doi:10.3799/dqkx.2011.021

A Lognormal Distribution of Metal Resources

Donald A. Singer

U. S. Geological Survey, 345 Middlefield Road, Menlo Park, California 94025, USA

Abstract: For national or global resource estimation of frequencies of metals, a lognormal distribution has commonly been recommended but not adequately tested. Tests of frequencies of Cu, Zn, Pb, Ag, and Au contents of 1 984 well-explored mineral deposits display a poor fit to the lognormal distribution. When the same metals plus Mo, Co, Nb_2O_3, and REE_2O_3 are grouped into 19 geologically defined deposit types, only eight of the 73 tests fail to be fit by lognormal distribution, and most of those failures are in two deposit types suggesting a problem with those types. Estimates of the mean and standard deviation of each of the metals in each of the deposit types are provided for modeling.

Key words: mineral deposit model; spatial rule; grade and tonnage model.

INTRODUCTION

Whether considering a country's possible future supply of minerals, planning the merits of exploring for certain kinds of mineral deposits, or examining possible global availability of some mineral materials, having a probability distribution of the amounts of minerals of interest would be invaluable (Singer and Menzie, 2010). One probability distribution that has been suggested as appropriate for mineral resources is the lognormal distribution. Recommendations of the lognormal distribution as an appropriate model of the frequency of ore deposits has waxed and waned over the years. Part of the change in views is due to variation in apparent fit of the distribution to case studies and part of the change may be due to variation in popularity of different techniques over time.

Much of the early research focused on the distribution of grades of mineral deposits or geochemical abundances (Rasumovsky, 1940; Ahrens, 1954; Matheron, 1959). These studies found an empirical and theoretical basis for believing that the lognormal distribution is an appropriate model for observed mineral deposit grades and values of trace elements in samples. Usefulness of the lognormal distribution was further documented by the development of its theoretical foundations and by the empirical evidence of its applicability in biology, sociology, astronomy and economics provided by Aitchison and Brown (1963).

Based on studies of mineral production, Allais (1957) selected a lognormal distribution to represent the values of mineral deposits thereby suggesting a lognormal distribution of metals. Slichter et al. (1962) documented a graphical fit of the lognormal distribution to the gross values of copper, lead, zinc, gold and silver mines in part of the Southwest of the United States. Gross values of sandstone-hosted uranium deposits in the Ambrosia Lake region of the United States were tested and shown to be well represented by the lognormal distribution (Griffiths and Singer, 1973). Economic effects on the fit of the lognormal distribution to diamond production were demonstrated by Sharp (1976). Zhang et al. (2004) found that copper equivalent grades of deposits in China could be represented by a lognormal distribution but metal content could only be represented by lognormal after separating the deposits and districts into different groups. Singer (1993) tested the distribution of ore tonnages and average grades of sixty-seven types of mineral deposits and found that most were not significantly different than the lognormal. On-

E-mail: singer@usgs.gov
Manuscript received July 15, 2010

ly five of the sixty-seven tonnage of ore distributions were significantly different from lognormal at the 1 percent level. Although it is commonly assumed that the distribution of metal amounts can be represented by lognormal distributions, the idea has seen little actual testing and no modern estimates of the parameters of these lognormal distributions of metals have been published.

In this study the ability of the lognormal distribution to fit the observed distribution of some metals is tested. In addition, estimates of the parameters of the lognormal distributions are provided where appropriate. Before testing the distributions consideration of the nature of mineral deposit data suitable for testing and the sources of these data are presented.

MINERAL DEPOSIT INFORMATION

Typically, a lognormal distribution can be used to model the observed distributions of homogeneous populations of variables representing weights, lengths, volumes, and grades of trace quantities. It tends to not fit grade distributions of elements that have grades greater than about 10 percent, such as Fe, Mn, and Al. So what is the proper homogeneous population that should be sampled to represent metals in mineral deposits?

The geological and mining literature contains many terms such as district, zone, ore body, lens, shaft, vein, bench, open-pit, underground, and mine that might be considered as possible sampling units. These terms are applied in different ways by different groups at different points in time, making them undesirable as our sampling unit. Grade-and-tonnage data are available to varying degrees for districts, deposits, mines, and shafts. In many cases, old production data are available for some deposits and recent resource estimates are available for other deposits. A common error is mixing old production data from some deposits with resource data from other deposits. It is extremely important that all data used in the model represent the same sampling unit because mixing data from deposits and districts or old production and recent resource estimates usually produce bimodal distributions representing non-homogeneous populations and it may introduce correlations among the variables that are artifacts of the mixed sampling units. Models constructed using data from mixed sampling units are of questionable value because the frequencies observed are directly related to the proportion of deposits from each sampling unit and are unlikely to be representative of the proportion in the undiscovered deposits being estimated.

For the analysis here of the frequency distributions of metal contents, data used in grade-and-tonnage models are used because they were specifically prepared for assessments to show the frequencies of different sizes and grades of each mineral deposit type based on data collected on thousands of well-explored deposits from around the world. For each deposit type, these models help define a deposit, as opposed to a mineral occurrence or a weak manifestation of an ore-forming process. Data utilized to construct these models include average grades of each metal or mineral commodity of possible economic interest and the associated tonnage based on the total production, reserves, and resources at the lowest possible cutoff grade. These data represent an estimate of the endowment of each of many known deposits. Well-explored in this report means completely drilled in three dimensions or completely mined out. Additionally these data were gathered using spatial rules in order to be consistent in what they represent.

DATA SOURCES AND SPATIAL RULES

For sediment-hosted zinc-lead deposits (Singer et al., 2009), all mineralized rock or alteration within two kilometers was combined into one deposit for these deposits. Thus, if the alteration zones of two deposits are within two kilometers of each other, they were combined. The two-kilometer rule was developed to try to insure that deposits in grade and tonnage and spatial density models correspond to deposits as geological entities. Rules such as the two-kilometer rule are essential in order to have an internally consistent assessment system where the estimate of number of undiscovered deposits is consistent with the grade-and-tonnage model. Sediment-hosted zinc-lead types include: CAam (carbonate-hosted amagmatic), CAig (carbonate-hosted igneous), CAme (carbonate-hosted metamorphic), SHam (shale-hosted amagmatic), Kipushi, SHig (shale-hosted igneous), MLig (mixed lithology-hosted igneous), MLme (mixed lithology-hosted metamorphic), and SSPb (sandstone-hosted Pb).

For the porphyry copper deposits used in this analysis (Singer et al., 2008), all mineralized rock or alteration within two kilometers was combined into one deposit. Thus if the alteration zones of two deposits are within two kilometers of each other, they were combined.

Sediment-hosted copper deposits (Cox et al., 2003) were combined into one deposit if they were within two kilometer of each other. Iron oxide Cu-Au deposits (Cox and Singer, 2007) were combined if they occurred within two kilometers of each other.

For the deposits in the carbonatite model (Berger et al., 2009), the following rule was used to determine which ore bodies were combined. All mineralized rock or altered rock within two kilometers was combined into one deposit. Some examples illustrate the effects of the application of this rule to combined deposits: (1) Salitre I and II deposits in Brazil and (2) Upper Fir and Fir in Canada.

For the volcanogenic massive sulfide deposits (Mosier et al., 2009), the following spatial rule was used to determine which ore deposits were combined. All mineralized rock within 500 m was combined into one deposit. A 500 m rule was used for volcanogenic massive sulfide deposits because of their smaller size and the scarcity of mapped alteration zones around these deposits. For example, in this report, Horne and Quemont in Quebec, Canada, are combined into one deposit, and Jerome in Arizona, United States, has been treated as two separate deposits, United Verde and United Verde Extension, because of the 500 m rule. The volcanogenic massive sulfide deposits were classed into three types. The felsic type (VMSFelsic) includes those volcanogenic massive deposits hosted in dominantly felsic or bimodal-felsic rocks. The bimodal-mafic type (VMSBimodal) includes those volcanogenic massive sulfide deposits dominantly hosted in mafic volcanic rocks with rhyolite to dacite constituting 10 to 40 percent of the host rocks. The mafic type (VMSMafic) of volcanogenic massive sulfide deposit is dominantly hosted in mafic volcanic rocks and associated pelitic rocks.

TESTS OF LOGNORMAL DISTRIBUTIONS

A total of 1 984 mineral deposits with reported tonnages and grades is available from the above listed sources. The copper content in the 1 591 copper-bearing deposits plotted against a standard normal distribution seems to deviate slightly from the expected normal line only in the upper copper values. By visual examination, many might conclude that the lognormal distribution adequately fits the observed copper contents on mineral deposits. Similar results are obtained for the 1 069 zinc-bearing deposits, the 786 lead-bearing deposits, the 1 061 silver-bearing deposits, and the 890 gold-bearing deposits. However, when the distributions of each of these metals were tested using the Shapiro-Wilk W test (Stuart and Ord, 1991), the probability that the observed distributions came from random samples of lognormal distributions was found to be 0.002 or less. Thus we reject the lognormal distribution as a model of the distribution of metal in these grouped deposits.

Perhaps the lack of fit is due to observed distributions not being from homogeneous populations of the metals. The data used in this analysis were compiled by deposit types, which are defined as deposits that occur in similar geological settings and have similar characteristics. Deposits within each type might represent the homogeneous popu-

Table 1 Copper content distributions by deposit type and tests of lognormality. Mean metric tons of contained copper ($\log_{10}$ data), standard deviation ($\log_{10}$ data), median observed copper content of all deposits, Shapiro-Wilk goodness-of-fit probability of lognormal distribution, number of deposits with reported grade, and total number of deposits with reported tonnage (in thousands t)

Deposit type	Mean	St. dev.	Median Cu (kt)	Prob.	Number deposits	Total number deposits
CAam	4.355 1	0.941 2	0	0.797	9	132
CAig	4.315 4	0.876 4	0	0.892	86	187
FeOxideCuAu	5.644 4	0.779 4	376	0.819	32	36
Kipushi	5.438 5	0.894 1	480	0.668	8	8
MLig	4.854 3	0.735 3	0	0.587	12	38
MLme	4.207 3	0.761 1	0	0.296	5	12
PorCu	6.008 6	0.703 1	1 030	0.391	422	422
RedbedCu	4.494 0	0.804 0	21	0.030	33	33
Reduced Cu	5.819 9	0.898 3	500	0.647	62	62
Revett Cu	5.224 0	0.931 5	125	0.934	14	14
SedHstCu	5.076 4	0.920 1	120	0.952	31	31
SHam	5.107 2	1.133 5	0	0.983	7	25
SHig	4.503 4	0.822 4	0	0.487	13	32
SSPb	3.830 8	1.625 7	0	0.042	5	22
VMSBimodal	4.376 7	0.929 2	26	0.666	267	272
VMSFelsic	4.480 2	0.920 0	36	0.001	411	421
VMSMafic	4.061 3	0.983 6	11	0.066	174	175

Table 2 Zinc content distributions by deposit type and tests of lognormality. Mean metric tons of contained zinc ($\log_{10}$ data), standard deviation ($\log_{10}$ data), median observed zinc content of all deposits, Shapiro-Wilk goodness-of-fit probability of lognormal distribution, number of deposits with reported grade, and total number of deposits with reported tonnage (in thousands t)

Deposit type	Mean	St. dev.	Median Zn (kt)	Prob.	Number deposits	Total number deposits
CAam	5.648 1	0.756 7	498	0.338	128	132
CAig	5.359 4	0.917 6	308	0.002	185	187
CAme	5.372 5	0.744 5	128	0.378	7	7
Kipushi	5.765 9	1.226 1	5	0.335	4	8
MLig	5.609 9	0.876 0	354	0.926	38	38
MLme	5.169 1	1.025 0	96	0.888	12	12
SHam	6.334 1	0.674 9	1 600	0.924	25	25
SHig	5.757 5	0.994 9	644	0.511	30	32
SSPb	4.628 2	1.543 8	8	0.264	16	22
VMSBimodal	4.716 6	0.923 1	30.5	0.407	217	272
VMSFelsic	5.043 7	0.892 7	75	0.000	349	421
VMSMafic	4.325 7	0.733 8	0	0.383	58	175

lations desirable for a good fit by the lognormal distribution.

Each of 1 984 deposits was classfied into one of 19 deposit types and the distributions of reported metals within type were each tested for lognormality with the Shapiro-Wilk W test (Tables 1－6). The probability that the observed distributions came from random samples of lognormal distributions was found to be less than 0.01 in only eight of the 73 tests. Thus we reject the lognormal distribution as a model of the distribution of metals in eight of these cases. Out of the eight rejections, five are from one deposit type, the volcanic-hosted massive sulfide felsic (VMSFelsic) type. Two other rejections are from the carbonate-hosted, igneous (CAig) related type. This strongly suggests that the VMSFelsic model contains more than one population and perhaps the CAig model does also.

Table 3　Lead content distributions by deposit type and tests of lognormality. Mean metric tons of contained lead ($\log_{10}$ data), standard deviation ($\log_{10}$ data), median observed lead content of all deposits, Shapiro-Wilk goodness-of-fit probability of lognormal distribution, number of deposits with reported grade, and total number of deposits with reported tonnage (in thousands t)

Deposit type	Mean	St. dev.	Median Pb (kt)	Prob.	Number deposits	Total number deposits
CAam	5.176 4	0.822 2	146	0.647	121	132
CAig	5.181 3	0.857 3	130	0.230	166	187
CAme	5.044 8	0.801 0	30	0.243	6	7
Kipushi	5.291 9	1.161 9	275	0.123	6	8
MLig	5.372 9	0.885 2	169	0.838	34	38
MLme	4.852 3	0.830 0	45	0.420	11	12
SHam	5.823 6	0.718 4	592	0.482	25	25
SHig	5.428 0	1.041 9	721	0.066	32	32
SSPb	5.167 4	0.852 6	204	0.691	22	22
VMSBimodal	4.058 6	0.844 2	0	0.006	77	272
VMSFelsic	4.497 6	0.982 0	7	0.005	273	421
VMSMafic	3.564 9	1.018 0	0	0.636 6	13	175

Table 4　Silver content distributions by deposit type and tests of lognormality. Mean metric tons of contained silver ($\log_{10}$ data), standard deviation ($\log_{10}$ data), median observed silver content of all deposits, Shapiro-Wilk goodness-of-fit probability of lognormal distribution, number of deposits with reported grade, and total number of deposits with reported tonnage (in t)

Deposit type	Mean	St. dev.	Median Ag (t)	Prob.	Number deposits	Total number deposits
CAam	2.275 0	0.759 8	0	0.875	58	132
CAig	2.651 2	0.939 1	144	0.147	138	187
CAme	2.045 0	0.927 6	26	0.210	5	7
FeOxideCuAu	1.918 2	1.155 2	0	0.621	10	36
Kipushi	2.692 5	0.968 3	193	0.970	6	8
MLig	2.688 5	0.951 3	106	0.928	26	38
MLme	2.588 2	0.702 0	32	0.565	6	12
PorCu	2.897 5	0.692 6	0	0.584	172	422
RedbedCu	1.544 1	1.142 0	0	0.698	10	33
Reduced Cu	2.678 0	1.053 5	0	0.274	16	62
Revett Cu	2.807 3	0.756 3	140	0.097	8	14
SedHstCu	1.797 9	0.988 4	0	0.153	7	31
SHam	3.067 3	0.783 8	277	0.540	18	25
SHig	2.760 3	0.999 9	786	0.056	28	32
SSPb	2.041 5	0.904 8	2	0.344	11	22
VMSBimodal	1.754 9	0.936 8	10	0.816	172	272
VMSFelsic	2.051 3	0.958 1	32	0.003	300	421
VMSMafic	1.442 0	1.058 8	0	0.399	70	175

The large number of tests of lognormality of these metals by deposit types that were not rejected indicates that the lognormal distribution is an appropriate model for metals when deposit types are used and when the data are consistently gathered. Figure 1 shows the copper content distributions plotted against standard normal distributions of several deposit types that seem to deviate slightly from the expected normal lines demonstrating how difficult it is to identify which distributions fit or do not fit a lognormal distribution graphically. Previous publications recommending the lognormal distribution for metal resources relied on graphical comparisons or on improper statistical tests such as correlations with cumulative distributions. The graphical results in Figure 1 should be compared to

Table 5 Gold content distributions by deposit type and tests of lognormality. Mean metric tons of contained gold ($\log_{10}$ data), standard deviation ($\log_{10}$ data), median observed gold content of all deposits, Shapiro-Wilk goodness-of-fit probability of lognormal distribution, number of deposits with reported grade, and total number of deposits with reported tonnage (in t)

Deposit type	Mean	St. dev.	Median Au (t)	Prob.	Number deposits	Total number deposits
CAam	0.985 7	0.402 7	0	0.424	3	132
CAig	0.676 6	1.026 3	0	0.002	70	187
FeOxideCuAu	1.313 5	0.797 8	12	0.563	27	36
MLig	0.408 9	0.766 9	0	0.039	7	38
PorCu	1.678 6	0.712 5	12	0.631	256	422
SHam	0.755 8	0.444 6	0	0.695	4	25
SHig	0.662 0	1.190 3	0	0.135	14	32
VMSBimodal	0.309 1	0.956 7	0	0.093	158	272
VMSFelsic	0.430 6	0.980 7	1	0.000	279	421
VMSMafic	0.058 6	1.042 8	0	0.121	72	175

Table 6 Molybdenum, cobalt, Nb_2O_5, and REE_2O_3 content distributions by deposit type and tests of lognormality. Mean metric tons of contained molybdenum (median observed Mo content in all deposits), cobalt (median observed cobalt content in all deposits), Nb_2O_5, and REE_2O_3 ($\log_{10}$ data), standard deviation ($\log_{10}$ data), Shapiro-Wilk goodness-of-fit probability of lognormal distribution, number of deposits with reported grade, and total number of deposits with reported tonnage (in thousands t)

Deposit type	Mean Mo (median Mo (kt))	Mean Co (median Co (kt))	Mean Nb_2O_3 (median Nb_2O_3 (kt))	Mean REE_2O_3 (median REE_2O_3 (kt))	St. dev.	Prob.	Number deposits	Total number deposits
PorCu	4.618 (28)				0.765 4	0.592	228	422
Reduced Cu		5.088 1 (0)			0.800 4	0.633	15	62
Carbonatite			5.325 8 (0)		1.102 5	0.737	39	55
Carbonatite				5.558 9 (68)	1.079 3	0.925	35	55

the probabilities of the fits to the lognormal in Table 1.

The tables also contain means and standard deviations of the metal contents for those deposits with reported grades for the metal and deposit type. Because many of the metals are byproducts, they are not always reported and in these cases it is reasonable to assume that they have quite low grades. It is for this reason that the numbers of deposits with reported grades are provided along with the total number of deposits for which there are tonnages. For example, in Table 1, the CAam deposit type has a total number of deposits with tonnages of 132, but only nine of these deposits have a reported copper grade and only these nine deposits were used for the mean and standard deviation contained metal estimates. Thus the mean and standard deviation copper contents are only representative of 9/132 or about 7 percent of the deposits of this type. Although the number of deposits that have reported copper is small in the CAam case, it provides useful information not otherwise available. For the copper contents of the CAam group and the CAig group, it is also important to remember that these types are representing zinc and lead deposits hosted in sediments and not principally copper deposits in these settings such are copper skarns which are not reported here. The column showing the median contained metal in the tables reflects the observed medians of all of the deposits including those without reported grades.

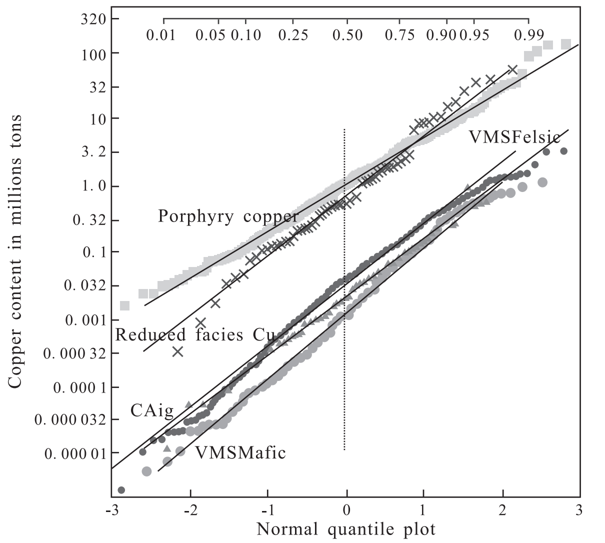

Fig. 1 The copper content of porphyry Cu, reduced facies Cu, carbonate-hosted igneous related, volcanic-hosted massive sulfide felsic-related, and volcanic-hosted massive sulfide mafic-related deposits plotted against each type's straight line representing its standard normal distribution

CONCLUSIONS

The lognormal distribution cannot be used to model the distribution of metals in mineral deposits when multiple deposit types are combined. Where the mineral deposit models are developed using consistent geological settings and the data represent well-explored deposits with consistently applied spatial rules, the lognormal distribution is a good representative of the frequencies of metal contents.

Out of the 73 tests of the adequacy of the lognormal distribution to fit the observed distributions of metals here, most of the inadequate fits were for metals in the volcanic-hosted massive sulfide, felsic setting type. In each of these cases, the lack of fit was because some of the largest content deposits had less metal reported than would be expected by the lognormal distribution. It is not known why this occurred, however, if the reason is poor reporting of tonnages or grade of these large deposits, then the lognormal distribution could be used to model the metals correctly. If the reason for the poor fit is that there is really another as yet unrecognized deposit type that is appropriate for these larger deposits, then the use of the estimated parameters of the lognormal distribution reported here would overestimate the amounts of metals. Fortunately, this deposit type is not the primary supplier of any of the metals, so an estimation error would probably not be fatal for global estimates of metals.

REFERENCES

Ahrens, L. H., 1954. The log-normal distribution of the elements (a fundamental law of geochemistry and its subsidiary). *Geochimica et Cosmochimica Acta*, 5: 49—73.

Aitchison, J., Brown, J. A. C., 1963. The lognormal distribution. Cambridge University Press, London, 177.

Allais, M., 1957. Method of appraising economic prospects of mining exploration over large territories—Algerian Sahara case study. *Management Science*, 3(4): 285—347.

Berger, V. I., Singer, D. A., Orris, G. J., 2009. Carbonatites of the world, explored deposits of Nb and REE: database and grade and tonnage models. U. S. Geological Survey Open-File Report 2009—1139, 17 and database. http://pubs.usgs.gov/of/2009/1139/.

Cox, D. P., Lindsey, D. A., Singer, D. A., et al., 2003. Sediment-hosted copper deposits of the world—deposit models and database. Revised 2007. U. S. Geological Survey Open-File Report 2003—107, v. 1.3. http://pubs.usgs.gov/of/2003/of03—107/.

Cox, D. P., Singer, D. A., 2007. Descriptive and grade-tonnage models and database for iron oxide Cu-Au deposits. U. S. Geological Survey Open-File Report 2007—1155. http://pubs.usgs.gov/of/2007/1155/.

Griffiths, J. C., Singer, D. A., 1973. Size, shape, and arrangement of some uranium ore bodies. 11th International Symposium on Computer Applications in the Mineral Industry, B82—B110.

Matheron, G., 1959. Remarques sur la loi de Lasky. Chronique des Mines d'Outre-Mer et de la Recherche Miniere, 27e annee, (282): 463—465.

Mosier, D. L., Berger, V. I., Singer, D. A., 2009. Volcanogenic massive sulfide deposits of the world—database and grade and tonnage models. U. S. Geological Survey Open-File Report 2009—1034. http://pubs.usgs.gov/of/2009/1034/.

Rasumovsky, N. K., 1940. Distribution of metal values in ore deposits. Acad. Sci. Comptes Rendus (Doklady) U. R. S. S. ,28:814－816.

Sharp,W. E. ,1976. A log-normal distribution of alluvial diamonds with an economic cutoff. *Economic Geology*, 71:648－655.

Singer,D. A. ,1993. Basic concepts in three-part quantitative assessments of undiscovered mineral resources. *Nonrenewable Resources*,2(2):69－81.

Singer,D. A. ,Berger,V. I. ,Moring,B. C. ,2008. Porphyry copper deposits of the world: database, map, and grade and tonnage models,2008. U. S. Geological Survey Open-File Report 2008－1155. http://pubs. usgs. gov/of/2008/1155/.

Singer,D. A. ,Berger,V. I. ,Moring,B. C. ,2009. Sediment-hosted zinc-lead deposits of the world: database and grade and tonnage models. U. S. Geological Survey Open-File Report 2009－1252. http://pubs. usgs. gov/of/2009/1252/.

Singer,D. A. ,Menzie,W. D. ,2010. Quantitative mineral resource assessments—an integrated approach. Oxford University Press,New York,232.

Slichter,L. B. ,Dixon,W. J. ,Meyer,G. H. ,1962. Statistics as a guide to prospecting. In:Proc. Symposium Mathematical and Computer Applications in Mining and Exploration. College of Mines, University of Arizona, Tucson,AZ,F1－1－F－1－27.

Stuart,A. ,Ord,J. K. ,1991. Kendall's advanced theory of statistics,v. 2,5th ed. . Oxford University Press,New York,1323.

Zhang,Q. L. ,Shoji,Tetsuya,Kaneda,Hiroaki,2004. Grade-tonnage models of copper deposits in China. Shigen-to-Sozai,The Mining and Materials Processing Institute of Japan,120:19－24.