

Probability Calculations for Three-Part Mineral Assessments

Chapter 15 of
Section C, Computer Programs
Book 7, Automated Data Processing and Computations

Techniques and Methods 7–C15

Probability Calculations for Three-Part Mineral Resource Assessments

By Karl J. Ellefsen

Chapter 15 of
Section C, Computer Programs
Book 7, Automated Data Processing and Computations

Techniques and Methods 7–C15

**U.S. Department of the Interior
U.S. Geological Survey**

U.S. Department of the Interior

RYAN K. ZINKE, Secretary

U.S. Geological Survey

William H. Werkheiser, Acting Director

U.S. Geological Survey, Reston, Virginia: 2017

For more information on the USGS—the Federal source for science about the Earth, its natural and living resources, natural hazards, and the environment—visit <https://www.usgs.gov> or call 1–888–ASK–USGS.

For an overview of USGS information products, including maps, imagery, and publications, visit <https://store.usgs.gov>.

Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Although this information product, for the most part, is in the public domain, it also may contain copyrighted materials as noted in the text. Permission to reproduce copyrighted items must be secured from the copyright owner.

Suggested citation:

Ellefsen, K.J., 2017, Probability calculations for three-part mineral resource assessments: U.S. Geological Survey Techniques and Methods, book 7, chap. C15, 14 p., <https://doi.org/10.3133/tm7C15>.

ISSN 2328-7055 (online)

Contents

Abstract.....	1
Introduction.....	1
Method.....	1
Conceptual Model.....	1
Data for the Probability Calculations.....	2
Background.....	2
Grade and Tonnage Model.....	2
Estimates of the Number of Undiscovered Deposits.....	2
Distribution for the Ore Tonnages.....	3
Distribution for the Mineral Resource and Gangue Grades.....	3
Distribution for the Number of Undiscovered Deposits.....	5
Simulation of the Undiscovered Deposits.....	5
Distribution for Total Ore and Mineral Resource Tonnage.....	6
Mean and Variance of the Total Ore and Mineral Resource Tonnage.....	7
Checking the Final Calculations.....	8
Calculations Summary.....	8
Properties of the Probability Calculations.....	8
Precision.....	8
Computation Time.....	9
Sensitivity Analysis.....	10
Discussion.....	11
Suitability of the Assumptions.....	11
Tonnage Models.....	12
Number of Simulations.....	12
Conclusions.....	12
New Features.....	12
New Understanding.....	13
New Findings.....	13
Acknowledgments.....	13
Software and Reproducibility.....	13
References Cited.....	13

Figures

1–7. Graphs showing:

1. Probability density function (histogram) that represents the ore tonnage in an undiscovered deposit.....3
2. *A*, *B*, and *C*, Histograms of the log-ratios of the grades from the joint probability density function that represents the grades in an undiscovered deposit.....4
3. *A*, Negative binomial probability mass function representing the number of undiscovered deposits in the permissive tract. *B*, Probability mass function recast as elicitation percentiles and compared to the estimated numbers of undiscovered deposits.....6

4. *A*, Univariate, marginal, probability density functions and *B*, univariate, marginal, complementary cumulative distribution functions for the total ore and mineral resource tonnages in all undiscovered deposits within the permissive tract7

5. *A-I*, Check of the final calculations.....9

6. *A-I*, Effect of the number of simulations on the simulation standard error.....10

7. *A*, Effect of the number of simulations on the computation time, for four probability mass functions (pmf). *B*, Effect of the mean of the pmf on the computation time, for five numbers of simulations11

Tables

1. First five entries in the grade and tonnage model that accompanies program MapMark4.....2

2. Estimates of the number of undiscovered deposits by the 10 members of an assessment team.....2

Conversion Factors

International System of Units to Inch/Pound

Multiply	By	To obtain
	Mass	
kilogram (kg)	2.205	pound avoirdupois (lb)

Probability Calculations for Three-Part Mineral Resource Assessments

By Karl J. Ellefsen

Abstract

Three-part mineral resource assessment is a methodology for predicting, in a specified geographic region, both the number of undiscovered mineral deposits and the amount of mineral resources in those deposits. These predictions are based on probability calculations that are performed with computer software that is newly implemented. Compared to the previous implementation, the new implementation includes new features for the probability calculations themselves and for checks of those calculations. The development of the new implementation lead to a new understanding of the probability calculations, namely the assumptions inherent in the probability calculations. Several assumptions strongly affect the mineral resource predictions, so it is crucial that they are checked during an assessment. The evaluation of the new implementation leads to new findings about the probability calculations, namely findings regarding the precision of the computations, the computation time, and the sensitivity of the calculation results to the input.

Introduction

Three-part mineral resource assessment is a methodology for predicting, in a specified geographic region, both the number of undiscovered mineral deposits and the amount of mineral resources in those deposits. (Singer and Menze, 2010). These predictions are based on probability calculations, which are implemented with computer software. A completely new version of this software and its documentation have been recently published as a software package called MapMark4 (Ellefsen, 2017). Although the documentation describes how the software is used, it does not describe, for example, how the software works. This and related information, which are needed to understand the probability calculations, are presented in this report.

The method of the probability calculations has been described in conceptual terms (Harris, 1984, p. 317 and 389; Singer and Menze, 2010, p. 148–158), and the probability calculations have been implemented in several different computer programs (Root and others, 1992, 1998; Duval, 2001, 2002, 2012; Brown and Friedel, 2011a, 2012b; Bawic

and Spanski, 2012). However, none of these authors have described the details of the probability calculations, described the properties of the calculations, and discussed the suitability of the assumptions.

These issues, as well as other issues, are addressed in the three major sections of this report. The first section presents the method of the probability calculations, including both the algorithmic details and the assumptions that are inherent in the method. The second section presents important properties of the probability calculations. The final section primarily discusses the suitability of selected assumptions.

This report focuses on the most common type of probability calculations, namely calculations involving inputs called “grade and tonnage models.” Calculations involving inputs called “tonnage models” are just a special case of grade and tonnage models, so they are discussed only briefly.

Method

Conceptual Model

All calculations involving known physical phenomena are based ultimately on a conceptual model that represents both the pertinent physical entities and the relations among those entities. A conceptual model is important because it is the key to mathematically formulating the calculations and to subsequently understanding the calculation results. The conceptual model for the probability calculations is not explicitly stated by Harris (1984, p. 317 and 389) and Singer and Menze (2010, p. 77–158), so the model that is described next is inferred from the cited books.

The conceptual model pertains to a permissive tract, which is defined as a geographic region within which there is some chance of finding mineralized rock that contains economically important mineral resources (Singer and Menze, 2010, p. 105–117). The permissive tract may have any shape and may even comprise noncontiguous geographic regions.

Some of the mineralized rocks in the permissive tract may be profitable to mine. If these potentially profitable, mineralized rocks have not been publicly reported, they are called “undiscovered deposits.” (It is possible that deposits have been discovered, but not publicly reported, by the exploration

company. Such deposits would still be classified as “undiscovered.”) Thus, the conceptual model is a permissive tract in which there are zero or more undiscovered deposits. This conceptual model is the basis for both the previous and the new implementations of the probability calculations.

Data for the Probability Calculations

Background

In the context of the conceptual model, the goal of a mineral resource assessment is to predict both the number of undiscovered mineral deposits and the amount of mineral resources in those deposits. The obvious difficulty is that both quantities are unknown. Nonetheless, Singer and Menze (2010, p. 77–158) developed procedures to estimate these two quantities. The procedures require two datasets that are not part of the conceptual model, and these two datasets are described in the next two sections.

Grade and Tonnage Model

To predict the mineral resource amounts in the undiscovered deposits, it is assumed that the undiscovered deposits are similar to previously discovered deposits. (This assumption, as well as all other assumptions in the probability calculations, are not explicitly stated by Harris [1984] and Singer and Menze [2010]. Nonetheless, all assumptions must be stated, so that the probability calculations can be understood by assessment geoscientists. To emphasize their importance, the text font is bold.)

Discovered deposits are defined as mineralized rocks for which a mineral inventory has been conducted and publicly reported. Although a mineral inventory comprises various different types of information about the mineralized rock (Sinclair and Blackwell, 2002, p. 1–28), only two types are pertinent to the probability calculations. The first type is the mass of the ore. Among economic geologists, mass usually is expressed in terms of metric tons (1,000 kilograms) and is called “tonnage.” Thus, economic geologists typically use the phrase “ore tonnage.” The second type is the concentration, by weight, of the economic mineral resources. Concentration refers to the average concentration for the entire deposit. Among economic geologists, concentration is usually expressed in terms of percent and is called “grade.”

Ore tonnages and the grades of discovered deposits are listed in a database called a “grade and tonnage model.” A portion of the grade and tonnage model that accompanies MapMark4 is shown in table 1. The first and second columns are just identifiers for discovered deposits; they can be ignored because they are irrelevant to the probability calculations. The remaining columns comprise the information that is needed for the probability calculations. For this grade and tonnage model, there are two mineral resources—copper and gold. Although this grade and tonnage model comprises computer-generated values for ore tonnages and mineral resource grades, it is very similar to grade and tonnage models that use actual values.

Most scientists and engineers define a model as a mathematical representation of some phenomenon. For example, Maxwell’s equations are a mathematical model for electromagnetic wave propagation. However, this widely accepted definition does not apply to a grade and tonnage model, which is just a database. Thus, the term “grade and tonnage model” is misleading.

Estimates of the Number of Undiscovered Deposits

The number of undiscovered deposits is fixed, but unknown. **It is assumed that uncertainty about the number of undiscovered deposits can be expressed quantitatively with probability.** Details of the procedure to make these probabilistic statements is presented in Singer and Menzie (2010, p. 132–135). The essential idea is that, during an assessment, each team member estimates the number of undiscovered deposits at three different probability levels. An example of such estimates is shown in table 2. The “Name” column lists either the names of assessment team members or other appropriate identifiers. The “Weight” column lists the weights associated with the estimates, which are described in the last part of this section. The N_{90} column lists, for each member, the estimated number of undiscovered deposits at a probability of 0.90. For example, person 6 estimates that there is a 0.90 probability of finding 3 or more undiscovered deposits in the permissive tract. Among assessment geoscientists,

Table 1. First five entries in the grade and tonnage model that accompanies program MapMark4.

[Cu, copper; Au, gold]

Identifier	Name	Ore tonnage (metric tons)	Cu grade (percent)	Au grade (percent)
DepositID1	DepositName1	1.47×10^8	0.52	4.9×10^{-5}
DepositID2	DepositName2	2.60×10^7	0.30	4.0×10^{-5}
DepositID3	DepositName3	2.21×10^8	0.60	1.7×10^{-5}
DepositID4	DepositName4	9.71×10^7	0.44	3.5×10^{-5}
DepositID5	DepositName5	2.14×10^9	0.59	3.2×10^{-5}

Table 2. Estimates of the number of undiscovered deposits by the 10 members of an assessment team (Zientek, 2014, p. 161). The meaning of the columns is explained in the text.

Name	Weight	N_{90}	N_{50}	N_{10}
Person 1	0.50	3	10	25
Person 2	1.00	2	5	10
Person 3	1.00	5	7	10
Person 4	1.00	2	10	20
Person 5	1.00	1	2	4
Person 6	3.00	3	10	20
Person 7	1.00	5	10	20
Person 8	1.00	3	5	7
Person 9	1.00	1	2	5
Person 10	0.01	10	20	60

this probability is called an “elicitation percentile of 90.” The N_{50} and N_{10} columns are similar and list, for each member, the estimated number of undiscovered deposits at elicitation percentiles of 50 and 10.

For most team members, the weight should be 1.00, which is the standard value. If a team member is an expert, then the weight for that member should be greater than 1.00 (for example, person 6). If the team member is a novice, then the weight for that member should be less than 1.00 (for example, person 1). If the estimates by a team member differ significantly from the estimates by other team members, then the weight should be small (for example, person 10).

Distribution for the Ore Tonnages

Recall that the amount of mineral resources in the undiscovered deposits must be predicted. One aspect of this prediction involves the ore tonnages of the undiscovered deposits. The undiscovered deposits are assumed to be similar to discovered deposits having similar geology, so **the combination of the undiscovered and the discovered deposits is assumed to be the set of all deposits**. In terms of probability, this set is called the “sample space.” There are two cases. First, the discovered deposits are only from the permissive tract that is being assessed. In this case, the sample space comprises only the undiscovered and discovered deposits within this permissive tract. Second, the discovered deposits are from permissive tracts throughout the world. For this case, the sample space comprises the undiscovered and discovered deposits within the permissive tracts throughout the world.

It is assumed that the discovered deposits are a simple random sample from this sample space. Because of this assumption, the ore tonnages of the discovered deposits can be used, in a straightforward way, to estimate a probability density function (pdf) for the ore tonnages in all deposits in the sample space. That is, **it is assumed this pdf represents the ore tonnages of all deposits, which includes the undiscovered deposits, of course. Furthermore, it is assumed that random samples from this pdf are independent of one another.** The task is to estimate this pdf.

To develop this pdf, consider the ore tonnages listed in table 1. For just these five tonnages, the range is two orders of magnitude; for the complete model, the range is three orders of magnitude. Such a large range is common in grade and tonnage models. Consequently, within program MapMark4, the ore tonnages are transformed with the natural logarithm, and the pdf is developed for the transformed ore tonnages. There are two choices for the pdf. One choice is the normal distribution, for which the mean and standard deviation are estimated with the sample mean and sample standard deviation. The other choice is a kernel density estimate (Hastie and others, 2009, p. 208–209; Shalizi, 2016, p. 319–321), for which the kernel is Gaussian. For both choices, random samples are drawn and then are transformed with exponentiation, yielding

random samples of ore tonnage. These random samples implicitly define the probability pdf for the ore tonnage in all deposits, which includes the undiscovered deposits.

An example of a pdf for the ore tonnage in an undiscovered deposit is shown in figure 1. The pdf, which is represented by the histogram, was generated from the complete grade and tonnage model, of which a portion is shown in table 1. The match between this pdf and ore tonnages from the grade and tonnage model is one way to check the estimated pdf; additional ways are described in Ellefsen (2017).

Distribution for the Mineral Resource and Gangue Grades

The argument for developing this distribution is very similar to that for the ore tonnage. Nonetheless, to minimize the chance of misunderstanding, the complete argument for the mineral resource and gangue grades is presented. Recall that the amount of mineral resources in the undiscovered deposits must be predicted. One aspect of this prediction involves the mineral resource and gangue grades of the undiscovered deposits. Because the undiscovered deposits are assumed to be similar to discovered deposits having similar geology, **the combination the undiscovered and the discovered deposits is assumed to be the sample space for all deposits**. This assumption is identical to the corresponding assumption in section “Distribution of Ore Tonnages.”

It is assumed that the discovered deposits are a simple random sample from the sample space. Again, this assumption is identical to the corresponding assumption in section “Distribution of Ore Tonnages.” Because of this assumption, the mineral resource and gangue grades of the discovered deposits can be used to estimate a pdf that applies to all deposits in the sample space. That is, **it is assumed this pdf represents the mineral resource and gangue grades of all deposits, which includes the undiscovered deposits, of course. Furthermore, it is assumed that random samples from this pdf are independent of one another.** The task is to estimate this pdf.

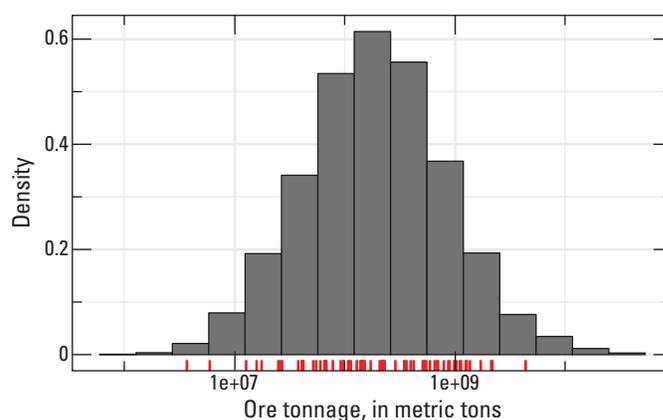


Figure 1. Probability density function (histogram) that represents the ore tonnage in an undiscovered deposit. The red vertical lines at the bottom represent the ore tonnages from the grade and tonnage model.

4 Probability Calculations for Three-Part Mineral Resource Assessments

To develop this pdf within program MapMark4, the gangue grade is calculated for each discovered deposit in the grade and tonnage model. It is simply 100 minus the sum of the mineral resource grades. The combined mineral resource and gangue grades are defined in a mathematical entity called the simplex. Defining pdfs within the simplex is difficult, so the combined grades undergo the isometric log-ratio transform (Pawlowsky-Glahn and others, 2015, p. 37), and the transformed grades are now in a real vector space, in which standard statistical methods can be applied. Thus, a standard pdf is now developed for these transformed grades; there are two choices for this pdf. One choice is the multivariate normal distribution, for which the mean vector and covariance matrix are estimated with the sample mean vector and sample covariance matrix. The other choice is a multivariate kernel density estimate (Hastie and others, 2009, p. 208–209; Shalizi, 2016, p. 319–321), for which the kernel is Gaussian. For both choices, random samples are drawn and then undergo the inverse isometric log-ratio transformation (Pawlowsky-Glahn

and others, 2015, p. 37), yielding random samples of mineral resource and gangue grades. These random samples implicitly define the probability distribution for the mineral resource and gangue grades in all deposits, which includes the undiscovered deposits.

An example of a pdf for the mineral resource and gangue grades is generated from the complete grade and tonnage model, of which a portion is shown in table 1. It is important to remember that the grades are defined in a simplex of dimension 3, which corresponds to the copper, gold, and gangue grades. Analyzing a pdf that is plotted within the simplex is often difficult. Instead, the log-ratios of the pdf grades are plotted as histograms. In this case, there are log-ratios for the copper and gold grades, for the copper and gangue grades, and for the gold and gangue grades; so there are three histograms (fig. 2). The match between the histograms and the corresponding log-ratios from the grade and tonnage model is one way to check the pdf, additional ways are described in Ellefsen (2017).

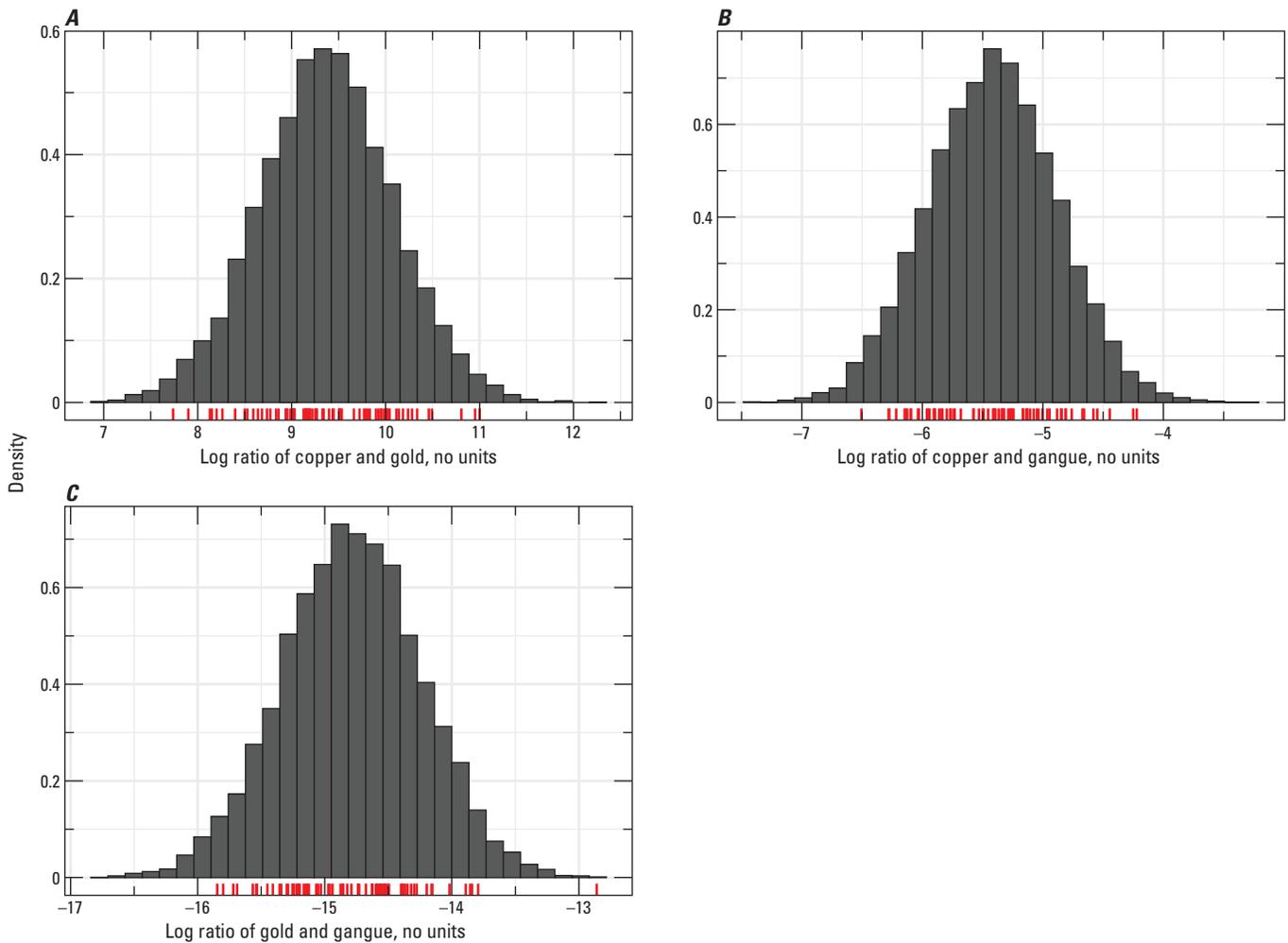


Figure 2. A, B, and C, Histograms of the log-ratios of the grades from the joint probability density function that represents the grades in an undiscovered deposit. The red vertical lines at the bottom represent the grades from the grade and tonnage model.

Distribution for the Number of Undiscovered Deposits

Recall that the number of undiscovered deposits must be predicted. Although estimates of this number have been made by the assessment team members (for example, table 2), these estimates are not in a form that is suitable for the probability calculations. Rather, the estimates must be recast as a probability mass function (pmf). **It is assumed that this pmf characterizes the number of undiscovered deposits in the permissive tract. Furthermore, it is assumed that the random samples of the number of undiscovered deposits are independent of one another.**

The pmf should satisfy at least two criteria. First, it should fit, as well as possible, the estimates of the number of undiscovered deposits that are generated by the assessment team members (for example, table 2). In other words, this criterion is to avoid underfitting the estimates. Second, it should not fit the estimates perfectly because the estimates are erratic. In other words, this criterion is to avoid overfitting the estimates.

A way to satisfy these two criteria is to pick a smoothly varying pmf that can be easily adjusted to have different shapes. The smoothness ensures that the problem of overfitting is minimized, and the variety of different shapes ensures that the problem of underfitting is minimized. Thus, within program MapMark4, the chosen pmf is the negative binomial. The task is to estimate the parameters of this pmf such that it fits, as well as possible, the estimates of the number of undiscovered deposits.

The best-fitting pmf can be estimated with optimization. The cost function c for the optimization is chosen to be:

$$c = \sum_{i=1}^M w_i \sum_{j=1}^3 |N_{ji} - N_j(\mu, \sigma)|^r. \quad (1)$$

Term N_{ji} is the estimated number of undiscovered deposits. Index j refers to the three elicitation percentiles, 90, 50, and 10; index i refers to an assessment team member. Hence term N_{ji} is the data from table 1, for example. Term $N_j(\mu, \sigma)$ is the number of predicted deposits for the percentile indexed by j . This prediction is calculated from the negative binomial pdf for which the mean is μ and the standard deviation is σ . The difference $N_{ji} - N_j(\mu, \sigma)$ represents the mismatch between the data and the prediction. The absolute value of the mismatch is raised to the power r , which is usually set to 1 so that optimization is robust. This quantity is summed over all three elicitation percentiles, and the scaled by weight w_i , which is the weight for team member i from table 1, for example. This scaled sum, which pertains to one team member, is then summed over all team members. The final sum is the cost function, which collectively represents the mismatches between all data and all predictions.

The optimal pmf is that which minimizes the mismatches between all data and all predictions—namely, minimizes the cost function. This minimization requires two steps. First, the cost function is calculated over a dense grid of points representing different values of the mean μ and the standard deviation σ . For this grid, the minimum and maximum values of μ are the weighted means of columns N_{90} and N_{10} , respectively. The minimum value of σ is $\sqrt{\mu}$, which is the smallest possible value of the standard deviation for a negative binomial pmf. The maximum value of σ is the difference between the weighted means of columns N_{90} and N_{10} . The value of the cost function c is examined at each grid point. At that grid point with the minimum value of c , the mismatches between all data and all predictions are almost as small as they can possibly be. Thus, the pmf that generates the predictions is close to the optimal pmf. The second step in the minimization is to determine the optimal pmf. The mean μ and the standard deviation σ that are associated with the minimum-value grid point are the starting values for the optimization algorithm (Nelder and Mead, 1965) that determines the optimal pmf. The result of the two-step minimization is an estimated pmf that is optimal in the sense specified by the cost function.

Equation 1 shows the importance of the weights. Large weights make the cost function larger than it would be otherwise. Conversely, small weights make the cost function smaller than it would be otherwise. Thus, the minimization of the cost function is more affected by those mismatches with large weights than by those mismatches with small weights. Hence, the optimal pmf is more affected by the experts than it is by both the novices and those team members making anomalous estimates.

An example of a pmf for the number of undiscovered deposits is shown in figure 3A. This pmf was optimized to fit the estimated numbers listed in table 2. To evaluate the fit, the pmf is recast as elicitation percentiles, which are plotted in figure 3B along with the estimated numbers from table 2. It is important to discount the estimated numbers from person 10, which correspond to the rightmost circle at each elicitation percentile in figure 3B. The recast pmf passes through the clusters of red circles at all three elicitation percentiles. The recast pmf exactly fits 4 estimates at elicitation percentile 50 and 3 estimates at elicitation percentile 10. Thus, this pmf best represents the estimates of the number of undiscovered deposits, in the sense specified by the cost function (eq. 1).

Simulation of the Undiscovered Deposits

Recall that the goal of the mineral resource assessment is to predict both the number of undiscovered deposits and the amount of mineral resources in those undiscovered deposits. This prediction is accomplished by simulation, using random samples from the pdf for the ore tonnages, the pdf for the mineral resource and gangue grades, and the pmf for the number of undiscovered deposits. **It is assumed that the ore tonnages,**

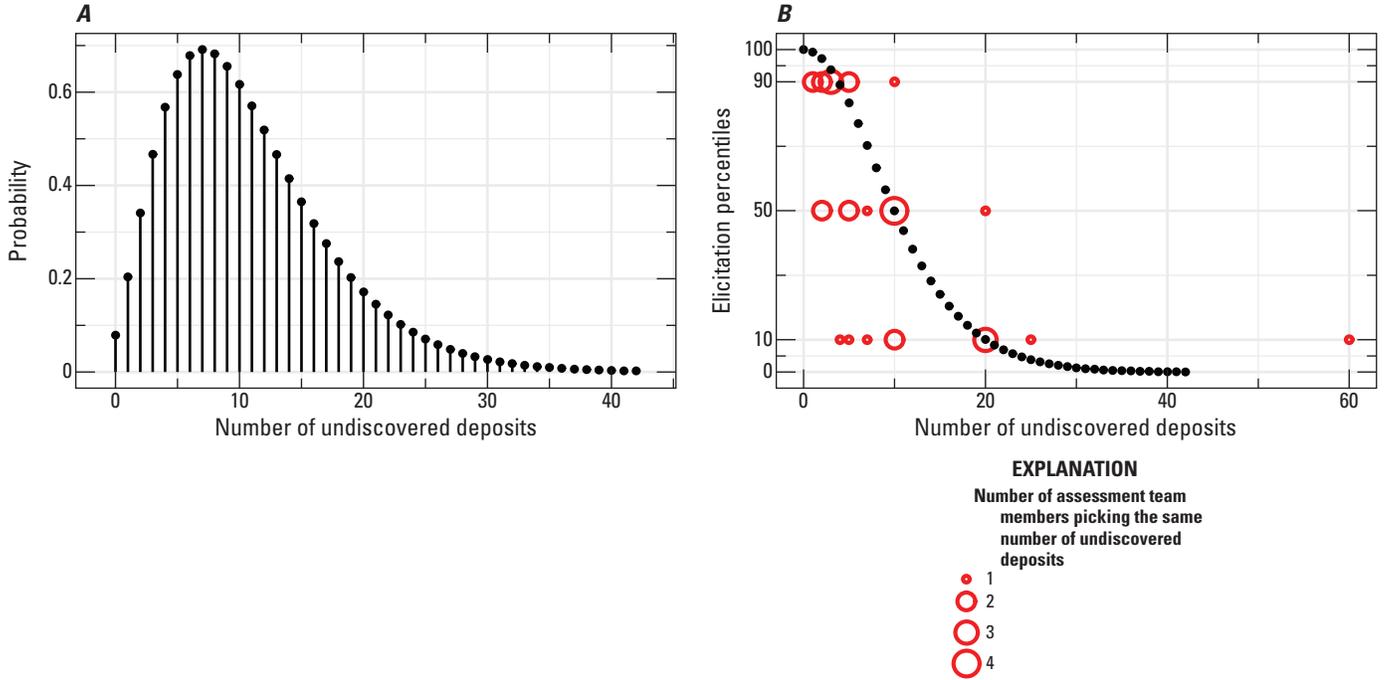


Figure 3. *A*, Negative binomial probability mass function representing the number of undiscovered deposits in the permissive tract. *B*, Probability mass function recast as elicitation percentiles (black dots) and compared to the estimated numbers of undiscovered deposits (red circles). The size of a red circle indicates how many assessment team members picked the same number of undiscovered deposits.

the mineral resource and gangue grades, and the number of undiscovered deposits are probabilistically independent of one another. Thus, random samples of ore tonnage are not coupled to random samples of mineral resource and gangue grades, which, in turn, are not coupled to random samples of the number of undiscovered deposits. Consequently, simulation with these random samples is somewhat simple.

A single simulation is performed on two levels. The first level is that of the permissive tract. A random sample N is drawn from the pmf for the number of undiscovered deposits; it represents the number of simulated undiscovered deposits within in the permissive tract. The second level of the simulation is that of the undiscovered deposit. For each of the N undiscovered deposits, a random sample of ore tonnage is drawn from its pdf, and a random sample of mineral resource and gangue grades is drawn from its pdf. These two random samples characterize the simulated undiscovered deposit. (There is a special case. If $N = 0$, then no random samples of ore tonnage and grades are drawn.) Many such simulations are performed, and the appropriate number of simulations is considered in the Discussion section.

Distribution for Total Ore and Mineral Resource Tonnage

Assessment geoscientists usually desire a summary of the simulated mineral resource amounts. To this end, the assessment geoscientist use summary statistics of total ore

and mineral resource tonnages in all simulated undiscovered deposits within the permissive tract. To understand how these total tonnages are calculated, recall how a single simulation is performed. For each of the N simulated undiscovered deposits, random sample of ore tonnage in simulated undiscovered deposit i is represented by O_i , and the grade of mineral resource j in simulated undiscovered deposit i is represented by G_{ij} . Index i ranges from 1 to N , and index j ranges from 1 to J , which equals the number of mineral resources. Gangue is omitted because gangue tonnage is not of interest to the assessment geoscientists. A mineral resource tonnage is the product of the ore tonnage and the mineral resource grade. Thus, the ore and mineral resource tonnages in simulated undiscovered deposit i are represented by random vector W_i :

$$W_i = (O_i, O_i G_{i1}, O_i G_{i2}, \dots, O_i G_{ij}) \quad (2)$$

The sum of the ore and mineral resource tonnages in the N simulated undiscovered deposits equals the total ore and mineral resource tonnages in the permissive tract, which is represented by random vector T :

$$T = W_1 + W_2 \dots + W_N, \quad (3)$$

(Harris and Carrigan, 1981). Suppose that 20,000 total simulations are performed. Then there are 20,000 samples of random vector T . Collectively, these 20,000 samples implicitly define a multivariate pdf for random vector T .

As an example, 20,000 simulations are performed using the pdf for ore tonnage (fig. 1), the pdf for mineral resource grades (fig. 2), and the pmf for the number of undiscovered deposits (fig. 3A). The simulations are summarized by the total ore and mineral resource tonnage—the multivariate pdf for T comprises the total ore tonnage, the total copper tonnage, and the total gold tonnage. One way to analyze this multivariate pdf is to plot just the univariate, marginal pdf (fig. 4A). In the upper left corner of the plot, the probability of zero tonnage is printed. This probability will be nonzero whenever the pmf for the number of undiscovered deposits has a nonzero probability for zero deposits (fig. 3). The univariate, marginal complementary cumulative distribution functions are shown in figure 4B. The upper asymptote for these functions equals one minus the probability for zero deposits. Another way to analyze this multivariate pdf is to plot each of the univariate and bivariate marginal pdfs (Ellefsen, 2017).

Mean and Variance of the Total Ore and Mineral Resource Tonnage

The mean and the variance of T can be calculated with analytic formulas. These two formula-calculated statistics are important for at least two reasons. They are used to check both the simulation and the distribution of total ore and mineral resource tonnages (section “Checking the final calculations”), and they are needed for sensitivity analyses (section “Sensitivity Analysis”).

The formula for the mean is most easily derived using conditional expectation. The equation for conditional expectation, expressed in terms of the previously presented variables, is $\mathbb{E}(T) = \mathbb{E}[\mathbb{E}(T|N)]$ (Grimmett and Stirzaker, 2001, p. 105). Using equation 3, $\mathbb{E}(T|N) = N\mu_w$ in which $\mu_w = \mathbb{E}(W)$. Consequently,

$$\mathbb{E}(T) = \mu_N \mu_w, \quad (4)$$

in which $\mu_N = \mathbb{E}(N)$. Equation 4 states that the mean of the total ore and mineral resource tonnages in the permissive tract is a product of two quantities: the mean number of simulated undiscovered deposits and the mean ore and mineral resource tonnages in a simulated undiscovered deposit. This simple result is intuitive.

The formula for the variance is most easily derived using conditional variance. The equation for conditional variance, expressed in terms of the previously presented variables, is $\text{var}(T) = \mathbb{E}[\text{var}(T|N)] + \text{var}[\mathbb{E}(T|N)]$ (Grimmett and Stirzaker, 2001, p. 69). To derive the first term on the right side, use equation 3: $\text{var}(T|N) = N\Sigma_w$ in which $\Sigma_w = \text{var}(W)$. (Recall that W_i and W_k ($i \neq k$) are independent of one another.) Consequently, $\mathbb{E}[\text{var}(T|N)] = \mu_N \Sigma_w$. To derive the second term on the right side, use the previously presented expression for $\mathbb{E}(T|N)$: $\text{var}[\mathbb{E}(T|N)] = \text{var}[N\mu_w] = \sigma_N^2 \mu_w \mu_w'$ for which $\sigma_N^2 = \text{var}(N)$ and the symbol $'$ indicates transpose. Using these derivations for the first and second terms, the variance is:

$$\text{var}(T) = \mu_N \Sigma_w + \sigma_N^2 \mu_w \mu_w'. \quad (5)$$

Equation 5 shows that the variance of the total ore and mineral resource tonnage in the permissive tract is a sum of two scaled variances. The variance of the ore and mineral resource tonnage in an undiscovered deposit Σ and the variance of the number of undiscovered deposits σ_N^2 .

Equations 4 and 5 involve four parameters that must be calculated. The mean number of undiscovered deposits μ_N and the associated variance σ_N^2 are calculated directly from the pmf. The two statistics for W are calculated from its random samples. Recall that both the pdf for the ore tonnage and the pdf for the grades are defined implicitly by random samples.

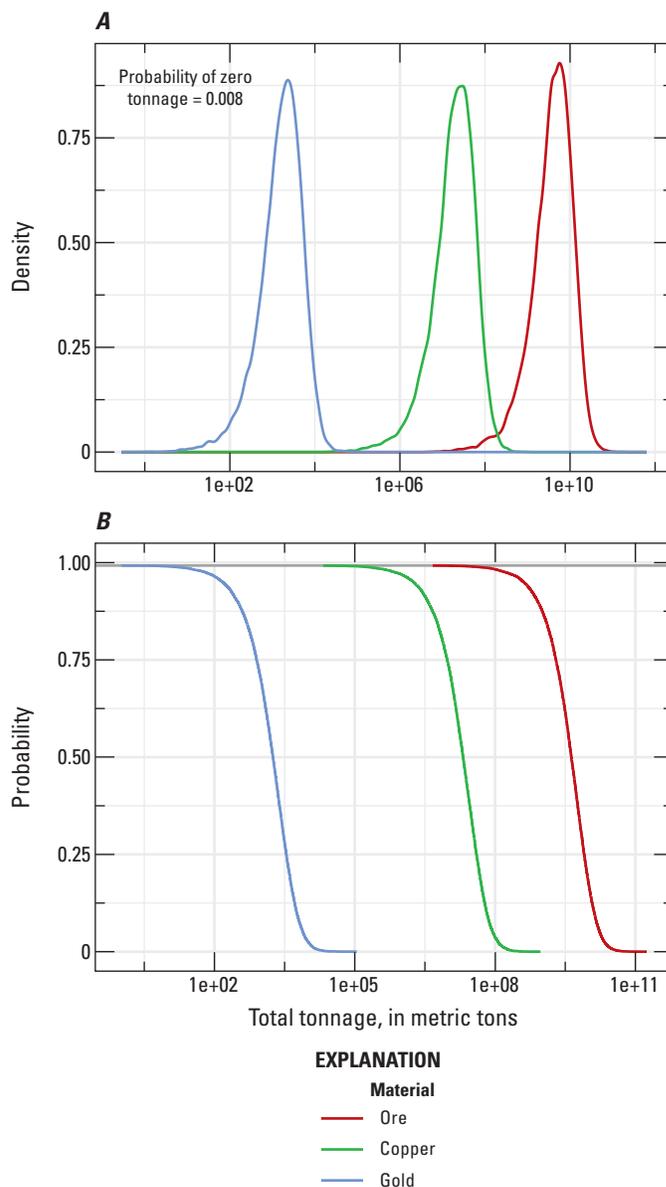


Figure 4. A, Univariate, marginal, probability density functions and B, univariate, marginal, complementary cumulative distribution functions for the total ore and mineral resource tonnages in all undiscovered deposits within the permissive tract.

These random samples are used in equation 2 to generate random samples of W . The mean of the random samples of W is μ_w , which is a random vector. The variance of the random samples of W is Σ_w , which is a random matrix. Consequently, $\mathbb{E}(T)$ is a random vector, and $\text{var}(T)$ is a random matrix.

Checking the Final Calculations

The first part of the probability calculations—the calculation of the pdf for the ore tonnage, the pdf for the grades, and the pmf for the number of undiscovered deposits—includes various graphical checks to ensure that these calculations are properly performed. However, a check of the final calculations—the simulation of undiscovered deposits and the calculation of the pdf for total ore and mineral resource tonnage—has not been presented. Such a check is needed to ensure that the mineral resource predictions are properly calculated.

The key to checking the final calculations is the analytic formulas for the mean and the variance of the total ore and mineral resource tonnage in the permissive tract. Recall that these formulas use statistics derived from the pdf for the ore tonnage, the pdf for the grades, and the pmf for the number of undiscovered deposits. Thus, the formulas are independent of the final calculations. Consequently, the mean and the variance from the analytic formulas are compared to the mean and the variance estimated from the pdf for total ore and mineral resource tonnage, and this comparison constitutes the check of the final calculations.

The mean calculated from the analytic formula and the mean calculated from the pdf are random vectors. Likewise, the variance calculated from the analytic formula and variance calculated from the pdf are random matrices. Samples of these statistics are generated by performing the probability calculations many times. As an example, the probability calculations presented in section “Method” are repeated 350 times, yielding 350 samples of each statistic. Comparison of the variance matrices is facilitated by decomposing each matrix into a standard deviation vector and a correlation matrix. Because a correlation matrix is symmetric, only the elements in its upper triangle are needed for the comparison.

Consider, for example, the mean ore tonnage. There are 350 samples of this statistic calculated with the analytic formula and 350 corresponding samples of this statistic calculated from the pdf. To determine their similarity, each formula-calculated sample is subtracted from the corresponding pdf-calculated sample; the difference is divided by the formula-calculated sample and multiplied by 100 to yield the relative difference in percent. The relative differences for all 350 samples are plotted as a histogram (fig. 5A). The same procedure is used for the other statistics (figs. 5B–I). For all nine statistics, the relative differences are centered at zero—the interpretation of this result is that the final calculations are correct.

Readers will observe that the spreads for the standard deviations (figs. 5D, E, and F) are much larger than the spreads for the other statistics (figs. 5A, B, C, G, H, and I). The reason is that standard deviations are harder to estimate than

the other statistics are. Readers will also observe that there are a few outliers in the relative differences for the standard deviations and the correlation coefficients (figs. 5D–I). The likely reason for these outliers is that the simulation includes some random samples of ore tonnage from the long right tail of its pdf, skewing the statistics. Thus, these outliers are not errors, and the probability calculations are performed properly.

The computation of the histograms in figure 5 requires many hours, so it is impractical to compute such distributions for routine probability calculations. Nonetheless, it is practical to compare corresponding pdf statistics and formula-calculated statistics for a single probability calculation. Such a comparison provides a quick check for final probability calculations. It is important to remember that, although the difference for a standard deviation or a correlation coefficient might be large in rare cases (fig. 5D–I), the final calculations are still performed correctly.

Calculations Summary

The probability calculations are somewhat complex, so the essential ideas underlying the probability calculations are summarized. They require five steps. The first step is estimating the pdf that represents the ore tonnage in an undiscovered deposit; the estimation uses information from the grade and tonnage model. The second step is estimating the pdf that represents the mineral resource and gangue grades in an undiscovered deposit; again, the estimation uses information from the grade and tonnage model. The third step is estimating the pmf that represents the number of undiscovered deposits in the permissive tract; the estimation uses information generated by the assessment team members. The fourth step is simulation—the information from the two pdfs and the pmf are combined to predict, in a probabilistic way, the ore tonnages and mineral resource grades in the undiscovered deposits within the permissive tract. The fifth and final step is to summarize the simulation by calculating the total ore and mineral resource tonnages in the undiscovered deposits within the permissive tract.

Properties of the Probability Calculations

Precision

Recall that the mean vector, the standard deviation vector, and the correlation matrix that are calculated from the pdf for the total ore and mineral resource tonnage are random quantities. Thus, the vector and matrix elements are random quantities that have distributions. The spreads of these distributions indicate the precision of the probability calculations. That is, if the spreads are small, then the calculations are precise. My anecdotal observation is that the precision is strongly affected by the number of simulations. Thus, the relation between precision and the number of simulations is investigated.

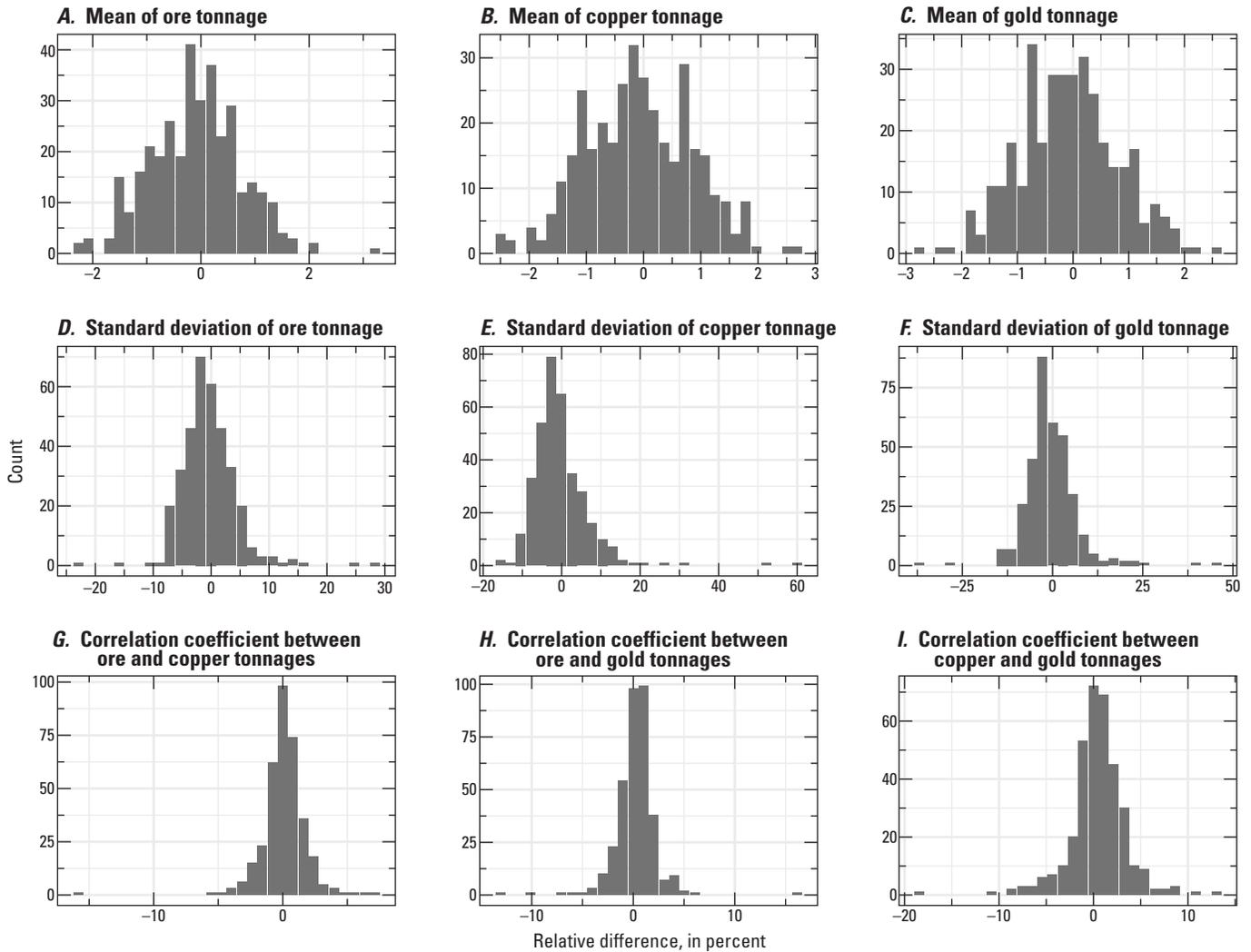


Figure 5. A–I, Check of the final calculations. Each histogram represents the relative difference between a statistic calculated from the pdf of total ore and mineral resource tonnage and the corresponding statistic calculated with an analytic formula.

Samples of the mean vector, the standard deviation vector, and the correlation matrix are generated by performing the probability calculations many times. As an example, the probability calculations presented in the Method section of this report are repeated 350 times, and the number of simulations for each repetition is 2,500. The result of the repeated calculations is 350 samples of each statistic, which constitute a distribution for that statistic. The spread of the distribution is quantified by its standard deviation, which is called “simulation standard error” in the context of a Monte Carlo simulation (DeGroot and Schervish, 2002, p. 704). This computation is repeated five times, and the five simulation standard errors for each statistic are summarized by the median. This procedure is repeated for 5,000; 10,000; 20,000; and 40,000 simulations.

The relation between the simulation standard error and the number of simulations is presented in figure 6. For all nine statistics, the simulation standard error decreases nonlinearly as the number of simulations increases. (The trend is somewhat irregular for the three standard deviations, [fig. 6D, E,

and F] because these estimates of the simulation standard error are especially noisy.) Thus, the precision of the calculations increases as the number of simulations increase.

Computation Time

Computation time is important because it is the time that an assessment geoscientist must wait to get the results of the probability calculations. Consequently, the computation time is investigated for two input parameters to the probability calculations—the pmf and the number of simulations. For this investigation, the probability calculations presented in the Method section of this report are timed for 2,500 simulations. There are 10 repetitions of these calculations because computation time varies somewhat, and the 10 computation times are summarized by their median. This procedure is repeated for 5,000; 10,000; 20,000; and 40,000 simulations. These times pertain to one pmf, which is called “Pmf 1” and is generated from the data in table 1.

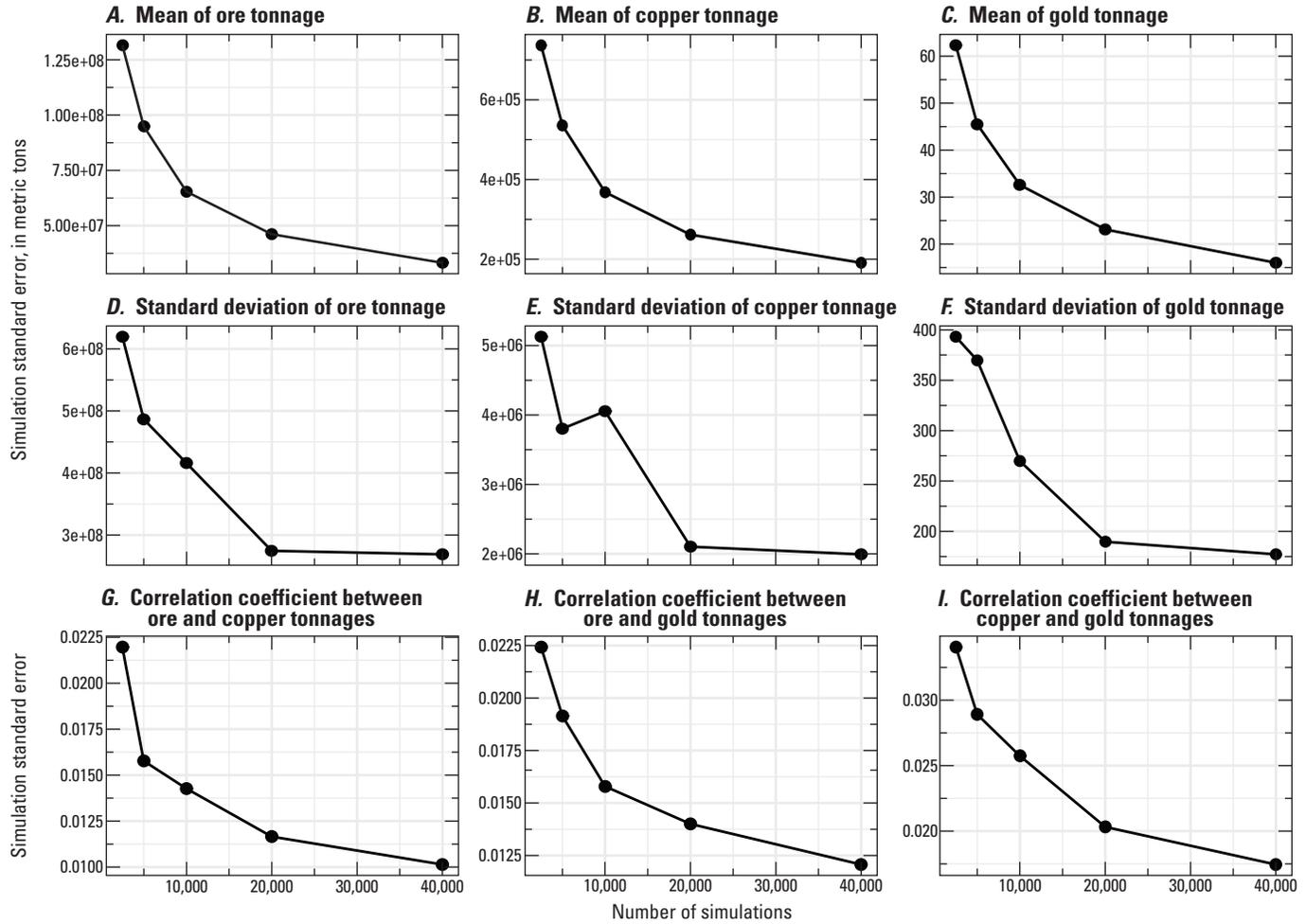


Figure 6. A–I, Effect of the number of simulations on the simulation standard error.

Within software package MapMark4 (Ellefsen, 2017), there are three other datasets like that in table 1, and these datasets are used to generate pmfs that are called “Pmf 2,” “Pmf 3,” and “Pmf 4.” The computation times for these three pmfs are calculated using the previously described procedure. The results for all four pmfs are summarized in figure 7A. Except for Pmfs 1 and 4 at 2,500 simulations, the graph shows that as the number of simulations increases, the computation time increases. The reason for the exception is that other factors have a greater effect on the computation time than the 2,500 simulations do.

The graph in figure 7A also shows that the computation times differ for the four pmfs. To understand this phenomenon, recall how a single simulation is performed. A random sample N is drawn from the pmf for the number of undiscovered deposits; for each of the N undiscovered deposits, a random sample of ore tonnage is drawn from its pdf, and a random sample of mineral resource and gangue grades is drawn from its pdf. Thus, as N increases, the computation time should increase.

A statistic that summarizes the possible values of the N is the mean of the pmf, and the means for Pmfs 1, 2, 3, and 4 are 10.7, 112, 6.36, and 1.11, respectively. The computation time as a function of this mean is plotted in figure 7B for different numbers of simulations. Except for 2,500 simulations, the graph shows that, as the mean increases, the computation time tends to increase. (Other graphs based on the variance, the mode and the maximum of the pmf show similar trends.) The bend in each line indicates that other factors affect the computation time—the mean only partially explains the computation time.

Sensitivity Analysis

It is important for the assessment geoscientists to understand how the inputs to the probability calculations affect the outputs. Determining these effects usually is called a “sensitivity analysis.” A simple way to conduct a sensitivity analysis involves the analytic formulas for the total ore and mineral resource tonnage (eqs. 4 and 5). The four parameters in these equations are related to the inputs to the probability

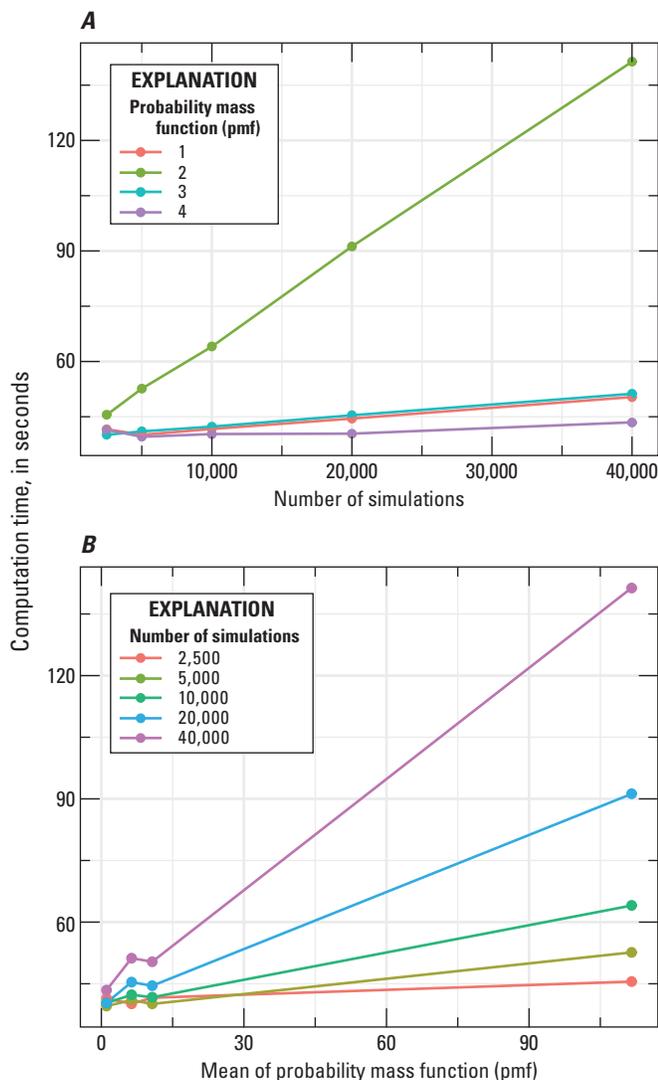


Figure 7. A, Effect of the number of simulations on the computation time, for four probability mass functions (pmf). B, Effect of the mean of the pmf on the computation time, for five numbers of simulations.

calculations, so four sensitivity analyses could be conducted. However, such a comprehensive analysis is beyond the scope of this report. Instead, because some assessment geoscientists believe that determining an appropriate pmf is difficult, a sensitivity analysis is conducted for its mean μ_N .

In equation 4, the mean μ_N is changed by amount $\delta\mu_N$, for which the lower bound is $-\mu_N \leq \delta\mu_N$. Consequently, the mean $\mathbb{E}(T)$ change by amount $\delta\mathbb{E}(T)$, for which the lower bound is $\mathbb{E}(T) \leq \delta\mathbb{E}(T)$. The relation incorporating these changes is $\mathbb{E}(T) + \delta\mathbb{E}(T) = (\mu_N + \delta\mu_N)\mu_W$. Equation 4 is subtracted from this equation, and the result is $\delta\mathbb{E}(T) = \delta\mu_N\mu_W$. Instead of using the random vectors $\delta\mathbb{E}(T)$ and μ_W for the sensitivity analysis, an individual vector element, which is referenced by index k , is used. Thus, the equation $\delta\mathbb{E}(T_k) = \delta\mu_N\mu_{Wk}$ contains only scalar quantities. This equation is divided by the corresponding scalar version of equation 4, and the result is:

$$\frac{\delta\mathbb{E}(T_k)}{\mathbb{E}(T_k)} = \frac{\delta\mu_N}{\mu_N}. \tag{6}$$

Because of the lower bounds for $\delta\mu_N$ and $\delta\mathbb{E}(T)$, the ratios on both sides of equation 6 have a lower bound of -1 .

The left side of equation 6 is interpreted as the relative change in the mean of a total ore or mineral resource tonnage, and the right side as the relative change in the mean of the number of undiscovered deposits. Clearly, the relative changes are equal. Thus, if the mean of the number of undiscovered deposits increases by, say, 50 percent, then the mean of a total ore or mineral resource tonnage will increase by 50 percent. This simple relation shows how an input to the probability calculations affects an output. The importance of this relation is most obvious, if the change in the input is an error in the pmf, which would be caused by an error in the estimates of the number undiscovered deposits that are provided by the assessment team members.

Discussion

Suitability of the Assumptions

The accuracy of a mineral resource prediction, which is conducted with the three-part method of mineral resource assessment, depends upon the suitability of the assumptions. That is, if the assumptions are unsuitable, then the mineral resource prediction will be inaccurate. The suitability of several assumptions, which are presented in section “Method,” require some discussion.

It is assumed that the undiscovered deposits are similar to previously discovered deposits. Similarity includes the geology, the mining technology, and the ore processing technology. If these factors change significantly, then the assumption may be unsuitable and the mineral resource predictions inaccurate.

It is assumed that the discovered deposits are a simple random sample from the sample space that comprises both the undiscovered and the discovered deposits. If the discovered deposits are a simple random sample, then the discovered deposit depends neither upon its ore tonnage nor its mineral resource and gangue grades. However, in some permissive tracts, large deposits tend to be discovered first (Chung and others, 1992; Stanley, 1992; Long, 1995). In such permissive tracts, the assumption is unsuitable (Bultman and Gettings, 1996). If the unsuitability is ignored and if the discovered deposits from the permissive tract are used for the grade and tonnage model, then the mineral resource prediction will be too high. Thus, this assumption must be carefully considered during an assessment.

It is assumed that random samples from the pdf that represents the ore tonnages of all deposits are independent of one another. In geologic terms, the implication is that the ore tonnage in an undiscovered deposit is independent, and

hence uncorrelated, with the ore tonnage in another undiscovered deposit, even if the two deposits are close to one another and are generated by the same physical processes. The corresponding assumption for the mineral resource and gangue grades has a similar implication. If proximity is important to the ore-forming process, then these two assumptions may be unsuitable.

It is assumed that the ore tonnages, the mineral resource and gangue grades, and the number of undiscovered deposits are probabilistically independent of one another. Independence of the tonnages, the grades, and the number of undiscovered deposits means that there is no relationship or coupling between them. However, for some mineral resources, tonnages and grades are coupled (Harris, 1984, p. 253–264). If the coupling is a negative correlation—as the tonnage increases, the grades tend to decrease—then neglecting the coupling will cause the mineral resource predictions to be too high. Thus, this assumption must be carefully considered during an assessment.

Tonnage Models

To understand a tonnage model, it is helpful to recall the grade and tonnage model that is described in section “Grade and Tonnage Model.” Each entry in a grade and tonnage model includes a deposit identifier, a deposit name, the ore tonnage, the grade of each mineral resource. However, a grade and tonnage model is inappropriate if the deposits are not characterized by ore tonnage and grade, but only by mineral resource tonnage. For such situations, a tonnage model is appropriate. Each entry in a tonnage model includes a deposit identifier, a deposit name, and the mineral resource tonnage. An example of a tonnage model is in Ellefsen (2017). For the current implementation of software package MapMark4, a tonnage model comprises only one mineral resource.

To perform probability calculations for a tonnage model, a pdf for the mineral resource tonnages in the undiscovered deposits must be developed. The argument for developing this pdf is identical to that in section “Distribution of the Ore Tonnages.” In this sense, the mineral resource tonnage for a tonnage model is the exact analog of the ore tonnage for a grade and tonnage model. Thus, all assumptions in section “Distribution of the Ore Tonnages” apply to the probability calculations for the tonnage model.

The next two steps in the probability calculations—developing a pmf for the number of undiscovered deposits and simulating the undiscovered deposits in the permissive tract—are identical to those in sections “Distribution of the Number of Undiscovered Deposits” and “Simulation of the Undiscovered Deposits.” Thus, all assumptions in those sections apply to the probability calculations for the tonnage model.

The final step in the probability calculations is generating the distribution for the total mineral resource tonnages in the undiscovered deposits in the permissive tract. This step requires only a slightly different interpretation of the random variables

in equation 2. Random variable O_i now represents the mineral resource tonnage in undiscovered deposit i , and random vector W_i now comprises just a single element, O_i . With this different interpretation of the variables, the total mineral resource tonnage is calculated with equation 3.

Every step of the probability calculations for a grade and tonnage model, except the step for the developing a pdf for the grades, is identical, or almost identical, to that for a tonnage model. To account for the few differences, a few modifications were added to software package MapMark4. Thus, one software package can be used for both models.

Number of Simulations

An important parameter in the probability calculations is the number of simulations. The choice of a suitable value for this parameter is influenced by two properties. As the number of simulations increases, the simulation standard error decreases (fig. 6) but the computation time increases (fig. 7). I am unaware of any simple rule for selecting a suitable number of simulations based on these two properties. Nonetheless, my experience is that roughly 20,000 simulations is a reasonable balance between a relatively small standard error and a relatively small computation time.

Conclusions

New Features

The new implementation of the probability calculations includes many new features. These features can be grouped into the two broad categories. The first category pertains to the implementation of the probability calculations, and within this category here are four, significant new features. First, the pdf for the ore tonnages may be generated with a kernel density estimate, which is a simple, well-established method for nonparametric distributions. Second, the pdf for the mineral resource and gangue grades is generated using the methods for compositional data analysis. It replaces the ad hoc method in the earlier versions of the software—a method that I have been unable to reproduce. The pdf is generated using either a multivariate normal distribution or a kernel-density estimate. Third, the pmf for the number of undiscovered deposits is a negative binomial distribution that is estimated directly from the elicitation percentiles provided by the assessment team members. It replaces the need for consensus elicitation percentiles from the assessment team. In addition, it replaces the geologically implausible pmf in earlier versions of the software—a thorough discussion of the reasons for geological implausibility are beyond the scope of this report. Fourth, the new implementation can perform probability calculations for a tonnage model.

The second category of new features pertains to checks of the probability calculations, and within this category there are four, significant new features. First, the pdf for the ore tonnages can now be checked with various plots and tables, to ensure that it adequately represents the ore tonnages in the grade and tonnage model. Second, the pdf for the mineral resource and gangue grades can be similarly checked. Third, the pmf for the number of undiscovered deposits can be checked to ensure that it adequately represents the elicitation percentiles provided by the assessment team members. Fourth, the deposit simulation and the distribution for total ore and mineral resource tonnage can be checked by comparing various statistics. This check ensures that both the deposit simulation and the distribution for total tonnage are correct.

New Understanding

This report presents the assumptions that are inherent in the probability calculations. In addition, this report discusses the suitability of selected assumptions and the constraints that they place on mineral resource assessments. Mineral resource assessment geoscientists must ensure that these assumptions are appropriate for their assessment area; otherwise, their mineral resource predictions will be inaccurate.

New Findings

This report presents analytic formulas for the mean and variance of the total ore and mineral resource tonnage. These formulas are needed to check the deposit simulations and the distribution of the total ore and mineral resource tonnages. In addition, these formulas are used to conduct a sensitivity analysis.

Three properties of the probability calculations are investigated. First, the precision of the probability calculations increases as the number of simulations increases. Second, the computation time increases as the number of simulations increases, and the computation times increases as the mean number of undiscovered deposits increases. Finally, a sensitivity analysis shows that the relative change in the mean of the number of undiscovered deposits equals the relative change in the mean of the total ore or mineral resource tonnage.

Acknowledgments

This work was funded by the Development of Assessment Techniques and Analysis Project within the Mineral Resources Program of the U.S. Geological Survey. J.D. Phillips provided guidance and feedback during frequent discussions on mineral resource assessments. D.V. Smith and M.A. Kass suggested improvements to this manuscript.

Software and Reproducibility

The calculations and the figures are generated with the R language scripts that are stored in file “CalculationScripts.R.” Readers can execute these scripts to reproduce and check the calculations and figures. Because of the random nature of the calculations, the reproduced results will closely, but not exactly, match the presented results.

References Cited

- Bawiec, W.J., and Spanski, G.T., 2012, Quick-start guide for version 3.0 of EMINERS—Economic mineral resource simulator: U.S. Geological Survey Open-File Report 2009–1057, 26 p., accessed August 2012, at <http://pubs.usgs.gov/of/2009/1057>.
- Brown, P.J., Friedel, M.J., 2011a, Test results comparing the new U.S. Geological Survey quantitative resource estimation simulation software to MARK3 *in* Brady, S.R., ed., 2011, Proceedings of the third USGS modeling conference, June 7–11, 2010, Broomfield, Colorado—Understanding and predicting for a changing world: U.S. Geological Survey Scientific Investigations Report 2011–5147, 86 p., accessed August 2012, at <http://pubs.usgs.gov/sir/2011/5147>.
- Brown, P.J., Friedel, M.J., 2011b, Overview of new Monte Carlo software for quantitative mineral resource estimation *in* Brady, S.R., ed., 2011, Proceedings of the third USGS modeling conference, June 7–11, 2010, Broomfield, Colorado—Understanding and predicting for a changing world: U.S. Geological Survey Scientific Investigations Report 2011–5147, 86 p., accessed August 2012, at <http://pubs.usgs.gov/sir/2011/5147>.
- Bultman, M.W., and Gettings, M.E., 1996, Quantitative mineral resource assessment of Coronado National Forest *in* du Bray, E.A., ed., Mineral resource potential and geology of Coronado National Forest, southeastern Arizona and southwestern New Mexico, U.S. Geological Survey Bulletin 2083 A–K, p. 183–202.
- Chung, C.F., Singer, D.A., and Menzie, W.D., 1992, Predicting sizes of undiscovered mineral deposits—An example using mercury deposits in California: *Economic Geology*, v. 87, p. 1174–1179.
- DeGroot, M.H., and Schervish, M.J., 2002, Probability and statistics (3d ed.): New York, Addison-Wesley, 816 p.

- Duval, J.S., 2001, A Microsoft Windows version of the MARK3 Monte Carlo resource simulator: U.S. Geological Survey Open-File Report 2000–415 (superseded), accessed August 2012, at <http://pubs.er.usgs.gov/publication/ofr00415>.
- Duval, J.S., 2002, EMINERS—An economic mineral resource simulator: U.S. Geological Survey Open-File Report 2002–380 (superseded), accessed August 2012, at <http://pubs.er.usgs.gov/publication/ofr02380>.
- Duval, J.S., 2012, Version 3.0 of EMINERS—Economic mineral resource simulator: U.S. Geological Survey Open-File Report 2004–1344, accessed August 2012, at <http://pubs.usgs.gov/of/2004/1344>.
- Ellefsen, K.J., 2017, User's guide for MapMark4—An R package for the probability calculations in three-part mineral resource assessments: U.S. Geological Survey Techniques and Methods, Techniques and Methods, book 7, chap. C14, 23 p., <http://dx.doi.org/10.3133/tm7C14>.
- Grimmett, G.R., and Stirzaker, D.R., 2001, Probability and random processes (3d ed.): New York, Oxford University Press, 596 p.
- Harris, D.P., 1984, Mineral resource appraisal—Mineral endowment, resources, and undiscovered supply—Concepts, methods, and cases: New York, Oxford University Press, 445 p.
- Harris, D.P., and Carrigan, F.J., 1981, Estimation of uranium endowment by subjective geological analysis—A comparison of methods and estimates for the San Juan Basin, New Mexico: *Economic Geology*, v. 76, p. 1032–1055.
- Hastie, Trevor, Tibshirani, Robert, and Friedman, Jerome, 2009, The elements of statistical learning—Data mining, inference, and prediction (2d ed.): New York, Springer Science + Business Media, LLC, 745 p.
- Long, K.R., 1995, Production and reserves of Cordilleran (Alaska to Chile) porphyry copper deposits, in Pierce, F.W., and Bolm, J.G., eds., *Porphyry copper deposits of the American Cordillera: Arizona Geological Society Digest 20*, p. 35–68.
- Nelder, J.A., and Mead, R., 1965, A simplex algorithm for function minimization: *Computer Journal*, v. 7, p. 308–313.
- Pawłowsky-Glahn, V., Egozcue, J.J., and Tolosana-Delgado, R., 2015, Modeling and analysis of compositional data: John Wiley and Sons, Ltd., 247 p.
- Root, D.H., Menzie, W.D., and Scott, W.A., 1992, Computer Monte Carlo simulation in quantitative resource estimation: *Nonrenewable Resources*, v. 1, no. 2, p. 125–138.
- Root, D.H., Scott, W.A., Jr., and Schruben, P.G., 1998, MARK3B, resource assessment program for Macintosh: U.S. Geological Survey Open-File Report 98–356, accessed August 2012, at <http://pubs.er.usgs.gov/publication/ofr98356>.
- Shalizi, C.R., 2016, Advanced data analysis from an elementary point of view: Cambridge, United Kingdom, Cambridge University Press, accessed April 2016, at <http://www.stat.cmu.edu/~cshalizi/ADAfaEPOV/>.
- Sinclair, A.J., and Blackwell, G.H., 2002, Applied mineral inventory estimation: Cambridge, United Kingdom, Cambridge University Press, 400 p.
- Singer, D.A., and Menzie, W.D., 2010, Quantitative mineral resource assessments—An integrated approach: Cambridge, United Kingdom, Oxford University Press., 219 p.
- Stanley, M.C., 1992, Statistical trends and discoverability modeling of gold deposits in the Abitibi greenstone belt, Ontario: in Kim, Y.C., ed., *Proceedings of the 23rd international symposium on the applications of computers and operations research in the mineral industry, APCOM 92, 7–11 April 1992: Tucson Arizona, Society for Mining, Metallurgy, and Exploration, American Institute of Mining, Metallurgical, and Petroleum Engineers*, p. 11–17.
- Zientek, M.L., Bliss, J.D., Broughton, D.W., Christie, Michael, Denning, P.D., Hayes, T.S., Hitzman, M.W., Horton, J.D., Frost-Killian, Susan, Jack, D.J., Master, Sharad, Parks, H.L., Taylor, C.D., Wilson, A.B., Wintzer, N.E., and Woodhead, Jon, 2014, Sediment-hosted stratabound copper assessment of the Neoproterozoic Roan Group, Central African Copperbelt, Katanga Basin, Democratic Republic of the Congo and Zambia: U.S. Geological Survey Scientific Investigations Report 2010–5090–T, 162 p., and spatial data, <http://dx.doi.org/10.3133/sir20105090T>.

Publishing support provided by:

Denver Publishing Service Center, Denver, Colorado

For more information concerning this publication, contact:

Center Director, USGS Crustal Geophysics and Geochemistry Science Center

Box 25046, Mail Stop 964

Denver, CO 80225

(303) 236-1312

Or visit the Crustal Geophysics and Geochemistry Science Center Web site at:

<https://crustal.usgs.gov/>

This publication is available online at:

<https://doi.org/10.3133/tm7C15>

