Contents

Figures

Appendix Figures

Conversion Factors

International System of Units to Inch/Pound

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

User's Guide for MapMark4—An R Package for the Probability Calculations in Three-Part Mineral Resource Assessments

By Karl J. Ellefsen

Abstract

MapMark4 is a software package that implements the probability calculations in three-part mineral resource assessments. Functions within the software package are written in the R statistical programming language. These functions, their documentation, and a copy of this user's guide are bundled together in R's unit of shareable code, which is called a "package." This user's guide includes step-by-step instructions showing how the functions are used to carry out the probability calculations. The calculations are demonstrated using test data, which are included in the package.

Introduction

MapMark4 is a software package that implements the probability calculations in three-part mineral resource assessments. The name is derived from the descriptive phrase "mineral assessment program mark4"—the name "mark4" is chosen because the previous, similar program was called "mark3" (Root and others, 1992). Functions within the software package are written in the R statistical programming language, which is called either "R language" or "R" in this user's guide. These functions, their documentation, and a copy of this user's guide are bundled together in R's unit of sharable code, which is called a "package." MapMark4 provides the essential R functions to perform the probability calculations but does not provide a graphical user's interface.

This user's guide provides an overview of the MapMark4 package, showing you how to use the functions to carry out the probability calculations. As an overview, the scope of this user's guide is limited. It does not provide detailed descriptions of the functions because this information is available in the package Help. It describes neither the underlying probability model nor the mathematics of the calculations, because this information is published in Ellefsen (2017).

We assume that you are familiar with the R language, statistical methods, compositional data analysis, and the method of three-part mineral resource assessment. That is, we assume that you are an advanced user. Furthermore, we assume that you are using the Windows operating system. If not, then you must modify those steps in the user's guide that are related to files and directories.

The goals of this user's guide are most readily achieved by showing the step-by-step calculations for an actual dataset. Thus, the package includes two datasets for your use. This user's guide includes the R-language scripts that carry out the step-by-step calculations on these datasets. We strongly encourage you to execute these scripts yourself because this effort will help you become familiar with the calculations.

This user's guide focuses primarily on probability calculations involving grade and tonnage models, which specify ore tonnage and mineral resource grades. However, probability calculations also can be performed for tonnage models, which specify contained metal tonnage. These calculations are described in Appendix 1.

In this user's guide, R language scripts, program variables, data structures, and so on are typeset using the Courier New font (for example, `setwd("F:\\tmp\\PT001")`).

Overview of Probability Calculations

The probability calculations are conveniently divided into six groups, which are implemented as software classes. The first class pertains to the metadata that describes the permissive tract. The metadata are encapsulated in an S3 class, which is one type of an R language class (R Core Team, 2015). This class is named "Metadata."

The second software class pertains to the mass of ore in an undiscovered deposit within the permissive tract. The units for mass that are commonly used by economic geologists are metric tons (1,000 kilograms). When using units of metric tons, mass traditionally is called "tonnage," and this term is used in the remainder of this user's guide. The ore tonnages in an undiscovered deposit are represented by a probability density function (pdf) that accounts for uncertainty. This pdf is estimated from known ore tonnages in discovered deposits. These and related computations are encapsulated in an S3 class that is named "TonnagePdf."

The third software class pertains to the grades of the mineral resources in an undiscovered deposit within the permissive tract. The units for grades that are commonly used by economic geologists are percentages. The grades in an undiscovered deposit are represented by a pdf that accounts for uncertainty. This pdf is generated from known grades in discovered deposits. These and related computations are encapsulated in an S3 class that is named "GradePdf."

The fourth class pertains to the number of undiscovered deposits within the permissive tract. This number is represented by a probability mass function (pmf) that accounts for uncertainty. The pmf is generated from estimates of the number of undiscovered deposits that are made by the assessment geoscientists. These and related computations are encapsulated in an S3 class that is named "NDepositsPmf."

The fifth class pertains to the simulation of undiscovered deposits within the permissive tract. The undiscovered deposits are simulated using information from classes TonnagePdf, GradePdf, and NDepositsPmf. The computations are encapsulated in an S3 class that is named "Simulation."

The sixth class pertains to summary statistics for the simulated undiscovered deposits. The summary statistics are the total ore and mineral resource tonnages in all simulated undiscovered deposits within the permissive tract. These total ore and mineral resource tonnages are represented by a pdf that accounts for uncertainty. These computations and related computations are encapsulated in an S3 class that is named "TotalTonnagePdf."

The general relations among the inputs to the probability calculations, the six software classes, and the output are conveniently summarized by the diagram in figure 1.

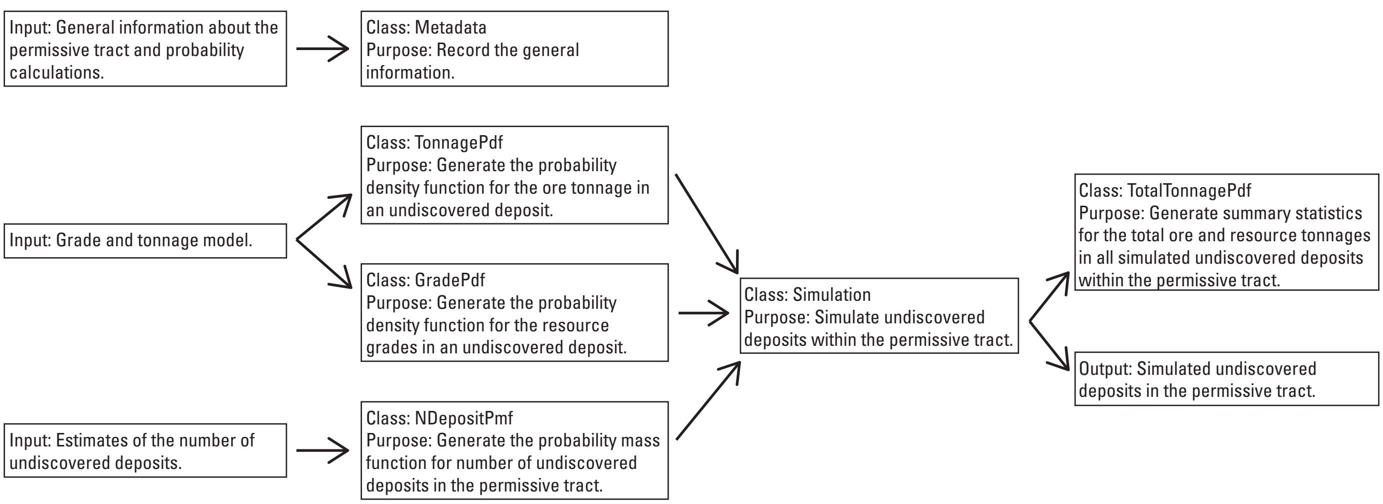

Figure 1. General relations among the inputs to the probability calculations, the six software classes, and the output. The arrows indicate major flows of information. This diagram pertains to a grade and tonnage model.

Data for Probability Calculations

Grade and Tonnage Model

For your probability calculations, you will need a grade and tonnage model that is pertinent to your mineral resource assessment. To understand the components of a grade and tonnage model, examine the example model that is provided with the package. To accomplish this task, type the following R scripts in the R console window:

```
library(MapMark4)
ExampleGatm
```

The grade and tonnage model is printed in the R console window. Here are the first five lines of the entire model:

	ID	Name	Ore	Cu	Au
1	DepositID1	DepositName1	147204645	0.5203261	4.911163e-05
2	DepositID2	DepositName2	25973547	0.3026978	3.975937e-05
3	DepositID3	DepositName3	220751458	0.6024545	1.726491e-05
4	DepositID4	DepositName4	97091334	0.4439770	3.505726e-05
5	DepositID5	DepositName5	2139367354	0.5893846	3.232766e-05

The first column lists unique identifiers for the discovered deposits. The second column lists the names of discovered deposits. In practice, two or more discovered deposits occasionally have the same name, so you should not rely on the name to identify the deposit. The third column lists the tonnage of the ore. The fourth and subsequent columns list the mineral resource grades. This computer-generated dataset comprises 71 discovered deposits, so there are 71 rows in `ExampleGatm`.

A grade and tonnage model, including this example model, is stored within an R session as an R language data frame. The information in the first and second columns is not used in MapMark4, so MapMark4 places no constraints on the information in these columns. However, this is not the case for the remaining columns. The third column, which lists the ore tonnages, must not have any missing values. Furthermore, all tonnages must be positive—that is, zero-valued and negative-valued tonnages are prohibited. The units for ore tonnage are metric tons. The fourth and subsequent columns, which list the mineral resource grades, must not have any missing values. The mineral resource grades are expressed with percent. All grades must be between 0 and 100—that is, they cannot equal 0 or 100.

Estimates of the Number of Undiscovered Deposits

For your probability calculations, you will also need the estimates of the numbers of undiscovered deposits within a permissive tract. These estimates are generated by the members of the assessment team. To see an example of such estimates, type the following R script:

```
ExampleDepEst1
```

The estimates are printed in the R console window:

	Name	Weight	N90	N50	N10
1	person1	0.50	3	10	25
2	person2	1.00	2	5	10
3	person3	1.00	5	7	10
4	person4	1.00	2	10	20
5	person5	1.00	1	2	4
6	person6	3.00	3	10	20
7	person7	1.00	5	10	20
8	person8	1.00	3	5	7
9	person9	1.00	1	2	5
10	person10	0.01	10	20	60

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 later. The N90 column lists, for each member, the estimated number of undiscovered deposits at a probability of 0.90. For example, person6 estimates that there is a 0.90 probability of finding three or more undiscovered deposits in the permissive tract. Among assessment geoscientists, this probability is called an "elicitation percentile of 90." The N50 and N10 columns are similar and list, for each member, the estimated number of undiscovered deposits at elicitation percentiles of 50 and 10. Some additional information about `ExampleDepEst1` is available in the package Help.

Corresponding statistics usually differ slightly. However, if the pdf is not truncated, then the corresponding minimum and the maximum statistics might differ a lot, as they do in this example. (Some readers may notice that the maximum percentage for gangue is 100, which appears to contradict the criteria listed in section “Grade and tonnage model.” The actual value is slightly less than 100, but has been rounded to 100 for the summary.)

Distribution for the Number of Undiscovered Deposits

Class `NDepositPmf` generates the pmf for the number of undiscovered deposits in the permissive tract. This pmf is calculated from the estimates of the number of undiscovered deposits, which is described in section “Data for probability calculations.” The estimates are read with this script:

```
depEst <- read.csv(getDepEstFilename(oMeta))
```

Function `getDepEstFilename` gets the name of the file containing the estimates from the metadata, and then function `read.csv` reads that file. Then class `NDepositPmf` is instantiated with this script:

```
oPmf <- NDepositsPmf("NegBinomial", list(nDepEst = depEst))
```

The function requires several seconds to complete its calculations. The first argument specifies that the pmf is a negative binomial distribution. The second argument is an R-language list that comprises the parameters for the pmf. Parameter `nDepEst` is an R-language data frame that contains the estimates of the number of undiscovered deposits.

A simple way to check the pmf is to plot it using this script:

```
plot(oPmf)
```

In the pmf (fig. 4*A*), the points represent the probabilities for the associated numbers of undiscovered deposits. For example, the probability is 0.00787 when the number of undiscovered deposits is 0, 0.0204 when the number is 1, and so on. To check the estimation of the pmf, it is recast as elicitation percentiles and then plotted with the estimated numbers of undiscovered deposits from the assessment team members (fig. 4*B*). The black dots are roughly near the red circles at the three different elicitation percentiles, with three exceptions. The three exceptions are the rightmost red circles at the three elicitation percentiles; these three exceptions are the estimates from one assessment team member and have a weight of 0.01. Thus, the estimated pmf adequately fits the estimated numbers of undiscovered deposits.

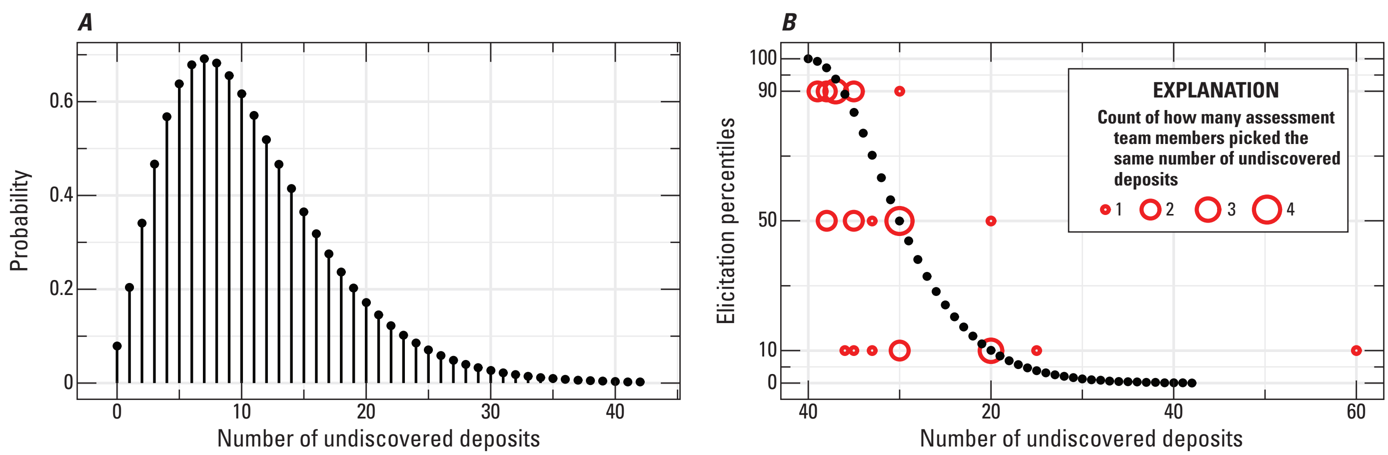

Figure 4. *A*, Negative binomial probability mass function representing the number of undiscovered deposits in the permissive tract. *B*, Probability mass function recast as elicitation percentiles (black dots) and compared to the estimated numbers of undiscovered deposits (red circles). The size of a red circle indicates how many assessment team members picked the same number of undiscovered deposits.

There is an additional, optional argument to function `plot` that is important. Argument `areLinesAdded` specifies whether a vertical line is added for each number of undiscovered deposit, making it easier to perceive the associated probability mass. The default value for `areLinesAdded` is `TRUE`. If the number of undiscovered deposits is large, (for example, roughly 80 or more) then the vertical lines will overwhelm the plot. In this case, argument `areLinesAdded` should be set to `FALSE`.

This script

```
summary(oPmf)
```

prints the following summary to the R console:

Appendix 1. Probability Calculations for a Tonnage Model

Overview

The probability calculations for a tonnage model are very similar to the probability calculations for grade and tonnage model, which are presented in the main part of this report. Consequently, this appendix focuses on the few differences. The tonnage model comprises only tonnages for a single contained metal—the model lacks grades. The implications are that a class for the grades is not needed and that the probability calculations for a tonnage model are just a special case of the probability calculations for a grade and tonnage model.

For a tonnage model, the general relations among the inputs to the probability calculations, the five software classes, and the output are conveniently summarized by the diagram shown in figure 1–1. The diagram is very similar to that in figure 1—the only significant difference is that it lacks a class of the grades.

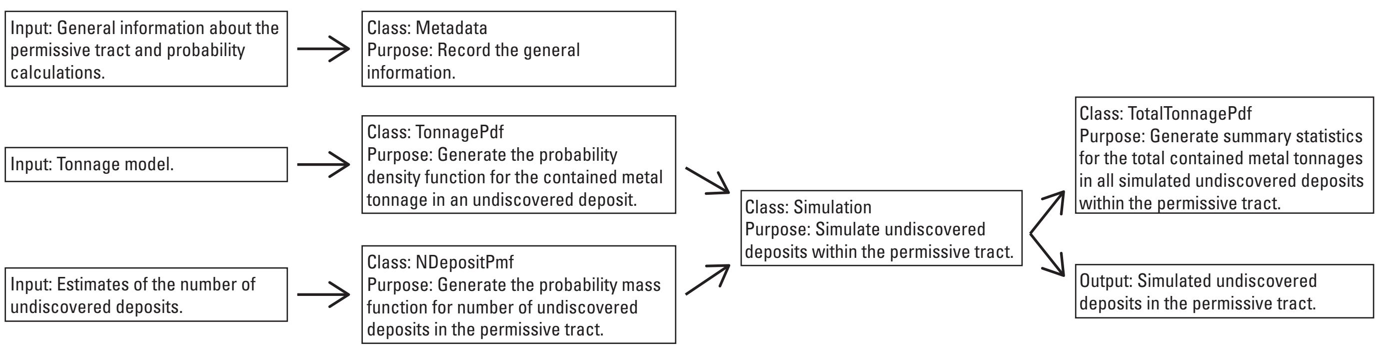

Figure 1–1. General relations among the inputs to the probability calculations, the five software classes, and the output. The arrows indicate major flows of information. This diagram pertains to a tonnage model

Data

To understand the components of a tonnage model, examine the example model that is provided with the package:

`ExampleTm`

The first 5 lines of the entire model are:

	ID	Name	U
1	DepositID1	DepositName1	18.34589
2	DepositID2	DepositName2	20.24016
3	DepositID3	DepositName3	34.38724
4	DepositID4	DepositName4	18.64999
5	DepositID5	DepositName5	33.04701

The first column lists unique identifiers for the discovered deposits. The second column lists the names of discovered deposits. The third column lists the tonnage of the contained metal, which is uranium for this example. The units for the tonnage are metric tons. Thus, a tonnage model and a grade and tonnage model are identical, except for the grades.

The other data that are needed for the probability calculations are the estimates of the number of undiscovered deposits, which is already described.

Preparatory Steps, Probability Calculations, and Archive of the Results

The scripts that are needed for the preparatory steps, the probability calculations, and the archive of the results are almost identical to the scripts for the grade and tonnage model. The only significant difference is that the class for the grades is not instantiated and is not passed to function `Simulation` and `TotalTonnagePdf`. Because of the identicalness, scripts are not repeated here—they are found in file "CalculationScripts_Tm.R," which accompanies this report. The file includes numerous comments that describe the scripts.

Appendix 2. Custom Distribution for the Number of Undiscovered Deposits

In some situations, a specific pmf for the number of undiscovered deposits is required. This need can be satisfied by class `NDepositPmf` using the following script, for example:

```
oPmf <- NDepositsPmf("Custom", list(nDeposits=c(1), relProbabilities=c(1))).
```

The first argument indicates that a custom pmf is being specified. The second argument is an R language list that comprises the parameters for the custom pmf. Parameter `nDeposits` is an R language vector for which each element specifies a number of undiscovered deposits with a nonzero probability. In this example, there is one undiscovered deposit. Parameter `relProbabilities` is also an R language vector, for which each element specifies the probabilities associated with parameter `nDeposits`. In this example, the probability is 1 for 1 undiscovered deposit. The pmf is plotted with this script

```
plot(oPmf)
```

and is shown in figure 2–1.

Other pmfs can be generated by changing the second argument to function `NDepositPmf`. For example, if the second argument were `list(nDeposits=c(1, 2, 3), relProbabilities=c(1, 1.5, 0.75))`, then the pmf would have three nonzero probabilities at 1, 2, and 3. The elements of `relProbabilities` need not sum to 1.

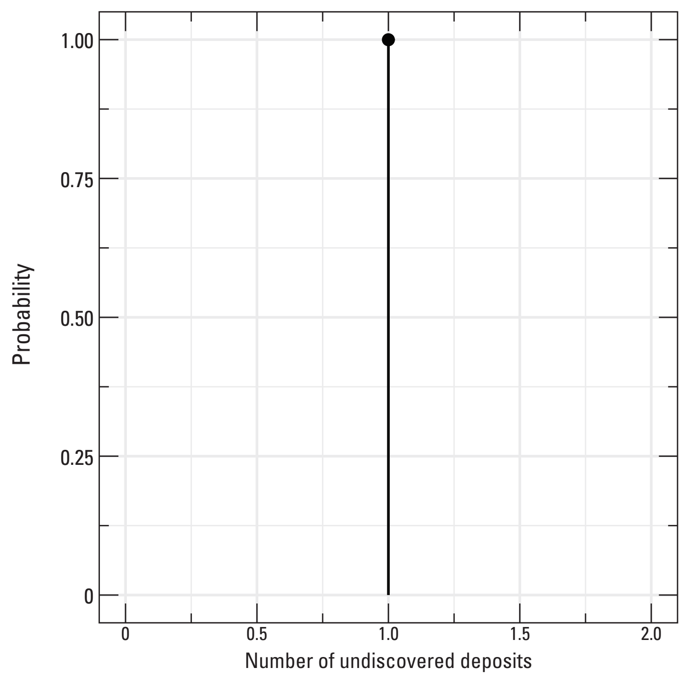

Figure 2–1. Example of a custom pmf for the number of undiscovered deposits in the permissive tract.

Appendix 3. Installation Instructions

Please see the file that accompanies this report, “InstallationInstructions.txt.”

Required Packages

Package MapMark4 requires packages compositions, mvtnorm, ks, and ggplot2. These packages are available from the R package repository and should be installed first. The latest versions of these packages are required for proper function of the package MapMark4. Also, Rtools (https://cran.r-project.org/bin/windows/Rtools/) is required during the install.

Installation Instructions for MapMark4 When Using RStudio

1. On the menu bar, click on “Tools” and then click on “Install Packages.” The Install-Packages window will appear.
2. Under item “Install from:” select “Package Archive File (.zip; .tar.gz).” Under item “Package archive,” use the Browse button to locate file “MapMark4_1.0.tar.gz.” If you use multiple libraries, then you must select the appropriate library for the installation, using item “Install to Library.” Otherwise, use the default value in this item. Finally, click on “Install.”

Installation Instructions for MapMark4 When Using RGui

1. In the R console window, type “setwd(directory)” where “directory” is the directory containing the file “MapMark4_1.0.tar.gz.” Two backslashes (“\\”) are needed to separate folders within argument directory. For example, if the directory is F:\temp\install, then argument directory should be “F:\\temp\\install”
2. Type “install.packages(‘MapMark4_1.0.tar.gz’)”, and the package will be installed in the default library.

Appendix 4. Calculation Scripts

Copies of the scripts in this User’s Guide are stored in files “CalculationScripts_Gatm.R” and “CalculationScripts_Tm.R” that accompany this report.

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/tm7C14

Ellefsen, K.J.—**User's Guide for MapMark4—An R Package for the Probability Calculations in Three-Part Mineral Resource Assessments**—Techniques and Methods 7–C14

ISSN 2328-7055 (online)
https://doi.org/10.3133/tm7C14