DOE/EIA-0103/31
Order No. 484

Technical Memorandum

A Reexamination of the Estimation of Undiscovered Oil Resources in the U.S.

April 1979

U.S. Department of Energy
Energy Information Administration
Assistant Administrator for Applied Analysis

Available from:

National Technical Information Service (NTIS)
U.S. Department of Commerce
5285 Port Royal Road
Springfield, VA 22161

Price:
Printed Copy: $4.00

Microfiche: $3.00

For sale by the Superintendent of Documents, U.S. Government Printing Office
Washington, D.C. 20402

Stock Number 061-003-00014-2

DOE/EIA-0103/31
Dist. Category UC 92

Technical Memorandum

A Reexamination of the Estimation of Undiscovered Oil Resources in the U.S.

TM/ES/79-03

April 1979

Prepared by:
Noel Uri
Oil and Gas Analysis Division
Office of Energy Source Analysis

U.S. Department of Energy
Energy Information Administration
Assistant Administrator for Applied Analysis
Washington, D.C. 20461

PREFACE

This Technical memorandum is an attempt at getting objective estimates of total discoverable and producible crude oil in the United States. By applying two functional specifications, a logistic curve and Gompertz curve, to historical data on discoveries and production in the United States, estimates are obtained of about 159 billion barrels for total producible oil.

The author is Noel D. Uri

For further information contact:

Charles Everett
Division Director
Oil and Gas Analysis
12th & Pa. Avenue, NW
Room 4447
Washington, D.C. 20461
(202) 633-9108

For copies of this report contact:

Energy Information Administration
Clearinghouse
1726 M Street, NW
Room 210
Washington, D.C. 20641
(202) 634-5694

SUMMARY

This paper is directed at estimating total producible oil in the United States. Two specific models, a logistic model and a Gompertz model, are suggested and estimated in a fashion consistent with the theoretical considerations. The results indicate that approximately 159 billion barrels are ultimately recoverable and producible of which 117 billion barrels have been produced through the end of 1978.

INTRODUCTION

Obtaining an accurate estimate of undiscovered oil resources in the United States has proved to be a most elusive proposition. There has recently been a revived interest in this subject (see Mayer, *et.al.* [6], for a short summary of the recent papers). These efforts have been directed at obtaining estimates of total producible reserves via fitting either a cumulative distribution or its first order derivative. Beyond the complexities of the estimation and hence the question of the robustness of the estimated values, the problem of serial correlation has not been satisfactorily handled. This paper is directed at simplifying the estimation sufficiently to yield robust estimates as well as deleting the problem of serial correlation and its consequent effects on the coefficient estimates (i.e., they will be efficient in the absence of serial correlation).

Before turning to the actual model specifications, it is useful to review the methodology for estimating cumulative discovery and cumulative production. This is the subject of the following section.

Estimating Maximum Recoverable Reserves

From the record of annual production, dQp/dt, the cumulative production, Qp, can be obtained. Then from the value of cumulative production and proved reserves, Qr, for any given year, the cumulative proved discoveries, Qd, may be defined by:

$$Qd = Qp + Qr \qquad (1)$$

That is, the oil whose discovery may be said to have been proved by any given year is the sum of the oil already produced plus the remaining proved reserves.

During the complete cycle of production, the curve of the rate of production dQp/dt, begins at zero and then after passing one or more maximum, ultimately returns to zero. Coincidentally, the cumulative production curve begins at zero and increases monotonically with time until it finally levels off asymptotically to the ultimate quantity Q^*, indicating total recoverable resources.

The relations between the curves of cumulative production, proved reserves, and cumulative proved discoveries for a single-cycle production history are shown in Figure 1.

Note that for a small area, the production curve for the complete cycle may involve more than one major cycle. However, in a large geographic area like a country, the production rate curve is the composite of the production from all its components, both old and new. In such an instance, production irregularities at the micro level tend to cancel out so that for such an area the production history shows every promise of giving a comparatively smooth single-cycle curve (Grunfeld and Griliches [3]).

There is a close resemblance between the cumulative discovery curve and the cumulative production curve except that in the mid-range the discovery curve precedes that of production by a nearly constant time interval Δt. Because of the similarity between the two curves, the discovery curves gives an approximate preview of the behavior of the production curve by the lead time interval Δt. That is, one may determine approximately how much oil will be produced Δt years hence by examining what the discovery curve is doing currently. This is demonstrated in Figure 2.

A third curve, that of the rate of increase of proved reserves, dQr/dt, is also of interest. There is a positive period representing the interval during which proved reserves are increasing and a negative period during which they are decreasing. The point at which the rate of increase is equal to zero coincides with the intersection between the rate-of-discovery and the rate-of-production curves. This can be seen from equation (1) by noting that

$$Qr = Qd - Qp,$$

the derivatives of which are

$$dQr/dt = dQd/dt - dQp/dt \qquad (2)$$

When proved reserves reach their maximum, the rate of increase of proved reserves is zero. That is,

$$dQr/dt = 0 \text{ and } dQp/dt = dQd/dt \qquad (3)$$

As can be seen, the curves in Figure 2 give little information about the magnitude of the complete cycle until the maximum value of the rate of increase of the proved reserves is reached. After that the three curves taken together given an increasingly accurate estimate of the degree of advancement reached over the complete cycle at any given time. In particular, after the peak in the rate of discoveries has been reached, the peak in proved reserves may be expected to occur at about $\Delta t/2$, and that in the rate of production at about Δt later.

In order to obtain analytical derivatives of the three curves, it is necessary to fit them to explicit functional relations. The logistic curve has been used effectively by others (Schanz [7]). The specific form of this equation is:

$$Qt = Q^{*}/(1+ ae^{-b(t-to)}, \qquad (4)$$

where Q denotes the cumulative quantity,

t-to denotes the time after some reference period t

e denotes the base of Napierian logarithms, and

Q*, a, and b are constants to be estimated.

Note that as time increases indefinitely (i.e., without limit), the quantity Q approaches the value Q* asymptotically. Hence, the curve of Q as a function of time begins at zero, rises initially exponentially, then slows down in its growth rate, passes its inflection point, and eventually levels off asymptotically to an ultimate maximum value Q*.

The derivative of equation (4) with respect to time, i.e., the rate of production (or discovery) from one year to the next is just

$$\frac{dQt}{dt} = a\, Q_t\, (Q^{*} - Q_t) \qquad (5)$$

where a=b/Q*. (This is easily shown by taking the derivative of Q_t in equation (4) and doing the requisite algebraic manipulations.)

One of the disadvantages of the logistic specification is the fact that it is symmetric with respect to time. This assumption has been extensively criticized by Mayer, et. al [6] and others. Consequently, is is useful to consider another functional form that is nonsymmetrical with the objective of comparing the robustness of the resulting estimates. Thus, the Gompertz curve which has a positive skew, will be used. Its cumulative distribution is given by

$$Q_t = Q^* \quad a^{b^{(t-to)}} \qquad (6)$$

where the terms are as previously defined.

The rate of production (or discovery) between periods will be given as

$$\frac{dQt}{dt} = B \; Q_t \quad (\log Q^* - \log Q_t) \qquad (7)$$

where B=-log b (log denoting logarithmic transformation to base e). With the Gompertz curve, growth in discovery or production rises rapidly to its maximum rate which occurs when actual discovery or production equals 37 percent of the maximum level. Thereafter, growth declines gradually so that the growth rate at any part above the maximum is greater than that equally distant point below the maximum. (Note the logistic curve reaches it maximum when actual discovery or production reaches 50 percent of the maximum level) (Lakhani[9]).

Either of these two functional specifications can be used to estimate the maximum producible oil in the United States. For comparative purposes, both will be used. Second, either the cumulative function or the first derivative of the cumulative function can be estimated. To simplify the estimation the latter will be opted for.

Empirical Approach

To estimate the two models, ordinary least squares with an adjustment for serial correlation could be employed. This, however, would not be using all of the available information. Specifically, the supposition is made that, regardless of the model, Q^* for both cumulative production and cumulative proved discoveries are the same. To impose this restriction (and at the same time test its appropriateness), parameter estimates are obtained via maximum likelihood estimation. 1/

Data

The data used in the estimation were obtained from the AGA, et. al. [1]. This is the conventional source. Data on cumulative production are available back to 1920 but reliable data on cumulative proved discoveries begin only in 1945. Therefore, since the equations are being estimated coincidentally, the years 1945 through 1977 serve as the sample period.

Estimation Results

The logistic model characterized in equation (5) and the Gompertz model characterized in equation (7) were estimated by maximizing the likelihood function. For both models the null hypothesis that Q* is equal for production and proved discoveries were tested and could not be rejected at the 95 percent level. The results are presented in Table 1 for the logistic model and in Table 2 for the Gompertz model imposing this restriction.

To be consistent with the earlier work of Hubbert [4], [5], t_o was taken to be 1900. All of the coefficient estimates for both models' specifications are significantly different from zero at the 95 percent level. Finally, with the adjustment for serial correlation, the Durbin-Watson statistic indicates the absence of that problem in both equations of both models.

When the value of Q* is computed from the Gompertz model (i.e., raise the estimated coefficient to the power of e), it is 158.02. This remarkably consistent with the estimated value from the logistic specification of 159.46. In fact, the two values are not statistically different at the 95 percent level.

What can one conclude from this? Focusing now on the estimate of Q*, the maximum cumulative proved discoveries and maximum cumulative production is about 159 billion barrels. This is approximately equal to the values computed by Mayer, et. al. [6].

A Cautionary Caveat

Both the logistic curve and the Gompertz curve have provided good fits to historical series that are approaching an asymptote. This has been adequately demonstrated in the foregoing analysis. One must be cautioned, however, not to infer that because the empirical fit is good, the functional specification has been validated. In spite of the fact that past discovery and production data has fit the suggested curves well, it is no guarantee that it will serve as an effective predictor of future behavior.

Estimating and forecasting any series is an extremely subjective process. The choice of the functional specification to a large degree predetermines what the future will be expected to look like. Further, the sample horizon (i.e., the length of the historical period) is crucial especially when the parameter estimates are not robust.

Finally, the use of discovery and production history profiles can be misleading for a number of reasons: (a) the assumption is implicit that past inter-dependencies will obtain in the future; (b) techno-logical innovation is assumed to be a stagnant factor; (c) it is assumed that no secondary peaks are possible at an aggregate level; and (d) hydrocarbon resources are assumed to be constrained physically but not necessarily institutionally (e.g., price incentives are of no import.) In each of these situtations, one might argue the assumption is untenable.

Conclusion

The foregoing analysis has been directed at more objectively estimating total producible oil in the United States. Two specific models are suggested and estimated in a fashion consistent with the theoretical considerations. The results suggest that approximately 159 billion barrels are ultimately recoverable and producible of which 117 billion barrels have been produced through the end of 1978.

In accepting this estimate, caution must be exercised by realizing that it is sensitive to a myriad of factors including the functional specification adopted, the estimating technique and the data.

TABLE 1

LOGISTIC MODEL PARAMETER ESTIMATES a/

	Coefficients α	Q^*	ρ b/
1. Production	0.00078 (0.0009)	159.426 (16.522)	-0.1837 (0.1129)
2. Discovery	0.0093 (0.0009)	159.426 (16.522)	-0.1756 (0.1111)

a/ Standard error of estimates in parentheses.

b/ Serial correlation adjustment coefficient.

TABLE 2

GOMPERTZ MODEL PARAMETER ESTIMATES a/

	Coefficients		
	α	Q*	ρ b/
1. Production	0.33005 (0.0235)	5.0627 (0.0318)	0.7077 (0.0485)
2. Discovery	0.11171 (0.01086)	5.0627 (0.0318)	-0.1817 (0.0675)

a/, b/ See Table 1 for footnotes.

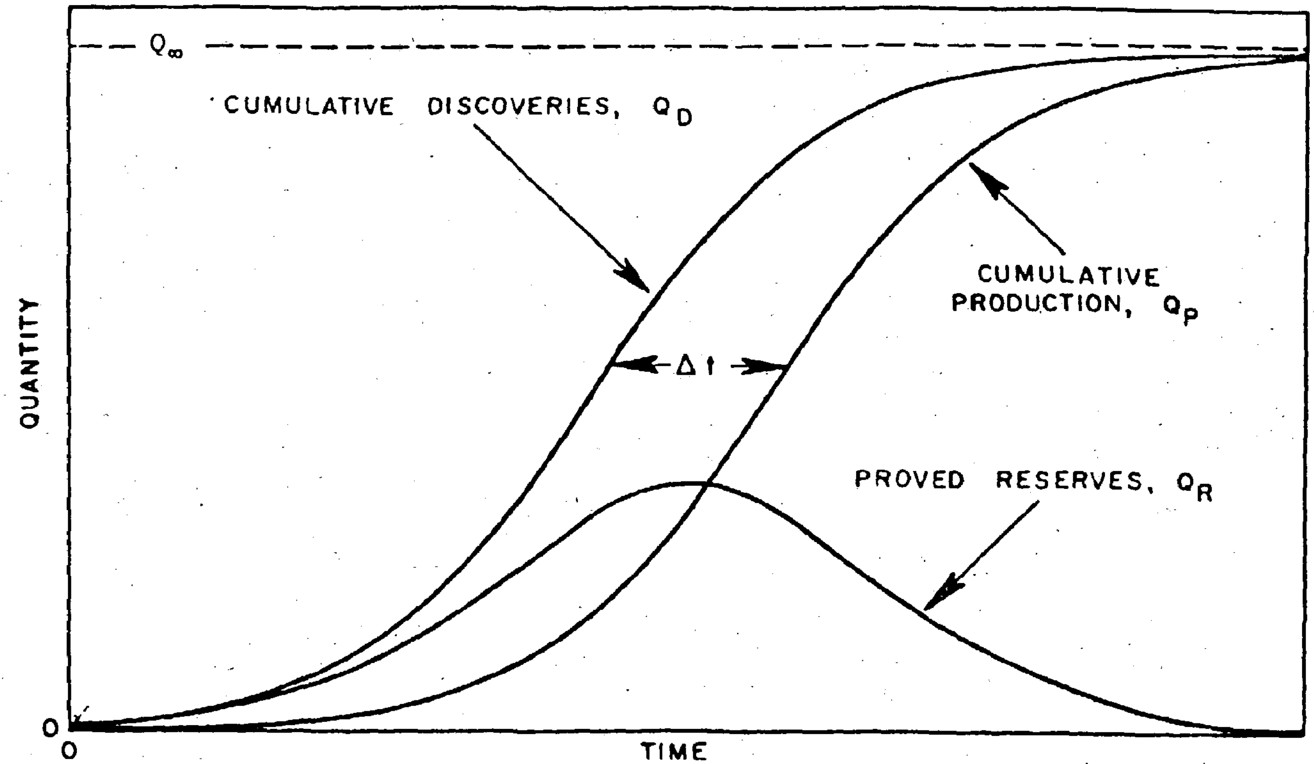

FIGURE 1: Variation with time of proved reserves, cumulative production, and cumulative proved discoveries

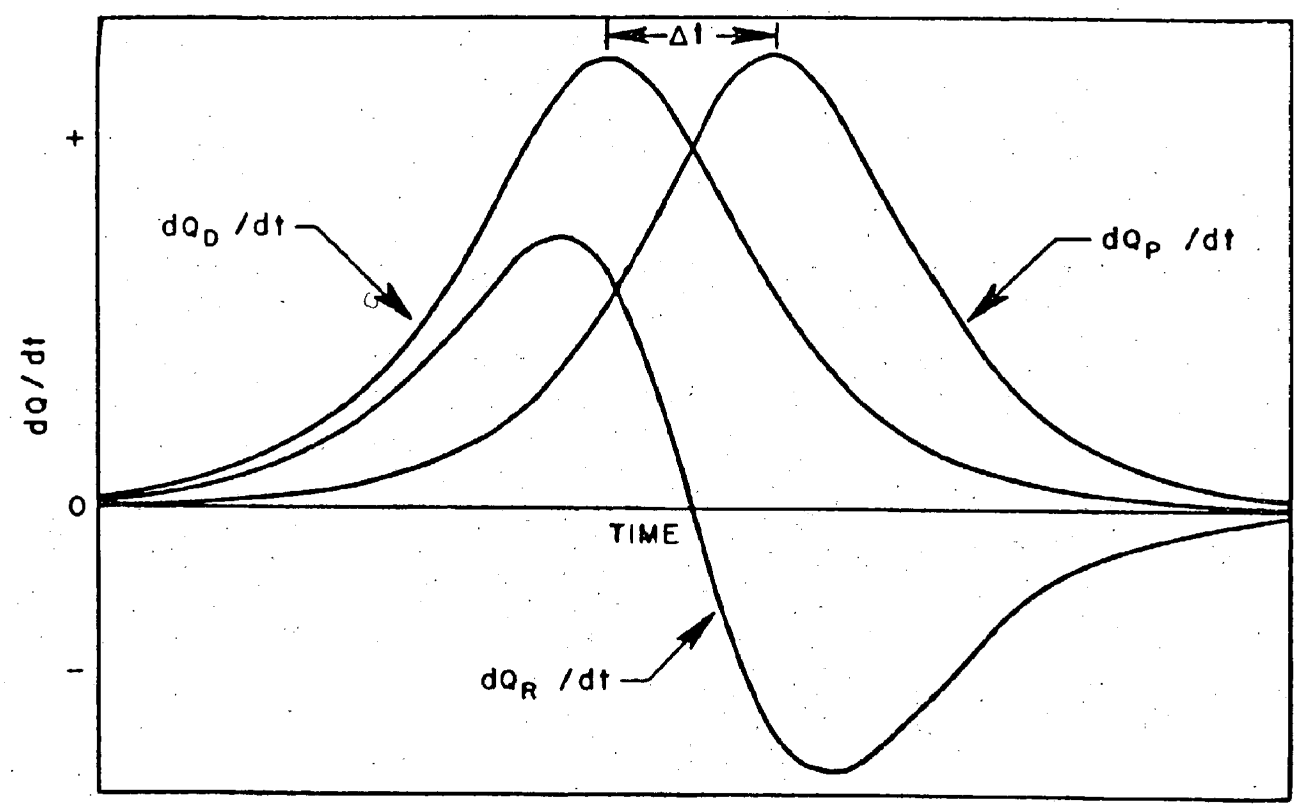

FIGURE 2: Variation of rates of production, of proved discovery, and of rate of increase of proved reserves

Footnotes

1. With the supposition that Q* for both cumulative production and cumulative proved discoveries are the same and since maximum likelihood estimates are obtained, it is possible to test this suppositions. Denoting the determinants of the unrestricted and restricted estimates of the disturbance covariance matrix by $|\Omega_u|$ and $|\Omega_r|$ when equations (5) and (7) are estimated, the likelihood ratio can be written

$$\lambda = (|\Omega_r| \; / \; |\Omega_u|)^{-n/2}$$

where n is the number of observations. The hypothesis is tested using the fact that -2 log e λ has a chi-squared distribution with degrees of freedom equal to the number of independent restrictions (one) being imposed (Goldfeld and Quandt [2]).

2. Mayer got value for various specification, of 154-178 billion barrels.

3. This section was taken from Schanz [7].

References

[1] American Petroleum Institute, American Gas Association, Canadian Petroleum Association, Reserves of Crude Oil, Natural Gas Liquids, and Natural Gas in the United States and Canada as of December 31, 1977, American Petroleum Institute, Washington, 1978.

[2] Goldfeld, S. and R. Quandt, Nonlinear Methods in Econometrics, North-Holland Publishing Company, Amsterdam, 1972.

[3] Grunfeld, Y and Z. Griliches, "Is Aggregation Necessarily Bad?", Review of Economics and Statistics, Volume 42 (February 1960), pp. 1-13.

[4] Hubbert, M.K., Energy Resources, A Report to the Committee on Natural Resources: National Academy of Science-National Research Council, Pub 1000-D, Washington, 1962.

[5] Hubbert, M.K. U.S. Energy Resources, A Review as of 1972, U.S. government Printing Office, Washington, 1974.

[6] Mayer, L.S., et. al., "Modelling the Rates of Domestic Crude Oil Discovery and Production," Resources Estimation and Validation Project, Princeton University, February 1969.

[7] Schanz, J., "Oil and Gas Resources, - Welcome to Uncertainty," Resources, Vol. 58, March 1978.

[8] Lakhani, H., "Diffusion of Environment-Saving Technological Change," Technological Forecasting and Social Change, Vol. 7 (1975), pp. 33-55.

*U.S. GOVERNMENT PRINTING OFFICE : 1979 0-281-128/603

United States
Department of Energy
Washington, DC 20585

Official Business
Penalty for Private Use, $300

Postage and Fees Paid
U.S. Department of Energy
DOE-350

FIRST CLASS MAIL

U.S.MAIL