J. S. Afr. Inst. Min. Metall., vol. 92, no.4.
Apr. 1992. pp.91–99.

A comparative study of three frequency-distribution models for use in ore evaluation

by H.S. Sichel*, W.J. Kleingeldt†, and W. Assibey-Bonsu*

SYNOPSIS

A frequency distribution model plays a fundamental role in ore evaluation, especially for the feasibility study of new deep gold mines.

The geological complexity of the reefs being mined at present in the Witwatersrand basin poses some difficulties regarding the modelling of the underlying ore-value distribution. This geological complexity has given rise to a need for more applicable distributional models. This paper places the geology of the Witwatersrand basin in perspective and draws some important conclusions from the genesis of the deposition. In addition to the three-parameter lognormal model, two more-general distributional models are proposed. These are the compound lognormal (CLD) and log-generalized inverse Gaussian (LN–GIG) distributions. A detailed, although preliminary, comparative study of these two more-general models with the three-parameter lognormal model is described, and the efficiencies of the estimates and modelling are analysed.

It is found that the broad assumption that gold grades follow a three-parameter lognormal law is not always the case, and the need to test for departures from lognormality should be emphasized so as to ensure the use of appropriate models.

SAMEVATTING

'n Frekwensieverdelingsmodel speel 'n fundamentele rol in ertswaardering, veral in uitvoerbaarheidstudies van nuwe diep goudmyne.

Die geologiese ingewikkeldheid van die tipe rif wat huidig in die Witwatersrand-kom gemyn word, lewer probleme wat betref die modellering van die onderliggende ertswaardeverdelingsmodel op. Hierdie geologiese ingewikkeldheid maak meer toepaslike verdelingsmodelle noodsaaklik. Hierdie referaat plaas die geologie van die Witwatersrand-kom in perspektief en daar word belangrike gevolgtrekkings oor die oorsprong van die afsetting gemaak. Benewens die drieparameterlognormaalmodel word daar twee meer algemene verdelingsmodelle, die saamgestelde lognormaal verdeling (compound lognormal (CLD)) en die logveralgemeende omgekeerde Gauss-verdeling (log-generalized inverse Gaussian (LN-GIG)) voorgestel. 'n Gedetailleerde, dog voorlopige, vergelykende studie van hierdie twee meer algemene modelle en the drie-parameter lognormaalmodel word uitgevoer en die doeltreffendheid van die skattings en die modellering word ontleed.

Daar is gevind dat die algemene aanname dat die goudgrade 'n drie-parameterlognormaalwet volg, nie altyd geld nie. Om dus te verseker dat toepaslike modelle gebruik word, word die noodsaaklikheid van afwykings van lognormaliteit te toets, benadruk.

INTRODUCTION

Research carried out over the past two years at the Anglo American/De Beers Ore Evaluation Department pertaining to gold-grade distributions, especially in regard to the Ventersdorp Contact Reef (VCR), has led to a questioning of the applicability of the lognormal and the associated three-parameter lognormal models for the modelling and estimation of these types of deposits. The major problems addressed are those of modelling the enrichment or loss of gold-grade values at the tail and upper end of the distributions (Figure 1). Research has shown that these occurrences are geologically related and, as such, cannot be made out to be artifacts. As a result, attention has been focused on the possible application of two new models, both of which originated from research carried out by H.S. Sichel. These new models are the compound lognormal (CLD) and the log-generalized inverse Gaussian (LN–GIG) distributions, which are discussed in this paper.

The previous extensive use of the three-parameter model as suggested by Krige[1] and of its associated t-estimation technique as developed by Sichel[2] was an indication of its practical usefulness and efficiency for the reefs in which it was being used. Although Sichel[3,4] and Link and Koch[5], commendably accept the practical efficiency of Krige's technique[1], they indicate that severe biases could generally result by the application of lognormal theory if, in fact, the underlying distribution of the variable of interest is not lognormal. They caution that there are a number of theoretical and empirical distributions that can mimic the lognormal model in as much as they exhibit unimodality, positive skewness, and a long tail to the right. Sichel[4] has given detailed theoretical explanations regarding these pseudolognormal models.

The departure of gold-grade distributions from two- or three-parameter lognormality can well be observed on a cumulative frequency-distribution plot on log-probability paper. The cumulative frequency-distribution graph will display distinct curvatures both at the low- and the high-grade categories (Figure 1). In such situations, the addition of a constant assists in eliminating the curvature in the lower-grade category, but the upper-grade curvature will persist.

The need for a more complex statistical model as mentioned above is related to the geological complexity of

* University of the Witwatersrand, P.O. Wits, 2050 Transvaal.
† Anglo American Corporation of South Africa Ltd, P.O. Box 61587, Marshalltown, 2107 Transvaal.

ISSN 0038-223X/3.00 + 0.00. Paper received August 1990.

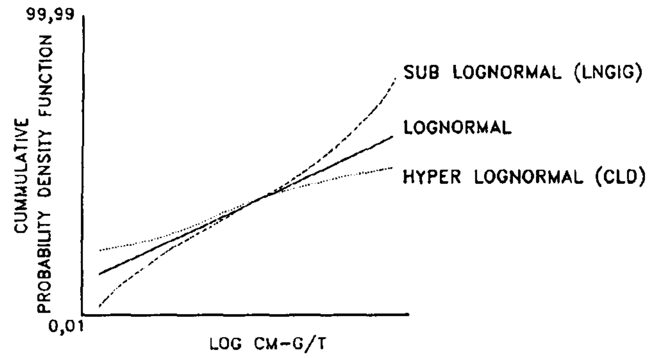

Figure 1—Probability density distribution functions

the type of reefs being mined at present in the Witwatersrand basin. The geology of the basin is as set out below.

GEOLOGICAL MODELS

Detailed discussion of the geology of the Witwatersrand basin is well presented by Minter[6,7], Antrobus[8,9], and Hallbauer[10]. However, for the purpose of this paper, reference should also be made to Krige *et al.*[11].

The gold-bearing sedimentary rocks of the Witwatersrand basin were discovered in 1886, and the amount of gold that has so far been produced from the basin makes it unique compared with similar gold formations. Figure 2 shows the basin and the major gold-producing areas.

In summary from the above-mentioned papers, it can be concluded that the gold- and uranium-bearing sedimentary rocks of the Witwatersrand basin are found in the Dominion Reef Group and the Witwatersrand and Ventersdorp Supergroups. The Witwatersrand Supergroup consists of the (lower) West Rand and the (upper) Central Rand Groups. The placer theory for Witwatersrand gold is accepted as the most valid genetic theory, and postulates that the sedimentary rocks containing the ancient placers were deposited by braided rivers on alluvial fans located at the margins of a yoked basin. The sediments in the basin can be subdivided into a number of geological units. These geological units are bounded by unconformities, which are a reflection of intermittent sedimentation in the basin. This historical intermittent sedimentation has been attributed to local and regional geological factors, as well as to geomorphological factors. In addition, regressive sedimentation and reworking of sediments led to varying gold concentrations in the different geological units, resulting in the development of gravel horizons (reefs) containing relatively high concentrations of gold.

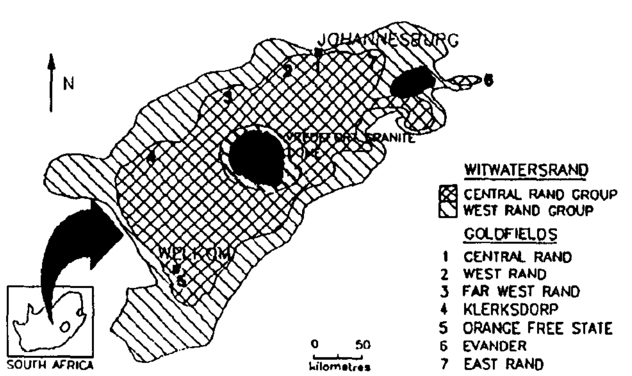

Figure 2—The Witwatersrand basin and goldfields

Most of the gold in the Witwatersrand basin is of a detrital nature, and the overall gold content of a reef is related to the amount of gold released in the provenance area and to the degree of reworking that took place. The distribution of gold within and between reefs shows different degrees of variability. On a micro-scale (as observed in closely spaced borehole deflections), the level of variance is high, and this variability increases on macro- and mega-scales as a result of the inclusion of more diversified features. Furthermore, the gold content within a reef can vary on a large scale owing to facies differences related to positions on the alluvial fans, or to overlapping fans that have different entry points and provenance areas.

The geological setting discussed above gives rise to the following salient features.

(1) The hydro-dynamic circumstances in an alluvial environment dictate that the river systems that distributed the placers could transport only certain sizes of sediments at different stages as the river systems gradually reduced their velocities downstream. Hence, as distances from the source of transportation increased, the heavier mineral particles would be deposited much earlier than the lighter particles. One would therefore expect a relationship (probably linear) to be set between the log-means and the log-variances of the gold variable as a function of distance from the stream source. One would further expect that, if gold samples are taken from different localities within the same horizon, or even from different overlapping reefs within the same horizon, the log-mean and log-variance parameters would be different, resulting in a mixture of lognormal distributions that will no longer follow the lognormal law. The compound lognormal model (CLD) has been developed to model these hyper-lognormal types of distributions. The CLD constitutes an *infinite* mixture of lognormal distributions. A simple *finite* mixture would not be flexible enough to meet the practical requirements and, furthermore, is much more complicated mathematically.

(2) Owing to the predominant erosional features of certain types of reefs, the high-grade values appear to be partly missing (sub-lognormality). This gives rise to association of the gold distribution with pockets, and the probability of finding gold depends on finding gold given a pocket or trap site, as well as the way in which the trap sites are distributed. In statistical terms, we then have

$$P(r) = P(r \mid \lambda).P(\lambda),$$

where $P(r)$ = probability of finding gold

$P(r \mid \lambda)$ = probability of finding gold given a trap site, and

$P(\lambda)$ = probability of finding a trap site.

Such types of genetic models have previously been explored for diamonds[12], and gave rise to the compound Poisson model, which is a subset of a more general model called the log-generalized inverse Gaussian distribution (LN–GIG), which is discussed in this paper.

Unlike the more general models, lognormal models presume the absence of any linear relationship between log-mean and log-variance parameters. As a result of the above discussions, it seems advisable that, whenever there are departures from lognormality, one of the more general models should be applied to avoid severe biases for obvious economic reasons.

This paper, therefore, is a comparative study of the new and the old gold-grade distribution models in the South African context. Data from some of the important reefs of the Witwatersrand basin have been used. The reefs that were considered are the Basal Reef, VCR, Vaal Reef, and Kimberley Reef. This is the first time that these more general models have been applied to represent gold-grade distributions, and the preliminary results are very successful.

THE MODELS IN PERSPECTIVE

Three-parameter Lognormal Distribution

The three-parameter density function is defined as follows:

$$f(z) = \frac{1}{\sqrt{2\Pi}\sigma}\frac{1}{(z+\beta)}\exp\left(-\frac{1}{2}\frac{[\ln(z+\beta)-\xi]^2}{\sigma^2}\right),$$

where z is the observed ore value in cm·g/t ($0 < z \le \infty$) ξ, σ and β are parameters to be estimated, β is the additive constant, ξ is the average of the $\ln(z+\beta)$ values, and σ^2 is the variance of the corresponding log values.

Compound Lognormal Distribution (CLD)

This distribution has the following mathematical representation:

$$\Omega(x) = \frac{s(1-c)^{v+\frac{1}{2}}}{2^v\sqrt{\pi}\,\Gamma\left(v+\frac{1}{2}\right)}\exp\left[\pm\sqrt{c}\,s(x-a)\right]$$

$$(s|x-a|)^v\,K_v(s|x-a|),$$

where $x = \ln(z)$, $(-\infty < x < +\infty)$, z = observed ore value in cm·g/t, and $K_v(.)$ is the modified Bessel function of the second kind of order v. The four parameters of the distribution are

- a = location parameter, $-\infty < a < +\infty$
- s = spread parameter, $s > 0$
- c = skewness parameter, $0 \le c < 1$
- v = kurtosis parameter, $v > -1/2$.

The positive sign before $\sqrt{c}$ indicates that the distribution is positively skew, whereas a negative sign shows negative skewness. The parameters a, s, c, and v were derived by Sichel[12].

A noteworthy feature of the shape of the CLD is as follows.

(i) For $1/2 < v < \infty$, the distribution has a conventional mathematical maximum with a horizontal tangent. (For $v \to \infty$ and $c = 0$, the distribution is normal.)

(ii) For $0 < v \le 1/2$, the distribution is continuous and has a mode in the form of a sharp cusp with two non-zero slopes.

(iii) For $-1/2 < v \le 0$, the distribution has no conventional mode. Instead, it has an infinite pole in the middle of its range; yet the area under the curve is still unity, and there is still a probability distribution.

The estimation of the four parameters is achieved *after* the log-transformation by the method of moments[12]. Some researchers may be interested in the moments *before* the log-transformation to find the shape coefficients $\sqrt{\beta_1(z)}$ and $\beta_2(z)$ of the original ore values z. These ratios can be found from the r'th moments *around the origin*:

$$\mu'_r(z) = \left[\frac{1-c}{1-\left(\frac{r}{s}\pm\sqrt{c}\right)^2}\right]^{v+1/2} e^{ra}.$$

As $\sqrt{\beta_1(z)}$ and $\beta_2(z)$ are very large, one should *not* estimate the four parameters from the untransformed z variables.

Log-generalized Inverse Gaussian Distribution (LN–GIG)

The four-parameter log-generalized inverse Gaussian distribution is defined as

$$\Lambda(x) = \frac{1}{2sK_\gamma(b)}\exp\left[\gamma\left(\frac{x-\xi}{s}\right) - b\cosh\left(\frac{x-\xi}{s}\right)\right],$$

where $x = \ln(z)$, $(-\infty < x < +\infty)$, z is the observed ore value in cm·g/t, and $K_\gamma(b)$ is the modified Bessel function of the second kind of order γ and argument b. The four parameters are

- ξ = location parameter ($-\infty < \xi < +\infty$)
- s = scale parameter $s > 0$
- b = kurtosis parameter $b > 0$
- γ = skewness parameter ($-\infty < \gamma < +\infty$).

Pearson's shape coefficients β_1 and β_2 are used to determine the specific model applicable for a given empirical distribution function, as shown in Figure 3, where

$\sqrt{\beta_1} = \mu_3/\sigma^3$

$\beta_2 = \mu_4/\sigma^4$.

$\sqrt{\beta_1}$ is the skewness coefficient, β_2 is the kurtosis coefficient, μ_3 and μ_4 are the third and fourth central moments of the logarithmically transformed variable x respectively, and σ is the corresponding standard deviation. The four parameters of the LN–GIG distribution are estimated *after* the log-transformation by the method of moments[13].

If $\sqrt{\beta_1(z)}$ and $\beta_2(z)$ are required for the original ore values z, one makes use of the r'th moments around the origin:

$$\mu'_r(z) = e^{r\xi}\frac{K_{\gamma+rs}(b)}{K_\gamma(b)}.$$

As the estimation of parameters based on the original z observations is very 'inefficient', one should first make the transformation $x = \ln(z)$ before estimating the parameters.

CASE STUDIES OF THE THREE DISTRIBUTIONS

Compound Lognormal and Log-generalized Inverse Gaussian Distributions

Applying the new distribution density functions (whichever are applicable), theoretical models were fitted to the empirical gold data (measured in sectional cm·g/t units) in 21 different cases. Typical results are shown in Figures 4 and 5 and in the Addendum. In almost all the cases, the hypothetical models describe the empirical data as observed by the 'test of the eye' (Figures 4 and 5).

Use was also made in all cases of a one-sample Kolmogorov–Smirnov (K–S) test of the goodness of fit based on[14]

$$n^{\frac{1}{2}} D(n) = n^{\frac{1}{2}} \, Sup \, | F_n(x) - F(x) |,$$

where $F_n(x)$ is the empirical distribution function from a population with continuous distribution function $F(x)$. A test of the hypothesis that the sample has arisen from $F(x)$ is provided by rejection at the P per cent level if

$$n^{\frac{1}{2}} D(n) \geq d(p).$$

At the 5 per cent level used in the study, $d(p) = 1{,}358$ for large n, implying that all calculated $d(p)$ values less than 1,358 pass the K–S test for goodness of fit. Almost all the K–S tests confirmed the above observation 'by the eye test' (Addendum). However, the chi-square test rejected the test for goodness of fit in almost all the cases. The latter indication is not unexpected because of the rigorous nature and super-sensitivity of the chi-square test, especially when dealing with large numbers of observations. Detailed results of the analyses are tabulated in the Addendum.

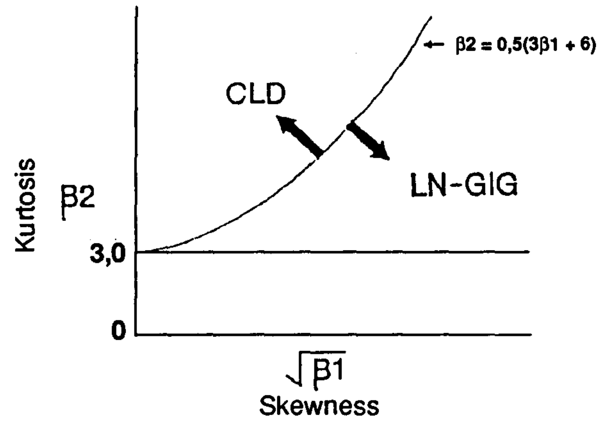

Figure 3—Regions of applicability of the various models

Furthermore, the new distributions have proved to give a more robust estimate of the population mean. Their efficient estimation of the population means is shown in the Addendum. When the chi-square minimization procedure[13] was combined, a bias of not more than 2,5 per cent in the population mean estimate was observed. In some cases, the chi-square minimization approach was adopted to improve the moment estimates of the parameters. This approach was not adopted in cases where it was not going to make much difference. The estimated parameters for the new distributions in the areas concerned are also shown in the Addendum. As indicated there, the location parameters of the new models seem to be stable in the respective reefs. Further analyses involving these parameters are being undertaken.

Three-parameter Lognormal Model

When Krige's technique[1] was applied, all the original empirical distributions were fitted after the respective

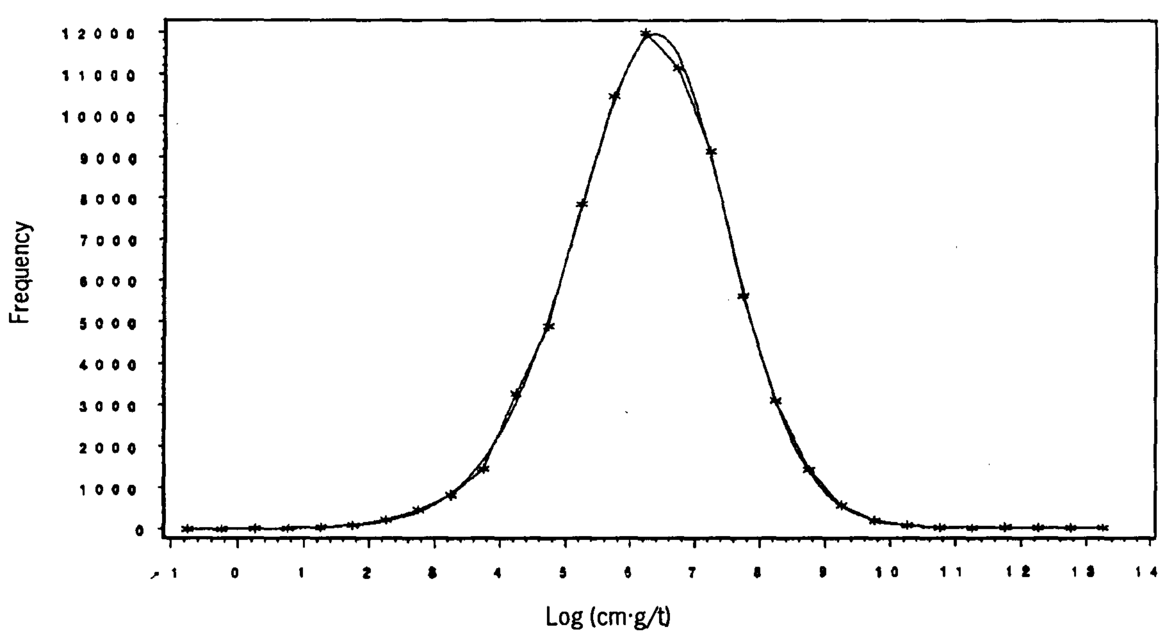

Figure 4—Compound lognormal models fitted to gold data area H, *N* = 72 766

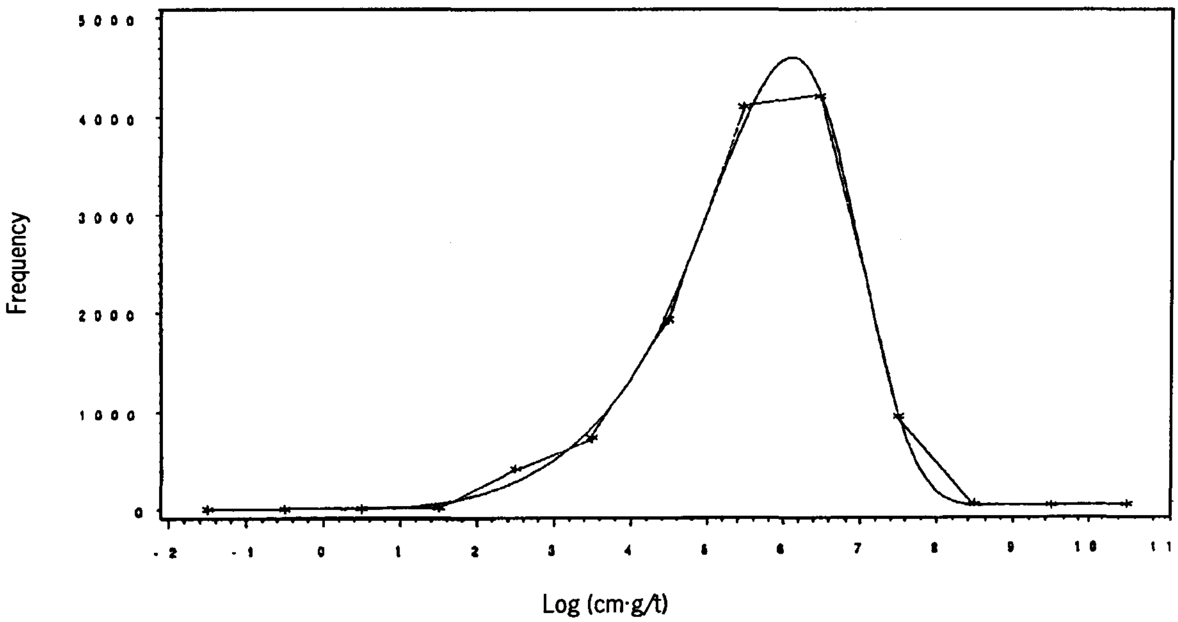

Figure 5— LN–GIG model fitted to gold data area J, N = 12 241

additive constants had been estimated. The analyses indicated a good agreement between the mathematical models, with the empirical data only beyond the logarithm of the additive-constant value. Plots of observed and expected frequencies are shown in Figures 6 and 7, and they pass the 'test of goodness of fit by the eye'. The K–S test of goodness of fit accepted almost all the goodness of fit test, although the chi-square test rejected almost all of them (the reason for this has already been given). In Klerksdorp area H (Vaal Reef), where 72 766 samples were analysed, the *a priori* additive constant of 50 gave a less efficient estimate of the population mean, although the bias was not practically significant. The overall analysis using this model showed less than 5 per cent bias in the estimation of the population mean. However, in more than 30 per cent of the cases, the negative area under the distribution curve was more than 1 per cent (Addendum).

SUMMARY OF THE RESULTS

In the above analysis, the three-parameter model for gold-grade distribution gave an efficient estimate of the population mean with less than 5 per cent bias. The new distributions showed similar efficient estimates with a bias of not more than 2,5 per cent (replacing the original results with chi-square minimization results wherever applicable). The three-parameter model showed biases that were comparable with those obtained from the new models. However, in all the cases where the departure from the three-parameter model was significant, and also where the additive constant was estimated inefficiently, the three-parameter estimates showed significantly higher biases.

The theoretical models of the three distributions fitted the empirical data, though to different extents. It should be noted that the minimum gold sample value, z (cm·g/t), is zero. This implies that, for the three-parameter model, the transformed gold value x ($x = \ln(z)$) is bounded by a finite lower size of $\ln(\beta)$ (indicated by the broken lines in Figures 6 and 7), where β is the additive constant. Any value under the mathematical model that is below this finite lower size will result in a negative gold value after back-transformation. Such negative gold values do not exist in reality. For more than 16 per cent of the cases, the three-parameter model indicated a negative gold proportion of more than 2 per cent, although, as suggested by Sichel[3], this area should not be more than 2 per cent.

In Klerksdorp area H (Vaal Reef), 72 766 gold values were analysed, and the compound lognormal model gave better results than the three-parameter model. This further confirms that, for large mining areas, where the lognormal distributions are likely to be mixed, the new distributions will give better estimates. In view of the large sample in this particular case, the use of twice the number of conventional intervals was possible; Figure 4 shows the excellent fit of the observations to the CLD.

ADVANTAGES AND DISADVANTAGES OF THE THREE MODELS

Three-parameter Lognormal Model

Advantages

- No complex mathematics is required for the estimates of the three parameters, and the distributions are simple to handle.
- Less computer time is generally required for the estimates of the three parameters.
- Only three parameters need to be estimated.
- Working permanence is accepted for the calculations of change of support.
- The estimate of the population mean is efficient.
- The third parameter can usually be accepted *a priori* from nearby data if such data exist.

Disadvantages

- It has the problem of truncation and bimodality, the

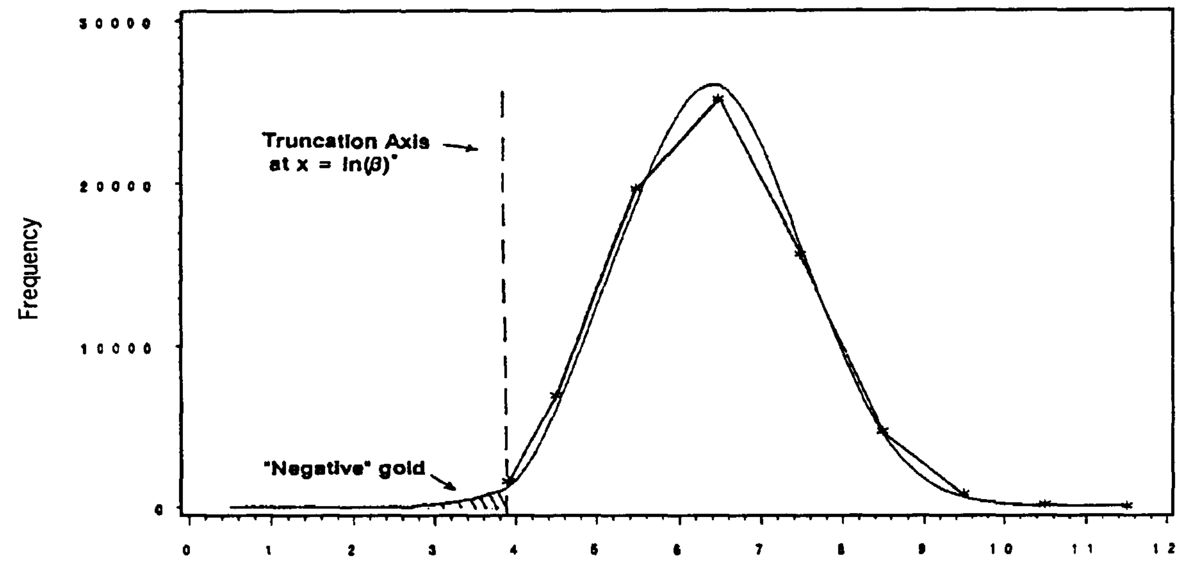

Figure 6—Three-parameter lognormal model fitted to gold data area H, *N* = 72 766 *β is the additive constant

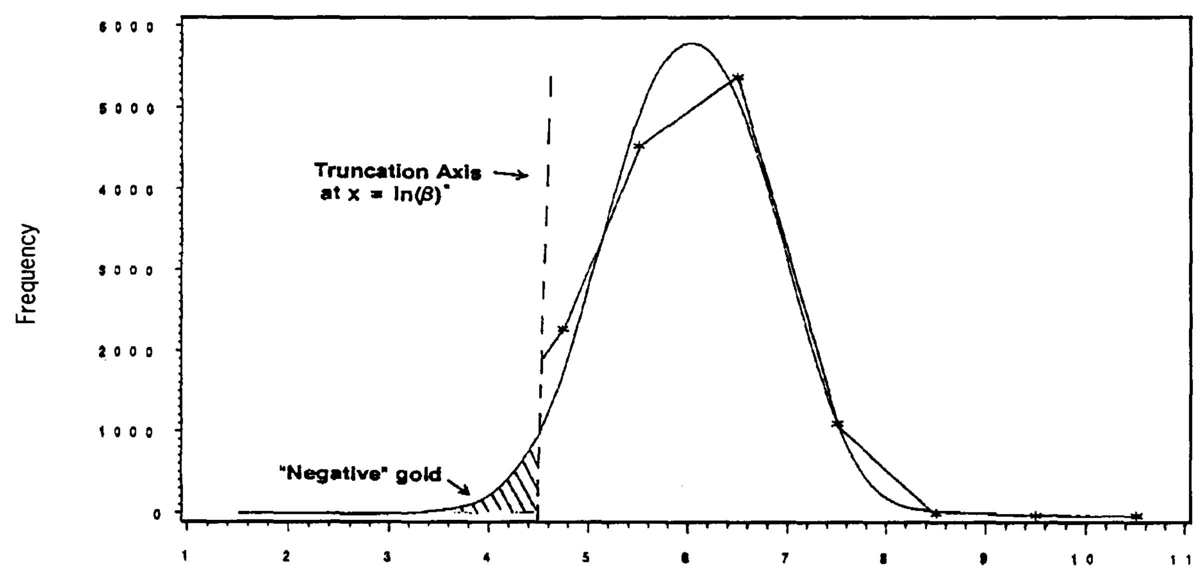

Figure 7—Three-parameter lognormal model fitted to gold data area J, *N* = 12 241 *β is the additive constant

squashing of the frequency distribution making it devoid of its naturality.

- It also has the problem that negative gold values are obtained on back-transformation.
- There is difficulty in obtaining an estimate for the additive constant, especially for small samples.
- Departure from three-parameter lognormality is possible.

Compound Lognormal and LN–GIG Models

Advantages

The new models have no truncation since the logarithm of the actual values ranges from $-\infty < x < +\infty$. There is therefore no issue of negative gold-value results.

- There is no inherent theoretical limitation.
- Since the new distributions deal with the logarithm of the gold values themselves, the frequency categories

look very natural.

- It can handle mixtures of lognormal distributions to a better extent.
- The estimates of the population mean are efficient.

Disadvantages

- The mathematics of the new distributions is more complicated.
- More computer time is generally required.
- Four parameters are required.
- The problem of change of support and small sampling theory is yet to be solved, and this serves as a limitation at present.

CONCLUSIONS AND RECOMMENDATIONS

The study showed that, although the three-parameter model is a practically reasonable model, especially in the estimation of the population mean of gold distributions, the usual assumption that gold-grade distributions follow a three-parameter law is not always true. As a result, the plot of the empirical cumulative frequency distribution is very necessary as a test of departure from lognormality. The study showed further that the new models also provide good estimates of the population means of gold distributions. Another outstanding feature that makes the new distributions more attractive is the fact that no inherent theoretical problems are expected. The new distributions also seem to give better estimates when larger mining areas are under consideration.

However, it is worth noting that change of support and small-sampling theory have to be resolved for the new models. One approach to the change-of-support problem would be to empirically verify a working hypothesis of permanence for the new distributions as it pertains to the three-parameter model. Since the three-parameter model is a limiting case of the new models, it would seem very likely that the permanence hypothesis is an attribute shared by this other family of models. The difference between this family of models at large support sizes (as against the chip-sample values on which these results are based) is another interesting area of research. It would also be useful for a researcher to draw small sample sets from the 72 766 Vaal Reef data for small-sample analysis and to apply the new models. Variability analysis for the new model parameters can also be of practical use, both for borehole valuation and for the analysis of local ore-reserve estimations. This study shows further that, in more than 95 per cent of the cases, the compound lognormal model can be used for the modelling of gold-grade distributions. This indicates that the compound lognormal model is more applicable in the modelling of gold-grade distributions in the Witwatersrand placers than the LN–GIG model.

ACKNOWLEDGEMENTS

The co-operation of the Anglo American, Anglo Vaal, JCI, Rand Mines, and Genmin mining groups is appreciated. A co-author, W. Assibey-Bonsu, made a major contribution to this paper based on some of the research work for his Ph. D. degree. Professor D.G. Krige also assisted with some useful suggestions.

REFERENCES

1. KRIGE, D.G. On the departure of ore value distribution from the lognormal model in South African gold mines. *J.S. Afr. Inst. Min. Metall.*, vol. 60. 1960. pp. 231-244.
2. SICHEL, H.S. The estimation of the means and associated confidence limits for small samples from lognormal populations. *Symposium on Mathematical Statistics and Computer Applications in Ore Valuation*. Johannesburg, South African Institute of Mining and Metallurgy, 1966. pp. 106-122.
3. SICHEL, H.S., and LINDER, A. Contribution to a paper given as reference (1). *J. S. Afr. Inst. Min. Metall.*, vol. 61. 1961. pp. 333-338.
4. SICHEL, H.S. Some advances in lognormal theory. *APCOM 87, vol. 3: Geostatistics*. Johannesburg, South African Institute of Mining and Metallurgy, 1987. pp. 3-8.
5. LINK, R.F., and KOCH, G.S. Some consequences of applying lognormal distributions. *Math. Geol.*, vol. 7. 1975. pp. 117-127.
6. MINTER, W.E.L. The Welkom Gold Field exhumed: Borehole predictions compared with thirty years of mining results. *Proceedings 12th CMMI Congress*. Johannesburg, South African Institute of Mining and Metallurgy, 1982. pp. 137-144.
7. MINTER, W.E.L. Sedimentology of the Witwatersrand basin. *Optima*, vol. 30. 1981–82. pp. 263-272.
8. ANTROBUS, E.S.A. *Witwatersrand gold–100 years*. Johannesburg, Geological Society of South Africa, 1986.
9. ANTROBUS, E.S.A. The Klerksdorp Goldfield. *Mineral deposits of Southern Africa*. Johannesburg, Geological Society of South Africa, 1986. pp. 549-596.
10. HALLBAUER, D.K. A geochemical identification of provenance controlled gold distribution patterns in the Ventersdorp Contact Reef placer of the Carletonville Gold Field. *Gold 100. Proceedings International Conference on Gold*, vol. 1, Gold Mining Technology. Johannesburg, South African Institute of Mining and Metallurgy, 1986. pp. 45-62.
11. KRIGE, D.G., KLEINGELD, W.J.,and OOSTERVELD, M.M. *A priori* parameter distribution patterns for gold in the Witwatersrand basin to be applied in borehole valuations of potential new mines using Bayesian geostatistical techniques. *APCOM 90, Berlin*, 1990. pp. 715-726.
12. SICHEL, H.S. Statistical valuation of diamondiferous deposits. *APCOM 72, Johannesburg*, 1972. pp. 17-24.
13. ASSIBEY-BONSU, W. New models for borehole valuation of new mining operations and routine valuation of ore reserve blocks. Ph.D.(Eng.) thesis, University of the Witwatersrand, Johannesburg, 1991.
14. LINDLEY, D.V., and SCOTT, W.F. *New Cambridge elementary statistical tables*. Cambridge, Cambridge University Press, 1984. 80 pp.

ADDENDUM: ANALYSIS OF EMPIRICAL AND THEORETICAL DISTRIBUTIONS

Refer to table, pp. 98 and 99.

Various areas	Reef	Reef area $m^2 \times 10^6$	No. of samples	Actual values from empirical distribution					Values assuming compound lognorm or LN–GIG* distribution		
				Arithmetic mean	Log variance	Log mean	Log skewness	Log kurtosis	Estimated mean	X^2 probability for goodness of fit† %	K–S test value
Orange Free State											
Area A	Basal	0,72	3962	812,0	1,428	6,154	-1,210	5,418	775,1	<,05 [204]	2,579
Area A1	Basal	0,24	2525	940,2	1,326	6,334	-1,393	6,581	894,5	<,05 [90]	1,997
Area A2	Basal	0,48	1433	586,1	1,459	5,838	-1,032	4,384	577,8*	<,05 [18]*	0,920*
Area AA‡	Basal	0,72	3962	812,0	1,433	6,154	-1,210	6,000	796,6	<,05 [39]	0,898
Area 11‡	Basal	0,24	2525	940,2	1,326	6,334	-1,393	7,360	926,1	<,05 [43]	0,703
Area 22‡	Basal	0,48	1433	586,1	1,459	5,838	-1,032	5,355	600,5	1,4 [13]	0,700
Area B	Basal	1,92	3295	666,4	1,079	6,057	-0,659	3,892	668,9	0,6 [10]	0,309
Area AB	Basal	2,64	7257	745,9	1,275	6,110	-1,003	4,937	730,3	<,05 [62]	1,309
Area C	Basal	1,44	1418	628,8	1,481	5,735	-0,240	3,343	620,8	40 [1,8]	0,257
Area D	Basal	1,44	1015	1239,9	1,116	6,665	-0,757	4,428	1245,8	63 [1,7]	0,125
Area E	Basal	3,52	2448	898,1	1,352	6,239	0,757	4,379	890,3	<,05 [21]	0,660
Area F	Basal	0,96	2209	1097,5	1,416	6,465	-0,970	4,695	1090,2	<,05 [29]	1,074
Area G	Basal	0,96	2804	1149,3	1,146	6,589	-0,905	4,920	1151,9	<,05 [28]	1,026
Klerksdorp											
Hartes											
Area H	V.R.	24,70	72766	1081,0	1,598	6,244	-0,266	3,466	1086,0	<,05 [91]	0,879
Stilfontein											
Area J	VCR	6,25	12241	447,9	1,268	5,618	-0,784	3,487	447,1*	<,05 [121]*	0,996*
S. Roodepoort											
Area K	VCR	0,19	1156	641,9	0,765	6,095	-0,386	5,274	657,1	0,07 [12]	0,823
Area L	VCR	0,18	1378	687,9	0,880	6,117	-0,241	3,615	693,6	0,25 [9]	0,483
Area M1	VCR	0,08	1067	635,3	0,702	6,176	-1,069	6,606	648,6	<,05 [39]	1,312
Area M2	VCR	0,08	1066	635,9	0,726	6,181	-0,852	5,107	644,6	0,06 [15]	0,658
Winkelhaak											
Area X	Kimb.	1,92	8439	752,1	1,897	5,890	-0,747	3,747	736,6*	<,05 [93]*	1,206*
Area XX‡	Kimb.	1,92	8439	752,1	1,897	5,890	-0,747	4,000	747,1	<,05 [99]	1,187
Area Y	Kimb.	1,92	8095	532,5	1,449	5,685	-0,683	3,926	528,5	<,05 [36]	0,801

* Results based on LN–GIG analysis
† [Chi-square values]
‡ Results based on minimization of chi-square (Areas: AA‡ for A; A11‡ for A1; A22‡ for A2; XX‡ for X)
§ Estimates

AND THEORETICAL DISTRIBUTIONS

	Values assuming three-parameter lognormal distribution								Moment estimate for the four parameters of the new distributions			
Additive constant	Skewness of log (values + constant)	Kurtosis of log (values +) constant)	Variance of log (values + constant)	Estimated mean	X^2 probability for goodness of fit†	K–S test value	Bias %	% of area in negative gold region	Skewness parameter c	Kurtosis parameter v§	Spread or scale parameter s§	Location parameter a
125	-0,001	3,421	0,601	782,7	<,05 [13]	0,671	-3,60	1,50	0,358862	1,704851	3,189133	7,445670
131	-0,008	3,649	0,580	893,9	<,05 [14]	0,499	-4,92	1,02	0,239773	0,860618	2,098297	7,168990
115	-0,044	3,050	0,579	587,8	0,3 [9]	0,581	0,29	2,28	0,750000*	0,032150*	0,797948*	3,378008*
N/A									0,180932	1,018951	1,931575	6,970482
N/A									0,148255	0,497549	1,543289	6,918064
N/A									0,153786	1,362535	2,027919	6,689711
90	-0,043	2,892	0,597	671,5	9 [5]	0,073	0,76	0,89	0,180394	4,604444	4,077832	7,353945
105	-0,019	3,180	0,617	735,0	<,05 [29]	0,803	-1,46	1,22	0,211816	1,941750	2,733739	7,152718
5	-0,063	3,034	1,355	625,0	26 [4]	0,404	-0,60	0,02	0,034022	9,370583	3,842793	6,715765
130	0,026	3,002	0,653	1241,1	2 [8]	0,501	0,10	0,62	0,120372	2,406130	2,745898	7,499414
70	0,040	3,225	0,804	882,4	<,05 [15]	0,655	-1,75	0,69	0,128085	2,550809	2,587789	7,207190
110	-0,044	2,812	0,767	1101,0	30 [2]	0,378	0,32	1,07	0,258092	2,424435	3,073016	7,767990
120	-0,030	3,049	0,664	1152,3	6 [6]	0,370	0,26	0,64	0,135175	1,717597	2,423696	7,366757
50	0,262	2,834	1,155	1034,7	<,05 [1047]	4,161	-4,30	0,99				
25	0,087	2,886	1,314	1066,3	<,05 [191]	1,135	-1,40	0,33	0,030255	6,672023	3,135737	7,064020
90	-0,068	2,581	0,554	456,4	<,05 [116]	1,453	1,90	2,02	0,702500*	0,092530*	0,784566*	4,037013*
24	0,010	3,940	0,629	634,1	1 [6]	0,569	-1,22	0,01	0,117279	0,879650	1,933119	6,251569
20	0,024	3,235	0,757	684,7	0,3 [11]	0,399	-0,47	0,01	0,017520	4,711047	3,533014	6,514262
75	-0,114	3,145	0,429	645,2	12 [2]	0,274	1,56	0,09	0,080660	0,561888	1,966968	6,509278
N/A									0,092135	1,363942	2,710510	6,640887
47	0,036	2,482	1,088	756,7	<,05 [139]	1,929	0,61	2,17	1,286000*	0,223000*	1,374096*	3,055494*
N/A									0,262266	4,475246	3,487997	7,870077
40	0,020	2,735	0,897	532,8	<,05 [25]	0,857	0,06	0,96	0,195387	4,511885	3,573866	7,225878