U.S. DEPARTMENT OF THE INTERIOR

U.S. GEOLOGICAL SURVEY

Fractal Lognormal Percentage Assessment of Technically Recoverable Natural Gas Resources in Continuous-Type and Coalbed (Unconventional) Plays, Onshore and State Waters of the United States

by

Robert A. Crovelli[1], James W. Schmoker, and Richard H. Balay

Open-File Report 95-647

This report is preliminary and has not been reviewed for conformity with U.S. Geological Survey editorial standards and stratigraphic nomenclature.

[1]U.S. Geological Survey, Box 25046, MS 971
Denver Federal Center, Denver, Colorado 80225 U.S.A.

1995

ABSTRACT

The 20/80 law is a heuristic law that has evolved over the years into the following rule of thumb for many populations: 20% of the population accounts for 80% of the total value. This law states quantitatively that often a relatively small portion of a population accounts for most of the total value of the population. This principle is certainly applicable in the case of unconventional natural gas resources of continuous-type and coalbed plays in the United States where the few largest plays have an overwhelming amount of the total resources.

The general $p100/q100$ law in statistical form is defined with the statistic q as a function of p where p is the population proportion and q is the proportion of total value. The general $p100/q100$ law in probabilistic form is defined with the parameter q as a function of p for any probability distribution that models the population distribution. Using the lognormal distribution, the $p100/q100$ law in lognormal form is derived with the lognormal q being a fractal, where q possesses the scale invariance property.

The $p100/q100$ law in lognormal form was applied to data on technically recoverable gas resources in continuous-type and coalbed (unconventional) plays in the onshore areas and adjoining State waters of the United States. These data were generated as part of the U.S. Geological Survey's 1995 National assessment of United States oil and gas resources. The theoretical percentages of total resources using the lognormal q are close to the empirical percentages from the continuous-type and coalbed data using the statistic q. For example, 20% of the 34 continuous-type plays account for 63.2% of the total continuous-type resources of the plays if we use the lognormal q, or for 59.3% if we use the statistic q; 20% of the 39 coalbed plays account for 65.2% of the total coalbed resources of the plays if we use the lognormal q, or for 64.1% if we use the statistic q.

INTRODUCTION

Vilfredo Pareto (1848-1923), an Italian engineer turned economist and sociologist, was a pioneer in econometrics, applying the statistical methods of the physical sciences to the social sciences. Pareto discovered that, generally, a relatively small portion of a population accounts for most of the total value of the population. That is, in any set of elements, the critical elements usually constitute a minority of the set (Douglass and Douglass, 1993). Pareto's principle is certainly applicable in the case of technically recoverable natural gas resources in continuous-type and coalbed (unconventional) plays in the United States, where the few largest plays have an overwhelming amount of the total resources (Gautier and others, 1995; U.S. Geological Survey National Oil and Gas Resource Assessment Team, 1995).

As originally proposed, Pareto's law (Pareto, 1897; Turla and Hawkins, 1983) was an empirical relation describing the distribution of income among the population of a country; viz.,

> 20 percent of the people in Italy owned 80 percent of the wealth.

From this principle, the heuristic "20/80 law" has evolved over the years into the following simple rule of thumb, which is applicable to many populations:

> 20% of the population accounts for 80% of the total value.

Numerous applications of this law exist in the management sciences. For example,

> 20% of the customers account for 80% of the sales;
>
> 20% of the employees account for 80% of the work; and,
>
> 20% of the components account for 80% of the cost.

The principle is observed in other disciplines as well. The objective of the present paper is to generalize "the 20/80 law" using the lognormal distribution and apply the generalization to data on continuous-type and coalbed resources to obtain for each: $p100\%$ of the plays account for $q100\%$ (q is a function of p) of the total resources of the plays. Fractal lognormal percentage theory is developed first; the continuous-type data are analyzed next; and, finally, the theory is applied to the coalbed data.

FRACTAL LOGNORMAL PERCENTAGE THEORY

The material of this section was first presented in Crovelli (1995) and is reproduced here for completeness.

THE $p100/q100$ LAW IN STATISTICAL FORM

Consider a population of N values for a random variable X arranged in increasing order of magnitude:

$$x_1, x_2, \ldots, x_N.$$

The total value of the population is:

$$\sum_{i=1}^{N} x_i .$$

Given a proportion p $(0 \le p \le 1)$ of largest values of the population, the partial sum of the $p100\%$ of largest values of the population is:

$$\sum_{i=[(1-p)N]+1}^{N} x_i \ ,$$

where $[\cdot]$ is the greatest integer value function.

Let the proportion of the total value of the population accounted for by $p100\%$ of the population be given by q $(0 \le q \le 1)$, where the statistic q is defined as:

$$q = \sum_{i=[(1-p)N]+1}^{N} x_i \Bigg/ \sum_{i=1}^{N} x_i .$$

An alternative form is:

$$q = \frac{p \displaystyle\sum_{i=[(1-p)N]+1}^{N} x_i \Bigg/ pN}{\displaystyle\sum_{i=1}^{N} x_i \Bigg/ N} .$$

Summarizing, p is a proportion (or fraction) of the population, and $p100\%$ is a percentage of the population; whereas, q is a proportion (or fraction) of the total value, $q100\%$ is a percentage of the total value, and $p100\%$ of the population accounts for $q100\%$ of the total value.

THE *p100/q100* LAW IN PROBABILISTIC FORM

Given a population of values of a random variable X having a probability distribution that models the population distribution, then $p100\%$ of the population values account for $q100\%$ of the total value, where the parameter q is called the proportion of total value and defined in terms of conditional expectation, $E(X \mid \cdot)$, as follows.

$$q = \frac{pE(X \mid X > x_p)}{E(X)}, \quad 0 \le p \le 1,$$

with $p = P(X > x_p)$ and x_p is called the $p100$th fractile.

Various probability distributions could be used as models for the population distribution. We will consider the lognormal probability distribution. The percentage theory based on the Pareto distribution has been developed and applied to petroleum field size data (Crovelli, 1995).

THE LOGNORMAL PROBABILITY DISTRIBUTION

The lognormal distribution is discussed in detail in Aitchison and Brown (1957), Johnson and others (1994), and Crow and Shimizu (1988). A nonnegative random variable X has a lognormal distribution if the random variable $Y = \ln X$ has a normal distribution.

The lognormal probability density function of X is

$$f(x) = \frac{1}{\sqrt{2\pi}\,\sigma x} \exp \frac{-(\ln x - \mu)^2}{2\sigma^2}, \qquad x > 0$$

where $\sigma > 0$ is a shape parameter and μ is a scale parameter (Law and Kelton, 1991, p. 337).

The expected value or mean of X is

$$E(X) = e^{\mu + \sigma^2/2}$$

The variance of X is

$$V(X) = e^{2\mu + \sigma^2}(e^{\sigma^2} - 1)$$

The scale parameter μ and the shape parameter σ of the lognormal distribution are computed from the following formulas:

$$\mu = \frac{1}{N}\sum_{i=1}^{N} \ln x_i$$

and

$$\sigma = \left[\frac{N\sum_{i=1}^{N} (\ln x_i)^2 - \left(\sum_{i=1}^{N} \ln x_i \right)^2}{N^2} \right]^{1/2} .$$

The following notation will be used throughout the remainder of the paper:

$X \sim$ lognormal (shape parameter σ, scale parameter μ), where the symbol ~ is read "is distributed as."

THE LOGNORMAL DISTRIBUTION HAS A FRACTAL PROPERTY

If X ~ lognormal (shape parameter σ, scale parameter μ), let

$$X' = cX, \qquad c > 0 .$$

The fractal property of X under a positive multiplicative constant results in:

$$X' \sim \text{lognormal (shape parameter } \sigma \text{, scale parameter } \mu + \ln c \text{).}$$

Note that the distributions of X and X' are identical (with the same shape parameter) except for scale. A probabilistic fractal is a probability distribution that is invariant except for scale (scale invariant or self-similar) under a given transformation (Mandelbrot, 1983, p. 343). Under a positive multiplicative constant, the lognormal distribution remains a lognormal distribution with fixed σ and changed μ.

THE *p*100/*q*100 LAW IN LOGNORMAL FORM

Consider the $p100/q100$ law in probabilistic form when the population distribution is modeled as a lognormal distribution; i.e.,

$$q = \frac{pE(X|X > x_p)}{E(X)}, \qquad 0 \le p \le 1$$

when X ~ lognormal (shape parameter σ, scale parameter μ).

Recall that the expected value of X is

$$E(X) = e^{\mu+\sigma^2/2} .$$

It can be shown (Johnson and others, 1994) that the conditional expectation is given by

$$E(X|X > x_p) = e^{\mu+\sigma^2/2} \frac{R((\ln x_p - \mu - \sigma^2)/\sigma)}{R((\ln x_p - \mu)/\sigma)}$$

where R denotes the complementary cumulative distribution function of a standard normal random variable Z, i.e., $R(z) \equiv P(Z > z)$.

We define the following notation:

$$R(z_p) = P(Z > z_p) = p$$

where z_p is called the $p100$th fractile of Z.

Because

$$R((\ln x_p - \mu)/\sigma) = R(z_p) = p$$

and

$$R((\ln x_p - \mu - \sigma^2)/\sigma) = R(z_p - \sigma)$$

we have

$$E(X \mid X > x_p) = e^{\mu+\sigma^2/2} \frac{R(z_p - \sigma)}{p}$$

Substituting into the general formula for q and simplifying, we get

$$q = R(z_p - \sigma), \quad 0 < p < 1$$

or

$$q = 1 - \Phi(z_p - \sigma), \quad 0 < p < 1$$

where Φ denotes the standard normal cumulative distribution function.

Observe that if $p \to 0$, then $q \to 0$, and if $p \to 1$, then $q \to 1$. Thus we define $q = 0$ if $p = 0$, and $q = 1$ if $p = 1$. The special case when $p = 0.5$ and $z_p = 0$ yields

$$q = R(-\sigma) = 1 - \Phi\,(-\sigma)$$

The parameter q has the very remarkable property of being scale-free; i.e., q does not depend upon the scale parameter μ. Because the parameter q is scale-free, q stays the same for all lognormal distributions with the same shape parameter σ, as in the case of the fractal property of the lognormal distribution under a positive multiplicative constant. Because it is scale invariant, the proportion q can be considered to be a parametric fractal.

The proportion q of the $p100/q100$ law in statistical form is called the statistic q, in probabilistic form it is called the parameter q, and in lognormal form it is called the lognormal q.

CONTINUOUS-TYPE AND COALBED DATA

The U.S. Geological Survey periodically makes appraisals of the oil and gas resources of the Nation. In its 1995 National Assessment (Gautier and others, 1995; U.S. Geological Survey National Oil and Gas Resource Assessment Team, 1995), the onshore areas and adjoining State waters of the Nation were assessed. The basic assessment unit was the play.

A new type of unconventional play was defined for the 1995 National Assessment in which hydrocarbons are not trapped in the conventional sense because accumulations do not result from the buoyancy of gas or oil in water (Schmoker, 1995). These unconventional accumulations are termed continuous because the reservoir rock within a play is everywhere charged with oil or gas.

For purposes of this report, continuous gas accumulations reservoired in sandstone, siltstone, shale or chalk are referred to as continous-type. Continuous gas accumulations reservoired in coal are identified by the adjective coalbed. For the 1995 National Assessment, these two types of plays were assessed using newly developed approaches (Crovelli and Balay, 1995; Rice and others, 1995; Schmoker, 1995).

Thirty-four continuous-type gas plays (Schmoker and Oscarson, 1995) were assessed as part of the 1995 National Assessment. Mean potential additions to technically recoverable resources for these plays total 303 trillion cubic feet of gas (Schmoker and others, 1995). The mean potential additions to technically recoverable resources for each of the 34 continuous-type plays constitute the data set used in the following section on Application of Theory to Continuous-Type Resources.

Thirty-nine coalbed plays were assessed as part of the 1995 National Assessment (Gautier and others, 1995). Mean potential additions to technically recoverable resources for these plays total 50 trillion cubic feet of gas (U.S. Geological Survey National Oil and Gas Resource Assessment Team, 1995). The mean potential additions to technically recoverable resources for each of the 39 coalbed plays constitute the data set used in the following section on Application of Theory to Coalbed Resources.

APPLICATION OF THEORY TO CONTINUOUS-TYPE RESOURCES

The fractal lognormal percentage theory is applied in this section to the continuous-type data. This new probabilistic methodology for percentage assessment lends itself as an ideal spreadsheet software application.

Table 1 gives the play names, numbers, and mean gas additions of the continuous-type plays in the 1995 National Assessment. A frequency distribution of the 34 continuous-type plays (in billion cubic feet mean gas) is displayed in Figure 1. Because the plotted random variable is the logarithm of mean gas, Ln (billion cubic feet mean gas), the bell-shaped distribution suggests that the continuous-type resources are approximately distributed as a lognormal distribution.

Spreadsheet software was used to make various calculations involving the continuous-type data. Table 2 is a spreadsheet with 34 rows corresponding to the 34 continuous-type plays. The data are listed and several calculations are made, including those necessary for the statistic q. The lognormal parameters were computed: scale parameter $\mu = 8.4570$ and shape parameter $\sigma = 1.1798$.

Table 3 is a spreadsheet of the calculations from the continuous-type data that are necessary for the lognormal q as a function of various values of p. The formula for the lognormal q in the case of the continuous-type data is the following:

$$q = R(z_p - 1.1798) = 1 - \Phi(z_p - 1.1798), \quad 0 < p < 1$$

Therefore, for any specified value of p, we can obtain: $p100\%$ of the continuous-type plays account for $q100\%$ of the total continuous-type resources of the plays. The fractal lognormal percentage assessment of continuous-type resources is summarized in Table 4. The corresponding graph of the summary is given in Figure 2.

Note that in Table 4 and Figure 2, the theoretical percentages of total continuous-type resources using the lognormal q are close to the empirical percentages from the continuous-type data using the statistic q. For example, 20% of the 34 continuous-type plays account for 63.2% of the total continuous-type resources of the plays if we use the lognormal q, or for 59.3% if we use the statistic q. Plays presumably vary in their degree of heterogeneity, and this can be quantitatively captured by this method. The method also forms a basis for comparing plays of different geologic styles, and similar geologic styles.

Table 1. Play names, numbers, and mean gas additions of the continuous-type plays in the 1995 National Assessment (Gautier and others, 1995).

	PLAY NAME	PLAY NO.	MEAN GAS (10^9 CF)
1	Columbia Basin - Basin Centered Gas	503	12200.100
2	Tight Gas Piceance Mesaverde Williams Fork	2007	4870.110
3	Tight Gas Piceance Mesaverde Iles	2010	4828.400
4	Tight Gas Uinta Tertiary East	2015	2141.220
5	Tight Gas Uinta Tertiary West	2016	514.479
6	Basin Flank Uinta Mesaverde	2018	3784.300
7	Deep Synclinal Uinta Mesaverde	2020	574.370
8	Dakota Central Basin Gas	2205	8211.280
9	Central Basin Mesaverde Gas	2209	9584.740
10	Pictured Cliffs Gas	2211	3264.040
11	Northern Great Plains Biogenic Gas - High Potential	2810	5436.740
12	Northern Great Plains Biogenic Gas - Moderate Potential (Suffield Block Analog)	2811	20479.700
13	Northern Great Plains Biogenic Gas - Low Potential	2812	15353.600
14	Southern Williston Basin Margin-Niobrara Shallow Biogenic	3113	1894.320
15	Greater Green River Basin - Cloverly-Frontier	3740	37250.700
16	Greater Green River Basin - Mesaverde (Almond)	3741	51707.800
17	Greater Green River Basin - Lewis	3742	19002.800
18	Greater Green River Basin - Fox Hills - Lance	3743	10223.600
19	Greater Green River Basin - Fort Union	3744	985.655
20	J Sandstone Deep Gas (Wattenberg)	3906	830.550
21	Cotton Valley Blanket Sandstones Gas	4923	6034.580
22	Antrim Shale Gas - Developed Area	6319	4928.110
23	Antrim Shale Gas - Undeveloped Area	6320	13941.900
24	Illinois Basin - New Albany Shale Gas	6407	1888.950
25	Devonian Black Shale Gas	6604	1388.550
26	Clinton/Medina Sandstone Gas High Potential	6728	24571.300
27	Clinton/Medina Sandstone Gas Medium Potential	6729	5740.270
28	Clinton/Medina Sandstone Gas Medium-Low Potential	6730	915.446
29	Upper Devonian Sandstone Gas High Potential	6733	9976.650
30	Upper Devonian Sandstone Gas Medium Potential	6734	3751.070
31	Upper Devonian Sandstone Gas Medium-Low Potential	6735	940.699
32	Devonian Black Shale-Greater Big Sandy	6740	9088.720
33	Devonian Black Shale-Greater Siltstone Content	6741	2772.530
34	Devonian Black Shale-Lower Thermal Maturity	6742	3452.750

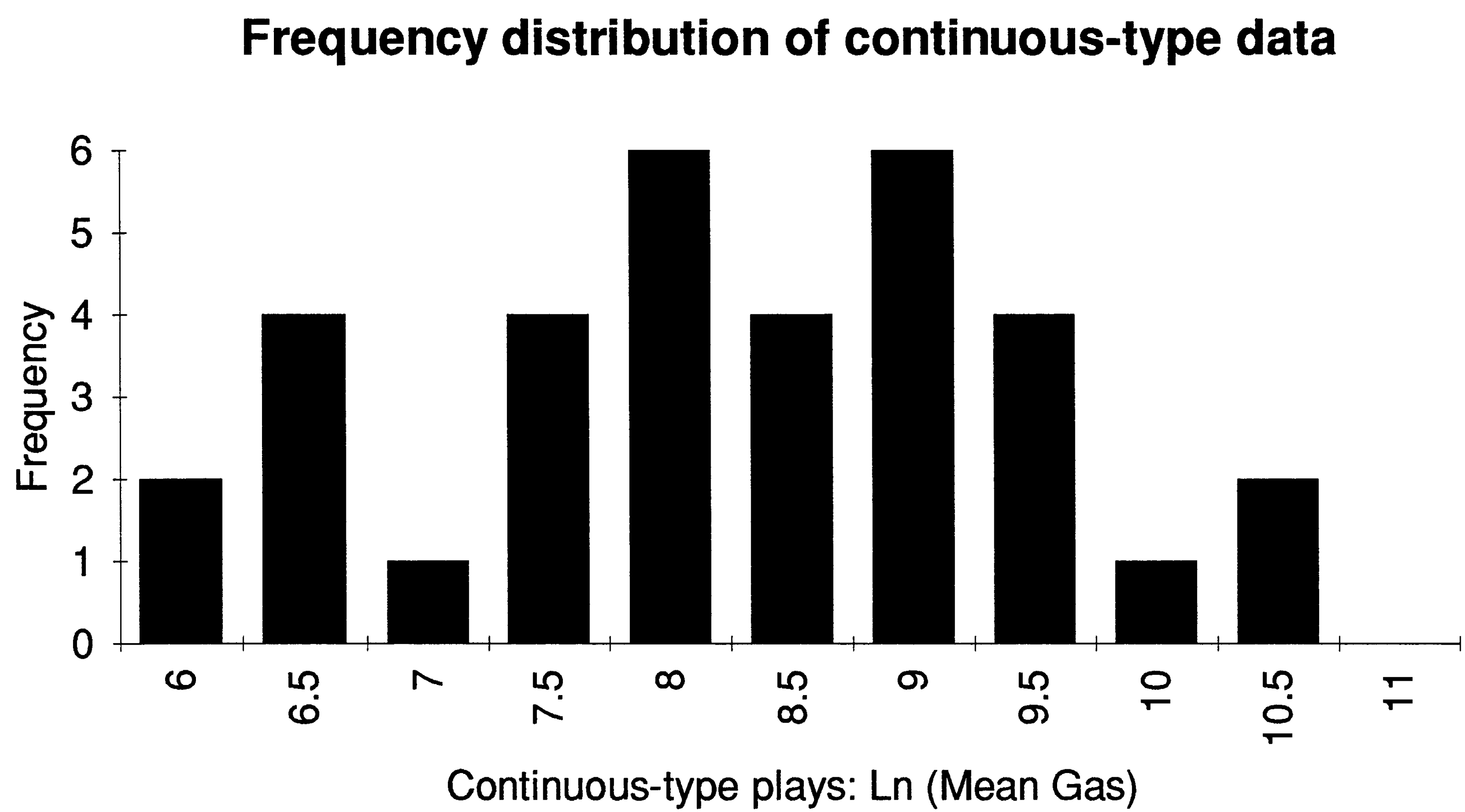

Figure 1. Frequency distribution of 34 continuous-type plays (in billion cubic feet mean gas), derived from data of Table 1. Data and bin frequencies are included with Table 2.

Table 2. Spreadsheet of continuous-type data and calculations for the statistic q.

n	Mean Gas (10^9 CF)	Ln of Mean Gas	Sorted Mean Gas	Partial sum Sorted	Statistic q Partial/Total
1	12200.100	9.4092	51707.800	51707.800	0.17092
2	4870.110	8.4909	37250.700	88958.500	0.29405
3	4828.400	8.4823	24571.300	113529.800	0.37527
4	2141.220	7.6691	20479.700	134009.500	0.44296
5	514.479	6.2432	19002.800	153012.300	0.50578
6	3784.300	8.2386	15353.600	168365.900	0.55653
7	574.370	6.3533	13941.900	182307.800	0.60261
8	8211.280	9.0133	12200.100	194507.900	0.64294
9	9584.740	9.1679	10223.600	204731.500	0.67673
10	3264.040	8.0907	9976.650	214708.150	0.70971
11	5436.740	8.6009	9584.740	224292.890	0.74139
12	20479.700	9.9272	9088.720	233381.610	0.77143
13	15353.600	9.6391	8211.280	241592.890	0.79857
14	1894.320	7.5466	6034.580	247627.470	0.81852
15	37250.700	10.5254	5740.270	253367.740	0.83750
16	51707.800	10.8534	5436.740	258804.480	0.85547
17	19002.800	9.8523	4928.110	263732.590	0.87176
18	10223.600	9.2325	4870.110	268602.700	0.88785
19	985.655	6.8933	4828.400	273431.100	0.90381
20	830.550	6.7221	3784.300	277215.400	0.91632
21	6034.580	8.7053	3751.070	280966.470	0.92872
22	4928.110	8.5027	3452.750	284419.220	0.94014
23	13941.900	9.5427	3264.040	287683.260	0.95092
24	1888.950	7.5438	2772.530	290455.790	0.96009
25	1388.550	7.2360	2141.220	292597.010	0.96717
26	24571.300	10.1093	1894.320	294491.330	0.97343
27	5740.270	8.6553	1888.950	296380.280	0.97967
28	915.446	6.8194	1388.550	297768.830	0.98426
29	9976.650	9.2080	985.655	298754.485	0.98752
30	3751.070	8.2298	940.699	299695.184	0.99063
31	940.699	6.8466	915.446	300610.630	0.99366
32	9088.720	9.1148	830.550	301441.180	0.99640
33	2772.530	7.9275	574.370	302015.550	0.99830
34	3452.750	8.1469	514.479	302530.029	1.00000

Bin	*Frequency*
6	2
6.5	4
7	1
7.5	4
8	6
8.5	4
9	6
9.5	4
10	1
10.5	2
11	0

5.14E+02	6.2432 <== Min
5.17E+04	10.8534 <== Max
3.03E+05	287.539 <== Total
8.90E+03	8.4570 <== Mean (μ)
1.09E+04	1.1798 <== Standard Deviation (σ)

Table 3. Spreadsheet of calculations from continuous-type data of Figure 1 for the lognormal q as a function of various values of p. NORMSINV is the spreadsheet function that computes z; NORMSDIST is the function that computes $1-q$.

p Pop. Prop.	$(p)(N)$ $(p)(34)$	$1-p$	z NORMSINV	$z-\sigma$ $(z-1.1798)$	$1-q$ NORMSDIST	lognormal q (theoretical)	statistic q (empirical)
0.1	3.4	0.9	1.2816	0.1018	0.5405	0.4595	0.4023
0.2	6.8	0.8	0.8416	−0.3382	0.3676	0.6324	0.5934
0.3	10.2	0.7	0.5244	−0.6554	0.2561	0.7439	0.7160
0.4	13.6	0.6	0.2533	−0.9265	0.1771	0.8229	0.8105
0.5	17	0.5	0.0000	−1.1798	0.1190	0.8810	0.8718
0.6	20.4	0.4	−0.2533	−1.4331	0.0759	0.9241	0.9213
0.7	23.8	0.3	−0.5244	−1.7042	0.0442	0.9558	0.9583
0.8	27.2	0.2	−0.8416	−2.0214	0.0216	0.9784	0.9806
0.9	30.6	0.1	−1.2816	−2.4614	0.0069	0.9931	0.9924

Table 4. Summary of fractal lognormal percentage assessment of continuous-type resources.

Percentage of continuous-type plays $p100\%$	Percentage of total continuous-type resources $q100\%$ Theoretical (lognormal q)	Empirical (statistic q)
0	0.00	0.00
10	45.95	40.23
20	63.24	59.34
30	74.39	71.60
40	82.29	81.05
50	88.10	87.18
60	92.41	92.13
70	95.58	95.83
80	97.84	98.06
90	99.31	99.24
100	100.00	100.00

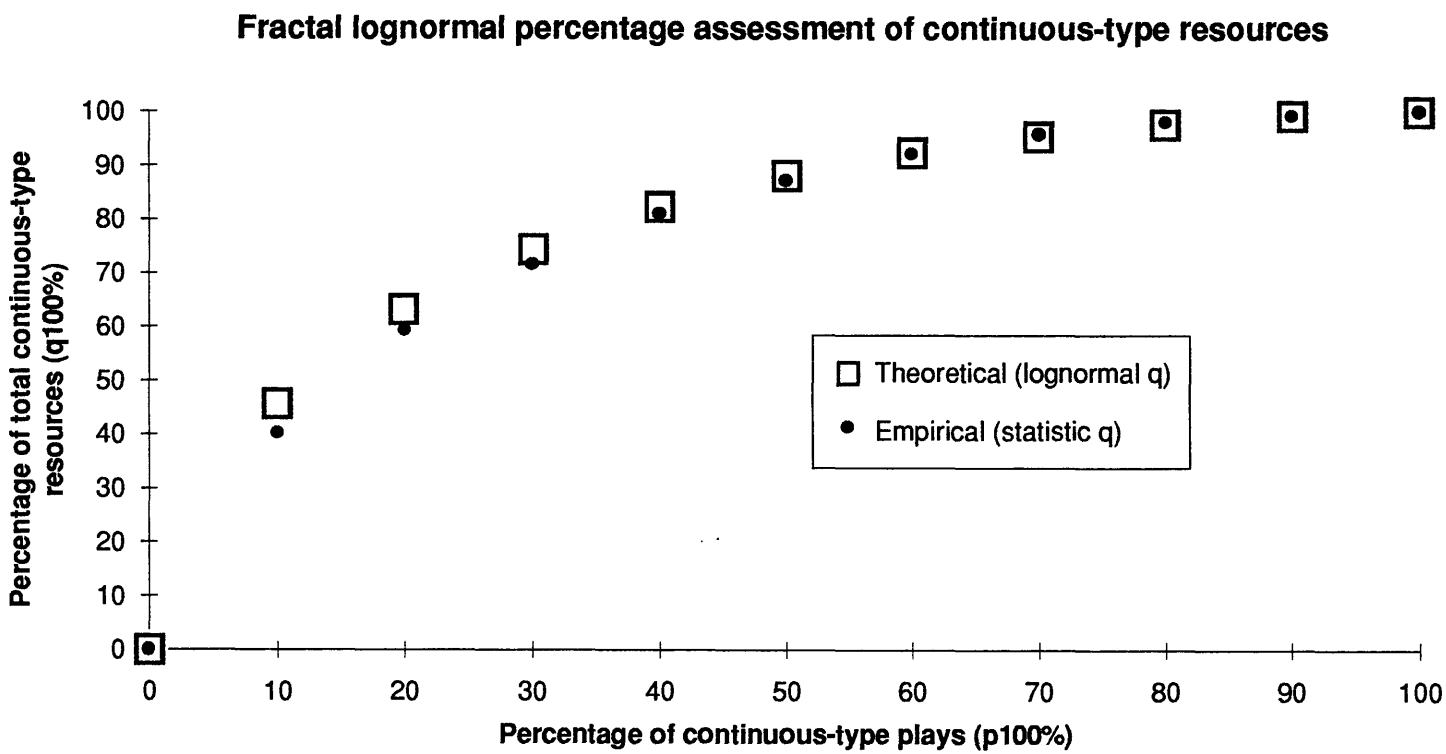

Figure 2. Theoretical percentage of total continuous-type resources using the lognormal q as a function of the percentage of continuous-type plays, along with the corresponding empirical values using the statistic q. Data are from Table 4.

APPLICATION OF THEORY TO COALBED RESOURCES

The fractal lognormal percentage theory is applied here to the coalbed data.

Table 5 gives the play names, numbers, and mean gas additions of the coalbed plays in the 1995 National Assessment. A frequency distribution of the 39 coalbed plays (in billion cubic feet mean gas) is displayed in Figure 3. Again, because the plotted random variable is the logarithm of mean gas, Ln (billion cubic feet mean gas), the bell-shaped distribution suggests that the coalbed resources are approximately distributed as a lognormal distribution.

Table 6 is a spreadsheet with 39 rows corresponding to the 39 coalbed plays. The data are listed and several calculations are made, including those necessary for the statistic q. The lognormal parameters were computed: scale parameter $\mu = 6.4606$ and shape parameter $\sigma = 1.2317$.

Table 7 is a spreadsheet of the calculations from the coalbed data that are necessary for the lognormal q as a function of various values of p. The formula for the lognormal q in the case of the coalbed data is the following:

$$q = R(z_p - 1.2317) = 1 - \Phi(z_p - 1.2317), \quad 0 < p < 1$$

The fractal lognormal percentage assessment of coalbed resources is summarized in Table 8. The corresponding graph of the summary is given in Figure 4.

Note that in Table 8 and Figure 4, the theoretical percentages of total coalbed resources using the lognormal q are close to the empirical percentages from the coalbed data using the statistic q. For example, 20% of the 39 coalbed plays account for 65.2% of the total coalbed resources of the plays if we use the lognormal q, or for 64.1% if we use the statistic q.

Table 5. Play names, numbers, and mean gas additions of the coalbed plays in the 1995 National Assessment (Gautier and others, 1995).

	PLAY NAME	PLAY NO.	MEAN GAS (10^9 CF)
1	Western Washington - Bellingham Basin	450	40.4637
2	Western Washington - Western Cascade Mountains	451	540.791
3	Western Washington - Southern Puget Lowlands	452	116.078
4	Uinta Basin - Book Cliffs	2050	1941.41
5	Uinta Basin - Sego	2051	521.184
6	Uinta Basin - Emery	2052	747.985
7	Piceance Basin - White River Dome	2053	329.864
8	Piceance Basin - Western Basin Margin	2054	6491.6
9	Piceance Basin - Grand Hogback	2055	241.301
10	Piceance Basin - Divide Creek Anticline	2056	397.005
11	Piceance Basin - Igneous Intrusion	2057	34.5965
12	San Juan Basin - Overpressured	2250	4165.41
13	San Juan Basin - Underpressured Discharge	2252	2143.84
14	San Juan Basin - Underpressured	2253	1223.78
15	Powder River Basin - Shallow Mining-Related	3350	681.485
16	Powder River Basin - Central Basin	3351	425.39
17	Wind River Basin - Mesaverde	3550	425.723
18	Greater Green River Basin - Rock Springs	3750	693.167
19	Greater Green River Basin - Iles	3751	377.3
20	Greater Green River Basin - Williams Fork	3752	1385.23
21	Greater Green River Basin - Almond	3753	794.8
22	Greater Green River Basin - Lance	3754	230.199
23	Greater Green River Basin - Fort Union	3755	407.799
24	Northern Raton Basin	4150	914.386
25	Raton Basin - Purgatoire River	4151	289.408
26	Southern Raton Basin	4152	571.388
27	Forest City Basin - Central Basin	5650	451.518
28	Cherokee Platform - Central Basin	6050	1913.95
29	Arkoma Basin, Anticline	6250	385.014
30	Arkoma Basin, Syncline	6251	2255.25
31	Illinois Basin - Central Basin	6450	1627.86
32	Black Warrior Basin Recharge	6550	253.435
33	Black Warrior Basin - Southeastern Basin	6551	1307.56
34	Black Warrior Basin - Coastal Plain	6552	683.58
35	Black Warrior Basin - Central and Western Basin	6553	58.8286
36	Northern Appalachian Basin - Anticline	6750	1068.97
37	Northern Appalachian Basin - Syncline	6751	10414
38	Central Appalachian Basin - Central Basin	6752	3067.95
39	Cahaba Coal Field	6753	294.758

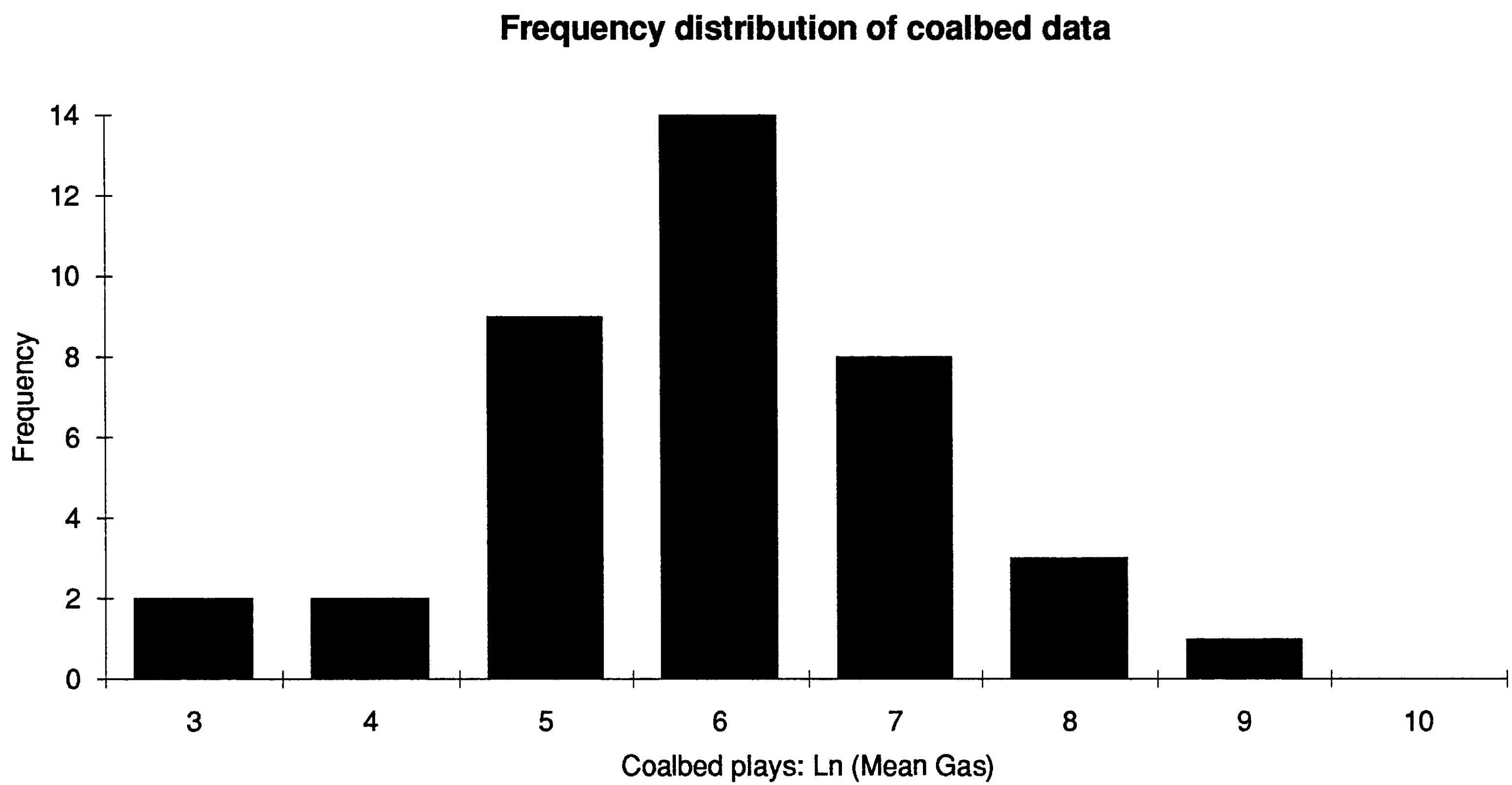

Figure 3. Frequency distribution of 39 coalbed plays (in billion cubic feet mean gas). Data and bin frequencies are included with Table 6.

Table 6. Spreadsheet of coalbed data and calculations for the statistic q.

n	Mean Gas (10^9 CF)	Ln of Mean Gas	Sorted Mean Gas	Partial sum Sorted	Statistic q Partial/Total
1	40.4637	3.7004	10414	10414.000	0.20864
2	540.791	6.2930	6491.6	16905.600	0.33869
3	116.078	4.7543	4165.41	21071.010	0.42214
4	1941.41	7.5712	3067.95	24138.960	0.48361
5	521.184	6.2561	2255.25	26394.210	0.52879
6	747.985	6.6174	2143.84	28538.050	0.57174
7	329.864	5.7987	1941.41	30479.460	0.61064
8	6491.6	8.7783	1913.95	32393.410	0.64898
9	241.301	5.4860	1627.86	34021.270	0.68159
10	397.005	5.9839	1385.23	35406.500	0.70935
11	34.5965	3.5438	1307.56	36714.060	0.73554
12	4165.41	8.3346	1223.78	37937.840	0.76006
13	2143.84	7.6704	1068.97	39006.810	0.78148
14	1223.78	7.1097	914.386	39921.196	0.79980
15	681.485	6.5243	794.8	40715.996	0.81572
16	425.39	6.0530	747.985	41463.981	0.83070
17	425.723	6.0538	693.167	42157.148	0.84459
18	693.167	6.5413	683.58	42840.728	0.85829
19	377.3	5.9330	681.485	43522.213	0.87194
20	1385.23	7.2336	571.388	44093.601	0.88339
21	794.8	6.6781	540.791	44634.392	0.89422
22	230.199	5.4389	521.184	45155.576	0.90466
23	407.799	6.0108	451.518	45607.094	0.91371
24	914.386	6.8183	425.723	46032.817	0.92224
25	289.408	5.6678	425.39	46458.207	0.93076
26	571.388	6.3481	407.799	46866.006	0.93893
27	451.518	6.1126	397.005	47263.011	0.94688
28	1913.95	7.5569	385.014	47648.025	0.95460
29	385.014	5.9533	377.3	48025.325	0.96216
30	2255.25	7.7210	329.864	48355.189	0.96877
31	1627.86	7.3950	294.758	48649.947	0.97467
32	253.435	5.5351	289.408	48939.355	0.98047
33	1307.56	7.1759	253.435	49192.790	0.98555
34	683.58	6.5273	241.301	49434.091	0.99038
35	58.8286	4.0746	230.199	49664.290	0.99499
36	1068.97	6.9745	116.078	49780.368	0.99732
37	10414	9.2509	58.8286	49839.197	0.99850
38	3067.95	8.0288	40.4637	49879.660	0.99931
39	294.758	5.6862	34.5965	49914.257	1.00000
	3.46E+01	3.5438	<== Min		
	1.04E+04	9.2509	<== Max		
	4.99E+04	245.505	<== Total		
	1.28E+03	6.4606	<== Mean (μ)		
	1.93E+03	1.2317	<== Standard Deviation (σ).		

Bin	*Frequency*
3	2
4	2
5	9
6	14
7	8
8	3
9	1
10	0

Table 7. Spreadsheet of calculations from coalbed data of Figure 3 for the lognormal q as a function of various values of p. NORMSINV is the spreadsheet function that computes z; NORMSDIST is the function that computes $1 - q$.

p Pop. Prop.	$(p)(N)$ $(p)(39)$	$1 - p$	z NORMSINV	$z - \sigma$ $(z - 1.2317)$	$1 - q$ NORMSDIST	**lognormal q (theoretical)**	**statistic q (empirical)**
0.1	3.9	0.9	1.2816	0.0499	0.5199	0.4801	0.4775
0.2	7.8	0.8	0.8416	−0.3901	0.3482	0.6518	0.6413
0.3	11.7	0.7	0.5244	−0.7073	0.2397	0.7603	0.7527
0.4	15.6	0.6	0.2533	−0.9784	0.1639	0.8361	0.8247
0.5	19.5	0.5	0.0000	−1.2317	0.1090	0.8910	0.8777
0.6	23.4	0.4	−0.2533	−1.4850	0.0688	0.9312	0.9171
0.7	27.3	0.3	−0.5244	−1.7561	0.0395	0.9605	0.9492
0.8	31.2	0.2	−0.8416	−2.0733	0.0191	0.9809	0.9758
0.9	35.1	0.1	−1.2816	−2.5133	0.0060	0.9940	0.9952

Table 8. Summary of fractal lognormal percentage assessment of coalbed resources.

Percentage of coalbed plays $p100\%$	Percentage of total coalbed resources $q100\%$ Theoretical (lognormal q)	Empirical (statistic q)
0	0.00	0.00
10	48.01	47.75
20	65.18	64.13
30	76.03	75.27
40	83.61	82.47
50	89.10	87.77
60	93.12	91.71
70	96.05	94.92
80	98.09	97.58
90	99.40	99.52
100	100.00	100.00

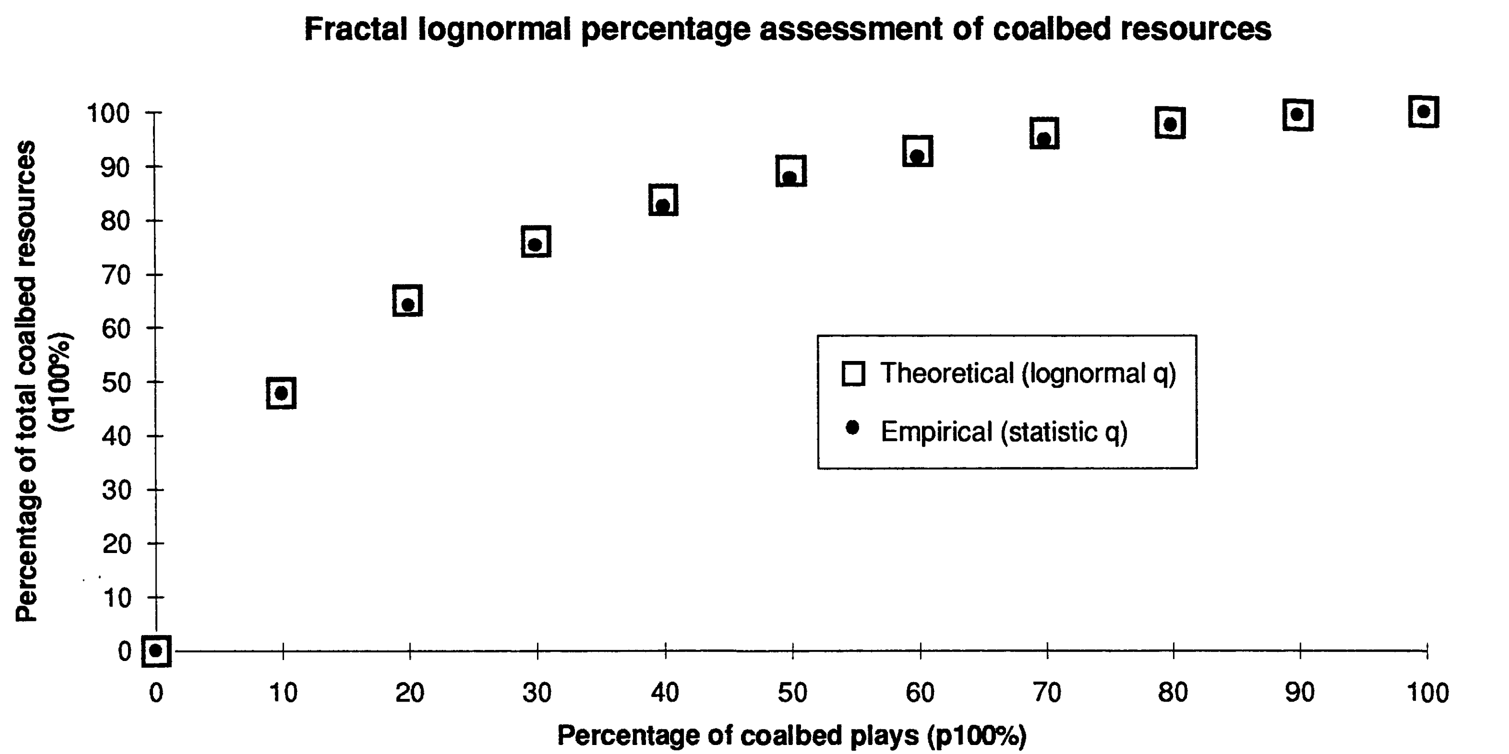

Figure 4. Theoretical percentage of total coalbed resources using the lognormal q as a function of the percentage of coalbed plays, along with the corresponding empirical values using the statistic q. Data are from Table 8.

SUMMARY AND CONCLUSIONS

1. "The 20/80 law" is generalized as the $p100/q100$ law in probabilistic form for any probability distribution that models the population distribution. The parameter q is called the proportion of total value and defined in terms of the population proportion p and conditional expectation.

2. When the population distribution is modeled with the lognormal distribution, the $p100/q100$ law in probabilistic form produces the $p100/q100$ law in lognormal form. The proportion of total value is derived, and the parameter q is a function only of p and the shape parameter σ. The parameter q is scale-free; i.e., it does not depend upon the lognormal scale parameter μ. Because it is scale invariant, the proportion q is a parametric fractal.

3. The $p100/q100$ law in lognormal form can be applied to unconventional natural gas resources in continuous-type and coalbed plays in the United States. The lognormal q is a function of p for each commodity such that $p100\%$ of the plays account for $q100\%$ of the total resources of the accumulations. In both cases the theoretical percentages of total resources using the lognormal q are close to the empirical percentages from the data using the statistic q. For example, 20% of the 34 continuous-type plays account for 63.2% of the total continuous-type resources of the plays if we use the lognormal q, or for 59.3% if we use the statistic q; 20% of the 39 coalbed plays account for 65.2% of the total coalbed resources of the plays if we use the lognormal q, or for 64.1% if we use the statistic q.

4. This new probabilistic methodology for percentage assessment of resources lends itself as an ideal application for spreadsheet software.

REFERENCES

Aitchison, J. and Brown, J.A.C., 1957, The lognormal distribution: Cambridge, Cambridge University Press, 176 p.

Crovelli, R.A., 1995, Fractal lognormal percentage assessment of petroleum field sizes in a play — Application of a generalized 20/80 law: U.S. Geological Survey Open-File Report 95-646, 17 p.

Crovelli, R.A., 1995, The generalized 20/80 law using probabilistic fractals applied to petroleum field size: Nonrenewable Resources, v. 4, no. 3, p. 233-241.

Crovelli, R.A. and Balay, R.H., 1995, Probabilistic methodology and computer programs for assessment of unconventional oil and gas resources, *in* Gautier, D.L., Dolton, G.L., Takahashi, K.I., and Varnes, K.L., eds., 1995 National Assessment of United States oil and gas resources - Results, methodology, and supporting data: U.S. Geological Survey Digital Data Series DDS-30 [CD-ROM].

Crow, E.L. and Shimizu, Kunio, eds., 1988, Lognormal distributions—Theory and applications: New York and Basel, Marcel Dekker, Inc., 387 p.

Douglass, M.E. and Douglass, D.N., 1993, Manage your time your work yourself: New York, AMACOM, 194 p.

Gautier, D.L., Dolton, G.L., Takahashi, K.I., and Varnes, K.L., eds., 1995, 1995 National Assessment of United States oil and gas resources - Results, methodology, and supporting data: U.S. Geological Survey Digital Data Series DDS-30 [CD-ROM].

Johnson, N.L., Kotz, S., and Balakrishnan, N., 1994, Continuous univariate distributions (2d ed.), v. 1: New York, John Wiley & Sons, Inc., 756 p.

Law, A.M. and Kelton, W.D., 1991, Simulation modeling and analysis (2d ed.): New York, McGraw-Hill, 759 p.

Mandelbrot, B.B., 1983, The fractal geometry of nature: San Francisco, W.H. Freeman and Co., 460 p.

Pareto, V., 1897, Cours d'Economie Politique: Lausanne and Paris, Rouge and Cie.

Rice, D.D., Young, G.B.C., and Paul, G.W., 1995, Methodology for assessment of technically recoverable resources of coalbed gas, *in* Gautier, D.L., Dolton, G.L., Takahashi, K.I., and Varnes, K.L., eds., 1995 National Assessment of United States oil and gas resources - Results, methodology, and supporting data: U.S. Geological Survey Digital Data Series DDS-30 [CD-ROM].

Schmoker, J.W., 1995, Method for assessing continuous-type (unconventional) hydrocarbon accumulations, *in* Gautier, D.L., Dolton, G.L., Takahashi, K.I., and Varnes, K.L., eds., 1995 National Assessment of United States oil and gas resources - Results, methodology, and supporting data: U.S. Geological Survey Digital Data Series DDS-30 [CD-ROM].

Schmoker, J.W., Crovelli, R.A., and Balay, R.H., 1995, Potential additions to technically recoverable resources for each continuous-type (unconventional) play of the U.S. Geological Survey 1995 National Assessment of United States oil and gas resources -- graphical and tabular presentations: U.S. Geological Survey Open-File Report 95-75E, 58 p.

Schmoker, J.W. and Oscarson, S.A., 1995, Descriptions of continuous-type (unconventional) plays of the U.S. Geological Survey 1995 National Assessment of United States oil and gas resources: U.S. Geological Survey Open-File Report 95-75B, 44 p.

Turla, P.A. and Hawkins, K.L., 1983, Time management made easy: New York, E.P. Dutton, 206 p.

U.S. Geological Survey National Oil and Gas Resource Assessment Team, 1995, 1995 National Assessment of United States oil and gas resources: U.S. Geological Survey Circular 1118, 20 p.