UNCERTAINTY IN PEAK OIL TIMING

Drs. M. Schoppers and N. Murphy

Prospective Modeling
Pasadena, California, USA
Marcel.Schoppers@earthlink.net

*Being careful with assumptions and statistical techniques, we find that **a**: least-squares fitting makes assumptions that are implausible in this case; **b**: global oil discovery can justifiably be regarded as stochastic, but **c**: global production is not stochastic, it could be a sum of bell-curves or simple exponential growth, and also shows 10-year cycles; **d**: fitting a sum of two bell-curves (vs only one) delays the production peak by 2–3 years; **e**: different types of bell-curves make peaks from 2004–14, with later peaks going higher and then declining faster; **f**: each such fit yields a 95% confidence interval that its peak is correct ± ½ year, thus reminding us that fitting does not make a model true (but fitting badly does matter); **g**: Gaussians are a poor fit to oil production, and we argue that the Central Limit Theorem does not apply; **h**: the Burr, Bass, Logistic, Weibull and exponential models fit well and suggest causal theories; **i**: the models' predictions can only be tested by time passing, so we can't be sure any model is right until the oil production peak is past, but that will be much too late to be useful; **j**: higher prices may occur for many reasons (such as dollar-devaluation or high demand) so do not prove the peak is imminent.*

I. Why Analyze Uncertainty

Jumping to conclusions is so easy, we often don't know we've done it. It avoids a lot of work, too.

The timing of the conventional-oil production peak depends *inter alia* on the size of the (global) Ultimately Recoverable Resource (URR). Back-of-envelope math (next section) reveals that ±1% error in the URR implies ±⅓ year in timing the peak. Unless the URR is known with unusual accuracy, and all other uncertainties are negligible, peak oil predictions that omit error bars run a very large risk of "crying wolf", even if they're approximately correct. On the scale of 150 years of oil production, "approximately correct" might mean 20 years. Responsibility to the public demands an explicit analysis of uncertainties. This is the task we set ourselves.

Royal Dutch/Shell lowered its reserves estimates 33% in one year (2004). The URR might be unreliable to a similar degree, and for similar reasons – it is probably much smaller than the USGS estimate of 2.6×10^{12} barrels. We appreciate that geologists like Campbell, Deffeyes and Laherrère have worked hard to correct various data, but how much uncertainty is left? If 1% error in the URR implies peaking ⅓ year earlier, "the peak" might even be behind us (but for the advent of horizontal drilling?). Without knowing the spread of possible times, a pin-point date is worthless. Mathematically, every real number (time) has zero chance of being correct.

It should also be obvious that we cannot locate the peak ahead of us with better accuracy than we can locate it behind us. If we believe that there's so much noise in the world that we won't know we've passed the peak until a year or two later, then with even less information now, "a year or two" would be the smallest possible uncertainty. (NB we said "If".)

II. Back-of-the-Envelope

Supposing global oil production is a bell-shaped function of time (a Gaussian), how much would the peak move if the URR were actually smaller/larger? On this model, the quantity of oil produced each year would be

$$p = \int \frac{URR}{\sqrt{2\pi}\,\sigma}\, e^{-(t-T)^2/2\sigma^2} dt$$

The quantity under the integral sign is instantaneous production, and the integral is over one year. Now if we vary the URR, the peak of production moves … not at all. We omitted to nail down the *beginning* of the bell-curve. We can accomplish this by making

$$\frac{\sigma_2}{\sigma_1} = \sqrt[3]{\frac{URR_2}{URR_1}} \quad \text{and} \quad T_1 - 3\sigma_1 = T_2 - 3\sigma_2$$

With $URR_2 = (1+\varepsilon)\times URR_1$ a Taylor series expansion yields $\sigma_2/\sigma_1 \approx 1+\varepsilon/3$ (for $\varepsilon \ll 1$) so the peak moves by

$$T_2 - T_1 = 3(\sigma_2-\sigma_1) = 3\sigma_1(\sigma_2/\sigma_1 - 1) \approx \sigma_1 \times \varepsilon.$$

Our Gaussian model of global production found $\sigma_1 \approx$ 30 years, so ±1% of URR → ±0.3 years, or 5.5 days per billion (10^9) barrels. Same answer as in [1].

That was easy, but only because we assumed a lot. Is global oil production a bell-curve? (We can't be sure.) Is it Gaussian? (Unlikely.) If not, can we still use σ_1? (No. Our models say 2×10^{12} barrels ±1% → peak ± 0.42 years, or 7.7 days per 10^9 barrels.)

III. Statistics Results

We have a relatively uncontroversial record of past global oil production. It reflects oil fields and technologies brought on-line along the way, contributions from marginal fields, fields already in decline, decline of the British Empire and rise of the American, two world wars, two oil shocks, and population growth. *We view our analysis as extrapolating all the uncertainties of the past century, into the future up to the peak, with no need to go further.*

A. Bell Curves ?

Goodness-of-fit tests (Wilcoxon-Mann-Whitney, Kolmogorov-Smirnoff, Kuiper) are not valid when applied to models whose parameters are estimated from the data. We did it anyway, and got "null" results: there is insufficient evidence to reject bell-curve models. This doesn't mean that some bell-curve model is correct.

"Pearson's r" test found no correlation between oil discoveries from one year to the next, i.e. discoveries appear to be random. But there is > 99% confidence that annual oil production is not a random process. (It looks like a bell-curve, not very noisy.)

After we had fitted our best models to the oil production data, the residuals (data – model) were not random noise (> 99% confidence from autocorrelation with Pearson's r), so none of our models are complete. The residuals hint of a 10-year cycle.

B. Least-squares And Chi-squares

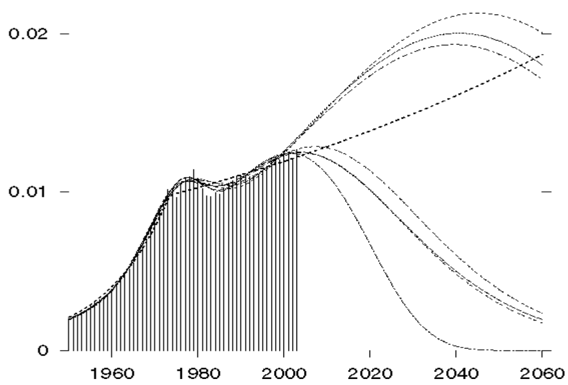
Least-squares fitting presumes that all data have equal significance, i.e. equal-size error-bars. This is highly unlikely, as annual production has grown from 0 to 25×10^9 barrels. More likely, the annual variance (σ^2) is proportional to annual production (no production, no uncertainty), and the cumulative variance is the sum of annual variances. Weighting the model's fit to each year's production p_i by a confidence $\propto 1/p_i$ leads to the "chi-square" goodness-of-fit criterion $\chi^2 \propto \Sigma\ (p_i\text{–fit}_i)^2/p_i$. (The proportionality constant is unknown, and is not needed to compare models.)

Despite the preceding point, we used both least-squares and χ^2 criteria to fit models to both annual and cumulative global oil production. Their scores correlate well.

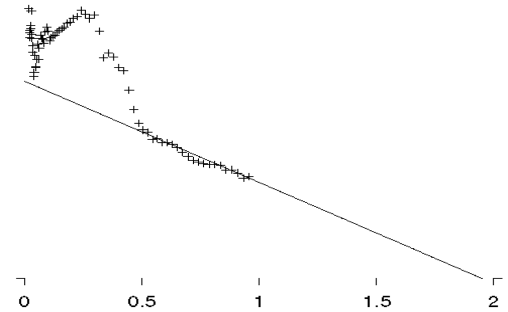
The models we fitted included single bell-curves, weighted sums of two bell-curves, and two exponential-growth curves laid head-to-toe. We used 8 key types of bell-curves, which behave like several dozen other types under suitable choices of their parameters. For example, Student's t distribution can behave like the Cauchy and Gaussian distributions (but we fitted a Gaussian also, to see how it scored).

C. Devilish Details

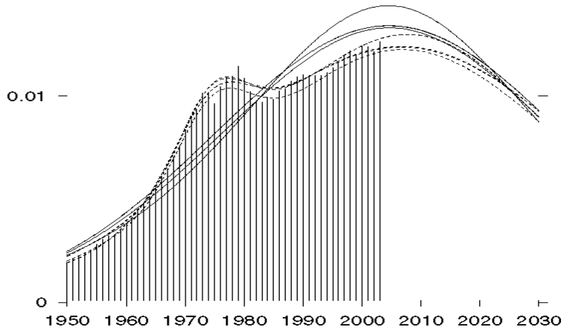
If the URR was allowed to vary as part of fitting, one model moved it to 1.5×10^{12} barrels and peaked around 1995; some models moved it to $\approx 5\times10^{12}$ and peaked around 2045 (see figure below)! The latter behavior arose for models with too-thin tails on the left side; it suggests that the Gaussian-related bell-curves (Gaussian, lognormal, gamma, Student's t) are not very good models of annual oil production.



To prevent such extremes, the URR had to be pinned down, so we charted y = annual production / cumulative production against x = cumulative production (see below). Laherrère had written that if the production curve was a Logistic, this would produce a straight line having the URR as its x-intercept. This turned out to be true of all the bell-curves we examined. The URR implied by production data is slightly less than 2.0×10^{12} barrels. We used 2.0×10^{12} to fit all our models.



The next difficulty was that when we fitted pure bell-curves to annual production data, the models' *cumulative* production far exceeded real cumulative production. This occurred because the production bump in the 1970s pulls the left side of the models upward, so that the models' annual production is too high for many years before and after that bump (figure below). We fixed this by a) forcing all the models' cumulative production to equal real cumulative production in 2003, and b) modeling annual production as a weighted sum of two bell-curves, a big one ≈ 93%, plus a small one ≈ 7% for the 1970s bump.



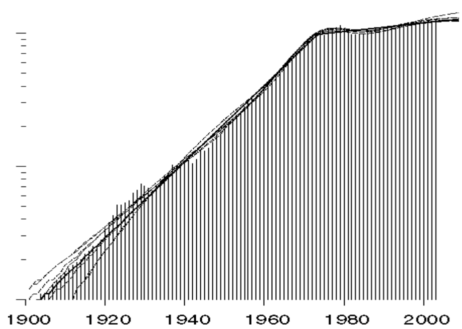
Solid lines show 1-bell models, dotted lines show corresponding weighted-sum-of-2-bells models.

Thus we discovered that the 1970s bump was not only pulling the models upward, it was pulling them left-ward too: the peaks of the 1-bell-curve models came 2–3 years earlier than they deserved. Not only was this an easy mistake to make, it also shows how far peaks can move, over a technical detail.

D. Selected Insights

The best-fitting 2-bell models were (best first) the Burr, Bass, Logistic and Weibull models. The Bass [3] Product Diffusion Equation's p parameter went to zero so it equaled the Logistic. Their peaks came in early-2005, 2008, 2008, and 2014 respectively. That's quite a spread! The later they peaked, the higher they went and the faster they declined.

The worst-fitting models were the normal, log-normal, gamma and Student's t – all cousins. To figure out why, we made a semi-log chart of the annual data:



This revealed that the Gaussian *et al* left tails were too thin (below real production in the early 1900s).

This chart led to another insight: the data looks like two straight lines joined at the mid-1970s. When we fitted such a model (two exponential growth curves joined end-to-end) it scored a better χ^2 than the Gaussian! From this model we realized that

- Global oil production grew very steadily 6.5 % per year for all of 1900–75, but only 1.2 %/year since 1982. That would explain how China's 9 %/year growth and demand could pull up the price of oil.
- The current oil price rise could also be due to US dollar devaluation, or fear of terrorism, or… Peak Oil may or may not have anything to do with it.
- Contrary to popular belief, Gaussians are not good models for time series. The Central Limit Theorem applies to random walks through *controllable* dimensions. To apply to oil production, the theory would have to be that God dropped 2 trillion barrels directly above the year 2008, and the barrels scattered forwards and backwards through time from there. That's obviously silly.
- The theories behind the well-fitting Burr, Bass and Logistic models are economic. Indeed, Peak Oil can be regarded as a change of regime, from economic choice to natural physical limits.
- The (extreme-value) theory for the Weibull model is that we pump the easy oil first, until production becomes difficult everywhere and many major fields peak at about the same time. This model fits well, predicts zero production before the mid-1800s (without being told that it's so), is comfortable with a URR of 2 trillion barrels, and predicts the last peak (in 2014) and fastest decline.
- The model consisting of two exponential-growth curves joined at the 1970s implies that there must be a third curve eventually to model falling production, but the transition can occur anytime up to (a vertical drop in) 2040, no way to predict when. Yet we have no reason to discount this model.
- *A good fit neither makes a model true, nor implies a better prediction, but a bad fit is a contradiction.*
- *Therefore, a single model can neither prove nor disprove any URR number. But* all *our bell-curve models validate the nearly-straight line that projects to* $< 2{\times}10^{12}$ *barrels of "conventional" crude.*

References

[1] A.Bartlett, "An analysis of U.S. and world oil production patterns using Hubbert-style curves." Mathematical Geology 32:1 pp 1–17 (2000).
[2] J.Laherrère, "Multi-Hubbert Modeling" http://www.oilcrisis.com/laherrere/multihub.htm 1997.
[3] F.Bass http://www.utdallas.edu/~mzjb/15