ELSEVIER

Energy Policy 35 (2007) 3074–3088

ENERGY POLICY

www.elsevier.com/locate/enpol

Testing Hubbert

Adam R. Brandt*

Energy and Resources Group, University of California Berkeley, 310 Barrows Hall, Berkeley, CA 94720-3050, USA

Received 30 June 2006; accepted 2 November 2006
Available online 22 December 2006

Abstract

The Hubbert theory of oil depletion, which states that oil production in large regions follows a bell-shaped curve over time, has been cited as a method to predict the future of global oil production. However, the assumptions of the Hubbert method have never been rigorously tested with a large, publicly available data set. In this paper, three assumptions of the modern Hubbert theory are tested using data from 139 oil producing regions. These regions are sub-national (United States state-level, United States regional-level), national, and multi-national (subcontinental and continental) in scale. We test the assumption that oil production follows a bell-shaped curve by generating best-fitting curves for each region using six models and comparing the quality of fit across models. We also test the assumptions that production over time in a region tends to be symmetric, and that production is more bell-shaped in larger regions than in smaller regions.

Keywords: Petroleum; Hubbert theory; Oil depletion

1. Introduction and context

Since the beginning of commercial exploitation of oil, there has been great interest in two related questions: how much oil exists in the world, and when will we run out of oil? This very old discussion has recently resurfaced, as interest in oil depletion has increased along with increasing oil prices. Recent projections of global oil production have been made using a set of methods commonly referred to as the "Hubbert theory" of oil depletion, but these projections have been rejected by those who doubt the effectiveness of the method. Importantly, however, the assumptions of the Hubbert method have never been tested against possible alternatives in a peer-reviewed format using a large, publicly available data set. This paper tests some aspects of the Hubbert theory against other plausible theories of how oil production varies over time.

1.1. The Hubbert theory of oil depletion

The Hubbert theory of oil depletion was developed by Hubbert (1956). Hubbert (1956) projected future United States oil production based on two estimates of the total amount of oil that would be produced in the United States. He did not provide a functional form for his prediction in this early paper, but instead fit past production to a bell-shaped curve in which the area under the curve was equal to his estimates of the amount of total oil available. Using this method, he arrived at two predicted dates for peak production, one in the mid-1960s, the other around 1970.

Hubbert later added other elements to his analysis (Hubbert, 1959). First, he specified a functional form for his prediction, the logistic curve, stating that cumulative production over time would follow a logistic curve, and thus that yearly production would follow the first derivative of the logistic curve, which is bell-shaped. He also analyzed patterns of discovery and production. He plotted cumulated discoveries alongside cumulated production and noted that the curves were similar in shape but shifted in time (Hubbert, 1959). With this paper, most of the major elements of modern Hubbert analysis were developed.

United States oil production peaked in 1970, and with this vindication the Hubbert theory became an important tool for those concerned about depletion of natural resources (Deffeyes, 2001). This success caused Hubbert and others to project global oil production. The recent

*Tel.: +1 510 642 8829.
E-mail address: abrandt@berkeley.edu.

0301-4215/$ - see front matter
doi:10.1016/j.enpol.2006.11.004

explosion of interest in the Hubbert theory started in the 1990s with Campbell's efforts to use it to predict global oil production (Campbell and Laherrere, 1998).

Modern Hubbert modeling is really a constellation of techniques, many of which were developed by Hubbert himself in his early papers, and some of which were not. The methods used vary widely by analyst, but the core techniques of modern Hubbert analysis are as follows:

Analysis of past discoveries: discovery data are plotted, and sometimes adjusted for reserve growth, and a best-fitting curve (typically logistic or Gaussian) is matched to discoveries.

Estimation of future oil discoveries: total amounts of oil to be found are extrapolated in a number of ways, including the "creaming curve" method, which estimates an asymptote for total discoveries when plotting cumulative discoveries by cumulative new field wildcat wells drilled (Campbell and Laherrere, 1998); by using a newer technique sometimes called "Hubbert linearization" to project ultimate recovery (Deffeyes, 2001); or by using a statistical relationship such as the parabolic fractal law to infer the size of undiscovered fields using the distribution of already-discovered field sizes (Laherrere, 1996).

Projection of future production: using discovery data in conjunction with estimated future discoveries, a curve (again, typically logistic or Gaussian) is fit to historical production data such that the area under the curve equals the sum of discovered and not yet discovered oil.

It should be emphasized that a key portion of the Hubbert methodology is the estimation of ultimate production. Indeed, estimation of ultimate production has a larger effect on the accuracy of projections than other aspects of the methodology because ultimate production, or the area under the production curve, strongly affects the path of production over time. However, we do not test the accuracy of previous estimates of ultimate production here, but seek instead to test other assumptions of the Hubbert model as commonly practiced.

A number of assumptions are commonly made in modern Hubbert modeling, although some of these were not developed by Hubbert himself, and different analysts relax some of these assumptions. Commonly used assumptions include the following: that production follows a bell-shaped curve over time; that production is symmetric over time (i.e. the decline in production will mirror the increase in production, and the year of maximum production, or peak year, occurs when the resource is half depleted); that production will follow discovery in functional form and with a constant time lag; and, lastly, that production increases and decreases in a single "up–down" cycle without multiple peaks.

1.2. Alternative models of oil depletion

A number of models of oil depletion have been used to forecast future oil production. The most simple of these models, and often not thought of as a "model" at all, is the reserve to production ratio (R/P), or simply the quantity of current reserves divided by current production. Criticisms of this methodology are too numerous to cite, but the general problem with this analysis is that neither reserves nor production are constant over time, making R/P nearly valueless as a forecasting technique.

Modified versions of the Hubbert methodology have been developed. These include a model by Hallock et al. (2004) which uses a modified version of the bell-shaped curve. This curve peaks at 60% of ultimate production instead of the typical 50%. This method implies an asymmetric shape to production and a steeper rate of decline than increase.

Another simple model is a linear oil depletion model, where production increases and decreases linearly. This model has never received much attention, but Hirsch (2005) notes that United States production in the period 1945–2000 fits a linear production profile better than a bell-shaped curve.

Exponential models are another possible simple model. Hubbert used an exponential fit in the 1956 paper where he first presents his method, plotting United States coal and oil production on a semi-logarithmic scale, noting the straight line over much of history, indicating exponential growth (Hubbert, 1956). Also, Wood et al. (2000) assumed a 2% exponential growth for world oil production, followed by a decline "at an R/P ratio of 10". This decline at a constant R/P of 10 is equivalent to exponential decline of 10% per year.

Laherrere (2005) has constructed multi-cycle models where production follows a number of discovery cycles with a constant shift in time. These curves have been prepared for regions such as France and Illinois, where there is a significant bimodal discovery trend that can be mapped onto the bimodal production trend.

Hirsch studied peaking rates of a small number of production regions, including the United States, Texas, the United Kingdom, and Norway, and noted that production peaks have tended to be steeper and sharper than predicted by the Hubbert theory (Hirsch, 2005). Some bottom-up modeling efforts, using models that simulate finding and extracting resources over time, suggest that production would be roughly bell-shaped, but not necessarily symmetric (Bardi, 2005; Reynolds, 1999). Bardi (2005) critiques the assumption of symmetrical production over time, stating that there is "no magic in the 'midpoint' of the production of a mineral resource" and that production can exhibit a decline rate greater than the rate of increase.

1.3. Problems with current depletion analysis

There are significant difficulties with current methods of predicting future oil production. Two classes of problems emerge: those resulting from poor data, and those resulting from uncertain terminology and methodologies.

The first class of problems stems from fundamental uncertainity about resource availability and poor access to oil industry data. There is still uncertainty with respect to remaining volumes of oil to be found. This uncertainty is necessarily decreasing over time as exploration continues. Andrews and Udall and Ahlbrandt have collected projections of estimated ultimate recovery (EUR) of oil, plots of which suggest that we are perhaps asymptotically approaching stable estimates of total recoverable conventional oil (Ahlbrandt, 2005; Andrews and Udall, 2003). Second, there is no access to data from many countries, including the producers that are most likely to influence the date of peak production, such as OPEC.

This lack of data manifests itself repeatedly and in multiple guises in Hubbert-type analysis. Most analysts using the Hubbert methodology have used proprietary datasets which are not accessible except at high cost (Campbell and Laherrere, 1998; Campbell and Sivertsson, 2003; Campbell, 1997). Unfortunately, most publicly available data sets only contain data back to the 1970s or 1960s (BP, 2005; EIA, 2005). This lack of access to complete data sets makes checking the work of Hubbert theorists impossible and invites skepticism of their results (Lynch, 2003).

Another manifestation of poor data availability is that most studies rely on a small number of cases, such that the United States is plotted many times, but other regions are not. Campbell (2005) plots dozens of curves, but his plots often do not overlay projections and historical data, and he does not discuss the quality of the model fit to data (Campbell and Laherrere, 1998; Campbell, 2005). This makes it difficult to estimate even qualitatively the goodness of fit of his figures.

Predictions of future oil production are also hindered by uncertainty about terminology and methodologies. Perhaps the most fundamental points of confusion are that the quantities being predicted are often unclear, and there is disagreement about what quantities are important to predict. Campbell is quite clear in his definitions (Campbell and Sivertsson, 2003), focusing on a specific definition of oil and separating these production data from deep offshore, heavy oil, etc. Others, more economically oriented, pay less mind to the distinction between these resources (Odell, 1999) and emphasize the substitutability of resources (Huber and Mills, 2005). In essence, these groups disagree about which quantity is "important". Hubbert modelers are interested in production and depletion of conventional or "cheap" oil, and often do not include alternatives such as low-quality petroleum resources. Economic observers are interested primarily in the transition to substitutes for conventional oil, and they find the question of the depletion of a specifically defined resource, such as conventional oil, uninteresting or unimportant. Not surprisingly, it is difficult to find agreement between the parties when the questions that they ask are different.

Hubbert-type analysis is generally based on a bell-shaped curve, most frequently a logistic curve, but sometimes a Gaussian curve (Deffeyes, 2001). This leads to another methodological difficulty: the use of the Central Limit Theorem (CLT) as the common justification for the use of the Gaussian curve in predicting future production. As has been argued previously, there is no theoretical basis for the assumption of a Gaussian production curve based on the CLT (Babusiaux et al., 2004). The CLT acts to create "bell-shaped" distributions when distributions that are *independent of one another* are *summed*. While individual production curves are summed, they are not independent. Production at a given oil field is determined at least in part by the decisions of the producers. These producers, across regions, nations, and even at a global level, respond to common stimuli. At a regional level, common stimuli include local transport costs, availability of nearby markets, and regulatory pressures (such as state or provincial environmental mandates), while national politics can force production up or down, particularly in nations with central control over production (e.g. OPEC). And, of course, both long and short-term trends influence producers simultaneously across the globe. Thus, there is no theoretical reason to expect Gaussian production profiles to fit all cases.

Given these difficulties with current methodologies, we seek to test certain assumptions made during Hubbert analysis.

2. Methods of analysis

In this paper we test three assumptions of the Hubbert theory. We first ask if bell-shaped models fit historical production data better than other simple models. We then ask if regional oil production curves have been historically symmetric. Lastly, we test a commonly made assertion about oil depletion: that the Hubbert model fits larger regions better than smaller regions, due to a "smoothing" behavior resulting from summing smaller pro duction curves. We emphasize that we do not test the *predictive* ability of the Hubbert model, but simply examine historical data to determine how accurate these assumptions are.

2.1. Data sets used

Data are collected at four scales: United States state-level, United States regional-level (created by summing state-level data), national-level, and multi-national-level (such as continental or sub-continental). Data are collected for 139 regions in total. The sources and years included in these data are shown in Table 1. The regions studied are listed in Table 2.

These data series are formed by joining two or more separate series, because early production data were only available in earlier reference volumes. In nearly all cases the transition between data sets is smooth (i.e. values were equal or extremely close in both data sets for overlapping years near the transition).

Table 1
Data sets used in analysis

Regional level	Source[a]	Years
US State level	API (1959)	1859–1946
	API (2004)	1947–1989
	DeGolyer and MacNaughton Inc. (2006)	1990–2004
US Regional level	API (1959)	1859–1946
	API (2004)	1947–1989
	DeGolyer and MacNaughton Inc. (2006)	1990–2004
National level	API (1971)	1859–1964
	DeGolyer and MacNaughton Inc. (2006)	1965–2004
World Regional level	API (1971)	1859–1964
	DeGolyer and MacNaughton Inc. (2006)	1965–2004

[a]Data for some regions, specifically Sudan, Morocco, and Equatorial Guinea, were collected from the EIA (2005) *International Energy Annual 2003* because these regions were not included in DeGolyer and MacNaughton Inc. (2006). Some states use API data until 2002 rather than DeGolyer and MacNaughton data from 1990 to 2004. Included in these states are Arizona, Louisiana, Nevada, Texas, and Washington. These regions suffered from inconsistent regional definitions that made them incomparable to the earlier data series.

Table 2
Regions analyzed[a]

United States state-level	United States regions and divisions	Nations	Nations cont.	Sub-continents and continents
Alabama	New England	Albania	Mexico	Middle Africa
Alaska	Middle Atlantic	Algeria	Morocco	Northern Africa
Arizona	East North Central	Angola	Netherlands	Western Africa
Arkansas	West North Central	Argentina	New Guinea	Caribbean
California	South Atlantic	Australia	New Zealand	Central America
Colorado	East South Central	Austria	Nigeria	South America
Florida	West South Central	Bahrain	Norway	Northern America
Illinois	Mountain	Bolivia	Oman	Central Asia
Indiana	Pacific	Brazil	Pakistan	Eastern Asia
Kansas	Northeast	Brunei/Malaysia	Peru	Southern Asia
Kentucky	Midwest	Bulgaria	Philippines	South-Eastern Asia
Louisiana	South	Burma	Poland	Western Asia
Michigan	West	Cameroon	Qatar	Eastern Europe
Mississippi	West of Mississippi River	Canada	Republic of Congo[b]	Northern Europe
Missouri	East of Mississippi River	Chile	Rumania	Southern Europe
Montana	Lower-48	China	Saudi Arabia	Western Europe
Nebraska		Columbia	Spain	Australia and NZ[c]
Nevada		Czechoslovakia	Sudan	Melanesia
New Mexico		Denmark	Syria	Africa
New York		Ecuador	Thailand	America
North Dakota		Egypt	Trinidad	Asia
Ohio		Equatorial Guinea	Tunisia	Europe
Oklahoma		FSU[d]	Turkey	Oceania
Pennsylvania		France	UAE[e]	World
South Dakota		Gabon	United Kingdom	
Tennessee		Germany	United States	
Texas		Greece	Venezuela	
Utah		Hungary	Yemen	
Virginia		India	Yugoslavia[f]	
Washington		Indonesia	Zaire[g]	
West Virginia		Iran		
Wyoming		Iraq		
		Italy		
		Japan		
		Kuwait		
		Libya		

[a]The regional definitions for the grouping of US state-level and world regional data are shown in Appendix A.
[b]The Republic of the Congo is listed simply as "Congo" in the data sets and any graphics from this research.
[c]Australia and New Zealand.
[d]Former Soviet Union. Data for the FSU are as follows: from 1859 to 1930 are API data for "Russia". From 1931 to 2004 are DeGolyer and MacNaughton data for "Former Soviet Union".
[e]United Arab Emirates.
[f]Yugoslavia in recent DeGolyer and MacNaughton data sets is listed as "Former Yugoslavia".
[g]The Democratic Republic of the Congo, sometimes known as Congo–Kinshasa, is referred to in the data set used in this research as "Zaire" because data collected from early sources use its older name of Zaire.

Table 3
Six studied models and their features[a]

Model	Number of parameters	Parameters fit by software
Hubbert	3	Maximum production, year of maximum production, standard deviation of production curve
Linear	3	Year of first production, year of maximum production, slope of increase and decrease
Exponential	3	Year of first production, year of maximum production, rate of increase and decrease
Asymmetrical Hubbert	4	Maximum production, year of maximum production, standard deviation of increasing side of production curve, standard deviation of decreasing side of production curve
Asymmetrical linear	4	Year of first production, year of maximum production, slope of increase, slope of decrease
Asymmetrical exponential	4	Year of first production, year of maximum production, rate of increase, and rate of decrease

[a]The mathematical formulation of each of these models is given in Appendix B.

Data are purposely collected at all production scales. Data for smaller regions are summed into larger aggregate regions to acknowledge that production statistics are collected for regions that are arbitrarily defined with respect to geology. Some states in the United States are the size of nations, while some nations are nearly the size of continents (e.g. Former Soviet Union). This aggregation of smaller regions into larger regions creates a smoother spectrum of regional sizes. The definitions of the regions are shown in Appendix A.

2.2. Methodology to determine best fitting model in each region

We first test whether a bell-shaped production curve fits past production data more accurately than other simple models. To this end, we fit six models to each data series. These six models are of two types: symmetric three-parameter models (Hubbert, linear, exponential) and asymmetric four-parameter models (asymmetric Hubbert, asymmetric linear and asymmetric exponential). The mathematical formulation of the models is shown in Appendix B. Two tests are conducted: first, the three symmetric models are compared, and all then six models are compared. These tests will be referred to as the three-model comparison and the six-model comparison.

These models were tested using the *non-linear modeling* function of the JMP statistical software package. For each of these models, there are a number of parameters that can vary, such as peak year, rates of change, and year of first production. These parameters are adjusted by the statistical software so as to minimize the sum of squared errors (SSE). The models studied are shown in Table 3, along with the parameters fit by the software in each of the models. Schematics of each of the models are illustrated in Fig. 1, while regions that are examples of a good fit by each of the six models are shown in Fig. 2.

2.2.1. Comparing models

After the non-linear fitting algorithm determines the best values of the model parameters, we can compare the quality of the fit across models to determine which of the models studied is most appropriate for each region. There is no single method to determine which model fits best, and it must be emphasized that it is impossible to determine, statistically or otherwise, which model is truly "correct", or even to definitively say which model fits best (Motulsky and Christopoulos, 2004). Given these uncertainties, the methodology used to assign a most appropriate model to each region is described below.

Our first step is to discard regions where the fitting process is fundamentally flawed. One flaw is insufficient data to make a meaningful fit (only New England and Washington state, with 0 and 4 years of production, respectively). Regions are also discarded because they do not conform to one of the basic assumptions of the models tested. The Hubbert model, as well as all other models tested, assume that production rises and falls in a single cycle. These models also implicitly assume that the production is in some sense predictable (i.e. not stochastic). Some regions do not conform to these expectations and have multiple peaks that are separated by multi-decade time periods, such as in Fig. 3(a). Other regions have production that is so chaotic that no one of the tested models can be seen, in practical terms, as more accurate than another, as in Fig. 3(b).

In total, 16 of the 139 regions are disqualified and labeled *nonconforming* because of these reasons, and are not analyzed in either the three-model comparison or the six-model comparison. The disqualified regions include: Arkansas, Illinois, Indiana, New York, Ohio, Pennsylvania, Virginia, Washington, New England, Middle Atlantic, Northeast, Burma, Iraq, France, Philippines, and Poland. In addition, six other regions are classified as borderline-nonconforming, but were still analyzed: Central Asia, FSU, Saudi Arabia, Southern Asia, Venezuela, and Western Asia.

2.2.1.1. Three-model comparison (comparing only symmetric models). For the three-model comparison, we first analyze the total amount of error over all points between the best-fitting curve produced by each of the three models and the data. The most basic numerical measures of overall fit are the SSE and the root mean square error (RMSE). Unfortunately, these measures do not properly account for the number of parameters in a model. Any model can be

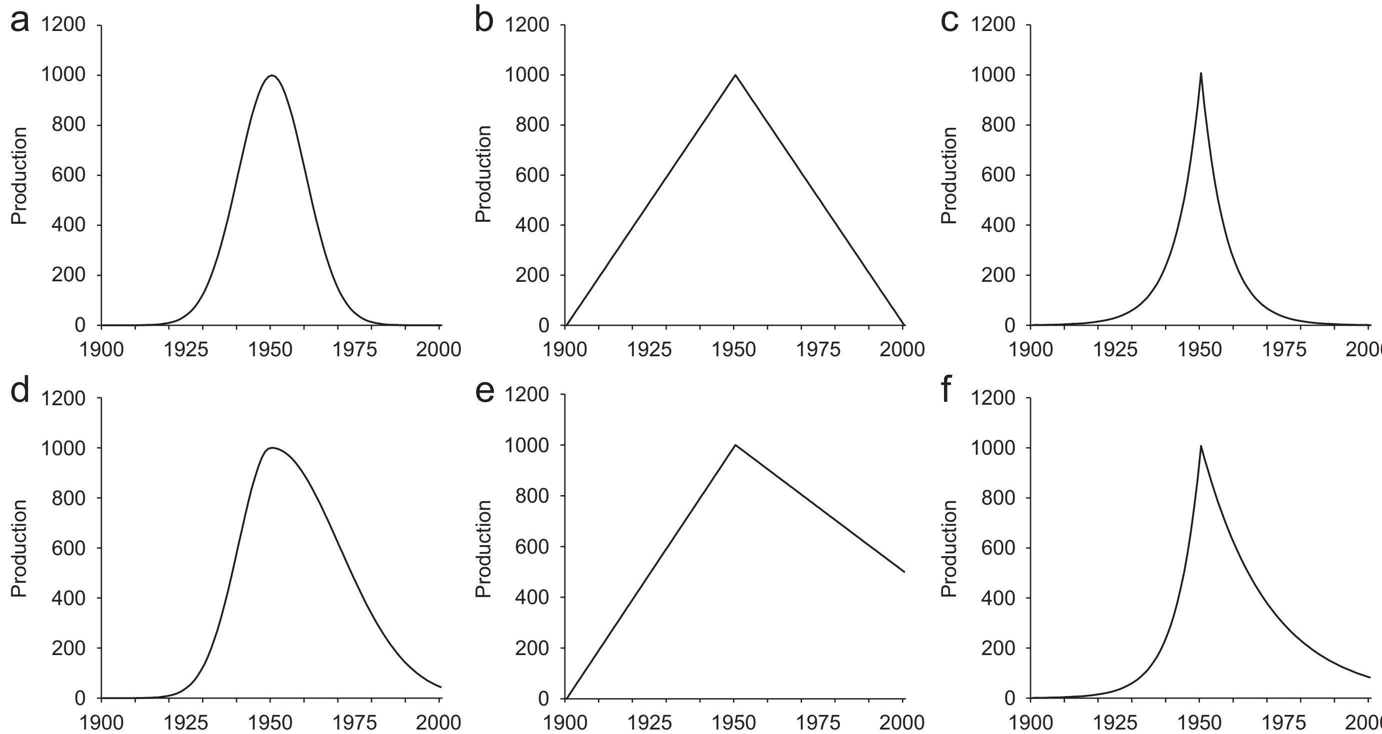

Fig. 1. Schematic illustrations of the six tested models: (a) Gaussian Hubbert production curve, (b) linear production curve, (c) exponential production curve, (d) asymmetric Hubbert production curve, (e) asymmetric linear production curve, and (f) asymmetric exponential production curve.

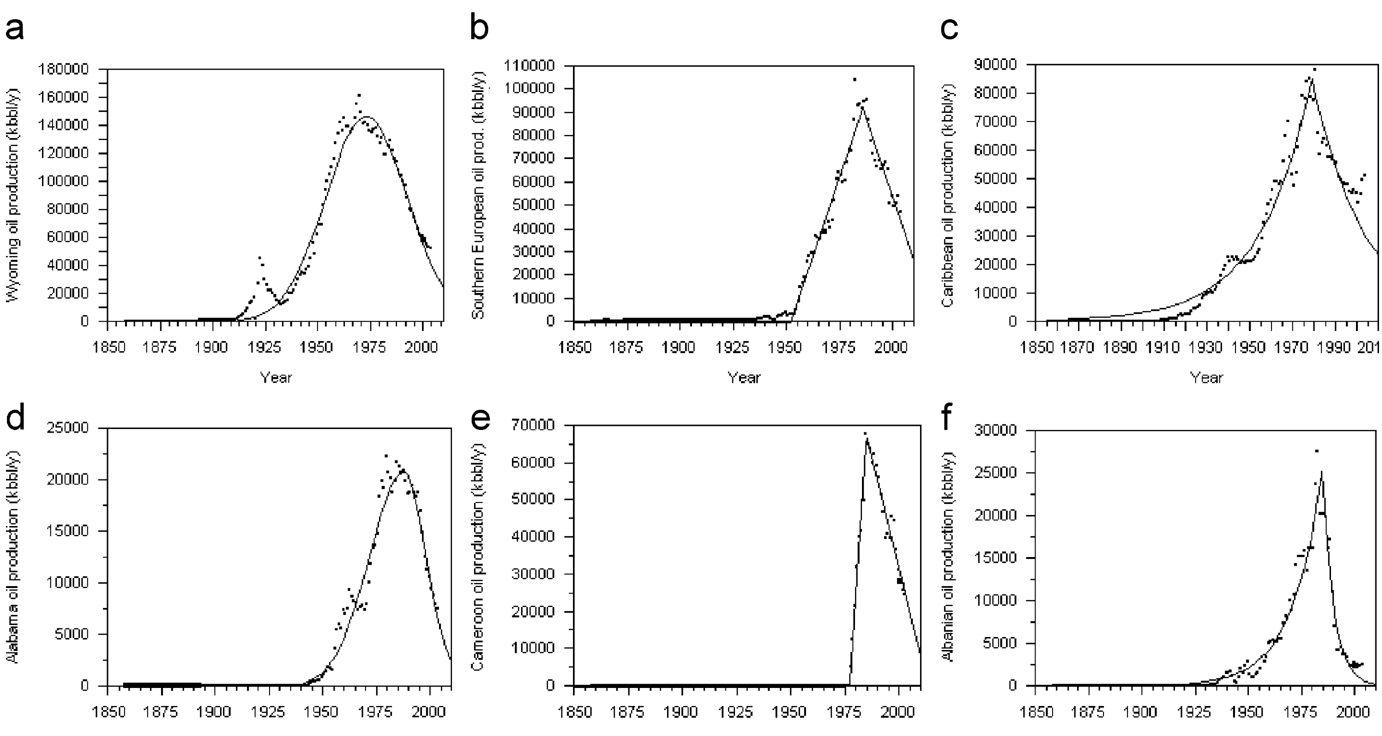

Fig. 2. Fit of models to six regions which the models fit well. In each graph, production data are in kbbl/y: (a) Hubbert fit to Wyoming production data, (b) linear fit to Southern European production data, (c) exponential fit to Caribbean production data, (d) asymmetric Hubbert fit to Alabama production data, (e) asymmetric linear fit to Cameroon production data, and (f) asymmetric exponential fit to Albanian production data.

made more complex by adding parameters, such as the differing rates of increase and decrease used in the asymmetric models in this analysis, and more complex models can in principle always fit better when measured by SSE because they are more flexible. However, the better fit of the more complex model may or may not be justified by

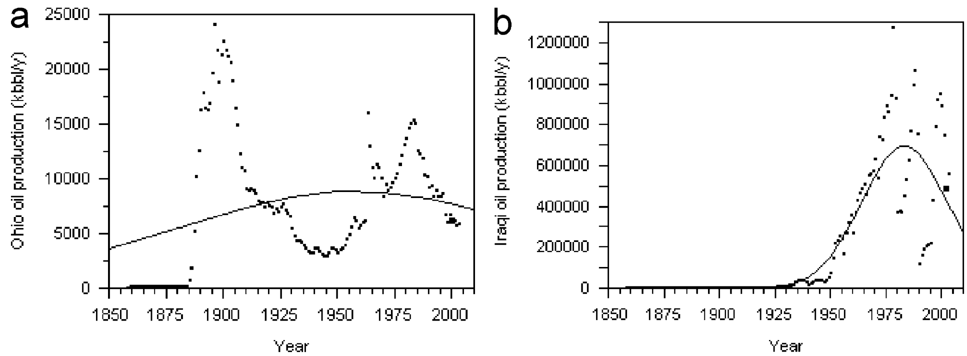

Fig. 3. These production curves, among others, were eliminated from the comparison process because of fundamental problems in determining a best fit: (a) Ohio production, a region with multiple temporally dispersed peaks, and (b) Iraq, a region with fundamentally chaotic production.

the amount of complexity created by additional parameters.

A number of approaches exist to deal with this problem (Motulsky and Christopoulos, 2004). We use Akaike's Information Criterion (AIC) because it allows us to compare models that are not mathematically "nested" (models are nested when one model can be written as a simplified version of the other). AIC allows us to compare models of different complexity while accounting for the advantage that a more complex model has in fitting (Motulsky and Christopoulos, 2004). The *AIC* score (actually the corrected *AIC* score, the AIC_c score)[1] is given as follows:

$$AIC_c = N \ln\left(\frac{SSE}{N}\right) + 2K + \frac{2K(K+1)}{N-K-1}, \quad (1)$$

where AIC_c is the corrected *AIC* score, N the number of data points in data series, *SSE* the sum of squared errors, and K the number of model parameters.

The model with the smallest AIC_c score is the most likely to be the best-fitting model (Motulsky and Christopoulos, 2004). We also calculate how much more one model is likely to be correct compared to another by using the difference in AIC_c scores. Because AIC is based on information theory, not statistics, we cannot correctly "reject" or "accept" a model as statistically significantly better than the others. We can however, determine the probability that one model is correct when compared to another, given by the equation

$$\text{Probability} = \frac{e^{-0.5\cdot\Delta AIC}}{1 + e^{-0.5\cdot\Delta AIC}}, \quad (2)$$

where ΔAIC is the (AIC_c of best-fitting model) − (AIC_c of second best-fitting model).

In the results below, we consider a best-fitting model to have "strong evidence" of being the correct model if it has >99% chance of being the correct model when compared, using Eq. (2), to the second best-fitting model.

[1]The AIC_c is corrected to account for errors that can occur with AIC when the number of datapoints is small, and it is considered superior to the AIC, even with data sets as large as the ones studied here.

We also visually inspect the models for goodness of fit. This is because there are regions where the total numerical error is minimized by a given model, but there is systematic divergence between the model and the data, or the best-fitting model is nonsensical (e.g. negative decline rates resulting in ever-increasing post-peak production). If the best fitting model in each region does not have strong evidence, as defined above, the *residuals* of the models are compared. The residual for each data point is the difference between what the best-fitting model predicts and the actual value. When comparing models, a better model fit results in residuals that are: (a) smaller in magnitude, (b) more evenly spread around zero (normally distributed), and (c) have fewer trends (i.e. fewer long runs of consecutive points above or below zero) (Motulsky and Christopoulos, 2004).

For some regions it is difficult or impossible to judge if one model or another fits better. In these cases, the AIC_c is sufficiently close between two models so as to not provide strong evidence (probability <99%), and inspection of the residuals provides no obvious choice. In these cases the best fitting model is left *undetermined.*

As an example, see Figs. 4(a) and (b), which show the Hubbert and linear fit for Hungary, a region that is among the most difficult to determine the best-fitting model. Also see Figs. 5(a) and (b), which show the residuals to these fits. AIC favors the Hubbert model over the linear model, but not strongly (80% probability). In favor of the linear model, we see smaller maximum error and less consistent error, as well as a somewhat better fit at the peak. Arguments in favor of the Hubbert model include the better overall numerical fit (lower AIC_c score), and that the curve does have some "rounded" characteristics. In this paper we select the Hubbert model, but arguments could be made for classifying this region as undetermined or linear.

2.2.1.2. Six-model comparison (symmetric and asymmetric models). We compare all six models using a similar method to that used for the comparison for the three symmetric models. Recall that the AIC_c accounts for the increased complexity of the asymmetric models and compares this to the reduction in error produced.

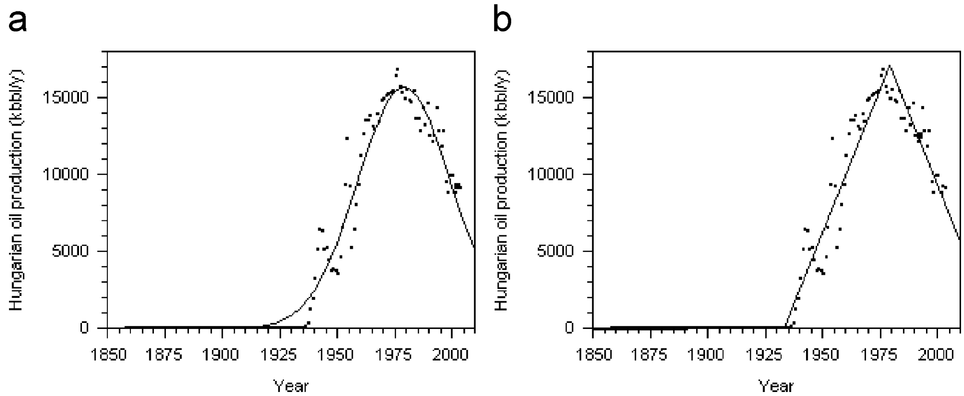

Fig. 4. Fit of two models to production data from Hungary: (a) Hubbert fit, and (b) linear fit. See residuals in Fig. 5.

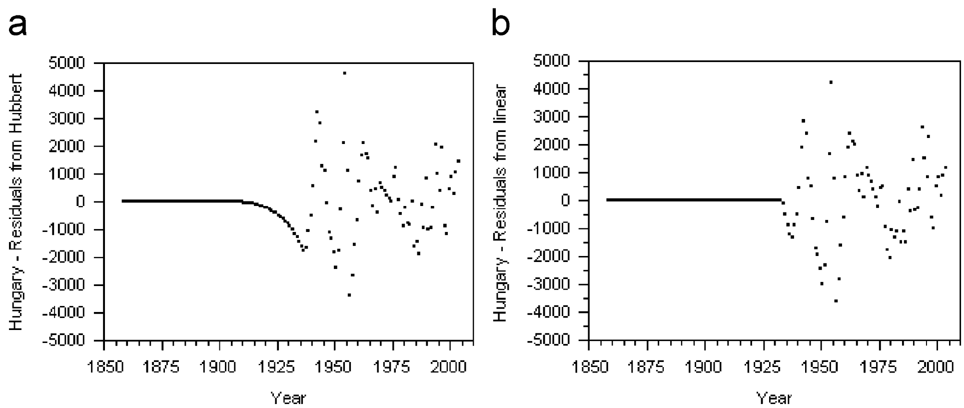

Fig. 5. Residuals from fitting two models to Hungarian production: (a) residuals from the Hubbert fit, and (b) residuals from linear fit.

In addition to the 16 regions that are disqualified as nonconforming from both the three-model and six-model comparisons, additional regions are disqualified from the six-model comparison. Because the asymmetric models include the rate of decline of the production curve as a parameter, the software cannot determine values for these parameters without sufficient data beyond a peak in production (in practice the fitting process is erratic or the confidence intervals on the decline rate parameter are very large). In total, 49 regions are not included in the six-model comparison for this reason. Consequently, the six-model comparison is performed with 74 regions (36 state and multi-state, 38 nations and groups of nations).

2.3. Methodology to test the symmetry of regional oil production

Another important assumption of Hubbert modeling that has not been rigorously tested is that oil production is assumed to be symmetric over time.

In order to test this assumption, we compare best-fitting incline and decline rates. Only the 74 regions that were fit with asymmetric models are included. It is likely that at least a few of these slopes, fit here as post-peak downslopes, will in the future be realized as temporary declines in production and not as a final decline. However, the choice of regions included is also somewhat conservative, such that regions that are thought to have recently peaked, including the United Kingdom and Norway, are not included because they do not have enough post-peak data for adequate fit.

We conduct two simple tests to gain insight into the best-fitting rates of increase and decrease. First, we examine the distribution of best-fitting rates of exponential increase and decrease in the asymmetric exponential model, r_{inc} and r_{dec}. We then study the distribution of a quantity we will call the *rate difference*, which is the difference between the best-fitting rates of exponential increase and decrease for each region. Thus, rate difference (Δr) is defined as

$$\Delta r_{exponential} = r_{inc} - r_{dec}. \quad (3)$$

2.4. Methodology to test the quality of the Hubbert fit across regions of different sizes

Some Hubbert theorists have suggested that larger regions may fit the Hubbert model more closely, due to a "smoothing" behavior, whereby noise in production from differing regions, as well as divergences from a bell-shaped model, will cancel each other out when summed. In order to test this, we compare the quality of fit across regions of different "size". We test using two definitions of size: the area of regions (km^2), and total production to date (cumulative bbl).

In order to compare fit across regions of different size, we use a "scale-invariant" measure of fit. Proper comparison requires a measure of fit that is comparable across regions as diverse as Arizona (peak production 3370 kbbl/y) and Asia (peak production 14,419,387 kbbl/y). SSE and RMSE are insufficient because the absolute amount of error will increase as the scale of production increases. For this reason, a normalized RMSE was used: the RMSE for each region was divided by the mean production from that region, thus normalizing the RMSE and making regions comparable.[2]

3. Results

3.1. Best-fitting model results

One important general result is that 16 regions were disqualified from comparison due to extremely poor fit, and six more were classified as borderline nonconforming. The disqualified regions are not a significant portion of global production (about 3% of 2004 production), but the borderline-nonconforming regions represent 36% of global production. Thus, fully a third of global production is not well characterized by models with a single up–down cycle.

3.1.1. Three-model comparison: best fit between symmetric models

In the three-model comparison we compare the Hubbert, linear, and exponential models. After the 16 nonconforming regions are disqualified, we compare 123 regions across the three symmetric models. The number of regions in which each model has the lowest AIC_c score is shown in Table 4. The number of these regions that are classified as having strong evidence from comparison of AIC_c scores are shown in the second column of Table 4. The final results, combining the AIC results with inspection of residuals, are shown in the last column of Table 4, with some models classified as "undetermined" if a clear best-fitting model did not exist.

These results suggest that in the three-model comparison, the Hubbert model fits production curves more frequently than the other two models. But, many regions are better represented by the other two models, suggesting that the Hubbert model is not dominant in its ability to fit production curves.

3.1.2. Six-model comparison: best fit between all models

Determining the best fit in the six model comparison is performed analogously to determining the best fit across the three symmetric models. After 16 nonconforming and 49 pre-peak regions are disqualified, we are left with 74 regions. The results of the AIC analysis are shown in Table 5. As can be seen, we again divide the results into models that are favored by comparison of AIC_c scores and models that are strongly favored by comparison of AIC_c scores. After visually accounting for fit in each region, one of six models is chosen as the best fitting model, or the region is left as undetermined if multiple models appear equally plausible. The final results of the six-model comparison are shown in the last column of Table 5.

Table 4
Results of three-model comparison

	Regions which AIC_c favors	Regions which AIC_c favors with strong evidence[a]	Regions in which model is best fitting[b]
Hubbert	63	48	59
Linear	36	23	26
Exponential	24	18	26
Undetermined	–	–	12
Nonconforming	16	–	16
Total	139	89	139

[a]"Strong evidence" is defined as a probability of being the correct model of greater than 99%. Those without strong evidence are still the most probable according to AIC.
[b]The best fitting regions are determined by combination of AIC analysis and inspection, as described in *Methods*.

Table 5
Results of six-model comparison

	Regions which AIC_c favors	Regions which AIC_c favors with strong evidence[a]	Regions in which model is best fitting
Hubbert	2	0	5
Linear	4	0	6
Exponential	7	1	7
Asymmetric Hubbert	16	11	14
Asymmetric linear	15	6	10
Asymmetric exponential	30	24	25
Undetermined	–	–	7
Nonconforming	16	16	16
Disqualified because of lack of post-peak data	49	49	49

[a]In many of the regions the asymmetric and symmetric versions of the same function are fit with the best and second best AIC_c score. In many of these cases, the asymmetric model was only slightly more probable than its simpler counterpart (60% vs. 40%, for example). In these cases the more complex model was discarded due to the inherent advantages given by a symmetric model.

[2]I would like to acknowledge Professor Jim Kirchner of the UC Berkeley Earth and Planetary Sciences department for his advice on normalizing the RMSE across regions of different size.

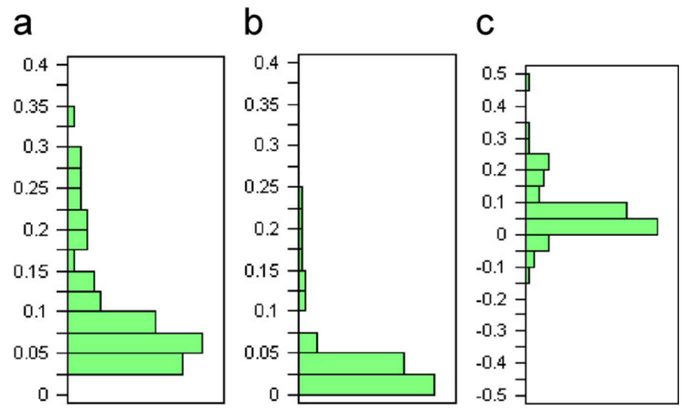

Fig. 6. Distributions of best-fitting exponential growth and decay rates for 74 regions: (a) distribution of r_{inc}, the best-fitting exponential rates of increase, $N = 74$, (b) distribution of r_{dec}, the best-fitting exponential rates of decrease, $N = 74$, and (c) distribution of rate difference, $N = 74$. Note that two outliers are off of the scale of the graph of r_{inc} at approximately 0.6 and 3 (Greece and Arizona, respectively).

Table 6
Properties of rate distributions, $N = 74$

	r_{inc}	r_{dec}	Rate difference
75th percentile	0.133	0.038	0.095
Median	0.078	0.026	0.052
25th percentile	0.056	0.016	0.019
Production weighted mean[a]	0.062	0.019	–
Mean[b]	0.148	0.041	0.108
Standard deviation	0.349	0.047	0.327

[a]This figure calculated be weighting each region's rate of increase and decrease by its share of cumulative production among post-peak regions.
[b]The mean is pulled upward for r_{inc} by a single very high value from Arizona of nearly 300% growth per year. Thus, the median is likely a more reliable value.

3.2. Symmetry of regional oil production

The distributions of rates of increase and decrease are shown in Fig. 6(a) and (b), respectively. The best-fitting values of r_{inc} and r_{dec} are collected into bins of width 0.025 (2.5%). The distributions of the best fitting values of r_{inc} and r_{dec} are quite different, with values of r_{inc} clustered between 0.025 and 0.1 and values of r_{dec} clustered between 0 and 0.05.

The distribution of the rate difference is shown in Fig. 6(c). Note that the values of the rate difference are highly concentrated above zero, reinforcing the conclusion that the typical rate of increase is higher in each region than the rate of decrease in that region. The median and a selected number of percentiles for these three measures are shown in Table 6.

3.3. Hubbert fit across regions of different size

We plot the results from the numerical measures of fit against two indexes of "size", namely cumulative production (kbbl) and area (km^2), shown in Figs. 7 and 8.

In Fig. 7 we plot the results of comparing the quality of fit to the amount of cumulative production. In Fig. 7(a) we see the very strong relationship in log–log space between an absolute measure of error (RMSE) and the level of cumulative production. The R^2 of this line in log–log space is 0.96.[3] This should be expected because as the magnitude of production increases a similar amount of relative error will result in larger absolute error.

In Fig. 7(b), however, we see the results of normalizing the error across regions. Note that after normalization most of the relationship is lost, creating a widely scattered plot. This line has an R^2 in semi-log space of only 0.037.[4] We can conclude that the normalized RMSE does not strongly scale with cumulative production, and thus that regions with greater cumulative production are not more correctly described by the Hubbert curve to any strong degree.

In Fig. 8 we plot the results of the Hubbert fit across regions of different area. In Fig. 8(a) we see the results of comparing RMSE to region area. We see some relationship between area of the region and quality of fit, and the best-fitting line in log–log space has an R^2 of 0.29. Some of this relationship between size and error is due to the correlation of area to oil production (larger regions have higher oil production on average). We see that when we normalize the error, as in Fig. 8(b), and thus remove this potential bias, nearly all of the relationship is lost, with an R^2 of only 0.013. It appears that regions of larger area do not adhere to the Hubbert model more strongly than smaller regions.

[3]The straight line in log–log space is give by the equation $\log(\text{RMSE}) = a + b\log(\text{cumulative production})$. The equation of the best-fitting line is $\log(\text{RMSE}) = -3.86 + 0.863\log(\text{cumulative production in kbbl})$.

[4]The straight line in semi-log space is given analogously by $\text{normalized RMSE} = a + b\log(\text{cumulative production})$. The equation of the best-fitting line is $\text{normalized RMSE} = 0.415 - 0.011\log(\text{cumulative production in kbbl})$.

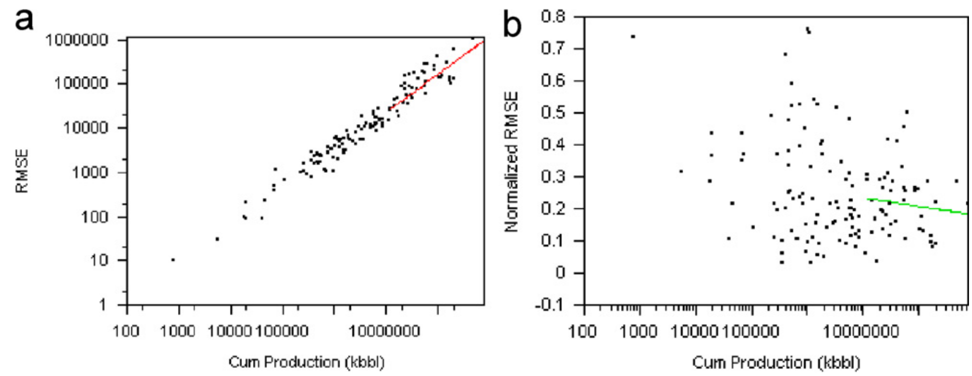

Fig. 7. Relationship of goodness of Hubbert fit to regional cumulative production: (a) RMSE by cumulative production. R^2 in log–log space = 0.96, and (b) normalized RMSE by cumulative production. R^2 in semi-log space = 0.037.

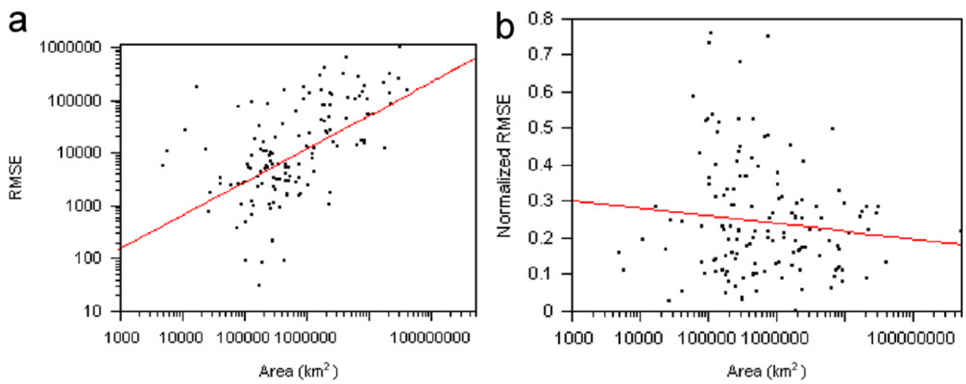

Fig. 8. Relationship of goodness of Hubbert fit to region area: (a) RMSE by area (km^2). R^2 in log–log space = 0.29, and (b) normalized RMSE by area (km^2). R^2 in semi-log space = 0.013.

4. Discussion and conclusion

It is clear from the results of this analysis that no simple, single cycle model fits all historical production curves from oil producing regions. We illustrated that when comparing the three symmetrical models, the Hubbert model is the most widely useful model, but that somewhat less than half of the regions are well-described by the linear and exponential models. We also showed that when asymmetry is allowed in our oil production curves, that the asymmetrical exponential model becomes the most useful model, and that no model dominates when we compare all six models.

When we allow for asymmetric models, we note two effects. First, the asymmetric models trump the symmetric models in most cases. This occurs even when accounting for the additional complexity of the asymmetric models. This, combined with the evidence of best-fitting rates of increase and decrease shown in Fig. 6, suggest significant asymmetry of production.

Thus, when attempting to understand *past* production, symmetric models are not satisfactory (we discuss *prediction* separately below, in which case symmetric models may well be more useful).

Second, we note that production is significantly asymmetric in one direction. As can be seen from Fig. 6(c) and column 3 of Table 6, the rate difference is overwhelmingly positive. In fact, the rate difference is positive in 67 of the 74 regions studied. The median rate of increase is 7.8% per year, while the median rate of decline is some 5% less at 2.6%. These data suggest that it is at least probable that regions where production peaks in the future will have more gentle decline rates than rates of increase. We reiterate: there is simply no evidence in the historical data that rates of decline will be generally sharper than rates of increase. This should be taken as comforting news for those concerned about an overly quick decline in production.

Hirsch's (2005) analysis, which suggests that the peaks may be sharper than suggested by the Hubbert model, is well substantiated by the large data set used in this analysis. A significant number of regions exhibit production peaks that are much sharper than the Hubbert model would suggest. This author estimates that 40–50 of the regions studied could be classified as significantly sharper around the peak than the Hubbert model suggests. This behavior partially explains the good performance of the exponential models in many cases.

This fact is important because the rates of adoption of alternatives to conventional oil will be governed by rates of change of conventional oil production. Further analysis should be performed to understand how the sharpness of the peak in production correlates with the amount of production in a region, because rapid rates of change are far more important if they occur in regions with high production.

We should note that this paper only analyzes three assumptions of the Hubbert method. As was argued earlier, a significant amount of the information contained in Hubbert predictions is provided by the EUR of oil used in the model. We did not seek to test the effectiveness of methods of predicting future volumes of oil to be recovered, but other analyses, such as those by Nehring (2006a,b,c) attempt to address these questions. In addition, the date of peak is surprisingly insensitive to EUR due to the power of exponential (or near exponential) growth in production. This is illustrated by the fact that cumulative global production from 1859 until the end of 1995 was 710 Gbbl (USGS, 2000), while production between 1996 and 2005 was 274 Gbbl (BP, 2006). Thus, approximately one quarter of total production over all time has occurred in the last 10 years. This overwhelming power of exponential growth ensures that Hubbert-like theories based on good estimates of ultimate recovery cannot be wrong by decades, regardless of the details explored here.

The application of the findings of this paper to prediction efforts is less certain. Asymmetric models are more difficult to fit to past production data than symmetric models, and are particularly so when no peak is evident. Attempting to make predictions with asymmetric models seems worse still, given that the decline rate of a region is most simply approximated under uncertainty as being equal to the rate of increase, if for no other reason than we do not have the information to justify more complex approaches. One possible methodology to take advantage of asymmetry would be to model the rate of decline in each region at a rate somewhat more gentle than the rate of incline.

This at last brings us to an important subject: oil depletion and the nature of predictability. This paper should not be misunderstood as testing the importance or existence of depletion of conventional oil, but should instead be understood as testing our methods of predicting this very real phenomenon. Also, it is important to note that Hubbert was, perhaps wisely, not wedded to his methodology. He states in *Nuclear energy and the fossil fuels*, p. 9, that

> For any production curve of a finite resource of fixed amount, two points are known at the outset, namely that at $t = 0$, and again at $t = \infty$...the production rate must begin at zero, and then after passing through *one or several maxima*, it must decline again to zero (emphasis added).

With this in mind, as well as the basic fact that the area under the production curve equals cumulative production, he suggested that we could draw a "family of possible production curves" (Hubbert, 1956).

Perhaps it is useful to consider that, had history progressed differently and Hubbert used a linear model rather than a bell-shaped model, he would likely still have been hailed as correct due to the extremely good fit of US production to the linear model as well as the Gaussian model. Or, had Hubbert analyzed a different region than the US, he would have almost certainly been less correct, simply because US production is quite symmetric compared to the global average (rate difference = 0.024 rather than the median of 0.05).

Methodological purity was not advocated by Hubbert, is not justified by the evidence presented here, and is, in the end, counter-productive. These data show that it is incorrect to emphasize that a "narrow" Hubbert methodology is correct, or to suggest that one's predictions have great accuracy given the uncertainties involved. Such a narrow methodology draws attention away from more important and fundamental points of contention between the more pessimistic Hubbert modelers and economists, including the nature of resource scarcity, energy resource substitution, and how energy system will or will not act to replace conventional oil after the inevitable peak in conventional oil production.

It would be more productive for Hubbert theorists to move from a "narrow" Hubbert methodology based solely on fitting symmetric Gaussian or logistic curves to production data to a more "broad" Hubbert methodology. A broad Hubbert methodology would present evidence that depletion of conventional oil is inevitable and becoming rapidly more important, without focusing its energies on a single functional form for production curves. Such a methodology would acknowledge the probability of a multitude of production curve shapes, use multiple types of evidence, and shun repeated attempts to project the year of peak production. Such a methodology would move us forward by focusing our attention on understanding and mitigating the social, economic, and environmental consequences of the inevitable transition away from conventional oil.

Acknowledgments

I would like to gratefully acknowledge the assistance of Anand Patil, Alex Farrell, Andrew Mills, and Jim Kirchner in the preparation and revision of this paper.

Appendix A. Definitions of regions

United States regional definitions and World regional definitions are given in Tables A.1 and A.2, respectively.

Table A.1
United States regional definitions

Name of division	Constituent states
US Census divisions[a]	
New England	Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, Vermont
Middle Atlantic	New Jersey, New York State and Pennsylvania
East North Central	Illinois, Indiana, Michigan, Ohio, and Wisconsin
West North Central	Iowa, Kansas, Minnesota, Missouri, Nebraska, North Dakota and South Dakota
South Atlantic	Delaware, Florida, Georgia, Maryland, North Carolina, South Carolina, Virginia and West Virginia
East South Central	Alabama, Kentucky, Mississippi, and Tennessee
West South Central	Arkansas, Louisiana, Oklahoma, and Texas
Mountain	Arizona, Colorado, Idaho, Montana, Nevada, New Mexico, Utah, and Wyoming
Pacific	Alaska, California, Hawaii, Oregon and Washington
Name of region	Constituent divisions
US Census regions[b]	
Northeast	New England and Middle Atlantic divisions
Midwest	East North Central and West North Central divisions
South	South Atlantic, East South Central, and West South Central divisions
West	Mountain and Pacific divisions
Name of region	Constituent states
Author-defined regions	
West of Mississippi River	All states west of the Mississippi River
East of Mississippi River	All states east of the Mississippi River
Lower-48 States	All states except Alaska and Hawaii

[a]These regions are United States *divisions*, as given in the US Census Bureau (2005) Geographical Reference Manual, Chapter 6.
[b]These regions are United States *regions*, as given in the US Census Bureau (2005) Geographical Reference Manual, Chapter 6.

Appendix B. Mathematical formulation of six tested models

B.1. Symmetric models

B.1.1. Hubbert

The Gaussian Hubbert model is defined as follows:

$$P(t) = P_{max} e^{(-(t-T_{peak})^2/2\sigma^2)}, \tag{B.1}$$

where $P(t)$ is the production in year t, P_{max} the maximum production (peak production), T_{peak} the year of peak production, and σ the standard deviation of the production curve.

B.1.2. Linear

The linear function is defined as follows:

$$\text{for } t \leqslant T_{peak}, \quad P(t) = S(t - T_{start}), \tag{B.2}$$

$$\text{for } t > T_{peak}, \quad P(t) = P|_{T_{peak}} - S(t - T_{peak}), \tag{B.3}$$

where $P(t)$ is the production in year t, T_{start} the date of first production, T_{peak} the year of peak production, and S the slope of production curve (units per year).

B.1.3. Exponential

The exponential function is formulated as follows:

$$\text{for } t \leqslant T_{peak}, \quad P(t) = e^{r\cdot(t-T_{start})}, \tag{B.4}$$

$$\text{for } t > T_{peak}, \quad P(t) = P|_{T_{peak}} e^{-r\cdot(t-T_{peak})}, \tag{B.5}$$

where $P(t)$ is the production in year t, T_{start} the year of first production (year in which production = 1 bbl), T_{peak} the year of peak production, and r the rate of change (both increasing and decreasing, percent per year).

B.2. Asymmetric models

B.2.1. Asymmetric Hubbert

This function is based on the same Gaussian curve as used above, but allows a different standard deviation on the increasing and decreasing sides of the production curve.[5] This model is defined as a compound function, with the basic Gaussian function intact in function $P(t)$:

$$P(t) = P_{max} e^{-(t-T_{peak})^2/2f(t)^2}, \tag{B.6}$$

where $f(t)$ is the sigmoid function that changes the standard deviation in the vicinity of $t = T_{peak}$:

$$f(t) = \sigma_{dec} - \frac{\sigma_{dec} - \sigma_{inc}}{1 + e^{k(t-T_{peak})}}, \tag{B.7}$$

where $P(t)$ is the production in year t, P_{max} the maximum production (peak production), T_{peak} the year of peak production, σ_{inc} the left side standard deviation (width of increasing side of production curve), σ_{dec} the right side standard deviation (width of decreasing side of production

[5]I would like to acknowledge the assistance of Anand Patil in developing the asymmetric Hubbert model.

Table A.2
World regional definitions

Name of region	Constituent oil producing nations[b]
United Nations regions[a]	
Eastern Africa	*No oil producing nations in Eastern Africa*
Middle Africa	Angola, Cameroon, Equatorial Guinea, Gabon, Democratic Republic of the Congo, Republic of the Congo
Northern Africa	Algeria, Egypt, Libya, Morocco, Sudan, Tunisia
Southern Africa	*No oil producing nations in Southern Africa*
Western Africa	Nigeria
Caribbean	Trinidad and Tobago
Central America	Mexico
South America	Argentina, Bolivia, Brazil, Chile, Colombia, Ecuador, Peru, Venezuela
Northern America	Canada, United States of America
Central Asia	Former Soviet Union[c] (Azerbaijan, Belarus, Georgia, Kazakhstan, Russia, Tajikistan, Turkmenistan, Ukraine, and Uzbekistan)
Eastern Asia	China, Japan
Southern Asia	India, Iran, Pakistan
South-Eastern Asia	Brunei Darussalam, Burkina Faso, Indonesia, Philippines, Thailand, Vietnam
Western Asia	Bahrain, Iraq, Kuwait, Oman, Qatar, Saudi Arabia, Syrian Arab Republic, Turkey, United Arab Emirates, Yemen
Eastern Europe	Bulgaria, Czech Republic, Hungary, Poland, Romania
Northern Europe	Denmark, Norway, United Kingdom
Southern Europe	Albania, Greece, Italy, Spain, Yugoslavia
Western Europe	Austria, France, Germany, Netherlands
Australia and New Zealand	Australia, New Zealand
Melanesia	Papua New Guinea
Micronesia	*No oil producing nations in Micronesia*
Polynesia	*No oil producing nations in Polynesia*
Name of continent	**Constituent UN regions**
United Nations continents[a]	
Africa	Eastern Africa, Middle Africa, Northern Africa, Southern Africa, Western Africa
America	Caribbean, Central America, South America, Northern America
Asia	Central Asia, Eastern Asia, Southern Asia, South-Eastern Asia, Western Asia
Europe	Eastern Europe, Northern Europe, Southern Europe, Western Europe
Oceania	Australia and New Zealand, Melanesia, Micronesia, Polynesia

[a]These region and continental definitions are from the UN (2005).
[b]The nations included in these regions are the nations for which oil production data were available. Of course, other countries exist within these UN regions.
[c]In UN regional definitions, Russia is considered part of Eastern Europe. Data for Russian/FSU historical oil production are available for most of the 20th century as the Soviet Union, and thus Russian data cannot be separated from data from the nations of Central Asia. Because of this, Russia is included as part of the Central Asia region.

curve), and k the rate of change from left-side to right-side standard deviation.

As can be seen by inspection from Eq. (B.7), if t is much smaller than T_{peak}, $f(t) \to \sigma_{inc}$, and that as t gets much larger than T_{peak}, $f(t) \to \sigma_{dec}$. At $t = T_{peak}, f(t)$ is equal to the average of the two values, or $f(t) = \frac{1}{2}(\sigma_{inc} + \sigma_{dec})$.

The asymmetric Hubbert model has either four or five parameters to be fit, depending on if the value k is to be fit. For all tests in this paper, k is fixed at the value 1, making it a four-parameter function, like the other asymmetric functions.

B.2.2. Asymmetric linear

The asymmetric linear is defined similarly to the symmetric linear model, but with a flexible downslope:

$$\text{for } t \leqslant T_{peak}, \quad P(t) = S_{inc}(t - T_{start}), \tag{B.8}$$

$$\text{for } t > T_{peak}, \quad P(t) = P|_{T_{peak}} - S_{dec}(t - T_{peak}), \tag{B.9}$$

where $P(t)$ is the production in year t, T_{start} the year of first production, T_{peak} the year of peak production, S_{inc} the slope on increasing side of production curve (units per year), and S_{dec} the slope on decreasing side of production curve (units per year).

B.2.3. Asymmetric exponential

The asymmetric exponential model is formulated analogously to the symmetric exponential curve:

$$\text{for } t \leqslant T_{peak}, \quad P(t) = e^{r_{inc}(t - T_{start})}, \tag{B.10}$$

$$\text{for } t > T_{peak}, \quad P(t) = P|_{T_{peak}} e^{-r_{dec}(t - T_{peak})}, \tag{B.11}$$

where $P(t)$ is the production in year t, T_{start} the year of first production (year where production = 1 bbl), T_{peak} the year of peak production, r_{inc} the rate of increase (percent per year), and r_{dec} the rate of decrease (percent per year).

References

Ahlbrandt, T.S., 2005. Global overview of petroleum resources. In: Workshop on Trends in Oil Supply and Demand and Potential for Peaking of World Oil Production. National Academy of Sciences, Washington, DC.

Andrews, S., Udall, R., 2003. Oil prophets: looking at world oil studies over time. In: Campbell, C.J. (Ed.), International Workshop on Oil Depletion. ASPO, Paris, France.

API, 1959. Petroleum Facts and Figures, centenial ed. American Petroleum Institute, New York.

API, 1971. Petroleum Facts and Figures, vol. 1971. American Petroleum Institute, New York.

API, 2004. Basic Petroleum Data Book: Petroleum Industry Statistics. American Petroleum Institute, New York.

Babusiaux, D., Barreau, S., Bauquis, P., et al., 2004. Oil and Gas Production: Reserves, Costs, and Contracts. Editons TECHNIP, Paris.

Bardi, U., 2005. The mineral economy: a model for the shape of oil production curves. Energy Policy 33 (1), 53–61.

BP, 2005. BP statistical review of world energy. Technical Report, British Petroleum.

BP, 2006. BP statistical review of world energy. Technical Report, British Petroleum.

Campbell, C.J., 1997. The Coming Oil Crisis. Multi-science Publishing, Brentwood, Essex, UK.

Campbell, C.J., 2005. Oil Crisis. Multi-science Publishing, Brentwood, Essex, UK.

Campbell, C.J., Laherrere, J., 1998. The end of cheap oil. Scientific American 278 (6).

Campbell, C.J., Sivertsson, A., 2003. Updating the depletion model. In: Second International Workshop on Oil Depletion, Paris, France.

Deffeyes, K.S., 2001. Hubbert's Peak: The Impending World Oil Shortage. Princeton University Press, Princeton, Oxford.

DeGolyer and MacNaughton inc., 2006. 20th century petroleum statistics, 2005. Technical Report.

EIA, 2005. International Energy Annual 2003.

Hallock, J.L., Tharakan, P.J., Hall, C.A.S., Jefferson, M., Wu, W., 2004. Forecasting the limits to the availability and diversity of global conventional oil supply. Energy 29 (11), 1673–1696.

Hirsch, R.L., 2005. The shape of world oil peaking: learning from experience. Technical Report.

Hubbert, M.K., 1956. Nuclear energy and the fossil fuels. In: Meeting of the Southern District, Division of Production, American Petroleum Institute. Shell Development Company, San Antonio, Texas.

Hubbert, M.K., 1959. Techniques of prediction with application to the petroleum industry. In: 44th Annual Meeting of the American Association of Petroleum Geologists. Shell Development Company, Dallas, TX, p. 43.

Huber, P.W., Mills, M.P., 2005. The Bottomless Well. Basic Books, New York.

Laherrere, J., 1996. Parabolic fractal distributions in nature.

Laherrere, J., 2005. Peak oil and other peaks. In: CERN Meeting, 3 October 2005, Geneva.

Lynch, M.C., 2003. Petroleum resources pessimism debunked in hubbert model and hubbert modeler's assumptions. Oil and Gas Journal 101 (27), 38–47.

Motulsky, H., Christopoulos, A., 2004. Fitting Models to Biological Data Using Linear and Non-linear Regression: A Practical Guide to Curve Fitting. Oxford University Press, New York.

Nehring, R., 2006a. How hubbert method fails to predict oil production in permian basin. Oil and Gas Journal 2006 (April 17th), 30–35.

Nehring, R., 2006b. Post-hubbert challenge is to find new methods to predict production, EUR. Oil and Gas Journal 2006 (April 24th), 43–51.

Nehring, R., 2006c. Two basins show hubbert's method underestimates future oil production. Oil and Gas Journal 2006 (April 3rd), 37–44.

Odell, P.R., 1999. Dynamics of energy technologies and global change. Energy Policy 27 (12), 737–742.

Reynolds, D.B., 1999. The mineral economy: how prices and costs can falsely signal decreasing scarcity. Ecological Economics 31 (1), 155–166.

UN, 2005. Composition of macro geographical (continental) regions, geographical sub-regions, and selected economic and other groupings (Updated 15 December 2005, Accessed January 20th, 2006).

U.S. Census Bureau, 2005. Last updated September 16th, 2005, Accessed November 11, 2005. Geographic Areas Reference Manual.

USGS, 2000. US geological survey world petroleum assessment 2000. Technical Report, USGS Geological Survey World Energy Assessment Team.

Wood, J.H., Long, G., Morehouse, D.F., 2000. Long term oil supply scenarios: the future is neither as rosy or as bleak as some assert. Technical Report, Energy Information Administration.