constituent fractions present in rock/ore/soil samples. Mineralization processes are extremely complex, nonhomogeneous, and nonlinear to be analyzed directly. Therefore, these must be partitioned into homogeneous subsystems having strong dependence within but very weak interactions among the populations (groups). Two types of univariate hypotheses include (i) sample mean that belongs to a given population (null or H_o) which is tested against alternative hypotheses (H_1) and (ii) variances that are homogeneous (equal, H_o) or not (H_1). Two kinds of errors made for any decision are (i) Type I or α error of rejection of H_o when true and (ii) Type II or error of acceptance of H_1 when false. It is prudent to test more powerful ($1–\beta$) error being maximized for a given Type 1 error; these procedures are extendable for MND models.

Multiple and polynomial regressions (correlations) are strictly univariate model having only one random error but generally included under multivariate model because matrix methods are essential to solve them as it is necessary in MND analyses. However, F tests cannot be reliable for polynomial regressions, since powers of variable cannot be Gaussian when the variable is Gaussian. However, any power function relation of dependent variable(Y) with predictor variable(s) (X) can be linearized using log(Y) pre-transformation for statistical analyses.

3.2. Multivariate statistical analysis

This is appropriate for analyzing multiple correlated measurements (random vectors) made on one or more samples and on one or more (homogeneous) populations/groups. If a p-variate ($p \times 1$) Gaussian random vector (X) is measured on N-independent samples of a population, then the mathematical model has multivariate normal distribution (MND) which is characterized as all linear compounds of the variables being also MND as the rank of vector X may/may not be equal to its order (p). MND methods are classified on the basis of the number of populations (one or more) and number of sets of random vectors (one or more).

Four classes thus form:

- One population and one set of variables: MND methods are principal component analysis (PCA), factor analysis (FA), and cluster analysis (CA).

- One population but more than one set of variables: MND methods are multiple regressions (correlations), polynomial regressions, and canonical correlation.

- One set of variables but more than one population: MND methods include MANOVA, discriminant functions (DF: linear, quadratic), and classification function (CF).

- More than one set of variables and populations: MANCOVA.

Constituents of rocks/ores/soils are constrained to 1.0 or 100% which form a mathematically induced dependence structure (not geologically interpretable or meaningful) but have a binomial/Poisson distribution with heteroskedastic (or unequal) variances. Gaussianization of marginal distributions are often performed, but this does not guarantee that the joint distribution of random vector belongs to MND. So, MND theory must be tested through which all linear combinations (esp. principal components, multiple/partial/canonical regression components) of the random measurements are MND. However, MND theory is very robust, and if N is fairly large, then the random vector may be accepted to have MND.

The parameters of a multivariate r.v. can be estimated for any homogeneous population by its mean vector (μ) and dispersion (correlation) matrix (D/R) using MLE. Null hypothesis of homogeneity of population mean vectors (H2) conditional on homogeneity of dispersion matrices is given by T^2 test:

$$T^2 = N \, (m - \mu)^T \, D - 1 \, (m - \mu), \text{ with } (N - p) T2/(N - 1) \tag{14}$$

is F, p,(N-p) distributed with m being sample mean. Homogeneity test for covariance matrices (H1) of different populations is more involved (Box test).

Matrix operations are essential for multivariate analysis. Matrix A is a rectangular array of numbers with p rows and q columns, where a (i,j) is its ijth element. Addition and scalar multiplication are straightforward, but matrix multiplication requires that the number of columns of the pre-matrix must be equal to the number of rows of the post-matrix; otherwise multiplication is not defined.

If AB = C exists, then the elements c(i,j) = sum over r (a(i,r) x b(r,j)).

But in general multiplication is not commutative and AB not equal to BA, so pre- or post-multiplication of the matrix must be specified. However, multiplication is associative: A (BC) = (AB)C = ABC. Transposed matrix A^T has the rows and columns interchanged in A, so (AB)^T = B^T A^T.

If X and Y are two conformable column vectors, then their inner product is given by X*Y.

If A is (m × n) matrix, then A.x is a column vector. A* is conjugate transpose of complex matrix A, then (AB)* = B* A*, and so on.

Rank of matrix is the number of independent columns (or rows) in the matrix. A square matrix of order m is nonsingular, if its rank (m) is less than its order (p). A unique inverse matrix A-1 exists if A is nonsingular, then AA^-1 = A^-1 A = I (identity matrix). If unique inverses of A and B exist, then (AB)^-1 = B^-1 A^-1; also (A)^-1 = AT and (A*)^-1 = (A^-1)*. Unitary matrix, A* has A*.A = I = A.A* and so, A* = A^-1. If A is a real nonsingular, square matrix is said to be

orthogonal if $A^T A = A A^T = I$, so $A^{-1} = A^T$. A Hermitian matrix has $A^{-1} = A^*$. Elementary operations on columns (rows) of a matrix can give simpler form to interpret and compute, but its rank is preserved.

Determinant of a square matrix A ($=a_{ij}$) can be obtained by expanding the element ($a_{(i,j)}$) of a row (column) by multiplying its cofactor and summing over all elements of the row (column) or by multiplying the eigenvalues of A. Generalized inverse of a singular matrix (rank less than its order) is denoted as A-, and then $A A^- A = A A^-$ is not necessarily unique.

$A^- A = H$ with $H^2 = H$ (idempotent).

If $A^- = A$, we get det $A = r(A) = r(H) = trace(H)$. If A^- exists, then $r(A^-) \geq r(A)$.

Quadratic form (Q) of matrix plays an important role in MND analysis and is given by $Q = X^T A X$ with $A = [(a_{(ij)} + a_{(ji)})/2]$ which is a symmetric matrix. If ($X^T A X$) is >0, it is positive definite (pd), =0 (null), and it is negative definite (nd) and semi-definite if 0 is included in the product. For any nonsingular linear transformation, Q remains definite and invariant. Every positive definite matrix A can be decomposed into CC^T where C^T is inverse(C^{-1}) of the linear transform matrix. A necessary and sufficient condition for A to be positive definite (p.d.) is that its determinant is positive. This summary on multivariate analysis is based on Sahu [16].

3.2.1. Principal component and factor analysis

This method is for single population and one set of random variables. Original vectors in p-dimensional space are linearly transformed to a smaller m-dimensional subspace of principal components which are orthogonal. Mathematically, a real symmetric covariance (correlation) matrix is diagonalized (all correlations become zero) such that the principal diagonal yields the eigenvalues (variances) along the orthogonal eigenvectors (directions). In factor analysis, some of the smaller nonsignificant eigenvalues are deleted as negligible error components without losing information. The retained eigenvectors are rotated orthogonally in the lower common factor space (m < <p), so the new correlations (loadings) become easily interpretable (either near 1 highly loaded/correlated or near zero loadings if uncorrelated) as rotated factors. Rotation of orthogonal factors in the lower space is made by standard varimax program [27]. Cluster analysis can be made to obtain homogeneous groups by using similarity or distance matrices, but this process is rather empirical and needs great care for accuracy.

Eigen-structure of correlation (dispersion) matrix (R or D) is achieved through powering the matrix to a very high index (say 64 or 128) so that the largest eigenvalue dominates over the rest and corresponding eigenvector is obtained by a few iteration. The effect of the first eigenvalue is subtracted from the matrix to obtain the residual matrix which is again powered to high index to get next eigenvalue and eigenvector. This sequence is continued till all information of R (or D) are extracted and residual matrix becomes zero matrix. However, before running the eigen-structure analysis, null hypothesis R = I must be tested for statistical significance by a chi-square test with p(p-1)/2 d.f. at 0.05 level.

$$\text{The test quantity is given by} - \big[(N-1) - (1/6)\,(2p+5)\big]\,\ln\,(|R|). \tag{15}$$

Spectral decomposition of matrix gives.

$$R = \Lambda 1 \; v1 \, (v1T) + \Lambda 2 \; v2(v2T) + \ldots + \Lambda \, p \; vp(vpT) = \Sigma \; Rj \text{ over all pj.} \qquad (16)$$

If m components are found to be statistically significant, then the rest (p-m) components are noise and are deleted. So, total variance explained is sum R(j) of first m components, and rank of R is now m (<< p).

Multiplication of all eigenvalues gives |R| and sum of all eigenvalues is called trace of R.

Principal factors (f(j); j = 1 to m) are computed dividing the retained eigenvectors by the square root of their eigenvalues. Thus each factor becomes equally important as the other with a variance of 1 for all j. Factor structure $S = V (\Lambda)^{\wedge}$ -1/2 and predicted R by all factors is $S*S^{\wedge}T$; residual error is R- $S*S^{\wedge}T$. The number of significant principal components (m) retained as factors is the most important.

A chi-square test of determinant of residual matrix, res (A) with (p-m)(p-m-1)/2 d.f., is given by.

$$-\{(N-1)–1/6(2p+5) - 2/3 \, (m)\} \ln \big[|R|/\{\Pi \text{ of m eigenvalues } (p - \text{sum m eigenvalues}))/$$

$$(p-m)\}(p-m)]], \qquad (17)$$

which is tested at 0.05 level.

Another method is to plot jth eigenvalue vs. j to get inflection point giving m factors or to plot standard deviation of cum. Eigenvalues are computed on independent replicate samples of size N from the same population vs. j to get a minimum at which cum. Eigenvalue of 85% or more gives m. This second procedure, given in 1973 by the author, is a second-order criteria for deciding the common factor space (m) [16]. Varimax rotation is absolutely necessary to eliminate non-interpretable intermediate loadings in the range of 0.2–0.5 in any unrotated eigenvector of principal component.

Factor j is interpreted by the rotated loadings in the jth rotated eigenvector as follows:

i.   Absolute value of loadings close to unity is statistically significant and identifies the factor in terms of the input variables, and loadings near zero are nonsignificant and do not contribute to this factor (but may identify some other factor on which they are strongly loaded).

Correlation matrix can be computed over N samples to give R-mode R showing correlations among the random variables or over the p variables to give Q-mode R showing correlations among N samples. However, either R or Q correlation matrices have the same information and hence give finally the same inferences/decisions. But the order of R in R-mode is p << N; hence computationally R-mode analyses are preferred/cheaper. The rotated eigenvalues are different from the variances from corresponding eigenvalues, although the total variance (= Cumulative Eigenvalue) of m(<< p) retained factors is conserved by orthogonal rotations as can be easily demonstrated by matrix theory [16].

3.2.2. Multiple regression (correlation) and canonical correlation

Multiple (including polynomial) regression (correlation) yields linear prediction of dependent (criterion) variable(Y) from the knowledge of the predictors (X). The slope of regression line $b = (Var\ x)^{-1} Cov\ (x,y)$ if X and Y are scalars (univariate analysis), which is extended to vector random variables as $b = Cov(x)^{-1} Cov\ (x, y)$ if X is a vector random variable and Y is a scalar random variable. Multiple correlation exists if multiple correlation coefficient R is statistically significant, and R^2 indicates the sum of squares explained by predictors and $(1-R^2)$ indicates noise sum of squares.

F test can be made with (p-1) and (N-p) as degrees of freedom. However, since elements of X are mutually correlated (not independent), the effect of each element of X on Y is highly confounded and not possible to correctly interpret. Partial correlations remove the effects for other elements mathematically to give correct inferences for correlation of Y with ith element, x (i), of X, and hence, it is preferred over multiple correlations. In canonical correlation, two or more sets of variables are needed: one set is criterion, the other set predictor, and the third set control which can be kept mathematically constant. In contrast to principal component analyses, the eigen-structure is computed along the maximum covariances (not along maximum variances). The total correlation matrix R (with y as the pth r.v.) is partitioned into X of order (p-1), and hence we get the real nonsymmetric matrix as $R22-1\ R21\ R11-1\ R21$ which is the product of two real symmetric matrices: $B = R22$ and $A = R21\ R11-1\ R12$.

Mathematically we solve the eigen-structure of $(A - \Lambda B) = 0$ or of eigen-structure of $B-1\ A\ V = V (\Lambda)$. Eigen-structure of $B-1\ A$ can be done through two stages:

i.   Eigen-structure of real symmetric matrix B to give $\Lambda 1$ and obtain $B^{-(1/2)}$.

ii.   Eigen-structure of symmetric matrix $(B^{-1/2}\ A\ B^{-1/2})$ gives $\Lambda 2$ and eigenvector U2. Eigenvalues of $B^{-1}A$ are the same as that of $(B^{-1/2}\ A\ B^{-1/2})$, and hence sought vector, V, is given by $B^{-1/2}\ U2$.

Statistical significance of diagonal elements of canonical eigenvalues ($\Lambda 2$) can be assessed as follows:

i.   Proportion explained by $\Lambda j = \Lambda j / (trace\ \Lambda 2)$

ii.   Bartlett Lamda statistic = Product (j = 1 to p2) of $(1 - \Lambda j)$, where p2 is dimension of predictor vector

The null hypothesis that criterion and predictor sets are uncorrelated is assessed through chi-square with p(1) x p(2) d.f. as: $- [(N-1) - (1/2) (p(1) + p(2) + 1)]\ ln\ (\Lambda)$.

If null hypothesis of no correlation is rejected, then the effects of the first canonical root ($\Lambda 1$) is subtracted; the rest p(2) -1 canonical roots tested as:

Product (j = r + 1 to p(2)) $(1 - \Lambda j)$ as a chi-square with degrees of freedom, d.f. = (p(1) -1). (p(2)-1).

$$Chi - square = -\{(N - 1) + \tfrac{1}{2}\big(p(1) + p(2) + 1\big)\} \ln (\Lambda 1\ residual). \tag{18}$$

This test is continued until nonsignificance is achieved.

i.   A thumb rule: a canonical correlation < absolute 0.30 is statistically nonsignificant and hence dropped.

Multiple regression with standardized variables z can be written as z (hat) = b1 z1 + b2 z2 + … + bp zp and multiple correlation coefficients Rp. 1,2,…, (p-1) = R are similar to product–moment correlation coefficient (r) having the range of [−1,+1] for linear regression of scalars, but R has a range from 0 to 1.

R2 explains a major part of the variance of criterion and (1-R2) gives the error variance of regression. Therefore, F test with (p-1) df in numerator and (n-p) df as the denominator is applicable for the quantity R2 (N-p)/(1-R2)(p-1).The (p*p) correlation matrix R can be partitioned into R11 with order (p-1), and the last criterion (scalar) z(p) has a variance of 1.0. The multiple slope vector b = R11^-1 R12.

However, high values and high significance of any bj do not imply true importance of zj since other predictor z's confound the multiple correlation slopes. Hence partial correlation of criterion with a zj keeping all other predictors mathematically constant is absolutely necessary for any statistical/geological inference.

Polynomial regression is similar to multiple regression, but powers of predictor of random variables and the interaction terms are included. High degree of polynomial regressions is very difficult to interpret, and also if X is Gaussian, then its powers and interactions cannot be Gaussian, precluding the use of F test for the regression equations. So, unless theory dictates such polynomial regression, it should be avoided, and in any case, the degree should be as low as possible (say, second order).

Multiple partial correlation matrix R2.1 = {r21.jk} = residual r22.jk/(res r22.jj x r22.kk) 1/2.

(19)

For a trivariate-random variable, system res r22 = 1- r212 and res r23.j = res r23.j/(1 – r212)1/2. So, r21.3j = res r23j/(1 – r212)1/2, a well-known result in statistical theory. The output of partial correlation analysis can be arranged as: R = [(R21/ R32.1) (R21.3/R33)].

An example of partial correlation would clarify many of these concepts developed above.

The following random variables were measured in 33 thin sections from 33 sandstone samples.

•   The variables were X1 = phi long axis of grains (which has Gaussian distribution).

•   X2 = matrix percent.

•   X3 = porosity percent as reported by Griffiths in 1967, (p.468).

The multiple correlation matrix R was found with r12 = .8813**, r13 = −.7094**, and r23 = −.66771**. Here, ** means statistical significance at 0.01 level.

We compute partial correlation r13.2 = (r13 − r12 x r23)/(1− r122)1/2. Partial correlations between X1 X3, X1 X2, and X21.3 are similarly computed, and we get r21.3 = .6439**,

r31.2 = −.3862**, but r23.1 = −.2222 NS instead of −.6671**. [Here, superscript ** means statistical significance at 0.01 level.]

Therefore, r23.1 is nonsignificant indicating X2 and X3 are truly independent (uncorrelated) rather than correlated or possessing a negative multiple correlation. This fallacy of multiple correlation coefficients must always be noted and true inference must be sought through computation of partial correlations.

Comments: Although X1 has a Gaussian distribution, X2 and X3 possess closure constraints (ranging from 0 to 100% or 0 to 1 as fractions) and not Gaussian but binomial. X2 and X3 should be Gaussianized by the prior transformation log (xj %/ (100% − xj,%)) for j = 2,3. Multiple correlations should have been computed with original X1 and the new transformed-Gaussianized X2 and X3 variables.

3.2.3. MANOVA: discrimination and classification

MANOVA is similar to ANOVA for vector random variable X. In ANOVA (scalar r.v), two types of tests are necessary to test equality of main effects:

i.   When interactions are nonsignificant, the interaction variances are pooled with error variance, and a pooled error variance is calculated to yield the F test.

ii.   When interaction variance is significant, then its variance is used to test main effects by F test.

In MANOVA, treatment variance is divided by the pooled error variance to give F test since interaction variance is nonsignificant. But if interaction variance is significant, then MANCOVA methods are used to test main effects (F test) by dividing treatment variance by interaction variance (not error variance).

Populations (groups) are not necessarily homogeneous in mean vectors and covariance matrices. Two situations can arise.

a.   Covariance matrices are homogeneous, and testing is done to find homogeneity of mean vectors (H2 test; linear discriminant functions (LDFs) and MDFs as hyperplanes) or otherwise.

b.   At least one covariance matrix is different; we have to use nonlinear quadratic hypersurfaces (QDF) to delineate regions of each population. If both the mean vectors and covariance matrices are utilized together, then the procedure is called classification.

We decompose an ith vector of kth group $X_{(ki)}$ from grand mean m as $x_{(ki)} = X_{(ki)} − m = (m_{(k)} − m) + (X_{(ki)} − m_{(k)})$, where $m_{(k)}$ and m are the mean vectors for kth population and all populations, respectively. So any data is the sum of main effects (among-group) and within-group deviations $(X_{(ki)} − m_{(ki)})$.

The SSCP is then $\Sigma\, x_{(ki)}\, x_{(ki)}{}^T = \text{Sum}\,(m_{(k)} − m)\,(m_{(k)} − m)^T + \text{Sum}\,(X_{(ki)} − m_{(k)})^T$, summed over i = 1 to $N_{(k)}$ and k = 1 to g groups. Symbolically, T = A + W, where only two matrices are independent because of closure constraint. We get $W^{−1}T = W^{−1}A + I$, having only one independent matrix, $W^{−1}A$, for further analysis as identity matrix (I) is a constant. If covariances

(correlations) among the groups are equal (H_1 true), the dispersion among the groups is (D(A)) = A/(g-1), and dispersion within groups is D(W) = W/(N-g) where N is the total data over g populations.

The null hypothesis H_2: $\mu_{(k)} = \mu$ for all k = 1,…,g. and m = (sum X_{ki} over all I and k)/N. Rao [28] proposed F test as follows:

$$s = \left\{ \left(p^2(g-1)^2 - 4 \right) / \left(p^2 + (g-1)^2 - 5 \right) \right\}^{1/2}; n_{(1)} = p(g-1); n_{(2)}$$
$$= s\left[(N-1) - (p*(s-1)+1)/2 \right] - (p(s-1)-2)/2.$$
(20)

Let y = $(|W|/|T|)^{-s}$. Then, $F(n_{(1)}, n_{(2)}) = ((1-y)/y) (n_{(2)}/n_{(1)})$ and tested for statistical significance.

H_2 true, if F test is nonsignificant, means all mean vectors are equal.

3.2.4. Linear discriminant function

For two groups, g-1 = 1; hence there can be only one LDF, linear discriminant function (hyperplane). But for multi-groups g-1 is more than one, so we can have several LDFs, some of which may not be significant (should be dropped), but we also need the angles between the accepted (significant) LDFs (hyperplanes).

The retained LDFs form a subspace within the original p-dimensional space, and samples may be projected onto this subspace for visual studies. Optimal solution is to maximize the ratio W-1A (nonsymmetric real matrix) in the common discriminant subspace defined by vector v s.t.; the ratio of (Λ) = (vT A v/ VTWV) is maximized with the constraint vTv.

The maximum values are the eigenvalues of W^{-1} A: that is we solve ($W^{-1}A$) V = $V\Lambda$.

Since W is full rank, W^{-1} is unique and can be decomposed as U $\Lambda 1$ U^T, so W-1/2 = U($\Lambda 1^{-1}$ U^T).

Then, eigenvalues of W^{-1} A = eigenvalues of $W^{-1/2}$ A $W^{-1/2}$, but B has a different eigenvector U2.

Since B is symmetric, its eigen-structure is $U2\Lambda2U2^T$, and the eigenvector matrix V of W-1A is obtained as V = $W^{-1/2}$ U2 and has eigenvalue matrix $\Lambda2$.

The number of LDFs to be retained are obtained by statistical significance tests for elements $\Lambda2,j$ where j = 1 to (g-1) or p whichever is minimum (= rank of W-1A matrix). The importance of jth discriminant function (if retained as significant) can also be judged by the ratio of $\Lambda2,j$/trace $\Lambda2$ where this ratio ranges from 0 to 100%.

Also, each $\Lambda2,j$ can be tested as a canonical correlation of discriminant vector vj with any population (group) as the criterion (Y). The eigenvectors in V should be normalized (to vector with magnitude 1), and the angle between the ith and jth linear discriminants J(i) and J(k) is given by ($\theta(i,j)$) = Cos-1(v(ik). v(jk)). These angles are not necessarily orthogonal since W-1A is a nonsymmetric matrix. Discriminant scores which can be computed as v(i)T x(jk) for each retained eigenvector ji and xjk are the kth sample of jth group.

These scores can be projected onto the common discriminant subspace for visual perusal. A chi-square test of significance of discrimination amount for remaining m-k discriminants after accepting the first k significant discriminants can be assessed and tested as $-(N-(p + g)/2)\ln \Lambda^*$ with df = (p-k) (g-k-1), and Λ^* is the product of $1/(1 + \Lambda j)$ for j = (k + 1 to m).

This chi-square test should be nonsignificant to stop analysis. Usually two discriminants are most useful for visual representation of projection of LDFs as straight lines in the discriminant space, but 3D projections can be made if three discriminants are significant and required [16, 29].

3.2.5. Quadratic discriminant function (QDF)

If at least one covariance matrix is unequal among the groups, then pooling of covariance matrices to give a common (homogeneous) covariance matrix is inadmissible. Then, discriminant is nonlinear and hypersurface given by $\mu_{(1)} D_{(1)}^{-1} \mu_{(1)}^T - \mu_{(2)} D_{(2)}^{-1} \mu_{(2)}^T$ which reduces to LDF if $D_{(1)} = D_{(2)} = D$ and QDF = LDF = $(\mu_{(1)} - \mu_{(2)}) D^{-1} (\mu_{(1)} - \mu_2)^T$ as was derived under LDF theory. If a number of samples $N_{(1)}$ and $N_{(2)}$ are large, LDF is sufficiently robust for applications. Also, for QDF,F, test is inapplicable to find its significance.

MANCOVA methods (not discussed here) are more involved but necessary and proven to be useful for multi-element ores and for multiple populations (groups) in order to discriminate and/or classify.

4. Some applications

Statistics in the technology of twenty-first century and along with current capability of computers will be essential, beneficial, and most useful to mining and mineral processing industries. A log $(x/(1 - x))$ pre-transformation of fractional concentrations (x; 0 < x < 1) yields the desired independence and Gaussian pdf of each constituent in the rock or ore. This is necessary for characterization/estimation of parameters of each pdf and for hypotheses of tests and inference. Univariate and multivariate statistical models are used for single and multiple pre-transformed random variables, respectively. These models are useful for geochemical exploration, mining, mine planning, mineral processing, and beneficiation and for marketing such that maximal profits with minimum environmental damage can be achieved to obtain sustainable economic and societal growth.

Some applications of statistical (Gaussian) technology to mineral industry are listed below (not exhaustive) for getting the feel of different scenarios involved:

Exploration

a. Detection of positive anomalies for further intensive search

b. Detection of negative anomalies for use as sinks for toxic materials/elements

c.   Decision on pathfinders using factor model for single-element and canonical correlation for multiple-elemental ores

d.   Decision on feasibility of mining using open-pit or underground mining method

Mining

a.   Development of optimal mine plans using concentration contours and risk analysis

b.   Decision on lateral and vertical extensions to present mine plans

Beneficiation

a.   Decision on lower limit of assay for grinding

b.   Decision on optimal beneficiation process and system to be installed

c.   Locating high-grade ore zones for conservation for later blending and marketing

d.   Waste management decisions and related operation planning

Marketing

a.   Optimal classification of ore grades by separation and/or beneficiation/blending

b.   Marketing of nonmarketable ores in situ or in dumps by blending and/or beneficiation in the future

Example 1: Estimation of ore reserves and average assays.

In the past spatial distributions of assays and their pdfs were not accounted for, and simple calculations yielded these quantities based on geometry of ore body and arithmetic averaging of assays within mineable ores. The geology of syngenetic ore deposits produces uni- or multi-metal binomial/Poisson distribution having homogeneous variances, whereas epigenetic ore deposits are likely to possess bimodal (low assays in host rock and higher assays in ore zones) distributions that need separate treatments. A $\log (x/(1 - x))$ pre-transformation of fractional assays achieves linearity, Gaussianity, and homoscedasticity of variance with elimination of spurious negative correlations among the constituents [15, 16].

4.1. Syngenetic deposits

The lower limit of assay value for mineable ore is given by lower than 95% confidence limit of mean which should be marketable as well. Multi-metal ores are converted to single metal ore through addition of equivalent prices of these metals or by using a principal component of the mineable/marketable elements. The associated risk factor should be evaluated for use of lower than 95% confidence limit.

4.2. Epigenetic deposits

Bimodal Gaussian pdfs which can result as the mean in host rocks may be much less than in ore body. Univariate/multivariate discriminant function (LDF) easily separates these two

modes for separate calculation of reserves and average assays. Since mineable ore zones may lie within ore body, or partly in ore body and host rock, the geometry of mineable ore zone can be complicated. A 3D mine model is necessary to delimit mineable zones, and this may be achieved through a computer system. The associated risk factor for mining should also be computed as indicated above.

4.3. Spatially correlated samples on log(x/(1 − x)) basis

This situation often arises in development and production stages when a large number of geochemical data becomes available. The data is having signal (mineralized assay) and noise (random errors) and time (spatial) series model which separated the signal needed for average assay computation from the Gaussian noise giving the confidence limits to the average assay.

Models may be nonstationary (ARIMA (p,d,q) which is made stationary (ARMA(p,0,q) by differencing the data 'd' times. Integration of location data over the spatial domain gives the total volume and of signal gives the average assay. Such integration can be done block-wise to obtain block reserves and block average assays. More details of general time series modeling are given in Sahu [15]. Geo-statistical models are a special case of time series models belonging to the ARIMA (p,1,q) if assay values are linearized prior to analysis but otherwise are generally nonlinear since log (x/(1 − x)) pre-transformation of fractional concentrations was not performed.

Example 2: Identification of pathfinders

Pathfinders include minerals, molecules, elements, and isotopes and are very useful for exploration of uni- and multi-metal ores. These are characteristically easy to recognize, have higher concentrations than sought element and higher occurrence frequency, and can be analyzed cheaply and easily at much less time. Correlation matrix, R, of pre-transformed fractional constituent data is then computed using R-mode analysis (cheaper and faster than Q-mode). The correlation coefficients could be strongly +ve (r > .50), strongly –ve (r < −.50), and weak (−.50 < r < +.50).

Factor analysis of R (without weak correlations) provides rotated factors that are statistically significant to yield pathfinder(s) loaded strongly on high positive and/or negative correlations of constituents. Pathfinders can also be identified through the use of partial correlations, but this method is computationally more intensive, and decision may be confounded (nonunique). Multi-element or multi-mineral ores would require canonical correlation analysis with the mineable metals/minerals taken as criterion vector and other sets as predictor and/or control random vectors. FA in such cases may not be useful unless criterion vector has only one random constituent which is mineable. More details on finding geochemical pathfinders can be obtained from Sahu [16].

Example 3: Mine feasibility.

Sustainable mining operations must insure that the expected profits/year remains positive and substantial for meeting cash flow and other financial commitments. Mining operations with profits depends on high sale value of high-grade and beneficiated low-grade ores to

marketable grades. Mining costs include mine operation, transportation, beneficiation, and disposal of mine wastes with remediation.

Main geological factors are:

i.   Total reserve (W tons)

ii. Proportion of high-grade, W(H)

iii. Proportion of low-grade W(L)

iv. Proportion of gangue W(G) = 1- W(H)-W(L)

v.   Assay for high-grade ore A(H)

Of these five, three factors are independent but (iv) factor W(G) is not independent of W(H) and W(L). A(H) and A(L), assays of low-grade ores, are two additional independent factors (total of five factors).

There are ten economic factors including profit per year, P, which must be positive; life of mine (L = W/ PR (production rate per year in tons)); sale price per ton of marketed ore S(H); capital cost of mine operations, C(M); capital cost of beneficiation plant, C(B); rate of interest, r; efficiency of beneficiation, e; per ton running costs of mine, R(M); beneficiation, R(B); and waste disposal R(D).

Economic analysis for mine operations having beneficiation processing results in profits = P:

$$P = [W (W(H) + W(L)) S(H)/L]–(C(M) + C(B) (1 + r)/L)–(W R(M)/L)–(W − W(H).A(H)$$
$$− W(L) e (A(H)–(A(L)/L) R(D)–(W R(B)/L)$$

$$(21)$$

Above equation is a complex nonlinear one which cannot be linearized. If W(L) is negligible (zero), then mine-site beneficiation would be unviable and associated costs C(B) and R(B) would become zeros. Smaller mines with less low-grade ores would need pooling of nearby small mines for establishing a combined beneficiation plant to be operated jointly by these mines with proportional cost and profit sharing per year [14].

Example 4: 3D modeling and mine planning.

Fast, efficient, and up-to-date computer systems having links to end users are essential for this purpose.

The following items seem to be useful:

1. Fast transmission of databases, maps, and sections to central processor online and/or offline

2. Preparation of 3D maps using GIS technology and generation of desired sections for planning and mine operations. Offline transmission of this to all subcenters

3. Optimal plans for mine transport, blending and beneficiation, and timeframe of works

4. Marketing plans, expected profits vs. actual profits, doubling of assets in <5/6 years

5. Planning for new biddable targets and for extensions to present mines

6. Monitoring environmental damage and plan mitigation of such damages

7. Optimal computer architecture on faster transfer ratio in several separate computer systems acting in parallel to yield distinct outputs from a single input to the whole system

8. Mine closure plan, settlement of personnel, equipment disposal, mitigation of ground- and surface water damages, and corporate social responsibility (CSR) for the locality and country

Example 5: Mine sustainability.

Environmental hazards and associated mitigation costs are site-specific and hence, will depend on local geology, topography, and climate. Mineralogy and geochemistry would affect emissions on metallurgy, pollution by toxic metals, and leakages from dumps/tailings. Environmental degradation is due to poor production efficiency and poor innovations.

We can achieve sustainability in mining industry through the following:

1. Market incentives with pollution prevention, focusing on management system (MS)

2. Mine closure plans with Environmental Impact Assessment (EIA) and SIA (S = social) at all stages

3. Bonds which should be issued to clean up pollution after mine closure

4. Obtaining environmental and social performance indicators, risk assessment for environmental management (EM), and use of life cycle assessment (LCA), for technology choice and EMS

5. R&D efforts for resource augmentation

R&D efforts for exploration of new deposits, invention of natural and/or technological substitutes, and waste treatment/disposal are omitted. Since high-grade ores are comparatively rare in occurrence and will be exhausted in finite terminal time, T, in-house R&D efforts in mining industry should primarily concentrate on optimal utilization of associated lean-grade ores that are not directly marketable as well as on treatment/disposal of associated waste products. There are two ways to upgrade lean-grade ores and market the upgraded products through in-house R&D efforts as given below:

•   Optimal marketing of lean ores by blending with appropriate amount of (locally available or imported) high-grade ores having assay value of a1

•   Optimal marketing of lean ores by technological beneficiation with or without blending with high-grade ores having assay value of a1

5.1. Ore blending

We assume for convenience that the exploited ore body has only one type of marketable grade and only one type of nonmarketable lean-grade ores with associated waste materials of little market value. The analysis can be easily extended to two or more types of marketable ores and lean ores.

Let price per ton of ore be p and the assays and weight fractions of high-grade and lean-grade ores be a1, f1 and a2, f2, respectively, so that waste fraction is (1- f1 – f2). We have to crush the high-grade and lean-grade ores to suitable optimal size for blending so that the mix becomes homogeneous and can yield a stable grade, a* > a(m), the minimum marketable grade necessary for marketing purposes.

Then the average assay of blended mixture is.

$$a* = (W1\ a1 + W2\ a2)/(W1 + W2), \tag{22}$$

where W1 and W2 are, respectively, the weights (in tons) of high-grade and lean-grade ores in the mixture. The blended ore can be sold at a market price p per ton. The total sell value of blended ore would be (W1 + W2)p.

The blended assay, a1*, must be greater than the minimum marketable assay, a(m), and then the triangle law of proportional lengths of sides provides the minimum W1 value as (a(m)–a2)/a1. However to be safe, a1* should be made 5% more than required minimum, a(m), value so that blended ore is not rejected and we can use a minimum W1 value of (1.05 a(m) – a2)/a1). If this minimum amount of W1 of the high-grade ore is not available, then we have to optimize the mixing weights W1 and W2 for maximizing the profits by using constrained Lagrange multiplier or may have to import the deficient quantity of high-grade ore.

5.2. Ore beneficiation

Beneficiation of lean ores can be performed using physical, chemical, and biochemical methods, and the optimal technological parameters for providing maximum present value to marketed products should be found by several experiments and using appropriate response-surface experimental designs. The lean ores are to be finely grounded using jaw crushers, ball or rod mills, to optimal grinding size to maximally liberate the ore minerals for beneficiation [17].

The optimal mix of beneficiation products can be estimated by using Lagrange multipliers to maximize the net present value (defined as sell price minus cost price) of beneficiated products. Details of applicable optimal beneficiation schemes and the associated plant designs are specific to the mine, ore type, and ore characteristics and, hence, cannot be discussed as a general system theory. Therefore, the different ore beneficiation procedures are not discussed here.

5.3. Cutoff grade estimation

With continued rapid rate of utilization of minerals and metals for societal growth and explosive rise in world population, it has become imperative to mine minerals with lower grade and

at greater depths which induces increased costs of extraction and processing to make these marketable. Hence, there is a greater need for conservation of these invaluable nonrenewable finite mineral resources as well as for the preservation of fragile ecology and environment. Balancing these two opposing concepts of maximal utilization of ores and sustainable societal growth is the critical need of the hour [1–3, 5, 14, 25].

Mineral resources are characterized by their unique geological setting and genesis, as well as their spatial distribution which greatly influence the optimal extraction of these nonrenewable resources. Mining industry becomes sustainable with consistent long-term profit accruals over the life span of the mine.

This dictates that extracted ores can be marketed with reasonable profit, with sale price (s/ton) exceeding the cost of production (c/ton). The cost of production includes many factors such as mining, blending, beneficiation, transport of ores and wastes for their marketing, and safe disposals, respectively. Sale price (s/ton) of marketable ore is highly unpredictable due to volatile demand and supply of ores, government policy, technological innovations, substitute products, etc.

The cutoff assay, x(C), is defined as the fractional assay (x) of resource above which the extracted product is marketable (x(M)) and above the break-even assay (x(B)) defining the equality of sale (s/ton) and cost (c/ton) prices of the produce. Unfortunately, break-even assay, x(B), is not very useful as a cutoff grade since at this mining strategy profits become nil and mine becomes unsustainable. However, x(B) does provide the upper bound to the cutoff grade x(C) for ore extraction. Profits accrue if the extractable grade is reasonably above the market-able grade x(M) with the sale price (s/ton) greater than the production cost(c/ton) and with an extraction rate that maintains long-term sustainability of mine (at least till the end of mine life or ore exhaustion).

This strategy would induce a cutoff grade, x(C), much lower than the x(B) assay value but should be equal to the minimum assay value, x min, or near zero assay value or waste materials. The lower-grade materials with assays less than the assigned cutoff grade, x(C), are not mined and left in situ as un-mined blocks and pillars.

The optimal cutoff assay, x(C,O), therefore, must lie satisfying the following sequence:

0 < x min < x (C,O) < x (B,R) < x(B) or x(M) < x maximum <1.0. Under the static model, even though.

x(C,O) variation has a high range of assays (between x min and x (B)), it can be optimally estimated using two factors of pdf of lg (x/(1 − x)) or lg(x) which is Gaussian and the ratio of sales to cost prices (s/c).

Under a dynamic model, however, these parameters are time-varying and hence have to be estimated for each time period of mining extractions, and hence, the procedure of estimation of x(C,O) becomes more complex and time-consuming. Sale price is much more volatile than the cost of production as it depends on supply and demand position, vagaries of technological innovations, market substitutions, government interventions, and management policies. Dynamic x(CO) must account for these dynamic changes for optimizing the current profits

and the future expected profits. All profits must be brought to a comparable level using the standard techniques of reduction using net profit value (NPV) [3, 14] to the present state of time origin. Characterization of pdf of fractional assays in the ore body can be made by measuring assays of a large number (N > 50) of independent REV [30] samples/cells/blocks collected in the 3D space over which the resource exists.

From these sample data using the standard statistical methods, the arithmetic mean, median, variance, standard deviation, etc. can be easily obtained/computed. However, it is well known that the fractional assay pdf is lognormal [6, 9, 15, 16], and hence all statistical parameters and hypothesis tests must be made on the transformed Gaussian random variable, log(x), where log stands for common logarithm of assay value, x.

Computing on log(x) basis, we obtain the mean (u) which is median of log(x), and standard deviation of log(x) (σ) for cutoff, x(C), and optimal cutoff, x(C,O) estimations.

In the case of high-valued ores like diamonds, U, REE, Au, Ag, etc., the optimal cutoff is to x(B). In static analysis the geologic, assay distributional and economic factors are assumed to be constants over time, while in dynamic analyses, these parameters are time-varying and have to be estimated after each time unit of ore extraction (say, quarterly, half-yearly, or yearly as felt necessary). Dynamic modeling, although more involved, would yield much greater profits than the simpler static analysis.

5.4. Relation between break-even, x(B), and cutoff, x(C), assays

The main goal in mining is to maximize the profits by sale of mined and/or beneficiated ores [5]. Break-even assay, x(B), is very important in mining industry as it delimits the profitable ores from nonprofitable mineral resources.

This concept largely depends on the ratio of sales (s/ton) and cost (c/ton) prices of ore extraction of the marketable ore grade, x(M). Thus, using the pdf of fractional assays, x, if x(C) is zero, or x min, we obtain, for cdf F(x) of assay values, x as

$$F(x(B)) \, s = c \tag{23}$$

$$\text{or } F(x(B)) = 1 - (c/s) \tag{24}$$

At some positive cutoff grade x(C) > 0 and/or x min, we similarly get

$$(1 - F(x(B))) \, s = (1 - F(x(C))) \, c, \tag{25}$$

or

$$F(x(B)) = 1 - (1 - F(x(C))) \, (c/s). \tag{26}$$

Eq. (2) can be reversed to obtain F(x(C)), as a function of F(x(B)), as

$$F(x(C)) = 1-(1-F(x(B)))\ (s/c). \qquad (27)$$

Break-even grade, x(B), obtained by $F^{-1}x(B) = N^{-1}x(B)$ and, similarly, optimal cutoff grade, x(CO), can be obtained by inverting F(x(CO)) or cutoff grade x(C) by $F^{-1}x(C) = N^{-1}x(C)$.

In dynamic models forecasted values of (s/c) ratios are needed which is achieved by linearizing Eq. 3 and adding random error terms to sale (e(s)/ton) and cost (e(c)/ton) prices, as

$$\log\ (1-F(x(C)) = \log\ (1-F(x(B)) + \log\ (s/c) + \log\ (e(s)) - \log\ (e(c)). \qquad (28)$$

The predicted values of log (1–F(x(C)) = Y(hat) from time series model equation (Eq. (25)) can, then, be inverted to obtain x(C) as {1 – Exp (Y(hat))}. But in this paper, dynamic models (for forecasting s and c values to get Y (hat)) are not investigated in details and not pursued further. From Eq. (24), it is obvious that F(x(B)) is the upper limit to cdf of cutoff grade F(x(C)), and if (s/c) is 1.0, then F(x(C)) = F(x(B)).

We can obtain the lower bound to cdf of cutoff grade F(x(C)) as the minimum assay in the resource, x min, since fractional assay always lies above x min for any lognormal pdf.

Thus, F(x(C)) lies in the range from (x min) to (x(B)) in the following sequence:

0 < x min < x(C)/ x(C,O) < x(B,R) < x(B) < x(M) < x max<1.0; (x(M) being the marketable grade and x(B,R) the break-even grade for mining with risk factor at alpha confidence level). The optimal cutoff grade, x(C,O), will lie in the range from (0/x min) to (x(B))/(x(B,R) or x(M)); and x(C,O) strongly depends on the (s/c) ratio as well as the lower (x min or 0) and upper, x(B,R) or (x (B)) bounds of x(C).

Table 3 indicates some typical values of lower and upper bounds of x(C) for various (mineable/ economic) sale-to-cost price ratios (s/c) for a lognormal assay pdf, x, (i.e., log(x) is Gaussian).

Geologists and mining engineers have been using cutoff grade empirically, determined by their experience or by simple thumb rule and without using the concept of pdf of assay distribution being lognormal in the ores.

(s/c) ratios	1.15	10	100	1000	10,000	Remark
x min or zero	0.3	0.01	0.001	0.0001	0.00001	Lower bound
F(x(B))	0.84	0.99	0.999	0.9999	0.99999	Upper bound

Standardized random variable, z, can be calculated as z = [log(x) – μ]/σ; hence we obtain x min = Exp [μ – σ. z min] and x max = Exp (μ + σ.z max). The values, F(x(B)) and x min/zero grade, form the upper and lower bounds of cutoff grades, respectively. However, the upper bounds to F(x(C)) given in **Table 4** would be more useful, in practice, rather than those listed in **Table 3**.

Table 3. Feasible break-even grade, x(B), and x min or zero grade < x(C)/ x(C,O) assume lognormal (log-Gaussian) pdf for log(x) with mean, μ, and standard deviation, σ.

x	1	2	3	4	5	6	10	Remarks
F(x(B))	1	0.5	0.33	0.25	0.2	0.167	0.100	Upper bound
F(x(C))	0	0	0	0	0	0	0	Lower bound

Table 4. Values of upper bounds, F(x(B)), and lower bounds, x (min) or zero, to the cutoff grade, F(x(C)), for different economically feasible (s/c) ratios, greater than 1.0, as obtained by Eq. (2).

5.5. Profit maximization

Assuming sufficient proportion of marketable ores, (1 −F(x(B)), with sale price (s/ton) greater than the production cost (c/ton), exists for mining until exhaustion (mine life), economic viability is obtained if.

$$W.(1–F(x(B))\ s > w.c. \tag{29}$$

The optimal cutoff grade, x(C,O), should be searched for in the range of profitable cutoff grade range (i.e., between x min and x(B)), which can be equivalent to z (C,O) value within the range of z(x min) to z (x(B)) calculated on the basis of lognormal mean (μ) and standard deviation (σ) of the log(x) pdf. Then the total profit, P, can be calculated as.

$$P = W.(1 − z\ (x(C)).(s − c). \tag{30}$$

The optimal cutoff grade, x(C, O), would be less than x(B) and can be obtained through risk analysis (sale price of material in the range of x(B) to x(C,O) is equal to the cost of mining up to this lower optimal grade, x(C,O)) [5, 16]. The resulting new break-even grade, x(B,R), taking risk at alpha level, would be less than the original break-even grade, x(B), without taking any risk, that is, x(B,R) < x(B). Using cumulative unit normal, N(0,1), statistical tabular values, we then obtain [5, 16].

$$[N(z(x(B)) − N(z(x(C,O)))]/[N(z(x(C,O)) − z(x\ min)\ or\ z\ (−∞)] = (c/s), \tag{31}$$

which can be easily computed at the mining office for static or dynamic modeling to obtain x (C,O) value, as required by the mine management. Compare Eq. (31) with Eq. (27) for calculation of F(x(C) assuming mining of ores at zero assay value as the cutoff grade.

Dynamic model calculations must be updated from time to time after a fixed time period of mining operations. In addition, forecasting techniques can be used to predict future sale price (s/ton) and future cost price (c/ton) using time series modeling parameters estimated from corresponding past time series data [15] to obtain c/s or s/c ratios as needed.

These forecasting techniques are complex and model parameter dependent, and hence not pursued further here. In the case of large mines, blending and/or ore beneficiation processes are usually employed to upgrade lean ores to marketable grades for profits as well as for waste utilization to protect ecology/environment.

Dynamic optimization of profits is mine and ore specific and cannot yield general optimal cutoff grade for any specific ore type (say, iron ore, copper ore, gold ore, etc.). Hence, dynamic optimization must be done periodically in every large mine and for every block as needed. In addition, mining company must adhere to the National Mineral Policy and Mining Laws to have safe and uninterrupted operations.

6. Conclusions

i.   Fractional constituents of rocks/ores form multicomponent system having less than full rank to be statistically meaningful for analyses and geological inferences. A $\log(x/(1-x))$ pre-transformation of x data is shown to eliminate rank problem and Gaussianizes/ normalizes (independent (uncorrelated) of sample data) the random variables/vectors for linear statistical inferences, for geologically appropriate decisions, using univariate and multivariate statistical model procedures. More complex analyses such as dependence (correlated) samples in temporal, spatial, and spatiotemporal domains are not used here for simplicity and for lack of space.

ii. Geochemical concentrations (x; $0 < x < 1$) are rarely point data but are averaged over 1, 2, or 3D space and attached to the center of the sample. Also, measurement base is rarely numbers, as needed for statistical analyses, but the averaged sample values over length, area, or volume (weight) are used as random variables located at the center of sample volume (REV). Fractional constituents are spuriously as well as complexly negatively correlated because of closure constraint for the total measure (1.0 or 100%) and possess nonconstant variance over their mean value because these are binomial (multinomial) for major and minor levels which reduces to a Poisson distribution for trace components.

iii. Spatial patterns of these constituents are usually heterogeneous as mineralization intensities vary over local and regional scales. Homogeneity can be achieved through sampling at scales of representative elementary volume (or greater volume) and over local scales through NN clustering techniques.

iv. Anomaly detection is a very complex task as it depends on accurate determination of local or regional threshold levels. The associated risks for exploration of different types of anomalies as well as for specific individual targets must be evaluated.

v.   Mineralization involves highly complex nonlinear geological processes, so simple univariate statistical (linear) approaches with graphical methods are neither unique nor optimal. Multivariate approaches are also linear but are much better (however, becomes unrealistic if the sample data are spatially/temporally correlated). Nonlinear methods such as nearest neighbor, fuzzy logic (FL), genetic algorithm (GA), and soft computing (SC) techniques are more complex to use, not dealt here, but can be useful.

vi. During exploration stage useful multivariate methods could be multiple and partial correlation, factor (principal component) analyses, canonical correlation for identifying pathfinders and delineating anomaly zones, two-group linear discriminant functions,

and multi-group linear discriminant functions for identifying ore from wastes and for delineating ore zones.

But during the development and production stages, it is prudent to identify different categories of ore and zones such as

- Marketable high grades

- Blended marketable grades

(s/c)	1.0	1.1	1.2	1.25	1.3	1.4	1.5	1.6	1.7	1.8	1.9	2.0
F(x(B))	F(x(C))											
.005	.005											
.010	.010											
.015	.015											
.020	.020											
.025	.025											
.030	.030											
.035	.035											
.040	.040											
.045	.045	.000										
.050	.050	.000	0									
.100	.100	.010	0	0								
.150	.150	.070	0	0	0							
.200	.200	.120	.040	0	0	0						
.250	.250	.175	.100	.060	.025	0	0					
.300	.300	.230	.160	.120	.090	.020	0	0				
.350	.350	.285	.220	.190	.145	.070	.025	0	0	0		
.400	.400	.340	.280	.250	.220	.160	.010	.040	0	0	0	0
.450	.450	.395	.340	.310	.285	.230	.175	.120	.065	.010	0	0
.500	.500	.450	.400	.370	.350	.300	.250	.200	.150	.100	.050	0
.600	.600	.340	.280	.250	.220	.160	.100	.040	x	x	x	x
.700	.700	.230	.160	.125	.090	.020	x	x	x	x	x	x
.800	.800	.120	.040	x	x	x	x	x	x	x	x	x
.900	.900	.010	x	x	x	x	x	x	x	x	x	x
1.00	1.00	x	x	x	x	x	x	x	x	x	x	x

Beyond (s/c) ratio of 2.0, the F(x(C)) values will be essentially, either, x (minimum) or zero; hence these values have not been given in this table. The optimal cutoff grade (as given by Eq.(V.5,-2), F(x(CO))) lies between F(x(min)/zero) and F(x(B)) values.
Values of x(C) = F^-1(x(C)) = N^-1(lg(x(C)); similarly, we obtain x(CO) = F^-1 (x(CO)) = N^-1(lg(x(CO))).

Table 5. Cumulative value of fractional cutoff grade (assay), F(x(C)), as calculated by Eq. (27), for different ratios of sales/cost prices, (s/c) ratios, and for different cumulative values of fractional break-even grade, F(x(B)), as calculated by Eq. (26), is given in **Table 5**.

- Beneficiated marketable grades

- Low-grade ores forming future resource

- Waste material

To achieve these delineations, we use multi-group linear discriminant function approach.

Then, optimal mining, blending, beneficiation, and marketing operations would maximize profits, social aspirations, and ecological/environmental protection/remediation for greater sustainability.

vii. Optimal utilization of unmarketable lean-grade ores helps to accumulate additional capital for economic and social growths and reduction in costs of waste treatment and disposal, thereby improving the health of local inhabitants, as well as helps in the conservation of high-grade ores for better sustainability and for utilization in the future.

viii. Estimation of cutoff grade, $x(C)$, or optimal cutoff grade, $x(CO)$, involves many complex geological, spatial assay distributional, pdf of assays in rocks/ores, and economic factors such as sell price (s/ton) and production to marketing costs for ores plus disposal costs of wastes (c/ton).

Estimation is simpler for static models where these random variables (rvs) are Gaussianized/ normalized using pre-transformation, $\log(x)$, and are time-invariant constants but highly involved when these parameters vary with time and have to be updated at short intervals (dynamic models). Estimation of cutoff grade can be obtained by solving the nonlinear equation for static models (see **Table 5**) and inverting the obtained $F(x(c)$ or $F(x(CO))$ using the standard unit-normal cumulative statistical tables giving $N(0,1)$, i.e., $N^{-1}(0,1)$.

Dynamic models do not have a global solution for cutoff grade since it has to be updated at every time interval of mining operation but yield much greater overall profits to the mining industry.

Author details

Basanta K. Sahu

Address all correspondence to: bas.sahu37@gmail.com

Department of Earth Sciences, IIT Bombay, Mumbai, India

References

[1] Sahu BK. Geology applied to sustainable development of mineral resources (Key-paper). In: International Seminar, Souvenir Volume; ISM Dhanbad; 2013. pp. 77-83

[2] Sahu BK. Mining and sustainable development of non-renewable mineral resources. In: Proceedings of the Seminar on Sustainable Development on Mineral Resources and Earth Processes, New Delhi; 2014. pp. 4-7

[3] Sahu BK. Optimal mining and marketing of lean grade ores. SGAT Bulletin. 2015;**16**(1):1-6

[4] Sahu BK, editor. Proceedings of the Geomodelling Workshop; IIT, Bombay; 1983. 274p

[5] Sahu BK. Optimal cutoff grades (assays) in mining for profit optimisation. In: International Conference on Mineral Processing Technology (MPT, 2016), Pune; 2016 (submitted). 7p

[6] Sahu BK. Probability density function (pdfs) of minerals and/or chemical constituents in rocks and ores for sustainable growth. In: MPT 2013 International Seminar, Bhubaneswar; 2013 (in press)

[7] Aitchison J. The Statistical Analysis of Compositional Data. London: Chapman & Hall; 1986

[8] Sahu BK. Multivariate statistical methods for mineral exploration using geochemical data. In: Vistas in Geology. Bhubaneswar: Utkal University; 2013. pp. 7-16

[9] Ahrens LH. The lognormal distribution of elements. Geochimica et Cosmochimica Acta. 1954;**5**:41-73; 121-132

[10] Sahu BK. Mineral deposit modeling. Mineralium Deposita. 1982;**17**:99-107

[11] Sahu BK. Statistical inference for geochemical and petrographic data. In: Proc. Rec. Res. NW India; Vadodara, Gujarat: M.S. Univ.; 1995. pp. 59-64

[12] Dasgupta PS, Heal GM. Economics and Extraction of Exhaustible Resources. Cambridge, UK: Cambridge University Press; 1980. 501p

[13] Pindyck RS. Optimal exploration and production of nonrenewable resources. Journal of Political Economy. 1978;**86**:841-861

[14] Sahu BK. Geology and economics for mining low-grade ores. SGAT Bulletin, Bhubaneswar. 2012;**13**:72-74

[15] Sahu BK. Time Series Modeling in Earth Sciences. Lisse, Netherlands: Balkema; 2003. 284p

[16] Sahu BK. Statistical Models in Earth Sciences. Hyderabad: B.S. Publications; 2005. 211p

[17] Sahu BK. Optimal liberation of ore minerals for ore beneficiation. In: International Seminar on Mineral Process Technology (MPT2010); Jamshedpur; 2010. pp. 3-8

[18] Sahu BK. Evaluation of iron ore prospect in BIF setting. In: Proc. Intl. Sem. Iron Ores, SGAT, Bhubaneswar; 2008. pp. 2-11

[19] Sahu BK. Statistical methods for geochemical exploration of ore deposits. In: Vistas in Geology. Bhubaneswar: Utkal University; 2013. pp. 1-6

[20] Egozcue JJ et al. Isometric log-ratio transformation for compositional data. Mathematical Geology. 2003;**35**:279-303

[21] Carranza EJM. Analysis and modeling of geochemical anomalies using log-ratio transforms. Journal of Geochemical Exploration (JGE). 2011;**110**:167-185

[22] Le Cam L. Asymptotic Methods in Statistical Decision Theory. New York: Springer; 1986. 742p

[23] Le Cam L, Yang L. Asymptotics in Statistics. New York: Springer; 1990. 180p

[24] Hotelling H. The economics of exhaustible resources. Journal of Political Economy. 1931;**39**:137-175

[25] Sahu BK. Prediction of permeability in clastic reservoirs and aquifers using stochastic and wavelet theories for sustainable fluid production. In: Intl. Conf. Adv. Water Resources; CAS Panjab Univ; 2013. 17p

[26] Delarer ST. Delineation of mineralization using c-V fractal method. Journal of Geochemical Exploration (JGE). 2012;**118**:98-118

[27] Kaiser HF. The Varimax criterion for analytic rotation in factor analysis. Psychometrika. 1958;**23**:187-200

[28] Rao CR. Linear Statistical Inference and Applications. New York: Wiley; 1973. 625p

[29] Sahu BK. Multi-group discrimination of river, beach, and dune sands using sphericity statistics. Mathematical Geology. 1982;**14**:577-586

[30] Bear J. Dynamics of Fluid Flow Through Porous Media. New York: Elsevier; 1972. 372p