UNITED STATES
DEPARTMENT OF THE INTERIOR
GEOLOGICAL SURVEY

PROBABILISTIC METHODOLOGY FOR OIL AND GAS RESOURCE APPRAISAL

By

Robert A. Crovelli

Open-File Report 81-1151

1981

INTRODUCTION

The process of estimating the quantity of undiscovered crude oil or natural gas in a province is a complex problem that will be assumed to consist of four stages. This paper deals with the last stage. In order to get a complete picture of the whole process, the first three stages are briefly summarized.

The first stage is the information or data-gathering stage and is outlined as to the types of data compiled as follows.

A. Data Compilation

1. Petroleum geology.
2. Exploration history.
3. Production and reserves statistics.
4. Analog selection.
5. Seismic data.

The second stage involves the analysis of the data and is outlined as to the major categories of resource appraisal methods.

B. Data Analysis

1. Geological and geophysical interpretation.
2. Extrapolation of historical trends.
3. Areal or volumetric yield methods.
4. Geochemical material balance equations.
5. Play analysis.

The third stage is the making of a direct subjective estimate by a resource appraisal team of geologic experts using a Delphi type of approach which is a forecasting procedure. The third stage is outlined as to the steps that the team members go through.

C. Direct Subjective Assessment

1. Geology and results of methods of stage B are discussed.
2. Each member determines his own conditional probability curve of potential resources.
3. Team reviews all conditional resource estimates.
4. One or more revisions are possible.
5. Conditional resource estimates (curves) are averaged.

The resource assessment procedure results in the probabilistic estimation of two uncertain events: (a) the presence of the assessed hydrocarbon, thereby obtaining an estimate of the probability that the resource is present, and (b) its quantity, if present, which involves conditional probability, thereby obtaining conditional resource estimates.

Conditional probability is useful in practice for two reasons. First, in calculating a desired probability or expectation it is often extremely useful to first "condition" on some appropriate event, e.g., "oil is present." Second, we are at times interested in calculating probabilities and expectations when some partial information is available; hence, the desired probabilities and expectations are conditional ones. Therefore, conditional probability can be a means to an end (a tool) or an end in itself (a main interest).

The purpose of this paper is to present probabilistic methodology for processing the probabilistic assessments of undiscovered hydrocarbon resources which is the fourth stage. The problem is to determine the probability distribution of the quantity of resource from the estimates of the probability that the resource is present and the conditional probability distribution of the quantity of resource given that the resource is present. From this judgmental resource distribution and its properties, such as the mean, point and interval estimates of the quantity of undiscovered resource are obtained.

Estimates are made of the quantity of undiscovered crude oil and natural gas which is classified into two types.

Natural Gas Definitions

1. Non-associated gas is natural gas that is not in contact with significant quantities of crude oil in the reservoir.
2. Associated-dissolved gas is natural gas that occurs in a crude oil reservoir, either as free gas (associated) or as gas in solution in the crude oil (dissolved). Associated gas is often called "gas-cap gas."

Probability distributions and their properties are established for several petroleum resources in a province.

Petroleum Resources Estimated in

Part 1. Oil and non-associated gas.

Part 2. Associated-dissolved gas.

Part 3. Oil plus its associated-dissolved gas.

Part 4. Total gas (sum of non-associated gas and associated-dissolved gas).

Part 5. Petroleum resources in two or more provinces.

This paper includes nearly all of the probabilistic methodology used in Dolton and others (1981). Also included are some methods that were not used in that report. Computer programs and graphics routines are illustrated with examples from that study.

PART 1. OIL AND NON-ASSOCIATED GAS

It is assumed that oil and non-associated gas are separately and independently assessed in a province. Let

X: Quantity of undiscovered resource

where the resource of interest is either oil or non-associated gas. Treating X as a random variable with range $x \geq 0$, a probability distribution of X will be determined. From this judgmental resource distribution and its properties, estimates of the quantity of undiscovered resource are obtained.

Let the more-than cumulative distribution function (more-than cdf, also referred to as the complementary cdf) be denoted by

$$\bar{F}(x) = P(X > x)$$

which is the probability that the quantity of undiscovered resource is more than the amount x. Other equivalent phrases used instead of "more than" are "exceeding" and "greater than." The cumulative distribution function (cdf) is defined by

$$F(x) = P(X \leq x) = 1 - \bar{F}(x)$$

which is the probability that the quantity of resource is at most the amount x. The probability density function (pdf) is given by

$$f(x) = \frac{dF(x)}{dx} = -\frac{d\bar{F}(x)}{dx}$$

The following notation will be used:

μ_X : Mean of X

m_X : Median of X

σ_X^2 : Variance of X

σ_X : Standard deviation of X

Because μ_X or m_X can be used as "point" estimates of the quantity of undiscovered resource, they are called the mean estimate and median estimate, respectively.

Let x_α denote the amount of resource such that

$$\bar{F}(x_\alpha) = P(X > x_\alpha) = \alpha.$$

Examples:

a) $P(X > x_{0.95}) = 0.95$; $x_{0.95}$ is called the 95th fractile.

b) $P(X > x_{0.50}) = 0.50$; $x_{0.50}$ is called the 50th fractile.

c) $P(X > x_{0.05}) = 0.05$; $x_{0.05}$ is called the 5th fractile.

Note that $m_X = x_{0.50}$, i.e., the median is the 50th fractile. Also,

$$P(x_{0.95} < X < x_{0.05}) = 0.90.$$

Hence, resource estimates will be obtained for the province such that

$x_{0.95}$ represents a low resource estimate

and

$x_{0.05}$ represents a high resource estimate.

The range from low to high estimates forms an "interval" estimate of the quantity of undiscovered resource.

In order to determine the above quantities of interest, a conditional probability approach will be used. Because

$$P(X > x) = P(X > 0,\ X > x) \qquad x \geq 0,$$

we have the important relationship

$$P(X > x) = P(X > 0)P(X > x \mid X > 0),$$

where

$P(X > 0)$ is the probability that the resource is present, called traditionally the marginal probability,

$P(X > x \mid X > 0)$ is the conditional probability that the quantity of resource is more than the amount x given that the resource is present, called the conditional more-than cdf.

Therefore, the assessment procedure depends upon probabilistic estimation of two uncertain events: (a) the presence of the assessed hydrocarbon, and (b) its quantity, if present.

The Marginal Probability

Let the marginal probability be denoted by p, i.e.,

$$p = P(X > 0).$$

Note that

$$1 - p = P(X = 0)$$

which is the probability that the resource is <u>not</u> present. (The marginal probability is also referred to as the "adequacy chance" in the oil and gas industry.)

As an example, in assessing undiscovered recoverable oil, an assessment must be made as to its presence in commercial deposits within the province. Although the assumption of its presence may be made with confidence in already producing provinces, this assumption cannot be made with certainty in areas where no economically recoverable petroleum has been discovered. In frontier areas where there has been little or no drilling, there is a risk that no recoverable petroleum exists. Therefore, the likelihood of any recoverable resource being present is estimated. A marginal probability is estimated for the event "recoverable oil present" or for the event "recoverable non-associated gas present" depending upon the resource of interest. The reader should realize that instead of recoverable quantities of a resource, one could be interested in estimating in-place quantities; the methods are still applicable.

A more general marginal probability is written as

$$p = P(X > c)$$

where $c \geq 0$ is called a resource cut-off value.

The Conditional Probability Distribution

The conditional probability distribution represents the judgmental probability distribution of the quantity of undiscovered resource conditioned on the resource being present in the province. (This distribution is also referred to as the "unrisked" distribution in the oil and gas industry.) The conditional probability distribution is quite important in its own right and is not just a tool to obtain the probability distribution.

Let the conditional more-than cdf be denoted by

$$\bar{F}_c(x) = P(X > x|X > 0), \qquad x \geq 0$$

the conditional cdf by

$$F_c(x) = P(X \leq x|X > 0) = 1 - \bar{F}_c(x), \qquad x \geq 0$$

and the conditional pdf by

$$f_c(x) = \frac{dF_c(x)}{dx} = -\frac{d\bar{F}_c(x)}{dx}, \qquad x > 0.$$

Note that

$$\bar{F}_c(0) = 1 \quad \text{and} \quad F_c(0) = 0,$$

that is, the graph of the conditional more-than cdf has the value of 1 at the origin; the graph of the conditional cdf has the value of 0 at the origin.

The following notation will also be used:

X_c : Conditional quantity of undiscovered resource

μ_c : Conditional mean

m_c : Conditional median

M_c : Conditional mode

σ_c^2 : Conditional variance

σ_c : Conditional standard deviation

X_c is a continuous random variable with range $x > 0$. Let ${}_c x_\alpha$ denote the conditional amount of resource such that

$$\bar{F}_c({}_c x_\alpha) = P(X > {}_c x_\alpha | X > 0) = \alpha.$$

Examples:

a) $\bar{F}_c({}_c x_{0.95}) = 0.95$; ${}_c x_{0.95}$ is called the conditional 95th fractile.

b) $\bar{F}_c({}_c x_{0.05}) = 0.05$; ${}_c x_{0.05}$ is called the conditional 5th fractile.

Note that

$$P({}_c x_{0.95} < X_c < {}_c x_{0.05}) = 0.90.$$

The determination of a conditional probability distribution for a province is a major problem in itself. One approach to the problem will be given.

For illustrative purposes, the lognormal distribution will be used as a probability model for the conditional probability distribution in a province. Of course, other probability distributions are possible, e.g., the gamma and triangular distributions. The conditional pdf would then be expressed as

$$f_c(x) = \frac{1}{\delta x \sqrt{2\pi}} e^{-\frac{1}{2\delta^2}(\ln x - \xi)^2} \qquad x > 0$$

with parameters ξ and δ. (If a resource cut-off value is used, the probability model would become the three-parameter lognormal distribution.)

Conditional resource estimates are made for the province. For example, estimates of the following are made by a resource appraisal team of geologic experts:

a) ${}_c x_{0.95}$, resulting in a low conditional resource estimate

b) ${}_c x_{0.05}$, resulting in a high conditional resource estimate.

The range from low to high conditional estimates forms a conditional "interval" estimate. There are other possible pairs of conditional resource estimates that could be used to establish a lognormal distribution, e.g., the conditional mode M_c and ${}_c x_{0.05}$.

Given the conditional fractiles ${}_c x_{0.95}$ and ${}_c x_{0.05}$, it can be easily shown that the lognormal parameters can now be estimated by

$$\xi = \frac{\ln {}_c x_{0.05} + \ln {}_c x_{0.95}}{2},$$

and

$$\delta = \frac{\ln {}_c x_{0.05} - \ln {}_c x_{0.95}}{3.29}.$$

Knowing the parameters, any conditional fractile, ${}_c x_\alpha$, can now be calculated for a given α from the formula

$${}_c x_\alpha = e^{\xi + z_\alpha \delta}$$

where Z is a standard normal random variable and $P(Z > z_\alpha) = \alpha$.

The following quantities can also be calculated:

$$\mu_c = e^{\xi + \delta^2/2}$$

$$m_c = e^{\xi}$$

$$M_c = e^{\xi - \delta^2}$$

$$\sigma_c^2 = e^{2\xi + \delta^2}(e^{\delta^2} - 1) = \mu_c^2(e^{\delta^2} - 1)$$

$$\sigma_c = \sqrt{\sigma_c^2}$$

Note that $\mu_c > m_c > M_c$.

For an example from Dolton and others (1981), we will consider the Montana Overthrust Belt province for illustrative purposes. A USGS (U.S. Geological Survey) resource appraisal team made the following assessments of undiscovered recoverable oil:

1. At the time there had been no recoverable oil found in the province. The chance of oil being present in recoverable quantities was estimated to be 86 percent, i.e., the marginal probability of oil is $p = 0.86$.

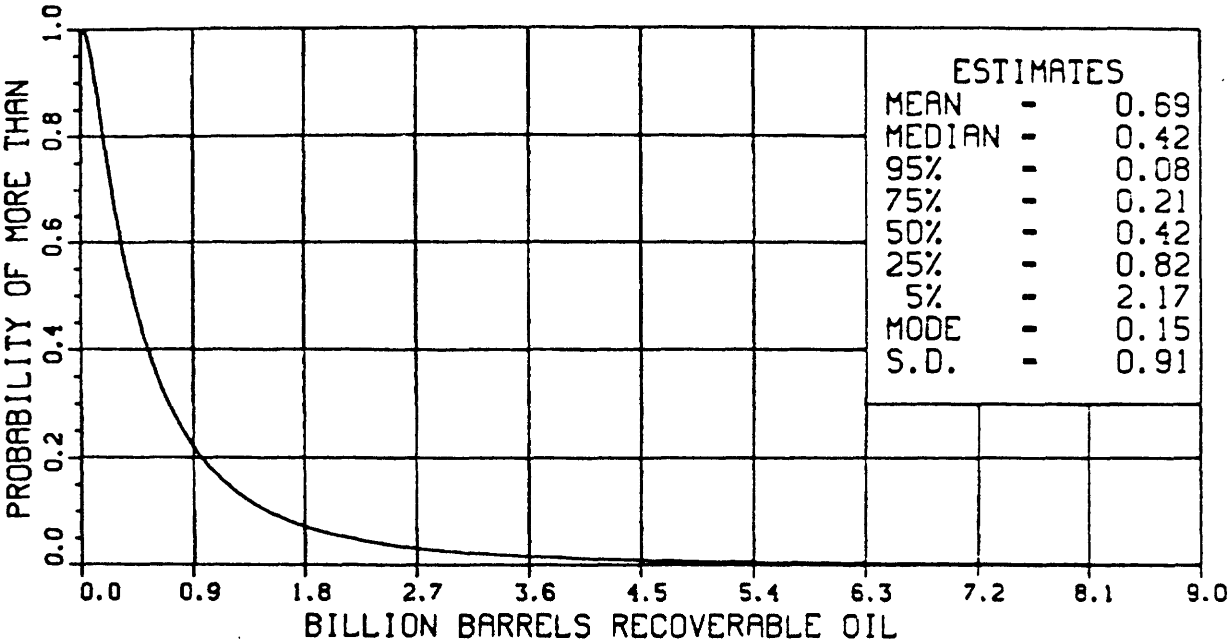
2. Conditional upon recoverable oil being present, conditional resource estimates were expressed in billion barrels

   a) ${}_{c}x_{0.95} = 0.08$, a low conditional resource estimate

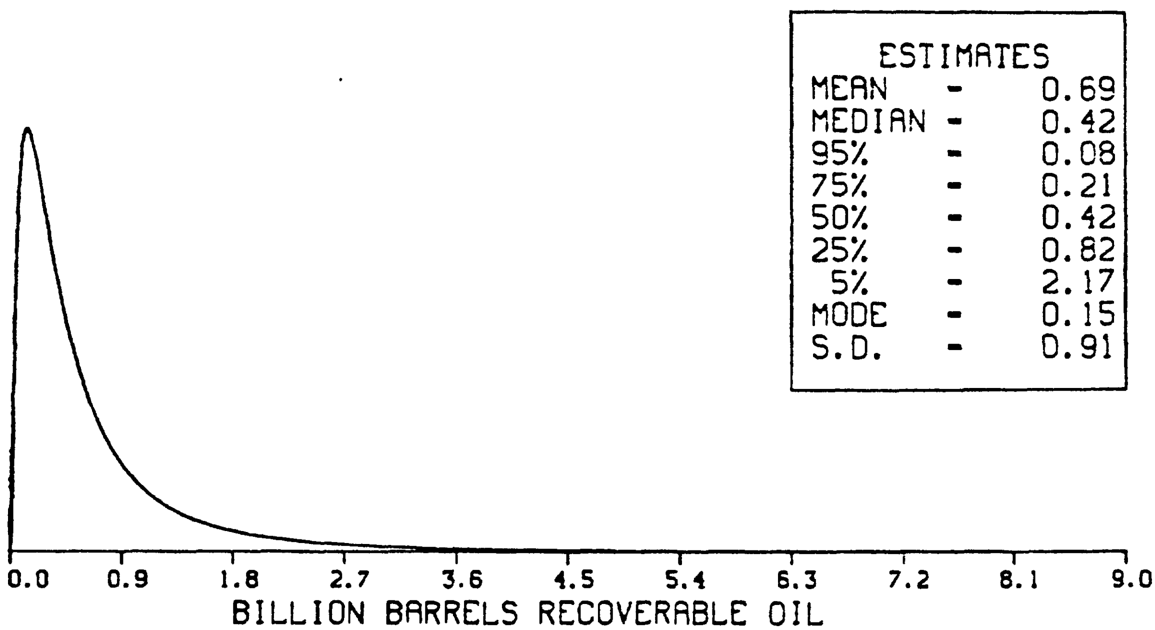
   b) ${}_{c}x_{0.05} = 2.17$, a high conditional resource estimate

The conditional probability distribution of the quantity of undiscovered recoverable oil for the Montana Overthrust Belt province is displayed in figure 1. The graph of the conditional more-than cdf (fig. 1a) has the value of 1 at the origin. The graph of the conditional pdf (fig. 1b) has the value of 0 at the origin. Some numerical characteristics in billion barrels (BB) of the lognormal distribution in figure 1 are the following:

${}_{c}x_{0.95} = 0.08$	$\mu_c = 0.69$
${}_{c}x_{0.75} = 0.21$	$m_c = 0.42$
${}_{c}x_{0.50} = 0.42$	$M_c = 0.15$
${}_{c}x_{0.25} = 0.82$	$\sigma_c^2 = 0.83\ (BB)^2$
${}_{c}x_{0.05} = 2.17$	$\sigma_c = 0.91$



(a) CONDITIONAL MORE-THAN CUMULATIVE DISTRIBUTION FUNCTION



(b) CONDITIONAL PROBABILITY DENSITY FUNCTION

Figure 1.-Conditional probability distribution of undiscovered recoverable oil for Montana Overthrust Belt province. Estimates are mean, median, mode, standard deviation (S.D.) and fractiles that correspond to percentages listed.

The Probability Distribution of X

The marginal probability of a resource for a province is applied to its corresponding conditional probability distribution to produce the probability distribution of the quantity of undiscovered resource. (This distribution is also referred to as the "unconditional" or "risked" distribution in the oil and gas industry.) More explicitly, recall the basic relationship

$$P(X > x) = P(X > 0)P(X > x|X > 0) \qquad x \geq 0$$

or

$$\bar{F}(x) = p\,\bar{F}_c(x),$$

which could also be expressed as

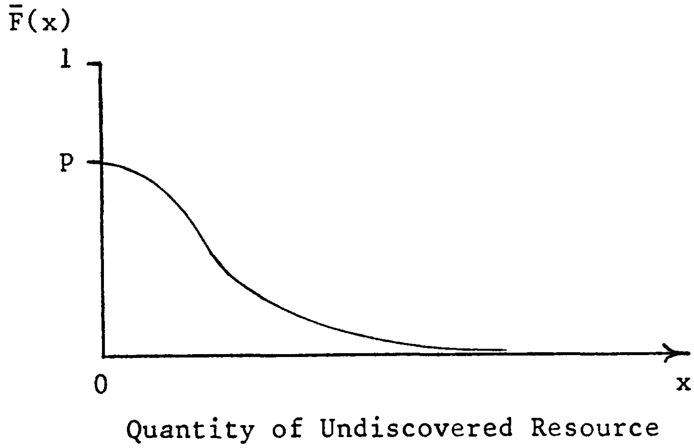
$$\bar{F}_c(x) = \frac{\bar{F}(x)}{p}.$$

More completely, the more-than cdf of X is given by

$$\bar{F}(x) = \begin{cases} 1 & \text{for } x < 0 \\ p & \text{for } x = 0 \\ p\bar{F}_c(x) & \text{for } x > 0 \end{cases}$$

Note that if $p = 0$, then $\bar{F}_c(x)$ is not defined, i.e., if the marginal probability is 0, then the conditional probability distribution is not defined. Also, if $p = 1$, then $\bar{F}(x) = \bar{F}_c(x)$, i.e., if the marginal probability is 1, then the probability distribution is equal to the conditional probability distribution.

The graph of a typical more-than cdf of X, $\bar{F}(x)$, is shown below with the curve having the value of the marginal probability, p, at the origin.



Determination of Various Fractiles of X, x_α

Given $p \neq 0$ and specified α, find x_α using

$$\bar{F}_c(x_\alpha) = \frac{\bar{F}(x_\alpha)}{p} = \frac{\alpha}{p} \qquad \text{for } \alpha < p$$

Case I: $\alpha/p < 1$, i.e., $\alpha < p$. Find x_α such that

$$\bar{F}_c(x_\alpha) = \frac{\alpha}{p} .$$

Therefore,

$$x_\alpha = {}_c x_{\alpha/p} .$$

Case II: $\alpha/p \geq 1$, i.e., $\alpha \geq p$. Then $x_\alpha = 0$.

The 95th and 5th fractiles, in particular, can now be determined. Recall that the range from low ($x_{0.95}$) to high ($x_{0.05}$) estimates forms an "interval" estimate of the quantity of undiscovered resource.

Examples:

1) Given $p = 0.20$

Case I: $\alpha < 0.20$, find x_α such that

$$\bar{F}_c(x_\alpha) = \frac{\alpha}{0.20} .$$

Therefore,

$$x_\alpha = {}_c x_{\alpha/0.20}$$

For $\alpha = 0.05$, find $x_{0.05}$ such that

$$\bar{F}_c(x_{0.05}) = \frac{0.05}{0.20} = 0.25.$$

Therefore,

$$x_{0.05} = {}_c x_{0.25}$$

The 5th fractile of X equals the conditional 25th fractile.

Case II: $\alpha \geq 0.20$, then $x_\alpha = 0$.

For $\alpha = 0.95$, we have $x_{0.95} = 0$.

2) Median of X, m_X. Recall that $m_X = x_{0.50}$, so that $\alpha = 0.50$.

Case I: $p > 0.50$, find $x_{0.50}$ such that

$$\bar{F}_c(x_{0.50}) = \frac{0.50}{p} .$$

Therefore,

$$x_\alpha = {}_c x_{0.50/p}$$

For $p = 0.75$, find $x_{0.50}$ such that

$$\bar{F}_c(x_{0.50}) = \frac{0.50}{0.75} = 0.67.$$

Therefore,

$$x_{0.50} = {}_c x_{0.67}$$

The median of X equals the conditional 67th fractile.

Case II: $p \leq 0.50$, then $x_{0.50} = 0$.

For $p = 0.10$, we have $x_{0.50} = 0$.

The random variable X is a <u>mixed type</u> of random variable. Because for the continuous part of the distribution

$$f(x) = -\frac{d\bar{F}(x)}{dx} = -p\,\frac{d\bar{F}_c(x)}{dx} = pf_c(x) \qquad x > 0,$$

we have the pdf of X

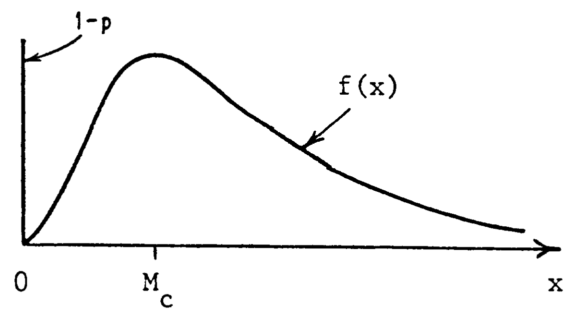
$$f(x) = \begin{cases} 1 - p & \text{for} \quad x = 0 \\ pf_c(x) & \text{for} \quad x > 0 \\ 0 & \quad \text{otherwise} \end{cases}$$

where $f(0) \equiv P(X = 0)$ for the discrete part of the distribution. For a more formal expression of f(x), the Dirac delta function can be used.

Note that

$$\int_{0+}^{\infty} f(x)dx = p \int_{0+}^{\infty} f_c(x)dx = p(1) = p.$$

The graph of a typical density function of X is shown below with a spike at the origin of probability weight 1 - p. Assuming a unimodal conditional distribution with mode M_c, then f(x) would also peak at M_c for x > 0. However, if the marginal probability is less than 1, the unconditional mode is not a meaningful concept.



Quantity of Undiscovered Resource

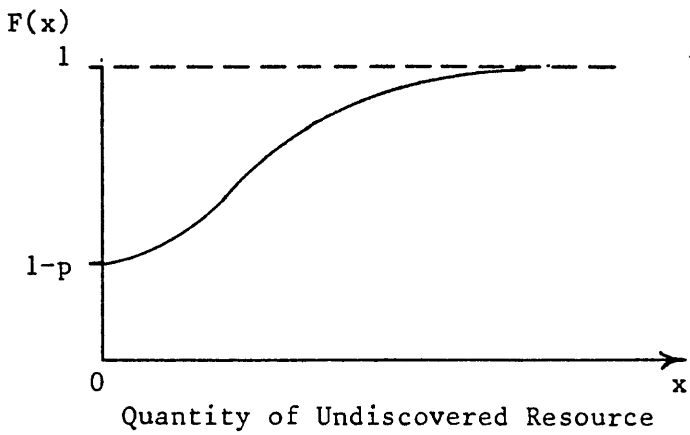
Because

$$F(x) = 1 - \bar{F}(x) \qquad x \geq 0$$
$$= 1 - p\bar{F}_c(x)$$
$$= 1 - p[\, 1 - F_c(x)\,]$$
$$= 1 - p + pF_c(x)$$

the cdf of X is given by

$$F(x) = \begin{cases} 0 & \text{for } x < 0 \\ 1 - p & \text{for } x = 0 \\ 1 - p + pF_c(x) & \text{for } x > 0 \end{cases}$$

The graph of a typical cdf of X, F(x), is shown below with the curve having the value of 1 - p at the origin.



The Mean of X, μ_X

Theorem $\mu_X = p\mu_c$

Proof: Define the indicator random variable I by

$$I = \begin{cases} 1 & \text{if the resource is present} \\ 0 & \text{if the resource is not present} \end{cases}$$

Note that $(I = 0) \equiv (X = 0)$ and $(I = 1) \equiv (X > 0)$

We will determine $\mu_X \equiv E(X)$ by conditioning on I.

An important property of conditional expectation is that

$$E(X) = \sum_i E(X|I = i)P(I = i).$$

Hence

$$E(X) = E(X|I = 0)P(I = 0) + E(X|I = 1)P(I = 1).$$

Because

$$E(X|I = 0) = E(X|X = 0) = 0 \text{ with } P(I = 0) = P(X = 0) = 1 - p$$

and

$$E(X|I = 1) = E(X|X > 0) = \mu_c \text{ with } P(I = 1) = P(X > 0) = p,$$

we have

$$\mu_X = (0)(1 - p) + (\mu_c)(p) = p\mu_c$$

Remark: $f_c(x) \equiv f_{X|I}(x|1)$

Example: Refer to the Montana Overthrust Belt province.

Because $p = 0.86$ and $\mu_c = 0.69$,

we have

$$\mu_X = p\mu_c = (0.86)(0.69) = 0.59 \text{ BB}$$

The Variance of X, σ_X^2

Theorem $\sigma_X^2 = p\sigma_c^2 + p(1 - p)\ \mu_c^2$

Proof: We will determine $\sigma_X^2 \equiv Var(X)$ by conditioning on the indicator random variable I defined previously. An important property of conditional variance is that

$$Var(X) = E\ [\ Var(X|I)\ ] + Var\ [\ E(X|I)\ ].$$

Evaluating the two terms on the right side, we get

$$\begin{aligned} E\ [\ Var(X|I)\ ] &= Var(X|I = 0)P(I = 0) + Var(X|I = 1)P(I = 1) \\ &= (0)(1 - p) + (\sigma_c^2)(p) \\ &= p\sigma_c^2 \end{aligned}$$

and

$$\begin{aligned} Var\ [E(X|I)\ ] &= E\{\ [\ E(X|I)\ ]^2\} - \{E\ [\ E(X|I)\ ]\ \}^2 \\ &= E\{\ [\ E(X|I)\ ]^2\} - [\ E(X)\ ]^2 \\ &= \mu_c^2\ p - (p\mu_c)^2 \\ &= p(1 - p)\ \mu_c^2 \end{aligned}$$

recalling that

i	$E(X\|I = i)$	$P(I = i)$
0	0	1-p
1	μ_c	p

Therefore

$$\sigma_X^2 = p\sigma_c^2 + p(1 - p)\ \mu_c^2.$$

The standard deviation of X is given by

$$\sigma_X = \sqrt{p\sigma_c^2 + p(1 - p)\ \mu_c^2}$$

Example: Refer to the Montana Overthrust Belt province.

Because $p = 0.86$, $\mu_c = 0.69$, and $\sigma_c^2 = 0.83$,

we have

$$\sigma_X^2 = p\sigma_c^2 + p(1 - p)\mu_c^2$$

$$= (0.86)(0.83) + (0.86)(0.14)(0.69)^2$$

$$= 0.77 \quad (BB)^2$$

and

$$\sigma_X = 0.88 \quad BB$$

Example: Referring once again to the Montana Overthrust Belt province, the probability distribution of the undiscovered recoverable oil is displayed in figure 2. The graph of the more than cdf (fig. 2a) has the value of the marginal probability, p = 0.86, at the origin. The graph of the pdf (fig. 2b) has a spike at the origin of probability weight 1 - 0.86 = 0.14 which represents the chance of no oil being present in recoverable quantities. Some numerical characteristics (in billion barrels) of this distribution are the following:

$x_{0.95} = 0.00$

$x_{0.75} = 0.13$

$x_{0.50} = 0.34$

$x_{0.25} = 0.73$

$x_{0.05} = 2.02$

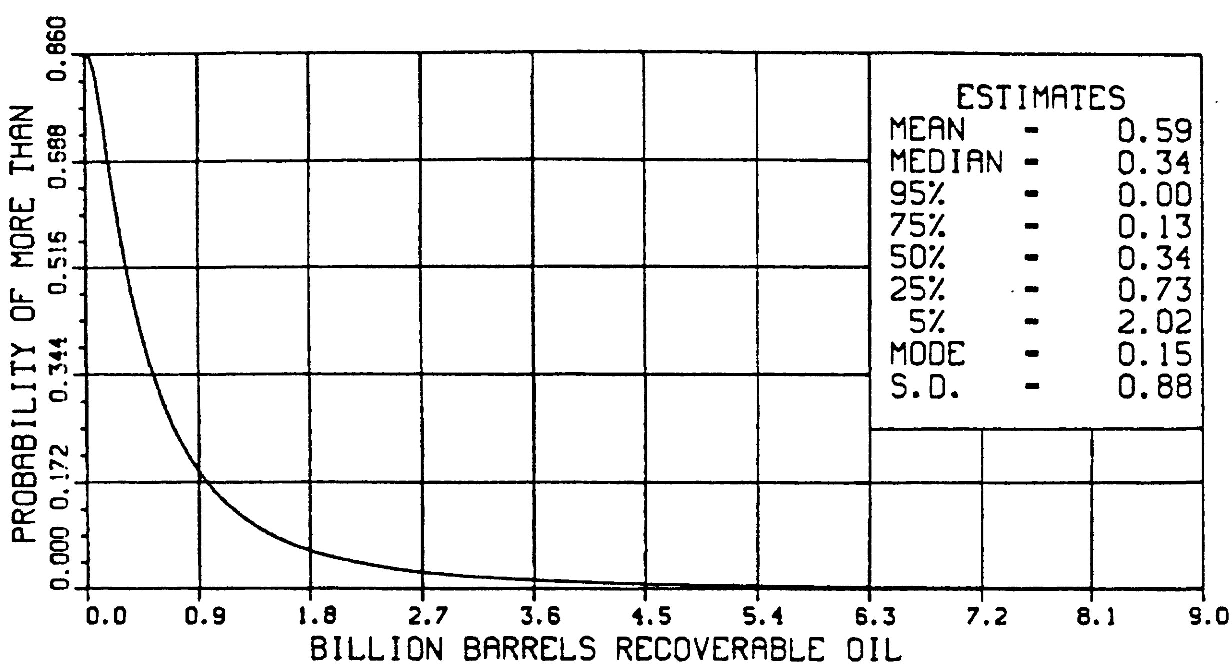
$\mu_X = 0.59$

$m_X = 0.34$

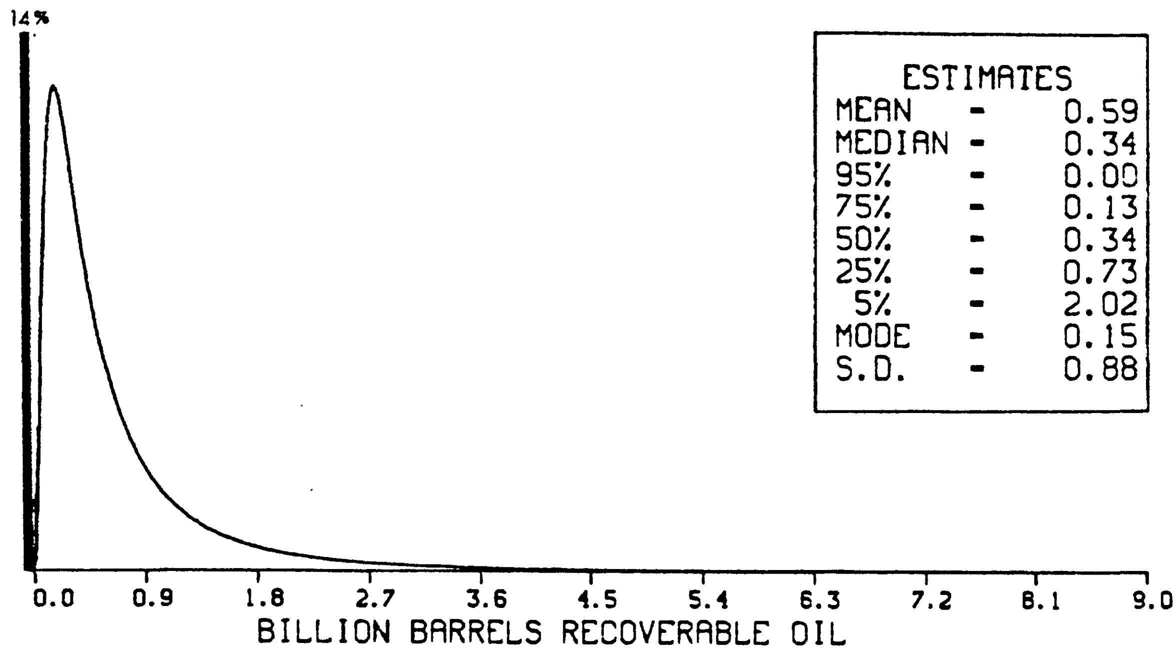
$\sigma_X^2 = 0.77 \quad (BB)^2$

$\sigma_X = 0.88$

The mean estimate of 0.59 or the median estimate of 0.34 can be used as "point" estimates of the quantity of undiscovered recoverable oil in the Montana Overthrust Belt province. The range from low resource estimate (0.00) to high resource estimate (2.02) forms an "interval" estimate.



(a) MORE-THAN CUMULATIVE DISTRIBUTION FUNCTION



(b) PROBABILITY DENSITY FUNCTION

Figure 2.-Probability distribution of undiscovered recoverable oil for Montana Overthrust Belt province. Estimates are mean, median, mode, standard deviation (S.D.) and fractiles that correspond to percentages listed.

Inequalities

One of the important uses of expectations of random variables lies in the collection of inequalities that involve the mean and standard deviation and that give bounds on the probabilities of particular events. It may happen that we can estimate the mean and standard deviation of a random variable without assuming its distribution. This knowledge often suffices to give upper or lower bounds on the probability of particular events, regardless of the exact definition of the distribution.

The Markov Inequality

Because X is a random variable that takes only nonnegative values, then for any value $x > 0$

$$\bar{F}(x) \leq \frac{\mu_X}{x} .$$

That is, the probability that the quantity of undiscovered resource is more than the amount x will be at most μ_X/x regardless of the distribution. Hence, the quantity μ_X/x is an upper bound for $\bar{F}(x)$.

The Chebyshev Inequality

Because X is a random variable with mean μ_X and standard deviation σ_X, then for any value $k > 0$

$$P(\mu_X - k\sigma_X < X < \mu_X + k\sigma_X) \geq 1 - \frac{1}{k^2} .$$

For $k = 2$,

$$P(\mu_X - 2\sigma_X < X < \mu_X + 2\sigma_X) \geq 0.75.$$

That is, the probability that the quantity of undiscovered resource lies within two standard deviations of the mean of any distribution is at least 0.75.

PART 2. ASSOCIATED-DISSOLVED GAS

The resource of interest to be estimated is the associated-dissolved gas in a province. We will derive an associated-dissolved gas distribution from its corresponding oil distribution in Part 1. From this judgmental resource distribution and its properties, estimates of the quantity of undiscovered associated-dissolved gas are obtained.

A gas-oil ratio (GOR) is established for the province. The GOR is the ratio of associated-dissolved gas to its oil for the particular province and varies from province to province. The gas-oil ratio could represent a consensus estimate of an appraisal team and based, when possible, on extrapolation of historic gas-oil ratios. Let

Y: Quantity of undiscovered associated-dissolved gas.

Assume that associated-dissolved gas is related to its oil by a linear model

$$Y = aX$$

where

a: Gas-oil ratio for the province, $a > 0$

X: Quantity of undiscovered oil, $x \geq 0$.

The GOR is taken to be an estimated constant, whereas X is a random variable from Part 1. Oil (X) and its associated-dissolved gas (Y) are obviously dependent random variables. Hence, treating Y as a random variable with range $y \geq 0$, a judgmental probability distribution of Y will be determined. Y is the same type of random variable as X, i.e., a <u>mixed type</u>.

More explicitly, the following will be determined

The more-than cdf of Y

$$\bar{F}_Y(y) = P(Y > y)$$

The cdf of Y

$$F_Y(y) = P(Y \leq y) = 1 - \bar{F}_Y(y)$$

The pdf of Y

$$f_Y(y) = \frac{d\ F_Y(y)}{dy} = -\frac{d\ \bar{F}_Y(y)}{dy}$$

Along with

μ_Y: Mean of Y

m_Y: Median of Y

σ_Y^2: Variance of Y

σ_Y: Standard deviation of Y

Let y_α denote the amount of associated-dissolved gas such that

$$\bar{F}_Y(y_\alpha) = P(Y > y_\alpha) = \alpha.$$

Examples:

a) $y_{0.95}$ is the 95th fractile of Y

b) $y_{0.50}$ is the 50th fractile of Y

c) $y_{0.05}$ is the 5th fractile of Y

Note that $m_Y = y_{0.50}$.

Resource estimates will be obtained for the province such that

$y_{0.95}$ represents a low resource estimate

and

$y_{0.05}$ represents a high resource estimate

The range from low ($y_{0.95}$) to high ($y_{0.05}$) estimates forms an "interval" estimate of the quantity of undiscovered associated-dissolved gas.

There are two approaches that can be used to determine the probability distribution of Y and its properties.

Approach I: Using the methodology of Part 1.

Approach II: Using the results of Part 1.

Approach I has an advantage from a computer programming point of view because the same program can be used as in Part 1. Approach II can be easily applied as a check on the work done by Approach I.

Approach I: Using the methodology of Part 1

The problem is approached exactly as it was in Part 1 using the marginal probability and the conditional probability distribution.

We once again have the important relationship

$$P(Y > y) = P(Y > 0)P(Y > y \mid Y > 0)$$

The marginal probability of Y is $P(Y > 0)$. Because

$$P(Y > 0) = P(aX > 0) = P(X > 0),$$

the marginal probability of Y is equal to the marginal probability of X. That is, under the assumed linear model, associated-dissolved gas has the same marginal probability as the oil with which it occurs.

The conditional more-than cdf of Y is $P(Y > y|Y > 0)$.

Let ${}_c y_\alpha$ denote the conditional amount of associated-dissolved gas such that

$$P(Y > {}_c y_\alpha \mid Y > 0) = \alpha.$$

Examples:

a) ${}_c y_{0.95}$ is the conditional 95th fractile

b) ${}_c y_{0.05}$ is the conditional 5th fractile

Recall that

$$P(X > {}_c x_\alpha \mid X > 0) = \alpha.$$

Because

$$P(Y > y|Y > 0) = P(aX > y|aX > 0) = P(X > y/a \mid X > 0),$$

we have

$$\alpha = P(Y > {}_c y_\alpha \mid Y > 0) = P(X > {}_c y_\alpha/a \mid X > 0).$$

Therefore, the conditional fractiles of Y are related to the conditional fractiles of X by

$${}_c y_\alpha = a \; {}_c x_\alpha$$

Examples:

a) ${}_c y_{0.95} = a \; {}_c x_{0.95}$

b) ${}_c y_{0.05} = a \; {}_c x_{0.05}$

If we let

X_c: Conditional quantity of undiscovered oil

Y_c: Conditional quantity of undiscovered associated-dissolved gas,

then

$$Y_c = aX_c$$

Of course, there are many relationships between X_c and Y_c.

The conditional more-than cdf of Y

$$\bar{F}_{Y_c}(y) = \bar{F}_{X_c}(y/a)$$

The conditional cdf of Y

$$F_{Y_c}(y) = F_{X_c}(y/a)$$

The conditional pdf of Y

$$f_{Y_c}(y) = \frac{1}{a} f_{X_c}(y/a)$$

The conditional mean of Y

$$\mu_{Y_c} = a\ \mu_{X_c}$$

The conditional median of Y

$$m_{Y_c} = a\ m_{X_c}$$

The conditional mode of Y

$$M_{Y_c} = a\ M_{X_c}$$

The conditional variance of Y

$$\sigma^2_{Y_c} = a^2\ \sigma^2_{X_c}$$

The conditional standard deviation of Y

$$\sigma_{Y_c} = a\ \sigma_{X_c}$$

The probability distribution of Y and its desired properties can be determined simply by multiplying the conditional oil estimates, ${}_{c}x_{0.95}$ and ${}_{c}x_{0.05}$, by the gas-oil ratio, a, and by using the methodology of Part 1. Recall that associated-dissolved gas has the same marginal probability as the oil with which it occurs.

The lognormal distribution was used in Part 1 as a probability model for the conditional probability distribution of X with parameters ξ and δ. In this case, it can be easily shown that the conditional probability distribution of Y is also lognormal but with parameters $\xi' = \xi + \ln a$ and δ.

The GOR could also be treated as a random variable and be estimated by a range of estimates, i.e., an interval estimate. Conditioned on associated-dissolved gas being present, we have

$$Y_c = A_c X_c$$

where a potential probability model would be to have the random variables A_c and X_c be independent and lognormally distributed. Then Y_c would also be lognormally distributed because the product of independent lognormals is lognormal. Once again the methodology of Part 1 is applicable.

Example: Refer to the Montana Overthrust Belt province.

1. The gas-oil ratio for this province was estimated to be

   $a = 2{,}200$ cu ft/bbl = 2.2 trillion cu ft/billion bbls

2. The marginal probability of associated-dissolved gas is $p = 0.86$ which is the same marginal probability as the oil with which it occurs.

3. Conditional upon recoverable associated-dissolved gas being present, conditional resource estimates are in trillion cu ft

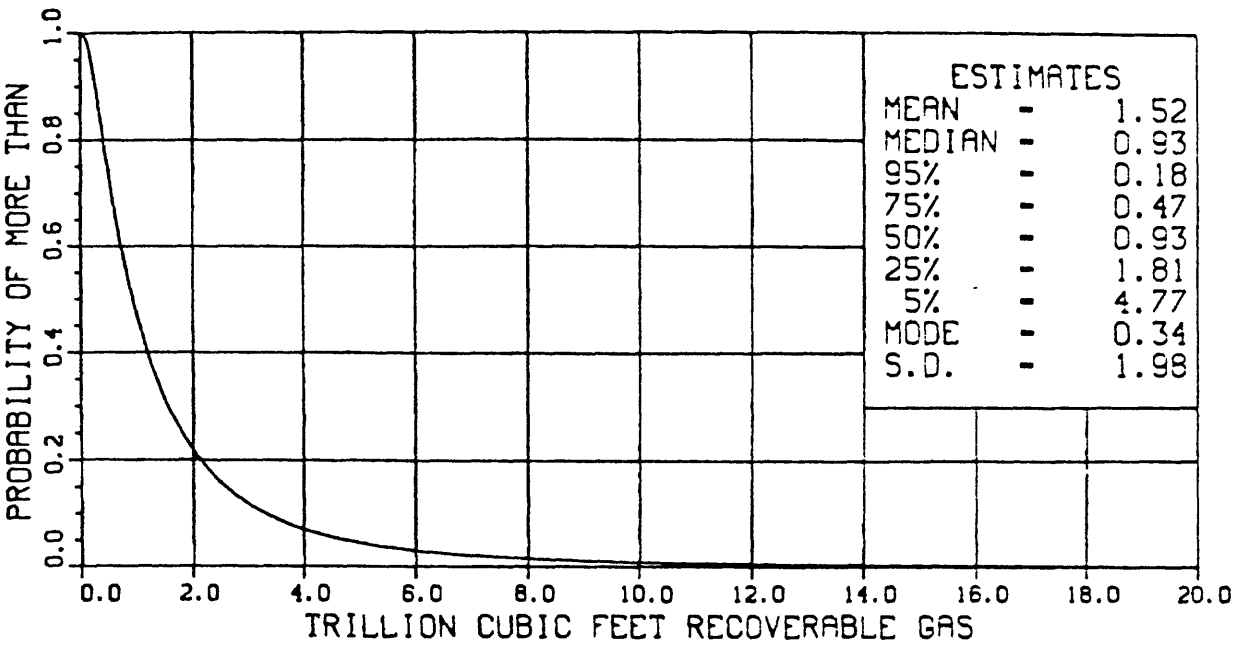
   a) ${}_cy_{0.95} = a\ {}_cx_{0.95} = 2.2(0.08) = 0.18$, a low resource estimate

   b) ${}_cy_{0.05} = a\ {}_cx_{0.05} = 2.2(2.17) = 4.77$, a high resource estimate.

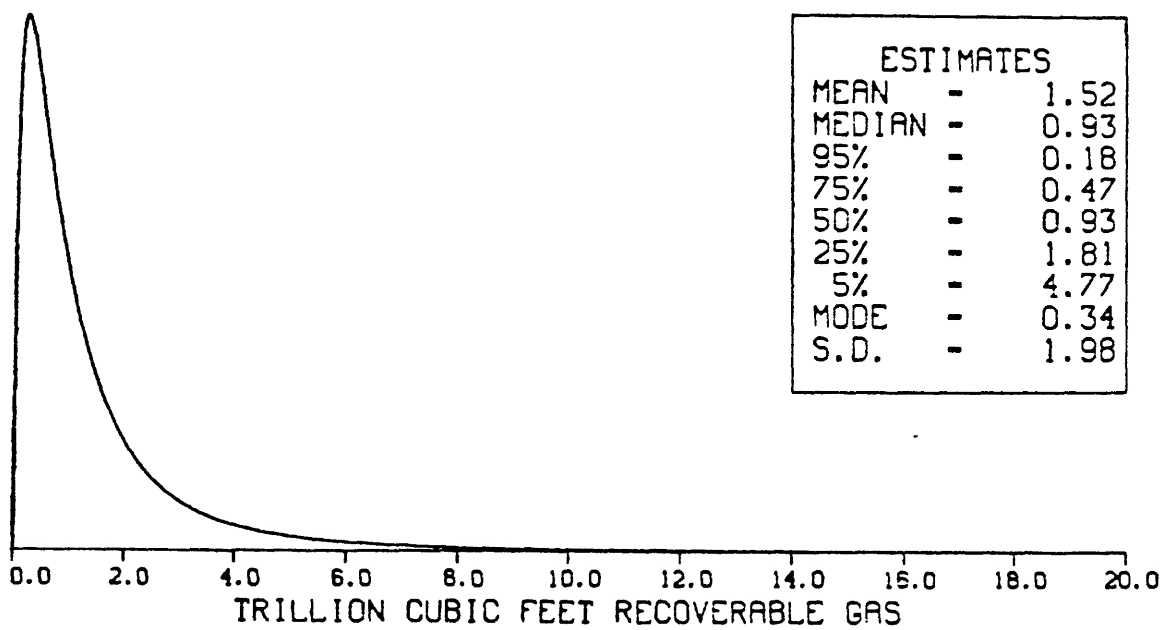
The methodology of Part 1 is now used. The conditional probability distribution of the quantity of undiscovered recoverable associated-dissolved gas for the Montana Overthrust Belt province is displayed in figure 3. The graph of the conditional more-than cdf is given in figure 3a and the graph of the conditional pdf in figure 3b. Some numerical characteristics in trillion cubic feet (TCF) of the lognormal distribution in figure 3 are the following:

$${}_cy_{0.95} = 2.2(0.08) = 0.18$$
$${}_cy_{0.75} = 2.2(0.21) = 0.46$$
$${}_cy_{0.50} = 2.2(0.42) = 0.92$$
$${}_cy_{0.25} = 2.2(0.82) = 1.80$$
$${}_cy_{0.05} = 2.2(2.17) = 4.77$$

$$\mu_{Y_c} = 2.2(0.69) = 1.52$$
$$m_{Y_c} = 2.2(0.42) = 0.92$$
$$M_{Y_c} = 2.2(0.15) = 0.33$$
$$\sigma^2_{Y_c} = 2.2^2(0.83) = 4.02\ (\text{TCF})^2$$
$$\sigma_{Y_c} = 2.2(0.91) = 2.00$$



(a) CONDITIONAL MORE-THAN CUMULATIVE DISTRIBUTION FUNCTION



(b) CONDITIONAL PROBABILITY DENSITY FUNCTION

Figure 3.-Conditional probability distribution of undiscovered recoverable associated-dissolved gas for Montana Overthrust Belt province. Estimates are mean, median, mode, standard deviation (S.D.) and fractiles that correspond to percentages listed.

As in Part 1 there are many similar relationships between the probability distribution of Y and its conditional distribution, i.e., between Y and Y_c.

The more-than cdf of Y

$$\bar{F}_Y(y) = \begin{cases} 1 & \text{for } y < 0 \\ p\ \bar{F}_{Y_c}(y) & \text{for } y \geq 0 \end{cases}$$

The cdf of Y

$$F_Y(y) = \begin{cases} 0 & \text{for } y < 0 \\ 1 - p + pF_{Y_c}(y) & \text{for } y \geq 0 \end{cases}$$

The pdf of Y

$$f_Y(y) = \begin{cases} 1 - p & \text{for } y = 0 \\ pf_{Y_c}(y) & \text{for } y > 0 \\ 0 & \text{otherwise} \end{cases}$$

where $f_Y(0) \equiv P(Y = 0)$ for the discrete part of the distribution.

The mean of Y

$$\mu_Y = p\ \mu_{Y_c}$$

The standard deviation of Y

$$\sigma_Y = \sqrt{p\sigma^2_{Y_c} + p(1 - p)\mu^2_{Y_c}}$$

Note that

$$\mu_Y = p\mu_{Y_c} = pa\mu_{X_c} = a(p\mu_{X_c}) = a\mu_X$$

Example: Refer to the Montana Overthrust Belt province.

The methodology of Part 1 is used as it was for oil. The probability distribution of the quantity of undiscovered recoverable associated-dissolved gas is displayed in figure 4. The graph of the more-than cdf (fig. 4a) has the value of the marginal probability, p = 0.86, at the origin. The graph of the pdf (fig. 4b) has a spike at the origin of probability weight 1 - 0.86 = 0.14 which represents the chance of no associated-dissolved gas being present in recoverable quantities. Some numerical characteristics (in trillion cu ft) of this distribution using a computer program based upon the methodology of Part 1 are the following:

$y_{0.95} = 0.00$ $\qquad \mu_Y = 1.31$

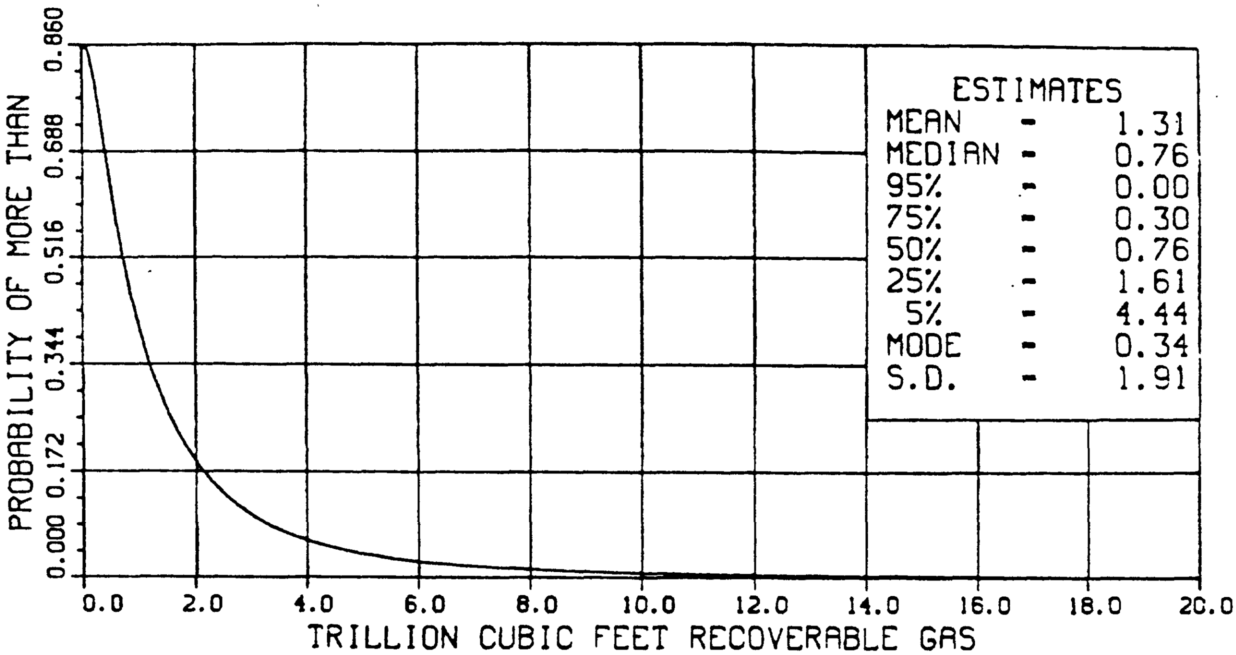
$y_{0.75} = 0.30$ $\qquad m_Y = 0.76$

$y_{0.50} = 0.76$ $\qquad \sigma_Y^2 = 3.65 \quad (\mathrm{TCF})^2$

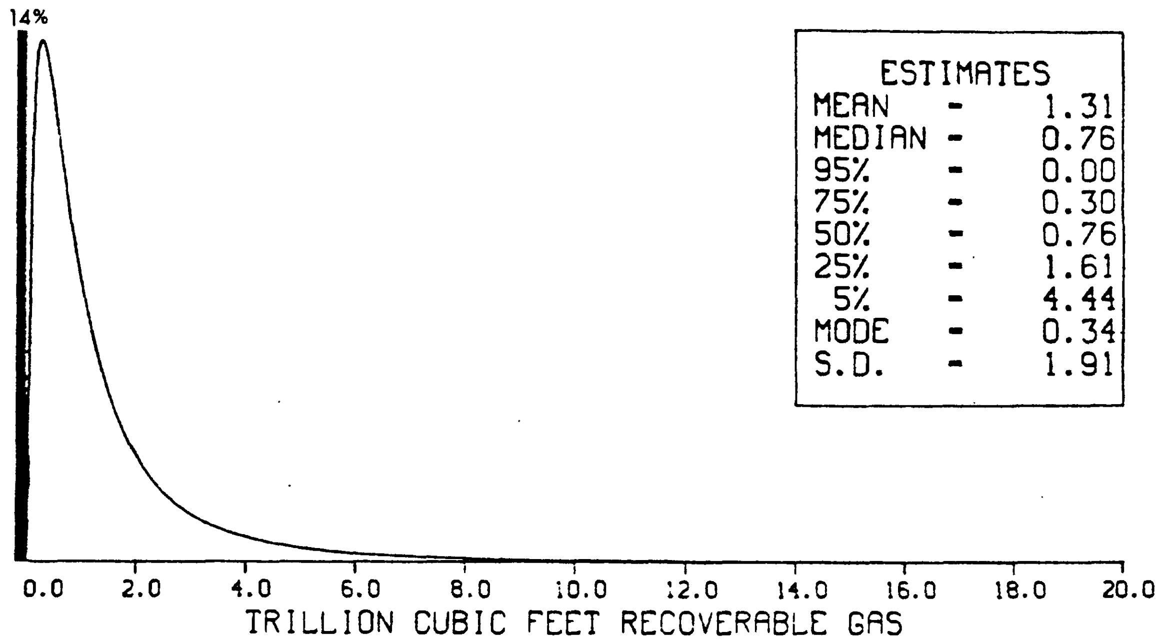
$y_{0.25} = 1.61$ $\qquad \sigma_Y = 1.91$

$y_{0.05} = 4.44$

The mean estimate of the quantity of undiscovered recoverable associated-dissolved gas for the Montana Overthrust Belt province is 1.31 TCF. The interval estimate is 0.00 to 4.44 TCF.



(a) MORE-THAN CUMULATIVE DISTRIBUTION FUNCTION



(b) PROBABILITY DENSITY FUNCTION

Figure 4.-Probability distribution of undiscovered recoverable associated-dissolved gas for Montana Overthrust Belt province. Estimates are mean, median, mode, standard deviation (S.D.) and fractiles that correspond to percentages listed.

Approach II: Using the results of Part 1.

The problem is now approached using the results of Part 1, i.e., the probability distribution of X. There are many well known relationships between the random variables X and Y:

The mean of Y

$$\mu_Y = a\ \mu_X$$

The median of Y

$$m_Y = a\ m_X$$

The variance of Y

$$\sigma_Y^2 = a^2\ \sigma_X^2$$

The standard deviation of Y

$$\sigma_Y = a\ \sigma_X$$

The more-than cdf of Y

$$\bar{F}_Y(y) = \bar{F}_X(y/a)$$

Because

$$P(Y > y) = P(aX > y) = P(X > y/a).$$

Hence, we have

$$\alpha = P(Y > y_\alpha) = P(X > y_\alpha/a).$$

Recalling that

$$P(X > x_\alpha) = \alpha,$$

the fractiles of Y are related to the fractiles of X by

$$y_\alpha = a\ x_\alpha$$

Examples:

a) $y_{0.95} = a\ x_{0.95}$

b) $y_{0.50} = a\ x_{0.50}$

c) $y_{0.05} = a\ x_{0.05}$

The cdf of Y

$$F_Y(y) = F_X(y/a)$$

since $P(Y \leq y) = P(aX \leq y) = P(X \leq y/a)$.

The pdf of Y

$$f_Y(y) = \begin{cases} 1 - p & \text{for } y = 0 \\ \frac{1}{a} f_X(y/a) & \text{for } y > 0 \\ 0 & \text{otherwise} \end{cases}$$

where $f_Y(0) \equiv P(Y = 0)$ for the discrete part of the distribution.

Note that

$$\int_{0+}^{\infty} f_Y(y)dy = \int_{0+}^{\infty} \frac{1}{a} f_X(y/a)dy = \int_{0+}^{\infty} f_X(x)dx = p.$$

The graph of a typical density function of Y would also have a spike at the origin of probability weight 1 - p.

Example: Refer to the Montana Overthrust Belt province.

The probability distribution of the quantity of undiscovered recoverable associated-dissolved gas is now considered using the results of Part 1, i.e., the probability distribution of the quantity of undiscovered recoverable oil. Some numerical characteristics (in trillion cu ft) of this distribution are the following:

$y_{0.95} = 2.2(0.00) = 0.00$

$y_{0.75} = 2.2(0.13) = 0.29$

$y_{0.50} = 2.2(0.34) = 0.75$

$y_{0.25} = 2.2(0.73) = 1.61$

$y_{0.05} = 2.2(2.02) = 4.44$

$\mu_Y = 2.2(0.59) = 1.30$

$m_Y = 2.2(0.34) = 0.75$

$\sigma_Y^2 = 2.2^2(0.77) = 3.73 \quad (\mathrm{TCF})^2$

$\sigma_Y = 2.2(0.88) = 1.94$

The difference in the answers between Approaches I and II in this illustrative example is due to round-off error. The computer calculations are much more precise. Nevertheless, the answers agree, and Approach II can be used as a quick check on Approach I.

Petroleum Resources in Portion of a Province

The methodology of Part 2 could also be used in an entirely different situation. Suppose we are interested in estimating the quantity of undiscovered resource in only part of a province or basin that has already been assessed. Let us make a resource appraisal of an area of interest from a province resource appraisal. The area itself is a fraction of the province. Whereas, the area resource is also a fraction of the province resource. The resource can be oil, non-associated gas or associated-dissolved gas. Let

X: Quantity of undiscovered resource in the province, i.e., the province resource.

Y: Quantity of undiscovered resource in the area, i.e., the area resource.

Assume that the area resource is related to the province resource by

$$Y = aX$$

where

a: The fraction of the province resource that is in the area, i.e., the area resource fraction; $0 < a \leq 1$.

Note that a is taken to be an estimated constant, whereas X is a random variable. When $a = 1$, the area resource appraisal would be the same as the province resource appraisal.

Example: Suppose it is estimated that 15% of the province resource is in the area. Then $a = 0.15$ and

$$Y = 0.15X.$$

Realize that under this model no matter what the province resource is, 15% of the province resource is in the area of interest.

Treating Y as a random variable with range $y \geq 0$, a judgmental probability distribution of Y can be determined in exactly the same way as it was done for associated-dissolved gas.

As in the case of the GOR, the area resource fraction could also be treated as a random variable and estimated by a "range estimate."

Example :

1. Consider an offshore province in the Gulf of Mexico called the Western Gulf Shelf province. A shelf province extends from shoreline to 200 meters water depth. Let

X: Quantity of undiscovered recoverable non-associated gas.

The marginal probability of non-associated gas was estimated to be 100 percent, i.e., p = 1.00. In this case, the probability distribution of X is the same as the conditional probability distribution of X.

Conditional upon recoverable non-associated gas being present, conditional resource estimates were made in trillion cu ft

a) ${}_{c}x_{0.95} = 16.36$, a low resource estimate

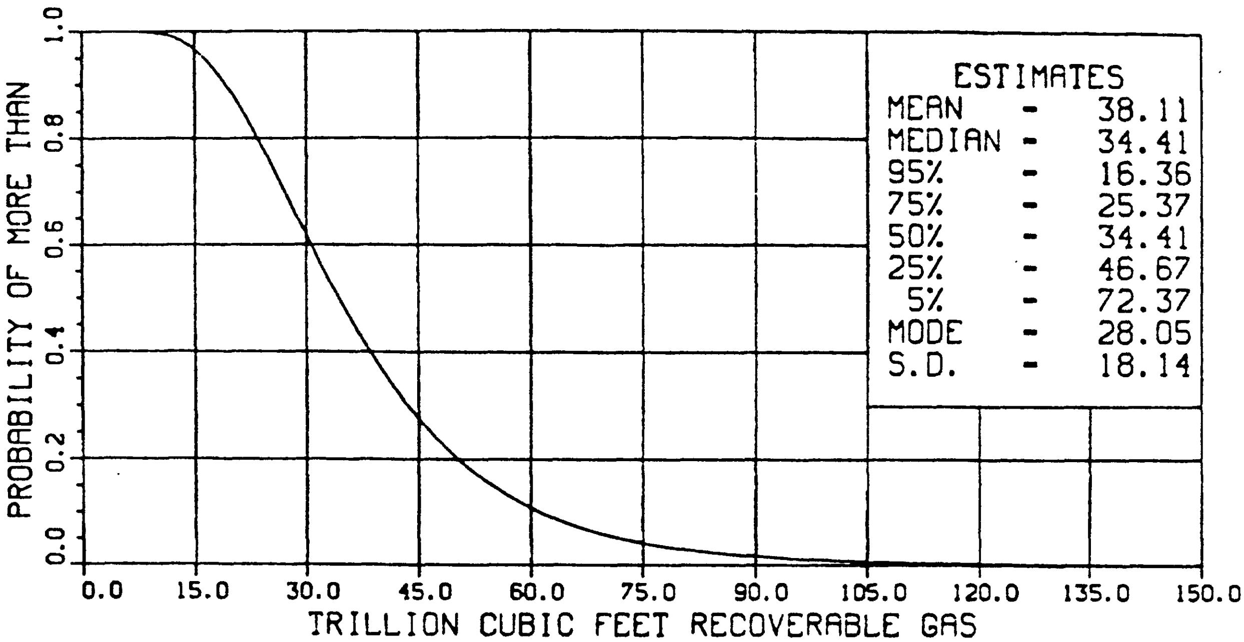
b) ${}_{c}x_{0.05} = 72.37$, a high resource estimate

The methodology of Part 1 was applied by means of a computer program. The probability distribution of the quantity of undiscovered recoverable non-associated gas for the Western Gulf Shelf province is displayed in figure 5. The graph of the more-than cdf (fig. 5a) has the value of the marginal probability, p = 1.00, at the origin. The graph of the pdf (fig. 5b) has no spike of probability weight at the origin. Some numerical characteristics (in TCF) of this lognormal distribution are the following:

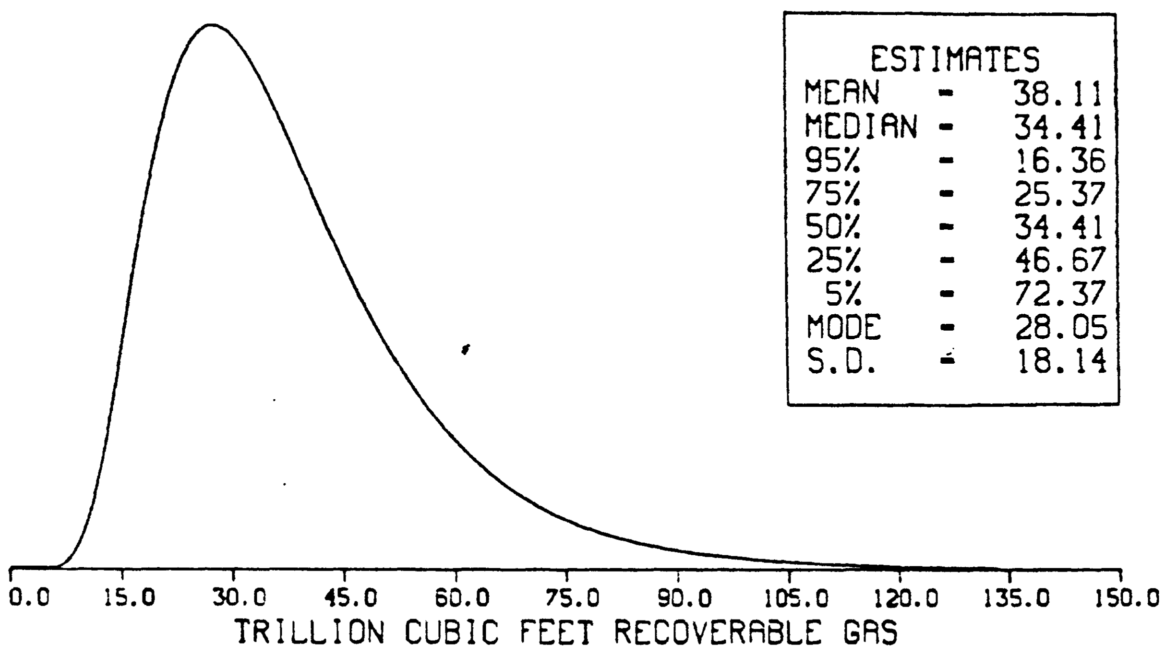
$x_{0.95} = 16.36$	$\mu_X = 38.11$
$x_{0.75} = 25.37$	$m_X = 34.41$
$x_{0.50} = 34.41$	$\sigma_X^2 = 329.06\ (\text{TCF})^2$
$x_{0.25} = 46.67$	$\sigma_X = 18.14$
$x_{0.05} = 72.37$	

Mean estimate of 38.11 TCF.

Interval estimate of 16.36 to 72.37 TCF.



(a) More-than cumulative distribution function



(b) Probability density function

Figure 5.-Probability distribution of undiscovered recoverable non-associated gas for Western Gulf Shelf province. Estimates are mean, median, mode, standard deviation (S.D.) and fractiles that correspond to percentages listed.

2. The Western Gulf Shelf province is divided into two areas of interest, called the Texas Shelf and the Louisiana Shelf. It was estimated that 23% of the undiscovered recoverable non-associated gas in the Western Gulf Shelf province was located in the Texas Shelf, and 77% in the Louisiana Shelf.

(a) Texas Shelf

The probability distribution of the quantity of undiscovered recoverable non-associated gas for the Texas Shelf is displayed in figure 6. Some numerical characteristics (in TCF) of this lognormal distribution are the following:

$y_{0.95} = 0.23(16.36) = 3.76$

$y_{0.75} = 0.23(25.37) = 5.83$

$y_{0.50} = 0.23(34.41) = 7.91$

$y_{0.25} = 0.23(46.67) = 10.73$

$y_{0.05} = 0.23(72.37) = 16.65$

$\mu_Y = 0.23(38.11) = 8.76$

$m_Y = 0.23(34.41) = 7.91$

$\sigma_Y^2 = 0.23^2(329.06) = 17.41\ (\text{TCF})^2$

$\sigma_Y = 0.23(18.14) = 4.17$

Mean estimate of 8.76 TCF

Interval estimate of 3.76 to 16.65 TCF

(b) Louisiana Shelf

The probability distribution of the quantity of undiscovered recoverable non-associated gas for the Louisiana Shelf is displayed in figure 7. Some numerical characteristics (in TCF) of this lognormal distribution are the following:

$y_{0.95} = 0.77(16.36) = 12.60$

$y_{0.75} = 0.77(25.37) = 19.53$

$y_{0.50} = 0.77(34.41) = 26.49$

$y_{0.25} = 0.77(46.67) = 35.94$

$y_{0.05} = 0.77(72.37) = 55.72$

$\mu_Y = 0.77(38.11) = 29.34$

$m_Y = 0.77(34.41) = 26.49$

$\sigma_Y^2 = 0.77^2(329.06) = 195.10\ (\text{TCF})^2$

$\sigma_Y = 0.77(18.14) = 13.97$

Mean estimate of 29.34 TCF

Interval estimate of 12.60 to 55.72 TCF.

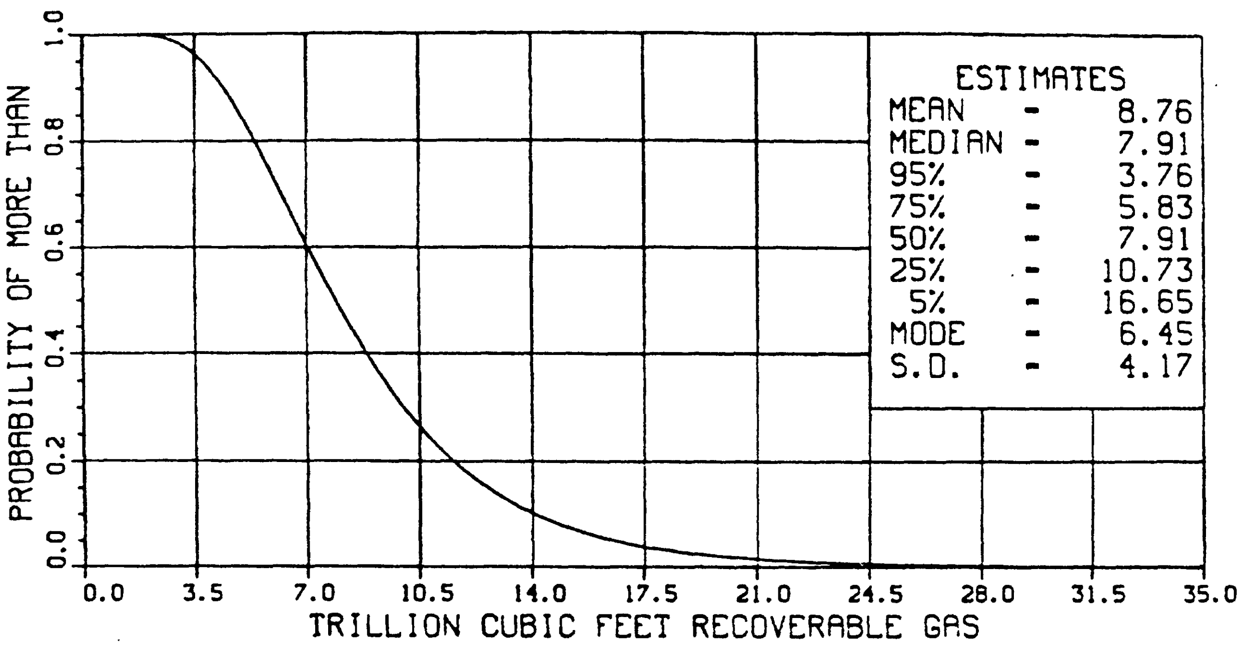


Figure 6.-Probability distribution of undiscovered recoverable non-associated gas for Texas Shelf. Estimates are mean, median, mode, standard deviation (S.D.) and fractiles that correspond to percentages listed.

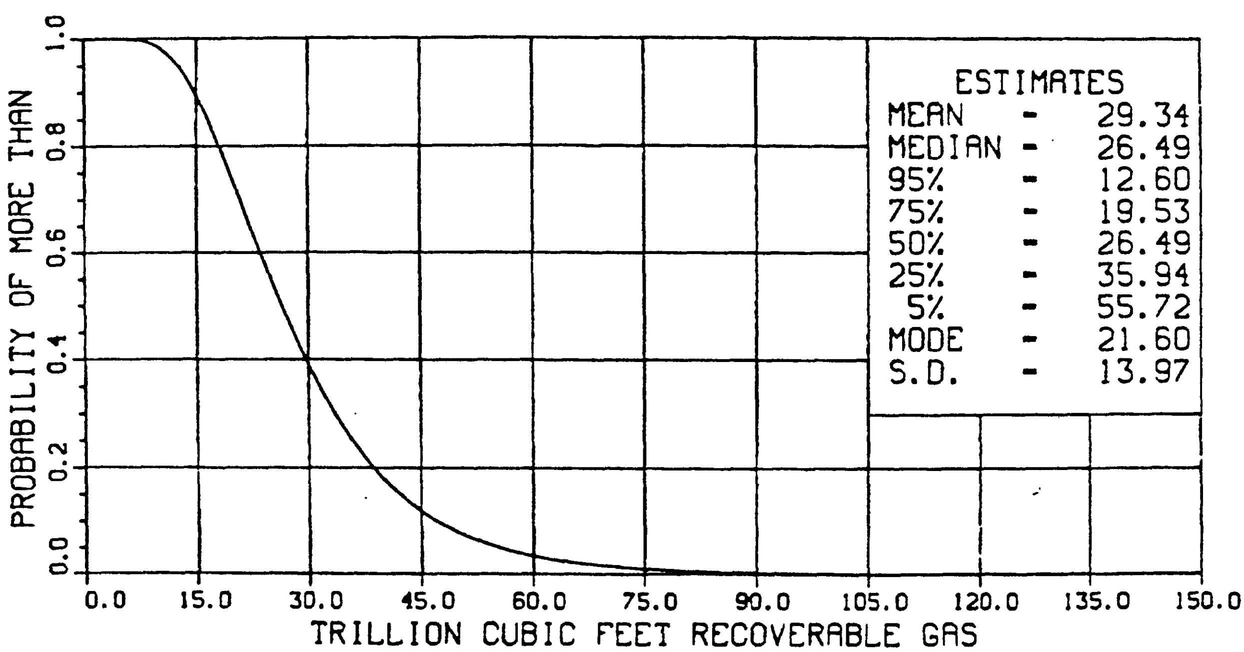


Figure 7.-Probability distribution of undiscovered recoverable non-associated gas for Louisiana Shelf. Estimates are mean, median, mode, standard deviation (S.D.) and fractiles that correspond to percentages listed.

PART 3. OIL PLUS ITS ASSOCIATED-DISSOLVED GAS

The resource of interest to be estimated is the undiscovered oil plus its associated-dissolved gas in a province. We will derive a judgmental probability distribution of oil plus its gas. From this judgmental resource distribution and its properties, estimates of the quantity of undiscovered oil plus its associated-dissolved gas are obtained.

In order to combine oil with its associated-dissolved gas, the gas is converted into an equivalent amount of oil with respect to energy, based on the BTU content of the gas to its oil. Thus, the conversion depends upon the quality of the gas to its oil. A typical range of conversions is 1 barrel of oil to 5,000-6,500 ft^3 of gas. A rich gas (high BTU content) would correspond to the 5,000 ft^3 of gas. An average conversion factor would be 1 bbl/6,000 ft^3. When gas is converted into an equivalent amount of oil, the unit reported is

BOE: Barrels of oil-equivalent.

Let

Y': Quantity of undiscovered associated-dissolved gas expressed in BOE's. Associated-dissolved gas is converted into an equivalent amount of oil by

$Y' = bY$

where

b: BOE conversion factor for the province, $b > 0$

Y: Quantity of undiscovered associated-dissolved gas, $y \geq 0$

Note that b is a constant, whereas Y is the random variable from Part 2. Because $Y = aX$, we get

$Y' = bY = b(aX) = abX.$

The unitless constant ab is the fraction of the quantity of undiscovered oil that would represent the quantity of undiscovered associated-dissolved gas in BOE. This linear model that relates Y' to X is the same type of model as in Part 2; the coefficient of X is ab instead of a. Therefore, the methodology of Part 2 applies in this case. Hence, treating Y' as a random variable with range $y' \geq 0$, a judgmental probability distribution of Y' could be determined.

Let

Z: Quantity of undiscovered oil plus its associated-dissolved gas expressed in BOE's.

We have

$$Z = X + Y'$$

where

X : Quantity of undiscovered oil, $x \geq 0$

Y′: Quantity of undiscovered associated-dissolved gas in BOE, $y' \geq 0$.

Because $Y' = abX$, we get

$$Z = X + Y' = X + abX = (1 + ab)X$$

which can be written as

$$Z = dX$$

where $d = 1 + ab$.

Note that d is a unitless constant, whereas X is a random variable from Part 1. The constant d is the fraction of the quantity of undiscovered oil that would represent the quantity of undiscovered oil plus its associated-dissolved gas in BOE. This linear model that relates Z to X is again the same type of model as in Part 2; the coefficient of X is now d instead of a. Therefore, the methodology of Part 2 also applies in this case; simply replace Y by Z and a by d. Hence, treating Z as a random variable with range $z \geq 0$, a judgmental probability distribution of Z can be determined. From this distribution, estimates of the quantity of undiscovered oil plus its associated-dissolved gas are obtained.

The reader should realize that only one computer program is necessary for Parts 1, 2 and 3. The difference lies in the input and its interpretation.

It is important when reporting estimates of oil plus its associated-dissolved gas in BOE, to also give the fraction that is oil and the fraction that is associated-dissolved gas in BOE. The fraction that is oil (or percent oil) is

$$\frac{X}{Z} = \frac{X}{dX} = \frac{1}{d}$$

and the fraction that is associated-dissolved gas (or percent gas) is

$$\frac{Y'}{Z} = \frac{abX}{dX} = \frac{ab}{d} = 1 - \frac{1}{d}$$

Example: Refer to the Montana Overthrust Belt province.

1. The gas-oil ratio for this province was estimated to be

$$a = 2{,}200 \text{ cu ft/bbl} = 2.2 \text{ trillion cu ft/billion bbls}$$

The BOE conversion factor for this province was estimated to be

$$b = 1 \text{ bbl}/6{,}000 \text{ cu ft} = 1 \text{ billion bbls}/6 \text{ trillion cu ft}$$

Hence, we get

$$ab = \left(2.2\ \frac{\text{trillion cu ft}}{\text{billion bbls}}\right)\left(\frac{1}{6}\ \frac{\text{billion bbls}}{\text{trillion cu ft}}\right) = 0.367$$

or

$$ab = \left(2{,}200\ \frac{\text{cu ft}}{\text{bbl}}\right)\left(\frac{1}{6{,}000}\ \frac{\text{bbl}}{\text{cu ft}}\right) = 0.367$$

Note that ab is a unitless number and that the quantity of undiscovered associated-dissolved gas in BOE is equal to 36.7% of the quantity of undiscovered oil.

Therefore, we have

$$d = 1 + ab = 1 + 0.367 = 1.367$$

The fraction of oil plus its associated-dissolved gas in BOE that is oil equals

$$\frac{1}{d} = \frac{1}{1.367} = 0.73$$

and the fraction that is associated-dissolved gas equals

$$1 - \frac{1}{d} = 1 - 0.73 = 0.27$$

i.e., 73% oil and 27 % gas.

2. The marginal probability of oil plus its associated-dissolved gas is $p = 0.86$, which is the same as the marginal probability of oil.

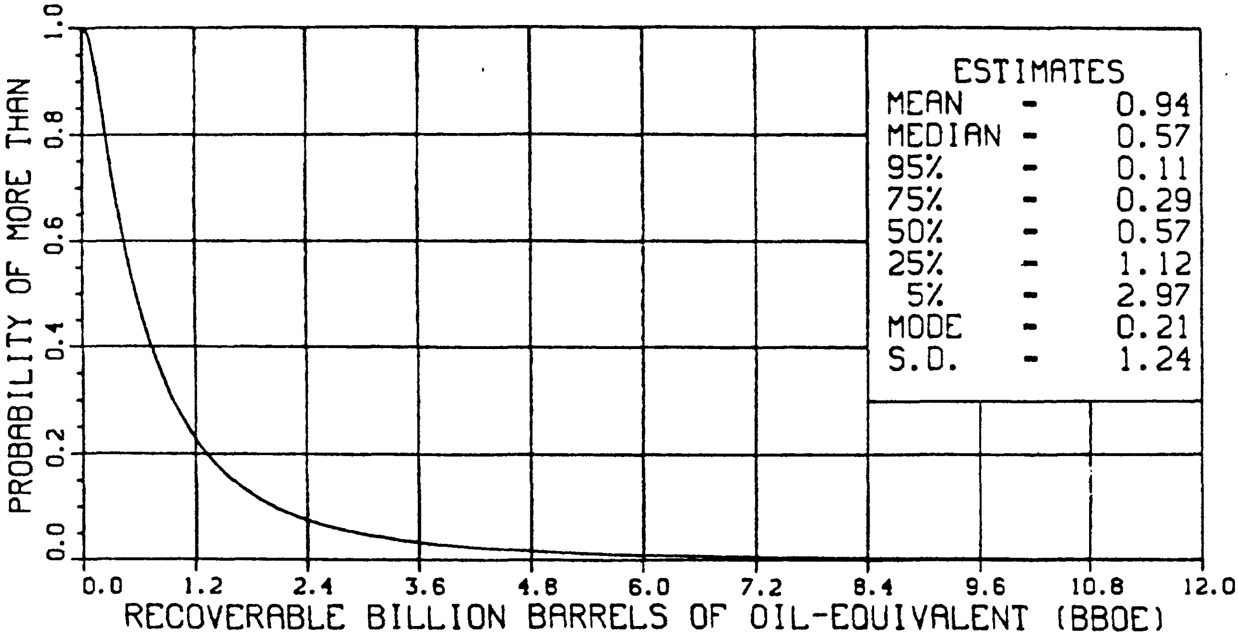
3. Conditional upon recoverable oil plus its associated-dissolved gas being present, conditional resource estimates are in billion BOE

   a) ${}_{c}z_{0.95} = d\,{}_{c}x_{0.95} = 1.367(0.08) = 0.11$, a low resource estimate

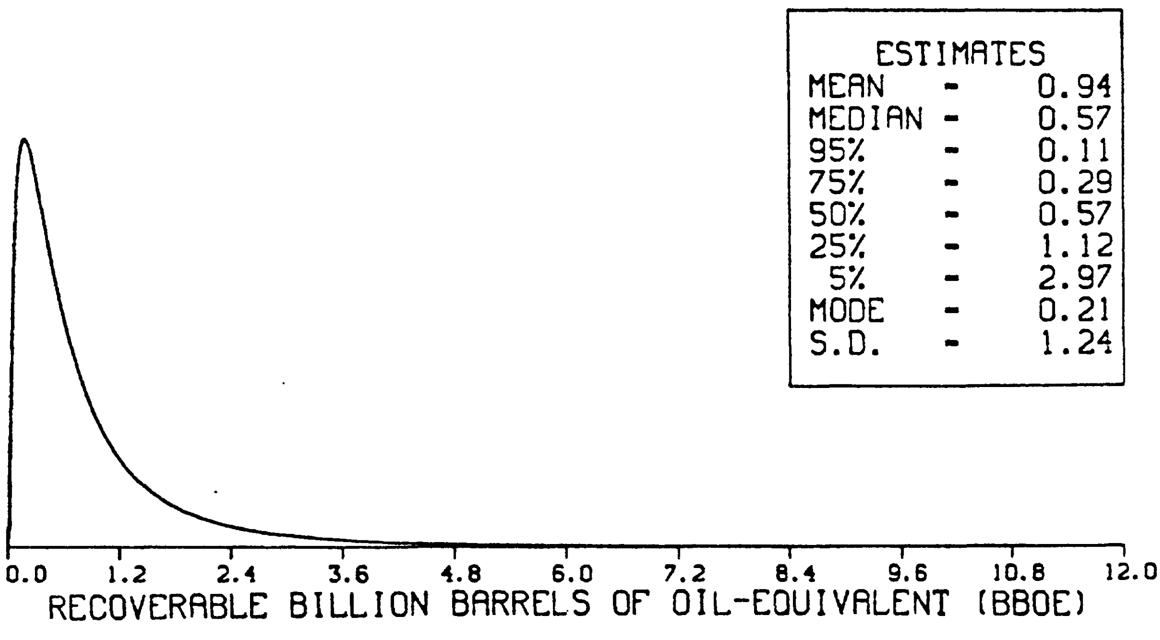
   b) ${}_{c}z_{0.05} = d\,{}_{c}x_{0.05} = 1.367(2.17) = 2.97$, a high resource estimate

The methodology of Part 1 is now used. The conditional probability distribution of the quantity of undiscovered recoverable oil plus its associated-dissolved gas in billion BOE (BBOE) for the Montana Overthrust Belt province is displayed in figure 8. The graph of the conditional more-than cdf is given in figure 8a and the graph of the conditional pdf in figure 8b. Some numerical characteristics (in billion BOE) of the lognormal distribution in figure 8 are the following:

${}_{c}z_{0.95} = 1.367(0.08) = 0.11$	$\mu_{Z_c} = 1.367(0.69) = 0.94$
${}_{c}z_{0.75} = 1.367(0.21) = 0.29$	$m_{Z_c} = 1.367(0.42) = 0.57$
${}_{c}z_{0.50} = 1.367(0.42) = 0.57$	$M_{Z_c} = 1.367(0.15) = 0.21$
${}_{c}z_{0.25} = 1.367(0.82) = 1.12$	$\sigma^2_{Z_c} = 1.367^2(0.83) = 1.55\ (\text{BBOE})^2$
${}_{c}z_{0.05} = 1.367(2.17) = 2.97$	$\sigma_{Z_c} = 1.367(0.91) = 1.24$



(a) CONDITIONAL MORE-THAN CUMULATIVE DISTRIBUTION FUNCTION



(b) CONDITIONAL PROBABILITY DENSITY FUNCTION

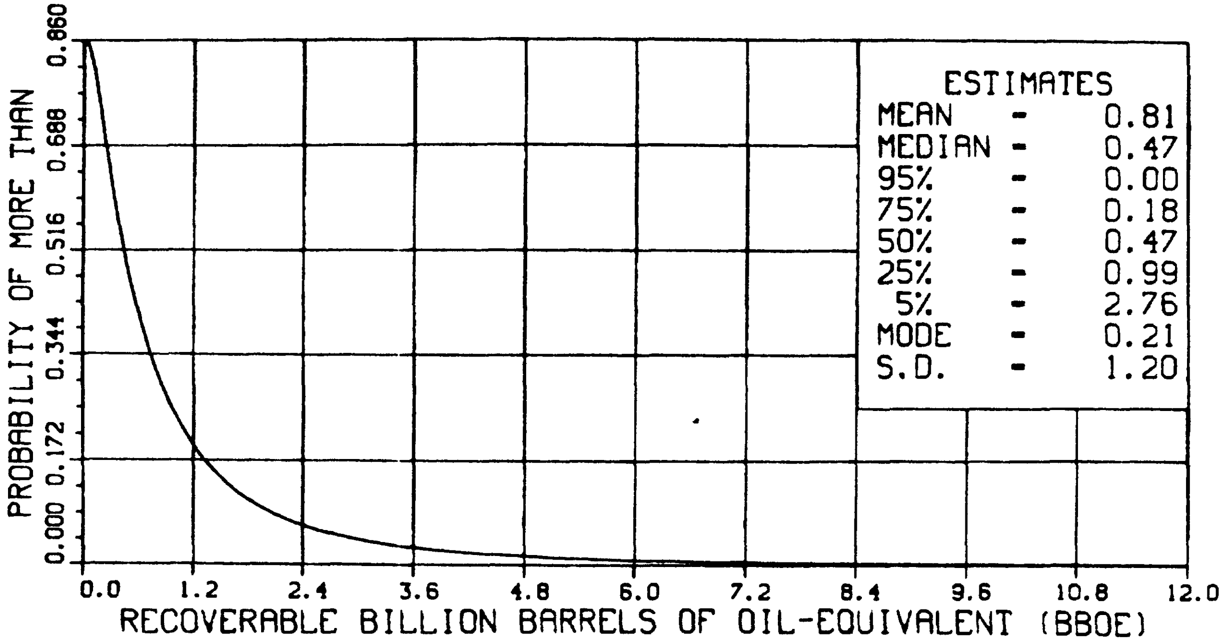
Figure 8.-Conditional probability distribution of undiscovered recoverable oil plus its associated-dissolved gas in BBOE for Montana Overthrust Belt province. Estimates are mean, median, mode, standard deviation (S.D.) and fractiles that correspond to percentages listed.

Example: Refer to the Montana Overthrust Belt province.

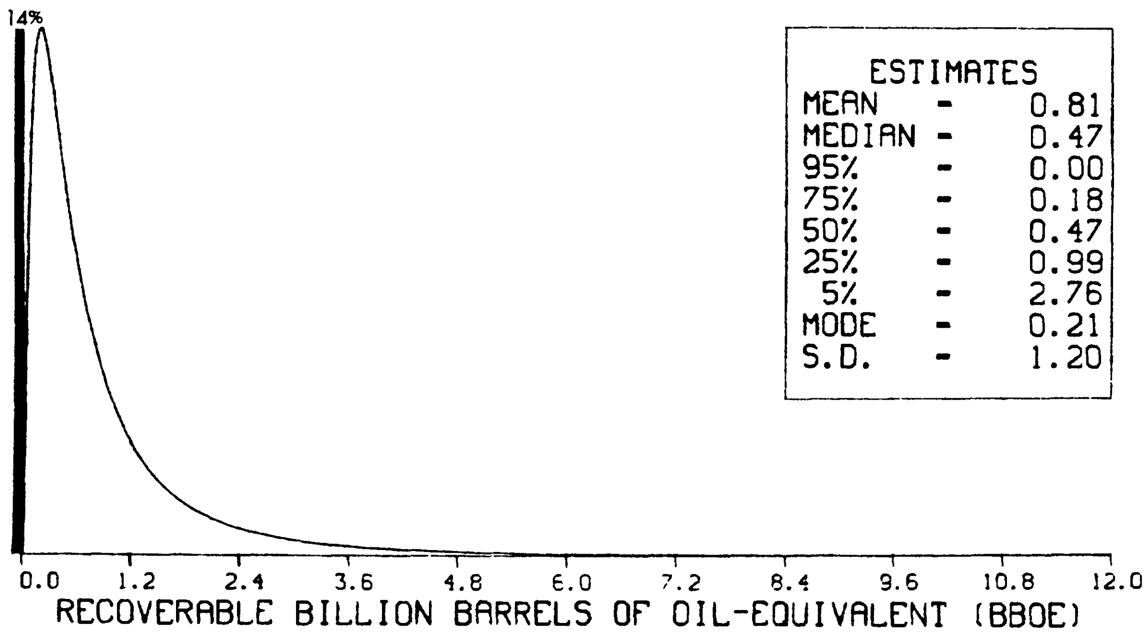
The methodology of Part 1 is used as it was for oil. The probability distribution of the quantity of undiscovered recoverable oil plus its associated-dissolved gas in billion BOE for the Montana Overthrust Belt province is displayed in figure 9. The graph of the more-than cdf (fig. 9a) has the value of the marginal probability, p = 0.86, at the origin. The graph of the pdf (fig. 9b) has a spike at the origin of probability weight 1 - 0.86 = 0.14, which represents the chance of no oil plus its associated-dissolved gas being present in recoverable quantities. Some numerical characteristics (in billion BOE) of this distribution using a computer program based upon the methodology of Part 1 are the following:

$z_{0.95} = 0.00$	$\mu_Z = 0.81$
$z_{0.75} = 0.18$	$m_Z = 0.47$
$z_{0.50} = 0.47$	$\sigma_Z^2 = 1.44 \quad (\text{BBOE})^2$
$z_{0.25} = 0.99$	$\sigma_Z = 1.20$
$z_{0.05} = 2.76$	

The mean estimate of the quantity of undiscovered recoverable oil plus its associated-dissolved gas for the Montana Overthrust Belt province is 0.81 billion BOE. The interval estimate is 0.00 to 2.76 billion BOE.



(a) MORE-THAN CUMULATIVE DISTRIBUTION FUNCTION



(b) PROBABILITY DENSITY FUNCTION

Figure 9.-Probability distribution of undiscovered recoverable oil plus its associated-dissolved gas in BBOE for Montana Overthrust Belt province. Estimates are mean, median, mode, standard deviation (S.D.) and fractiles that correspond to percentages listed.

Example: Refer to the Montana Overthrust Belt province.

The probability distribution of the quantity of undiscovered recoverable oil plus its associated-dissolved gas in billion BOE is now considered using the results of Part 1, i.e., the probability distribution of the quantity of undiscovered recoverable oil. Some numerical characteristics (in billion BOE) of this distribution are the following:

$z_{0.95} = 1.367(0.00) = 0.00$

$z_{0.75} = 1.367(0.13) = 0.18$

$z_{0.50} = 1.367(0.34) = 0.46$

$z_{0.25} = 1.367(0.73) = 1.00$

$z_{0.05} = 1.367(2.02) = 2.76$

$\mu_Z = 1.367(0.59) = 0.81$

$m_Z = 1.367(0.34) = 0.46$

$\sigma_Z^2 = 1.367^2(0.77) = 1.44 \text{ (BBOE)}^2$

$\sigma_Z = 1.367(0.88) = 1.20$

The slight difference in the answers between Approaches I and II in this illustrative example is due to round-off error. Nevertheless, the answers agree, and Approach II can be used as a quick check on Approach I.

PART 4. TOTAL GAS

The resource of interest to be estimated is the undiscovered total gas in a province. Total gas is all the undiscovered natural gas in a province, i.e., the total of the non-associated gas and the associated-dissolved gas. Total gas is also referred to as the aggregation of non-associated gas and associated-dissolved gas. We will derive a probability distribution of total gas. From this judgmental resource distribution and its properties, estimates of the quantity of undiscovered total gas are obtained.

Let

T : Quantity of undiscovered total gas

Thus, we have by the definition of total gas

$$T = X + Y$$

where

X : Quantity of undiscovered non-associated gas

Y : Quantity of undiscovered associated-dissolved gas

Note that X is a random variable from Part 1, and Y is the random variable from Part 2. Assume that X and Y are independent random variables because we will take the position that oil and non-associated gas are separately and independently assessed in a province.

Example : Refer to the Montana Overthrust Belt province.

1. Quantity of undiscovered non-associated gas, X.

The marginal probability of non-associated gas was estimated to be 100 percent, i.e., $p_X = 1.00$. In this case, the probability distribution of X is the same as the conditional probability distribution of X.

Conditional upon recoverable non-associated gas being present, conditional resource estimates were made in trillion cu ft

a) ${}_c x_{0.95} = 1.78$, a low resource estimate

b) ${}_c x_{0.05} = 20.50$, a high resource estimate

The methodology of Part 1 was applied by means of a computer program. The probability distribution of the quantity of undiscovered recoverable non-associated gas for the Montana Overthrust Belt province is displayed in figure 10. Note that the graph of the more-than cdf has the value of the marginal probability, $p_X = 1.00$, at the origin. Some numerical characteristics (in trillion cu ft) of this lognormal distribution are the following:

$x_{0.95} = 1.78$

$x_{0.75} = 3.66$

$x_{0.50} = 6.04$

$x_{0.25} = 9.97$

$x_{0.05} = 20.50$

$\mu_X = 7.96$

$m_X = 6.04$

$\sigma_X^2 = 46.65 \text{ (TCF)}^2$

$\sigma_X = 6.83$

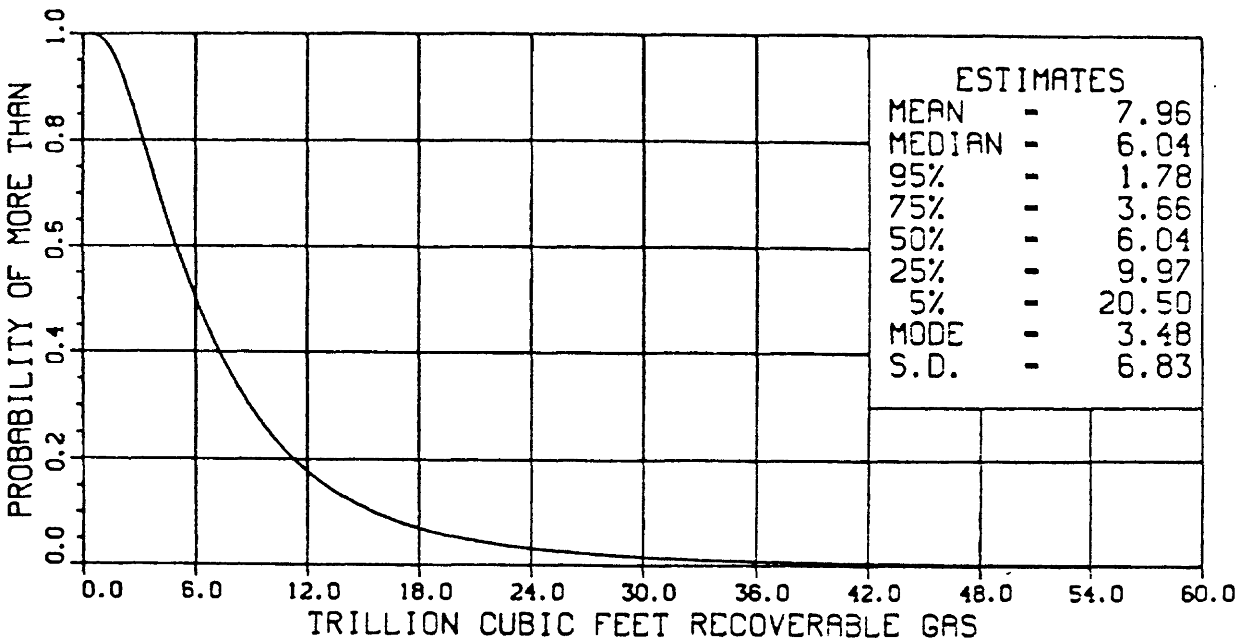


Figure 10.-Probability distribution of undiscovered recoverable non-associated gas for Montana Overthrust Belt province. Estimates are mean, median, mode, standard deviation (S.D.) and fractiles that correspond to percentages listed.

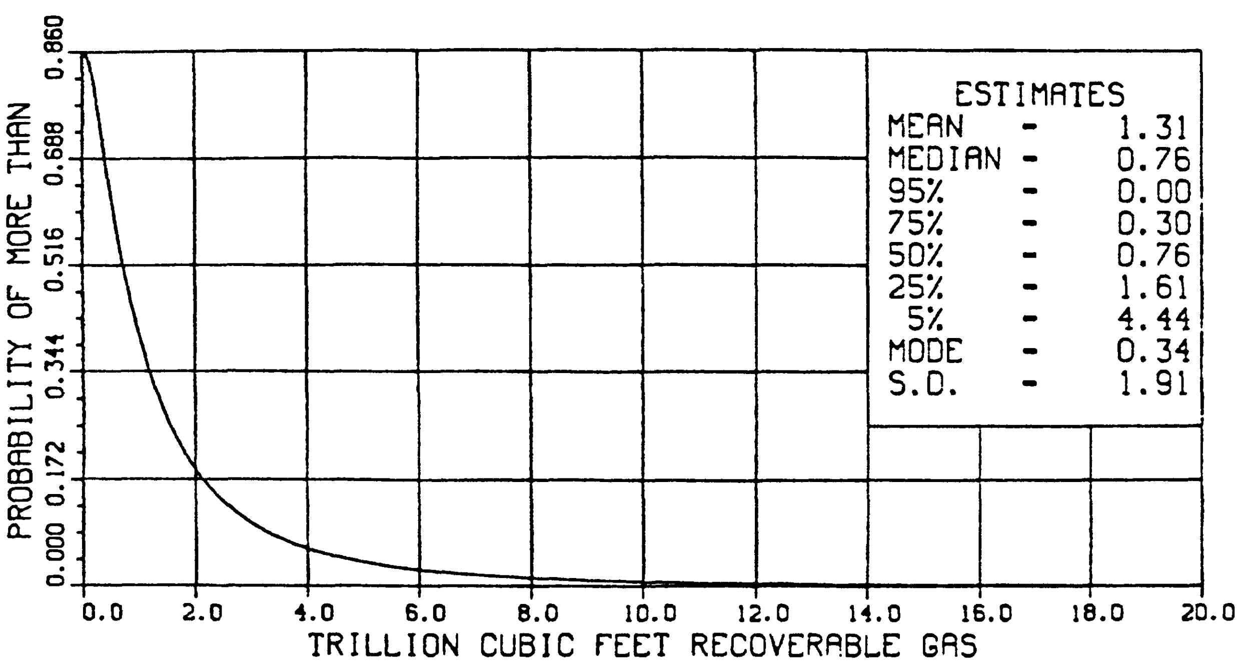


Figure 11.-Probability distribution of undiscovered recoverable associated-dissolved gas for Montana Overthrust Belt province. Estimates are mean, median, mode, standard deviation (S.D.) and fractiles that correspond to percentages listed.

2. Quantity of undiscovered associated-dissolved gas, Y.

The probability distribution of Y was determined in Part 2. However, the probability distribution of the undiscovered recoverable associated-dissolved gas for the Montana Overthrust Belt province is again displayed in figure 11 (same as figure 4a) for easy reference. Recall that the graph of the more-than cdf has the value of the marginal probability, $p_Y = 0.86$, at the origin. Some numerical characteristics (in trillion cu ft) of this distribution are the following:

$y_{0.95} = 0.00$	$\mu_Y = 1.31$
$y_{0.75} = 0.30$	$m_Y = 0.76$
$y_{0.50} = 0.76$	$\sigma_Y^2 = 3.65\ (\text{TCF})^2$
$y_{0.25} = 1.61$	$\sigma_Y = 1.91$
$y_{0.05} = 4.44$	

Treating T as a random variable with range $t \geq 0$, a judgmental probability distribution of T will be determined. The probability distribution of T is also referred to as the aggregate probability distribution of X and Y. The phrase "aggregating two components" is also used.

The marginal probability of total gas is the probability of any gas being present in the province. The phrase "gas is present" means either non-associated gas <u>or</u> associated-dissolved gas <u>or</u> both are present. Hence, the marginal probability of total gas can be written as

$$p_T = P(T > 0)$$

A property that expresses p_T in terms of p_X and p_Y is as follows.

Property $\quad p_T = 1 - (1 - p_X)(1 - p_Y)$

Proof:
$$\begin{aligned} p_T &= P(T > 0) = 1 - P(T = 0) \\ &= 1 - P(X + Y = 0) \\ &= 1 - P(X = 0,\ Y = 0) \\ &= 1 - P(X = 0)P(Y = 0) \text{ since X and Y are independent} \\ &= 1 - [1 - P(X > 0)][1 - P(Y > 0)] \\ &= 1 - (1 - p_X)(1 - p_Y) \end{aligned}$$

Note that if p_X or p_Y equals 1, then $p_T = 1$. Also, if $p_T = 1$, the probability distribution of T is the same as the conditional probability distribution of T.

Because the random variable T is the sum of two random variables X and Y, i.e., T = X + Y, there are many well known relationships between T and the random variables X and Y:

The mean of T

$$\mu_T = \mu_X + \mu_Y$$

The variance of T

$$\sigma_T^2 = \sigma_X^2 + \sigma_Y^2 \quad \text{if X and Y independent}$$

The standard deviation of T

$$\sigma_T = \sqrt{\sigma_X^2 + \sigma_Y^2} \quad \text{if X and Y independent}$$

Example : Refer to the Montana Overthrust Belt province.

Some numerical characteristics of the probability distribution of the undiscovered recoverable total gas for the Montana Overthrust Belt province are the following:

The marginal probability of total gas

$$p_T = 1 - (1 - p_X)(1 - p_Y) = 1 - (1 - 1.00)(1 - 0.86) = 1$$

The mean of T

$$\mu_T = \mu_X + \mu_Y = 7.96 + 1.31 = 9.27 \text{ TCF}$$

The variance of T

$$\sigma_T^2 = \sigma_X^2 + \sigma_Y^2 = 46.65 + 3.65 = 50.30 \text{ (TCF)}^2$$

The standard deviation of T

$$\sigma_T = \sqrt{\sigma_X^2 + \sigma_Y^2} = \sqrt{50.30} = 7.09 \text{ TCF}$$

The mean estimate of the quantity of undiscovered recoverable total gas in the Montana Overthrust Belt province is 9.27 trillion cu ft.

A Monte Carlo technique can be utilized to determine an approximate probability distribution of T in order to obtain the fractiles of T, t_α. This is a computer-oriented approach to the problem.

The idea behind the Monte Carlo aggregation computer program is as follows. A random value is selected for X, i.e., from the distribution of X in Part 1. A random value is selected for Y, i.e., from the distribution of Y in Part 2. The two random values are added together to give a random value for T. This operation is repeated a large number of times (called the sampling number) in order to generate an approximate probability distribution of T. The sampling number is in the order of thousands.

The more-than cdf of T is denoted by

$$\overline{F}_T(t) = P(T > t)$$

which is the probability that the quantity of undiscovered total gas is more than the amount t.

Example : Refer to the Montana Overthrust Belt province.

A Monte Carlo aggregation computer program was used to generate an approximate probability distribution of T. The number of samplings was 5,000. The approximate probability distribution of the quantity of undiscovered recoverable total gas for the Montana Overthrust Belt province is displayed in figure 12. Note that the graph of the more-than cdf has the value of the marginal probability, $p_T = 1.00$, at the origin. Some fractiles of T (in trillion cu ft) are the following:

$$t_{0.95} = 2.46$$

$$t_{0.75} = 4.78$$

$$t_{0.50} = 7.46$$

$$t_{0.25} = 11.79$$

$$t_{0.05} = 22.94$$

The interval estimate of the quantity of undiscovered recoverable total gas in the Montana Overthrust Belt province is 2.57 to 22.05 trillion cu ft.

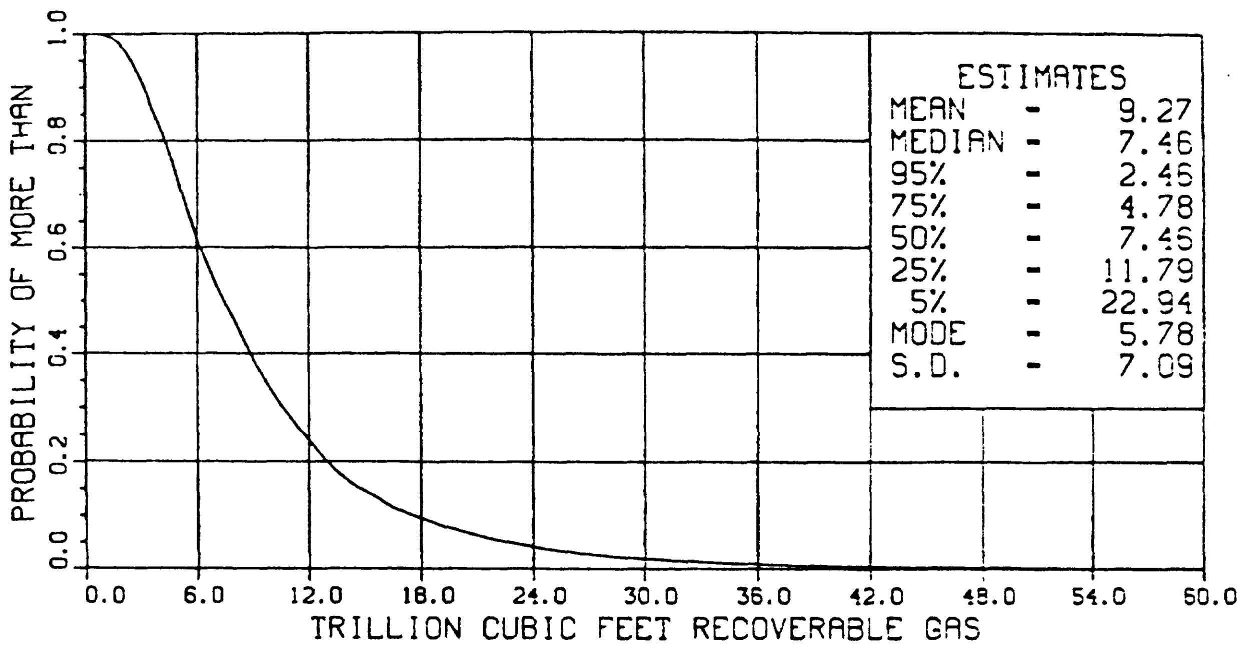


Figure 12.-Probability distribution of undiscovered recoverable total gas for Montana Overthrust Belt province. Estimates are mean, median, mode, standard deviation (S.D.) and fractiles that correspond to percentages listed.

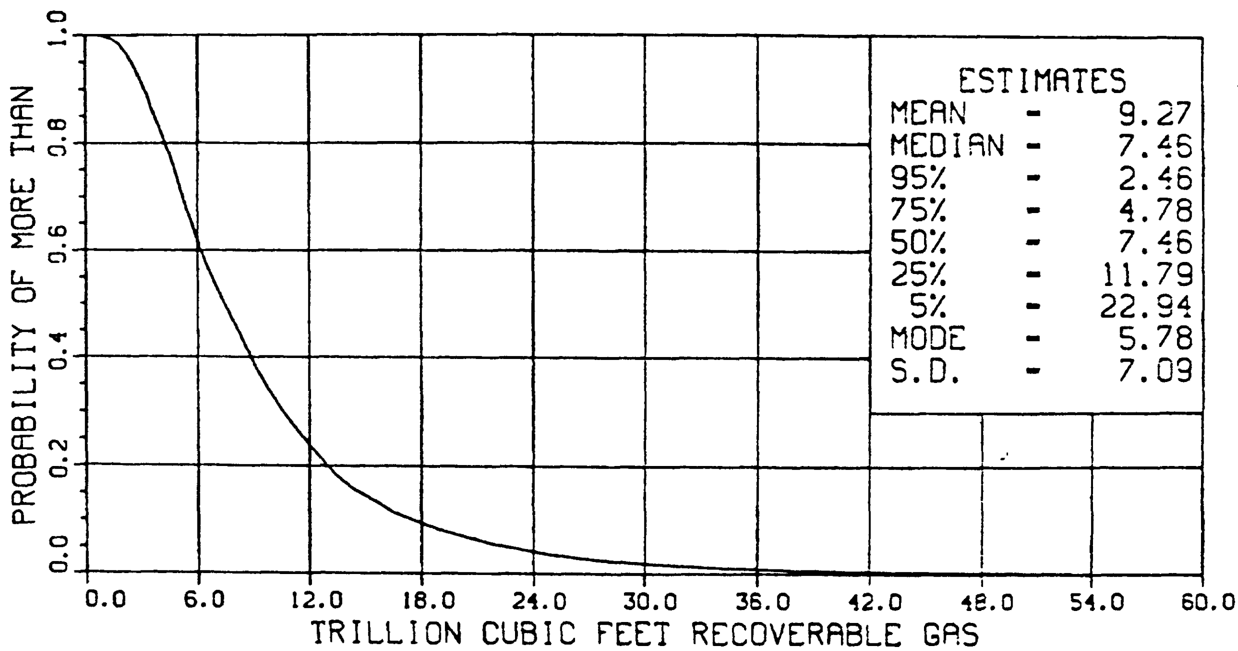


Figure 13.-Conditional probability distribution of undiscovered recoverable total gas for Montana Overthrust Belt province. Estimates are mean, median, mode, standard deviation (S.D.) and fractiles that correspond to percentages listed.

The conditional probability distribution of T can be determined from the probability distribution of T by using the methodology of Part 1 in reverse.

The conditional more-than cdf of T is denoted by

$$\overline{F}_{T_c}(t) = P(T > t \mid T > 0)$$

which is the probability that the quantity of undiscovered total gas is more than the amount t given that gas is present.

The following results are derived using the methodology of Part 1:

1) Because

$$\overline{F}_T(t) = p_T \overline{F}_{T_c}(t)$$

we have

$$\overline{F}_{T_c}(t) = \overline{F}_T(t)/p_T$$

2) Because

$$\mu_T = p_T \mu_{T_c}$$

we have

$$\mu_{T_c} = \mu_T / p_T$$

3) Because

$$\sigma_T^2 = p_T \sigma_{T_c}^2 + p_T(1 - p_T)\mu_{T_c}^2$$

we have

$$\sigma_{T_c}^2 = \sigma_T^2/p_T - (1 - p_T)\mu_{T_c}^2$$

Note that if $p_T = 1$, the formulas simplify to

1) $\overline{F}_{T_c}(t) = \overline{F}_T(t)$

2) $\mu_{T_c} = \mu_T$

3) $\sigma_{T_c}^2 = \sigma_T^2$

Example: Refer to the Montana Overthrust Belt province.

Because $p_T = 1$, the conditional probability distribution of T is the same as the probability distribution of T. Hence, the approximate conditional probability distribution of the quantity of undiscovered recoverable total gas for the Montana Overthrust Belt province in figure 13 is a duplicate of figure 12. However, some numerical characteristics (in TCF) of this distribution would be expressed as

$_ct_{0.95} = 2.46$ $\mu_{T_c} = 9.27$

$_ct_{0.75} = 4.78$ $m_{T_c} = 7.46$

$_ct_{0.50} = 7.46$ $\sigma^2_{T_c} = 50.30\ (TCF)^2$

$_ct_{0.25} = 11.79$ $\sigma_{T_c} = 7.09$

$_ct_{0.05} = 22.94$

PART 5. PETROLEUM RESOURCES IN TWO OR MORE PROVINCES

We are now interested in estimating the total resource in two or more provinces where the resource of interest is oil, non-associated gas, or associated-dissolved gas. Total resource is also referred to as the aggregation of the resource (e.g., oil) in the provinces. We will derive a probability distribution of total resource. From this judgmental resource distribution and its properties, estimates of the quantity of undiscovered total resource are obtained. From a methodology point of view, we can think of Part 5 as an extension of Part 4.

Let

W: Quantity of undiscovered total resource

with the resource of interest being oil, non-associated gas, or associated-dissolved gas in an area consisting of two or more provinces. Thus, we have by the definition of total resource

$$W = \sum_{i=1}^{n} X_i$$

where

X_i: Quantity of undiscovered resource in province i, $x_i \geq 0$

n : Number of provinces composing the area, $n \geq 2$

Note that X_i is a random variable from Part 1. Assume that the X_i's are independent random variables because we will take the position that the resource is separately and independently assessed from province to province.

The following notation will be used:

p_i: Marginal probability for province i

μ_i: Mean of X_i

m_i: Median of X_i

σ_i^2: Variance of X_i

σ_i: Standard deviation of X_i

Example : Oil in two offshore provinces

1. The North Atlantic Shelf province

The resource of interest is oil. Let

X_1 : Quantity of undiscovered recoverable oil in province 1

The chance of oil being present in recoverable quantities was estimated to be 42 percent, i.e., a marginal probability of $p_1 = 0.42$.

Conditional upon recoverable oil being present, conditional resource estimates were made in billion barrels

a) ${}_c x_{0.95} = 0.18$, a low conditional resource estimate

b) ${}_c x_{0.05} = 3.14$, a high conditional resource estimate

The methodology of Part 1 was applied by means of a computer program. The probability distribution of the quantity of undiscovered recoverable oil for the North Atlantic Shelf province is displayed in figure 14. Note that the graph of the more-than cdf has the value of the marginal probability, $p_1 = 0.42$, at the origin. Some numerical characteristics (in billion barrels) of this distribution are the following:

$x_{0.95} = 0.00$	$\mu_1 = 0.45$
$x_{0.75} = 0.00$	$m_1 = 0.00$
$x_{0.50} = 0.00$	$\sigma_1^2 = 0.88 \ (BB)^2$
$x_{0.25} = 0.59$	$\sigma_1 = 0.94$
$x_{0.05} = 2.07$	

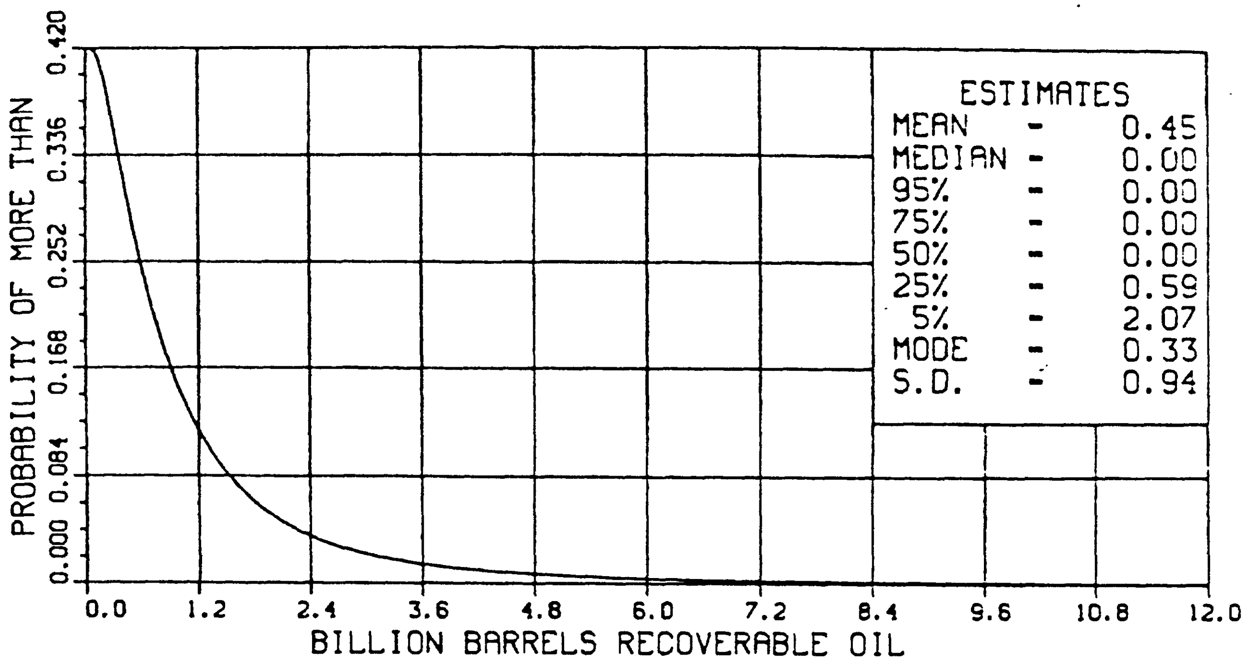


Figure 14.-Probability distribution of undiscovered recoverable oil for North Atlantic Shelf province. Estimates are mean, median, mode, standard deviation (S.D.) and fractiles that correspond to percentages listed.

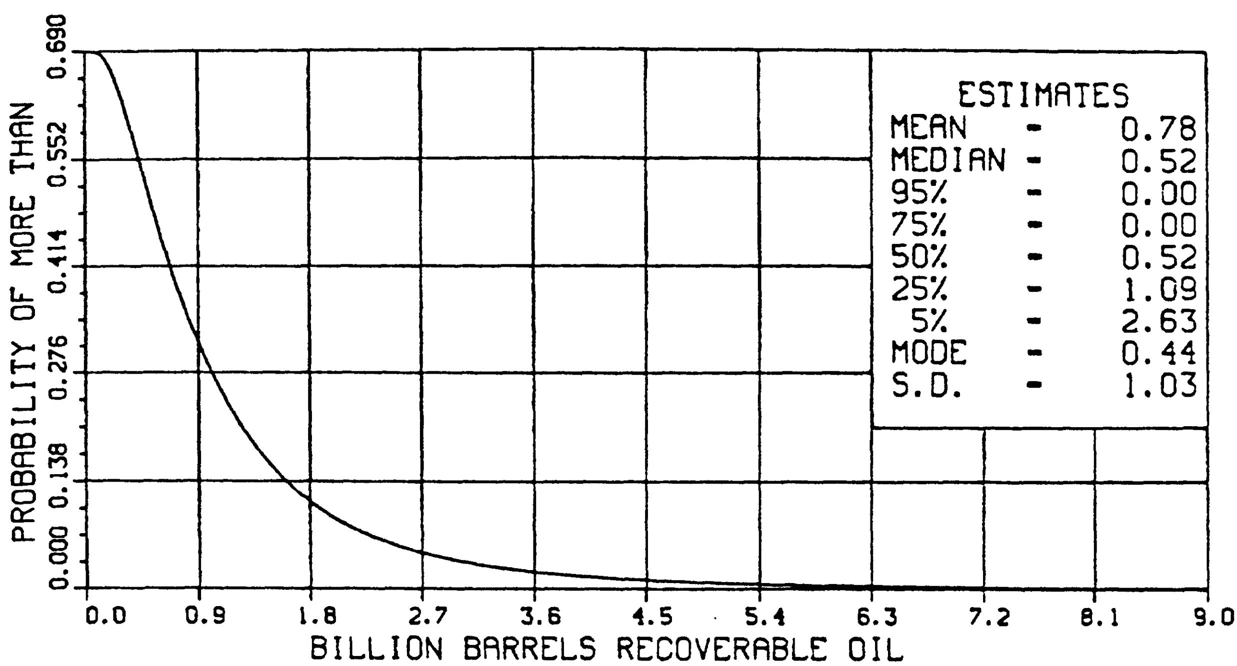


Figure 15.-Probability distribution of undiscovered recoverable oil for Mid-Atlantic Shelf province. Estimates are mean, median, mode, standard deviation (S.D.) and fractiles that correspond to percentages listed.

2. The Mid-Atlantic Shelf province

The resource of interest is oil. Let

X_2 : Quantity of undiscovered recoverable oil in province 2

The chance of oil being present in recoverable quantities was estimated to be 69 percent, i.e., a marginal probability of $p_2 = 0.69$.

Conditional upon recoverable oil being present, conditional resource estimates were made in billion barrels.

a) ${}_{c}x_{0.95} = 0.23$, a low resource estimate

b) ${}_{c}x_{0.05} = 3.05$, a high resource estimate

The methodology of Part 1 was applied by means of a computer program. The probability distribution of the quantity of undiscovered recoverable oil for the Mid-Atlantic Shelf province is displayed in figure 15. Note that the graph of the more-than cdf has the value of the marginal probability, $p_2 = 0.69$, at the origin. Some numerical characteristics (in billion barrels) of this distribution are the following:

$x_{0.95} = 0.00$	$\mu_2 = 0.78$
$x_{0.75} = 0.00$	$m_2 = 0.52$
$x_{0.50} = 0.52$	$\sigma_2^2 = 1.06\ (BB)^2$
$x_{0.25} = 1.09$	$\sigma_2 = 1.03$
$x_{0.05} = 2.63$	

Treating W as a random variable with range $w \geq 0$, a judgmental probability distribution of W will be determined. The probability distribution of W is also referred to as the aggregate probability distribution of $X_1, X_2, \ldots, X_n$. The phrase "aggregating n components" is also used.

The marginal probability of total resource is the probability of any resource of interest being present in the area. Therefore, the phrase "resource is present" means that the resource is present in at least one province. Hence, the marginal probability of total resource can be written as

$$p_W = P(W > 0)$$

A property that expresses the "aggregate marginal probability" p_W in terms of the province marginal probabilities p_i, $i = 1,2,\ldots,n$, is as follows.

Property $p_W = 1 - \prod_{i=1}^{n} (1 - p_i)$

The proof is an extension of the proof of the property for p_T in Part 4. Note that if any p_i equal 1, then $p_W = 1$. Also, if $p_W = 1$, the probability distribution of W is the same as the conditional probability distribution of W.

The components of an aggregation can have 95th fractiles of 0, and yet the aggregation has a 95th fractile greater than 0. This happens when the component marginal probabilities are at most 0.95, but the aggregate marginal probability is greater than 0.95.

Example:

	Marginal Probability	95th Fractile
Component 1	0.90	0
Component 2	0.90	0
Aggregate	0.99	>0

This example also illustrates that the 95th fractiles are not additive.

Because the random variable W is the sum of n random variables, i.e., $W = \sum_{i=1}^{n} X_i$, there are many well known relationships between W and the random variables X_i that are the extensions of similar results in Part 4:

The mean of W

$$\mu_W = \sum_{i=1}^{n} \mu_i$$

The variance of W

$$\sigma_W^2 = \sum_{i=1}^{n} \sigma_i^2 \quad \text{if } X_i \text{ independent}$$

The standard deviation of W

$$\sigma_W = \sqrt{\sigma_W^2} \quad \text{if } X_i \text{ independent}$$

Example: The combined North and Mid-Atlantic Shelf area.

Let

W: Quantity of undiscovered recoverable total oil in the combined North and Mid-Atlantic Shelf area. Thus, we have

$$W = X_1 + X_2$$

Some numerical characteristics of the probability distribution of W are the following:

The marginal probability of total oil

$$p_W = 1 - (1 - p_1)(1 - p_2) = 1 - (1 - 0.42)(1 - 0.69) = 0.82$$

The mean of W

$$\mu_W = \mu_1 + \mu_2 = 0.45 + 0.78 = 1.23 \text{ BB}$$

The variance of W

$$\sigma_W^2 = \sigma_1^2 + \sigma_2^2 = 0.88 + 1.06 = 1.94 \text{ (BB)}^2$$

The standard deviation of W

$$\sigma_W = \sqrt{\sigma_W^2} = \sqrt{1.94} = 1.39 \text{ BB}$$

The mean estimate of the quantity of undiscovered recoverable total oil for the combined North and Mid-Atlantic Shelf area is 1.23 billion barrels.

A Monte Carlo technique can be utilized to determine an approximate probability distribution of W in order to obtain the fractiles of W, w_α. This is the same computer-oriented approach to the problem as in Part 4. The same computer program is used except now the number of components being aggregated, n, is allowed to vary.

The more-than cdf of W is denoted by

$$\bar{F}_W(w) = P(W > w)$$

which is the probability that the quantity of undiscovered total resource is more than the amount w.

Example : The combined North and Mid-Atlantic Shelf area.

A Monte Carlo aggregation computer program was used to generate an approximate probability distribution of W. The number of samplings was 5,000. The approximate probability distribution of the quantity of undiscovered recoverable total oil for the combined North and Mid-Atlantic Shelf area is displayed in figure 16. Note that the graph of the more than cdf has the value of the marginal probability, $p_W = 0.82$, at the origin. Some fractiles of W (in billion barrels) are the following:

$$w_{0.95} = 0.00$$

$$w_{0.75} = 0.31$$

$$w_{0.50} = 0.87$$

$$w_{0.25} = 1.71$$

$$w_{0.05} = 3.76$$

The interval estimate of the quantity of undiscovered recoverable total oil for the combined North and Mid-Atlantic Shelf area is 0.00 to 3.76 billion barrels.

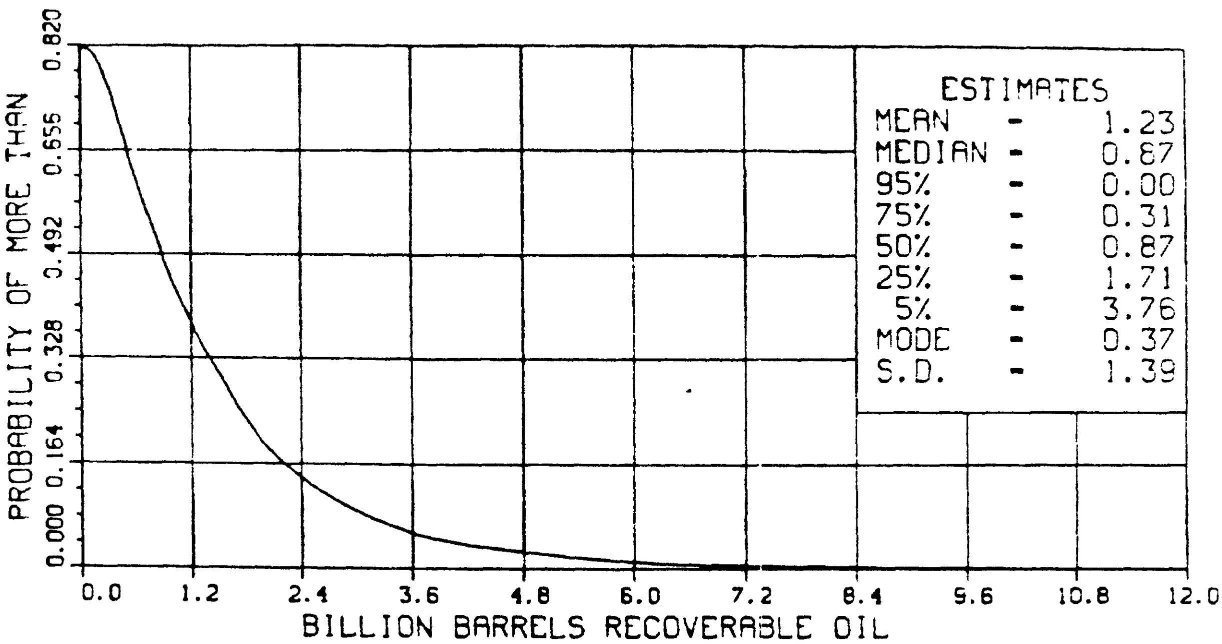


Figure 16.-Probability distribution of undiscovered recoverable total oil for combined North and Mid-Atlantic Shelf area. Estimates are mean, median, mode, standard deviation (S.D.) and fractiles that correspond to percentages listed.

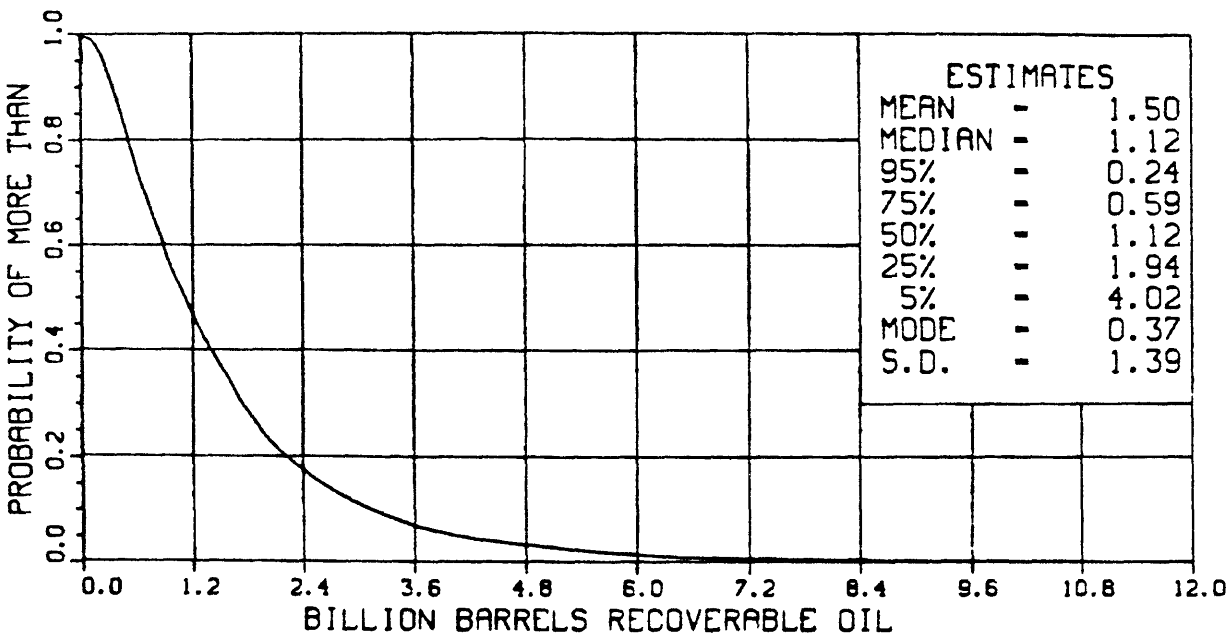


Figure 17.-Conditional probability distribution of undiscovered recoverable total oil for combined North and Mid-Atlantic Shelf area. Estimates are mean, median, mode, standard deviation (S.D.) and fractiles that correspond to percentages listed.

As in Part 4, the conditional probability distribution of W can be determined from the probability distribution of W by using the methodology of Part 1 in reverse.

The conditional more-than cdf of W is denoted by

$$\bar{F}_{W_c}(w) = P(W > w \mid W > 0)$$

which is the probability that the quantity of undiscovered total resource is more than the amount w given that the resource is present within the area of interest.

The following results are derived using the methodology of Part 1:

1) $\bar{F}_W(w) = p_W \bar{F}_{W_c}(w)$ implies $\bar{F}_{W_c}(w) = \bar{F}_W(w)/p_W$

2) $\mu_W = p_W \mu_{W_c}$ implies $\mu_{W_c} = \mu_W/p_W$

3) $\sigma_W^2 = p_W \sigma_{W_c}^2 + p_W(1 - p_W)\mu_{W_c}^2$ implies $\sigma_{W_c}^2 = \sigma_W^2/p_W - (1 - p_W)\mu_{W_c}^2$

Example: Conditional total oil in the combined North and Mid-Atlantic Shelf area.

Because $p_W = 0.82 \neq 1$, the conditional probability distribution of W is not the same as the probability distribution of W. The approximate conditional probability distribution of the quantity of undiscovered recoverable total oil for the combined North and Mid-Atlantic Shelf area is displayed in figure 17. Note that the graph of the conditional more-than cdf has the value of 1 at the origin. Some numerical characteristics (in billion barrels) of this distribution are the following:

${}_c w_{0.95} = 0.24$	$\mu_{W_c} = 1.50$
${}_c w_{0.75} = 0.59$	$m_{W_c} = 1.12$
${}_c w_{0.50} = 1.12$	$\sigma^2_{W_c} = 1.96\ (BB)^2$
${}_c w_{0.25} = 1.94$	$\sigma_{W_c} = 1.40$
${}_c w_{0.05} = 4.02$	

where

$$\mu_{W_c} = \mu_W / p_W = 1.23/0.82 = 1.50$$

$$\sigma^2_{W_c} = \sigma^2_W / p_W - (1 - p_W)\mu^2_{W_c} = 1.94/0.82 - (1 - 0.82)(1.50)^2 = 1.96$$

$$\sigma_{W_c} = \sqrt{\sigma^2_{W_c}} = \sqrt{1.96} = 1.40$$

Example: Total oil and total gas in the United States

1. Total Oil

In Dolton and others (1981) a Monte Carlo aggregation computer program was used to generate an approximate probability distribution of the quantity of undiscovered recoverable total oil for the United States that is displayed in figure 18. The United States was divided into 137 provinces that were aggregated; i.e., there are 137 components, n = 137. The number of samplings was 5,000. In one sampling, a random value was selected from each of the 137 oil probability distributions, and then the 137 random values were added together to produce a random value for total oil. Some numerical characteristics (in billion barrels) of this distribution are the following where fractiles in general are also denoted by $F_{\alpha 100}$:

Fractiles	
$F_{95} = 64.29$	mean $\mu = 82.59$
$F_{75} = 73.59$	median $m = 81.23$
$F_{50} = 81.23$	mode $M = 79.96$
$F_{25} = 89.81$	variance $\sigma^2 = 180.90\ (BB)^2$
$F_5 = 105.10$	standard deviation $\sigma = 13.45$

The mean estimate of the quantity of undiscovered recoverable total oil for the United States is 82.59 billion barrels. The interval estimate is 64.29 to 105.10 billion barrels.

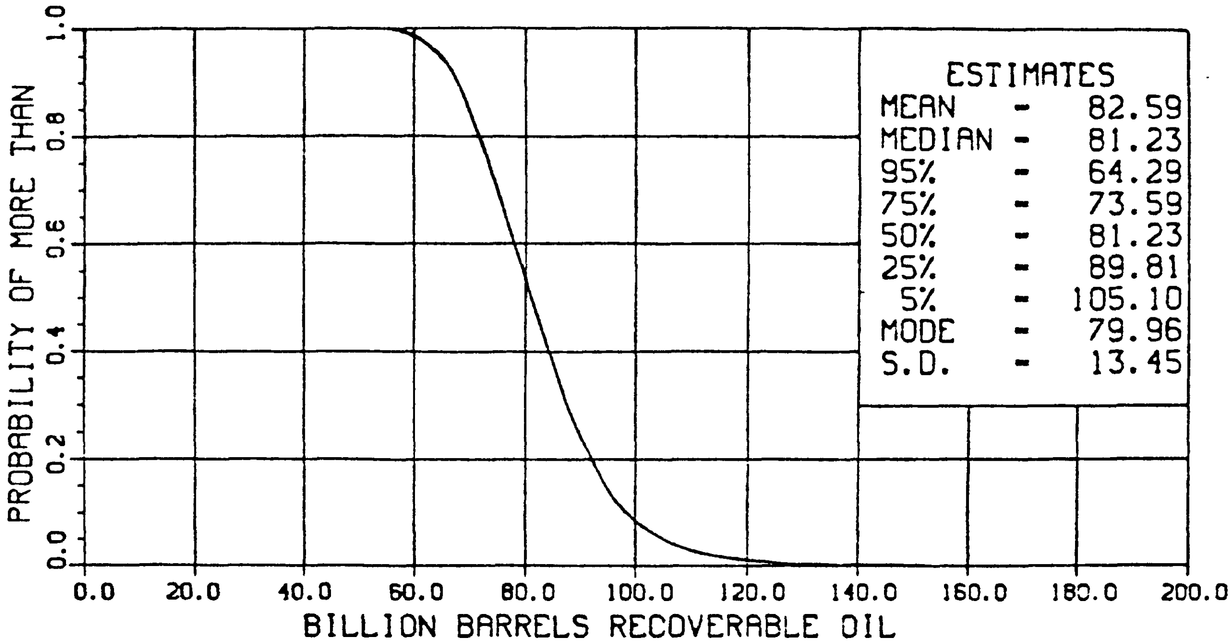


Figure 18.-Probability distribution of undiscovered recoverable total oil for United States. Estimates are mean, median, mode, standard deviation (S.D.) and fractiles that correspond to percentages listed.

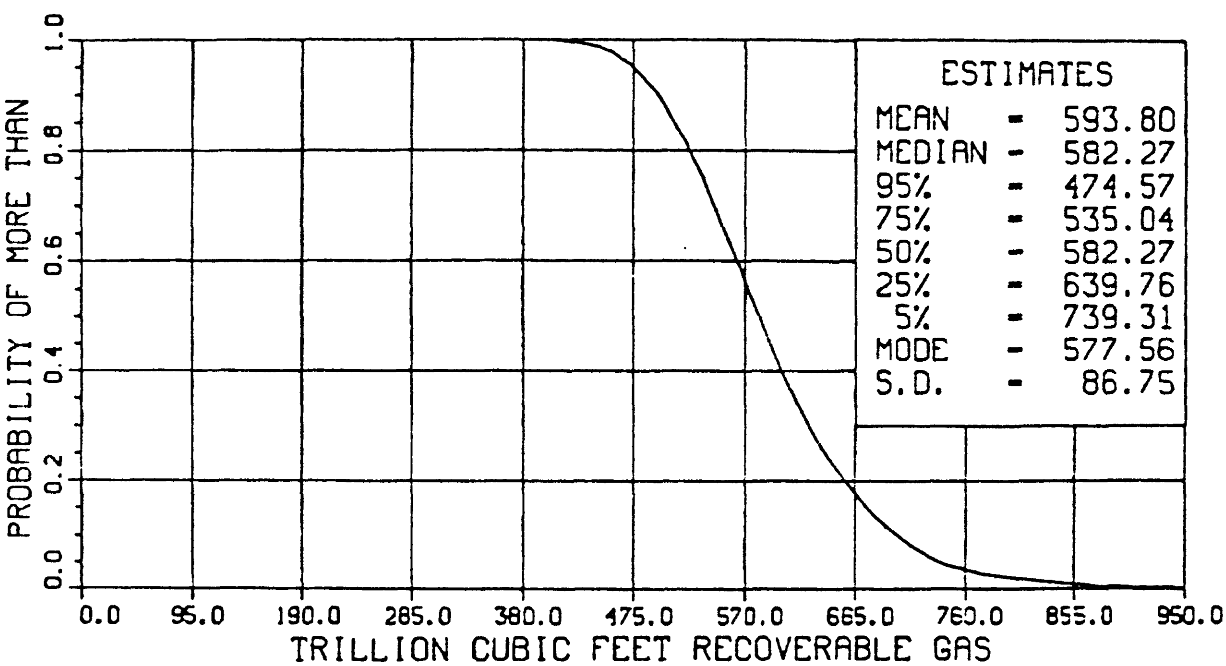


Figure 19.-Probability distribution of undiscovered recoverable total gas for United States. Estimates are mean, median, mode, standard deviation (S.D.) and fractiles that correspond to percentages listed.

2. Total Gas

In Dolton and others (1981) a Monte Carlo aggregation computer program was used to generate an approximate probability distribution of the quantity of undiscovered recoverable total gas for the United States that is displayed in figure 19. Because the United States was divided into 137 provinces, 274 components were aggregated, i.e., $n = 274$. This is because there are 2 components (non-associated gas and associated-dissolved gas) per province. The number of samplings was 5,000. In one sampling a random value was selected from each of the 274 gas probability distributions, and then the 274 random values were added together to produce a random value for total gas. Some numerical characteristics (in TCF) of this distribution are the following:

Fractiles	
$F_{95} = 474.57$	mean $\mu = 593.80$
$F_{75} = 535.04$	median $m = 582.27$
$F_{50} = 582.27$	mode $M = 577.56$
$F_{25} = 639.76$	variance $\sigma^2 = 7525.56\ (\mathrm{TCF})^2$
$F_5 = 739.31$	standard deviation $\sigma = 86.75$

The mean estimate of the quantity of undiscovered recoverable total gas for the United States is 593.80 TCF. The interval estimate is 474.57 to 739.31 TCF.

The Monte Carlo technique for generating an approximate probability distribution in order to obtain the fractiles can result in excessive cost and time, especially when the number of components to be aggregated (n) is large. There are further approximations that involve practically no cost or time, and no computer program. Realize that the mean μ and standard deviation σ are known.

a) Normal approximation

The normal distribution is considered as an approximate probability distribution when n is large (say at least 40). This is motivated by the central limit theorem for independent random variables. The normal fractiles can be calculated for a given α from the formula

$$F_{\alpha 100} = \mu + z_{\alpha}\sigma$$

where Z is a standard normal random variable and $P(Z > z_{\alpha}) = \alpha$. Some special normal fractiles are

$$F_{95} = \mu + z_{0.95}\sigma = \mu - 1.645\sigma$$
$$F_{75} = \mu + z_{0.75}\sigma = \mu - 0.675\sigma$$
$$F_{50} = \mu + z_{0.50}\sigma = \mu$$
$$F_{25} = \mu + z_{0.25}\sigma = \mu + 0.675\sigma$$
$$F_{5} = \mu + z_{0.05}\sigma = \mu + 1.645\sigma$$

b) The Chebyshev Inequality

The Chebyshev Inequality was covered in Part 1. The probability that the quantity of undiscovered total oil or gas resource lies within two standard deviations of the mean of any distribution is <u>at least</u> 0.75.

Example: Total oil and total gas in the United States

1. Total Oil

Because $\mu = 82.59$ and $\sigma = 13.45$, the normal approximation gives the following normal fractiles:

$$F_{95} = \mu - 1.645\sigma = 82.59 - 1.645(13.45) = 60.46$$

$$F_{75} = \mu - 0.675\sigma = 82.59 - 0.675(13.45) = 73.51$$

$$F_{50} = \mu = 82.59$$

$$F_{25} = \mu + 0.675\sigma = 82.59 + 0.675(13.45) = 91.67$$

$$F_{5} = \mu + 1.645\sigma = 82.59 + 1.645(13.45) = 104.72$$

The normal approximation approach compares favorably with the Monte Carlo technique regarding the fractiles.

2. Total Gas

Because $\mu = 593.80$ and $\sigma = 86.75$, the normal approximation gives the following normal fractiles:

$$F_{95} = \mu - 1.645\sigma = 593.80 - 1.645(86.75) = 451.10$$

$$F_{75} = \mu - 0.675\sigma = 593.80 - 0.675(86.75) = 535.24$$

$$F_{50} = \mu = 593.80$$

$$F_{25} = \mu + 0.675\sigma = 593.80 + 0.675(86.75) = 652.36$$

$$F_{5} = \mu + 1.645\sigma = 593.80 + 1.645(86.75) = 736.50$$

The normal approximation approach compares favorably with the Monte Carlo technique regarding the fractiles.

CONCLUSIONS

Part 1 is the foundation of the remaining Parts 2, 3, 4, and 5 with respect to the probabilistic methodology developed, and also the corresponding petroleum resources estimated. Assumptions of the probabilistic methods in the five parts are important, especially the assumed linear model when assessing petroleum resources in a portion of a province (in Part 2), and the independence assumption from province to province (in Part 5). Knowing when the methods are applicable requires prudence on the part of the user. For example, the large aggregations in Part 5 illustrated that a normal approximation can be quite adequate in practice.

Improvements and refinements can be made in the methods covered in this paper, and that, of course, is desirable. Nevertheless, an awareness of the complete picture of the whole resource appraisal process is essential in order to have a proper perspective. The probabilistic methodology should be viewed in light of the fact that the geologic experts of a resource appraisal team may have drastically differing views of the petroleum potential of a frontier province. Also, the output of any methodology is only as good as the input, no matter how sophisticated or "exact" the intervening methodology becomes, or put another way, "garbage in-garbage out." Therefore, whether a lognormal or gamma distribution is used as a probability model for the conditional probability distribution in a province may well be relatively insignificant within the total resource appraisal problem. The probabilistic methodology stage within the total resource appraisal problem is somewhat of a "science" within an "art" (once the assumptions are established).

The probabilistic methodology developed for oil and gas resource appraisal also could be used for resource appraisal of other types of natural resources, and even for other areas of application of subjective probability in decision making.

ACKNOWLEDGMENTS

The author thanks Joe Megeath and Norman Pence for their computer science expertise, Brent Lane for assisting in the preparation of the graphs, and Dianne Fox and Debbie Hobbs for typing the manuscript.

SELECTED REFERENCES

Dolton, G. L., Carlson, K. H., Charpentier, R. R., Coury, A. B., Crovelli, R. A., Frezon, S. E., Khan, A. S., Lister, J. H., McMullin, R. H., Pike, R. S., Powers, R. B., Scott, E. W., and Varnes, K. L., 1981, Estimates of undiscovered recoverable resources of conventionally producible oil and gas in the United States, a summary: U.S. Geological Survey Open-File Report 81-192, 17 p.

Dolton, G. L., Carlson, K. H., Charpentier, R. R., Coury, A. B., Crovelli, R. A., Frezon, S. E., Khan, A. S., Lister, J. H., McMullin, R. H., Pike, R. S., Powers, R. B., Scott, E. W., and Varnes, K. L., (in press), Estimates of undiscovered recoverable resources of conventionally producible oil and gas in the United States: U.S. Geological Survey Circular.

Hillier, F. S., and Lieberman, G. J., 1980, Introduction to Operations Research: San Francisco, Holden-Day, Inc., 3rd edition, 829 p.

Linstone, H. A., and Turoff, M., eds., 1975, The Delphi Method; Techniques and applications: Reading, Massachusetts, Addison-Wesley Publishing Company, 620 p.

Ross, S. M., 1972, Introduction to Probability Models: New York, Academic Press, Inc., 272 p.