

## The Shock Model (Part II)

By Khebab (Khebab@theoildrum.com)

This post is the second part of a review of the Shock Model that was introduced in [part I](#). The shock model was developed by [WebHubbleTelescope](#) and aims at modeling oil production based on the backdated oil discovery data. In the first part, we proposed to apply a bootstrap filter in order to estimate the shock function that was previously manually set by the user. We also observed that the predictive ability was limited because of a too conservative projection of future extraction rate values.

In this second part, I propose a modification of the extraction rate function in order to improve the predictive ability of the model. This modification is based on the observation that the extraction rate function is linearly dependent to the ratio of the cumulative production to the cumulative shifted discovery. The new formulation is similar to the logistic differential equation at the difference that the Ultimate Recoverable Resource (URR) is replaced by the cumulative shifted discovery.

I look also at the modelisation of reserve growth which is an important aspect of modern oil production that is often overlooked in the peak oil community.

The code in R language is provided at the end of this post.

[break]

### Unification of the Logistic Model and the Shock Model

The shock model can be summarized by the following equation:

$$P(t) = E(t) (Q_T(t) - Q(t)), Q_T(t) = \int_{t_0}^t T(u) du \quad (1)$$

where  $Q_T(t)$  is the cumulative filtered oil discovery,  $Q(t)$  is the cumulative production and  $E(t)$  the extraction rate function (or shock function). I note  $t_l$  the last year of our baseline (i.e. the last year of available production data). One issue with this formulation is that  $E(t)$  remains constant and equals to  $E(t_l)$  in predictive mode when  $t > t_l$ . We have shown last time (see [Figure 8](#)) that this behaviour makes the predicted production follow the reserve profile  $R(t)$  leading to an immediate peak most of the time.

For the logistic model, the extraction rate is proportional to the cumulative production and all the oil is available for extraction at time  $t_0$ :

$$P(t) = \frac{KQ(t)}{URR} (URR - Q(t)), URR = Q_T(\infty) \quad (2)$$

In equation (2) there is no notion of reserves, it implicitly assumes that all the available oil has been discovered and brought online during the first year ( $t=t_0$ ). By combining equations (1) and (2), we can derive a hybrid shock model by modifying the extraction rate:

$$P(t) = \frac{K(t)Q(t)}{Q_T(t)} (Q_T(t) - Q(t)) \quad (3)$$

The extraction rate (or shock function) is then linearly dependent on the cumulative production as a fraction of the cumulative gross reserve addition:

$$E(t) = \frac{K(t)Q(t)}{Q_T(t)} \quad (4)$$

Note that  $K$  is no longer a constant as in the logistic model (2) but becomes the shock function. In predictive mode ( $t > t_l$ ),  $E(t)$  will still evolve thanks to the ratio  $Q(t)/Q_T(t)$  giving a logistic-like behaviour to the forecast.

Is the relation (4) observed in reality? the response is yes as shown on Figure 1 for the years 1985-2006.

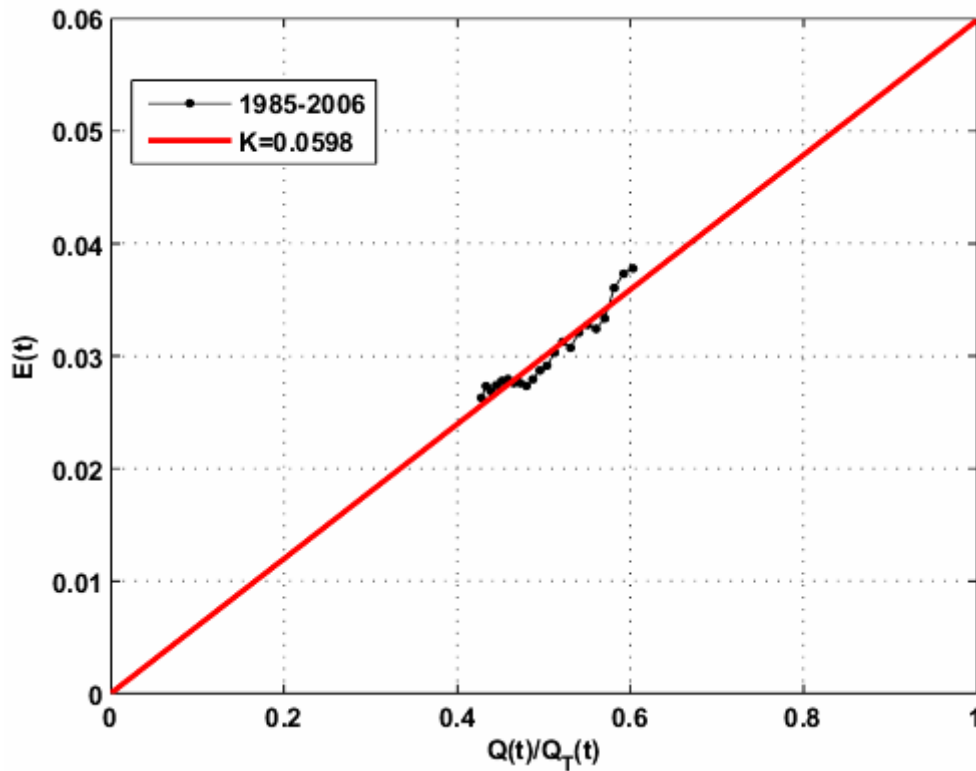


Fig 1. Estimated Shock function  $E(t)$  versus the cumulative production as a fraction of the cumulative gross reserve additions.

World production for crude oil + condensate from 1985 to 2006. The red line has a slope equals to  $K=0.0598$ .

I call Hybrid Shock Model (HSM) this model defined by the relation (3). If I repeat my last time experiment, we can see that predictions are better behaved. In particular, the forecast based on the 1900-1990 baseline gives a 2006 production close to the actual number.

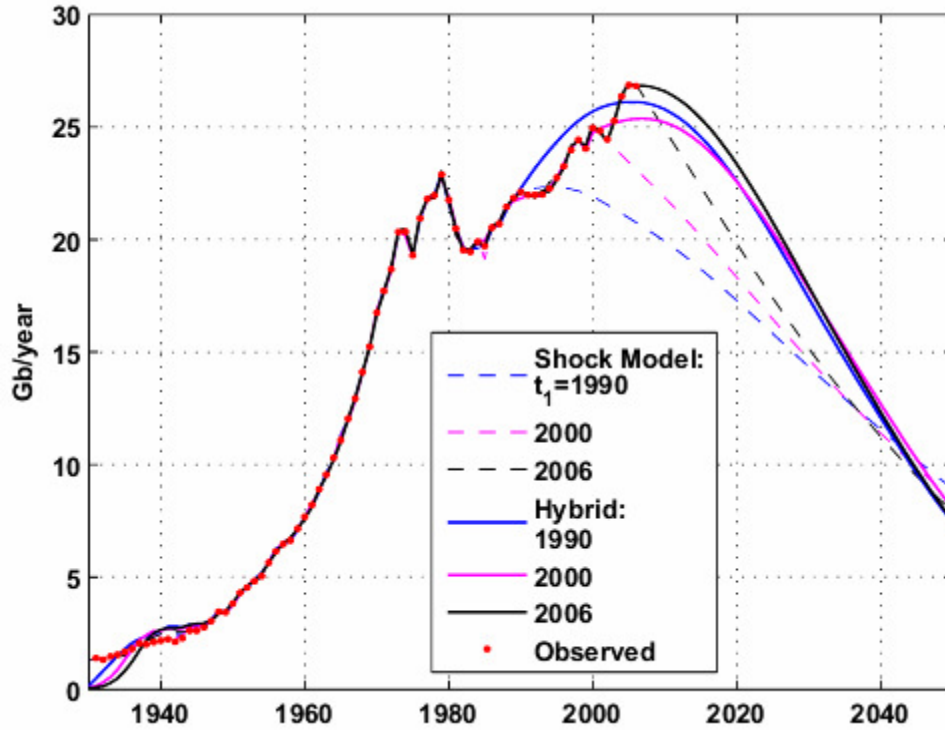


Fig 2. Comparison of the two models for different baselines.

Compared to logistic curves derived from the Hubbert Linearization for different time ranges (1983-1990,1983-2000,1983-2005). We can see that the HSM is better behaved mainly because the logistic curve is unable to account for the 1970-1980 production bump.

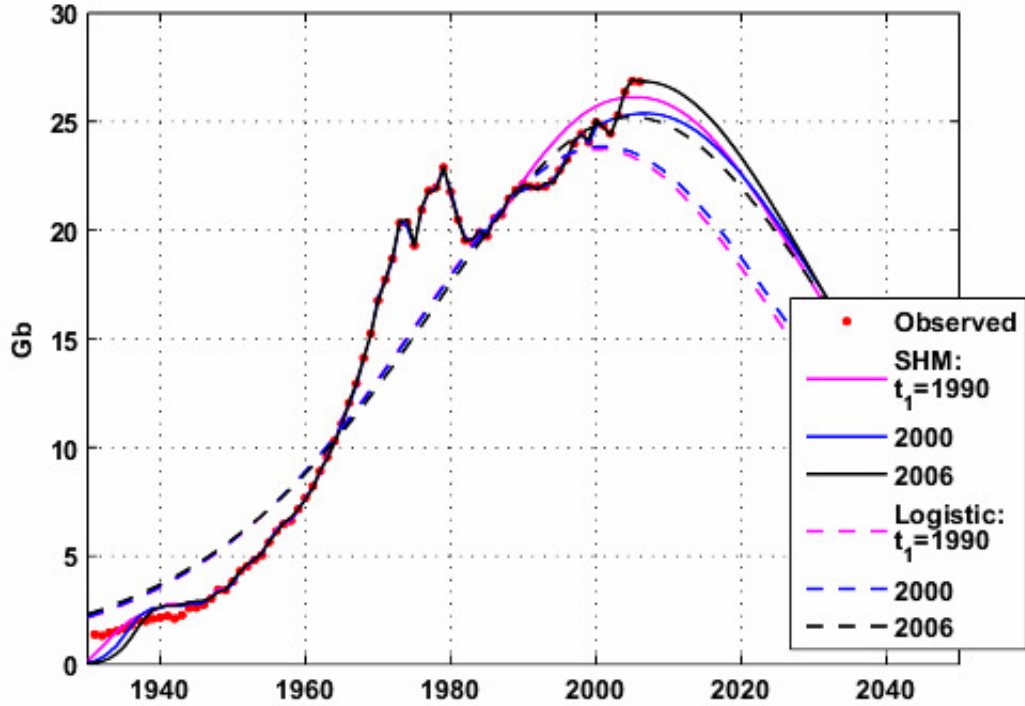


Fig 3. Comparison of the logistic curves and the HSM for different baselines.

On Figure 4, the estimated HSM shock function for the baseline 1900-2006 is close to the one given by the SM except in the predictive part where we can see clearly the convergence of  $E(t)$  toward  $K(t)$ .

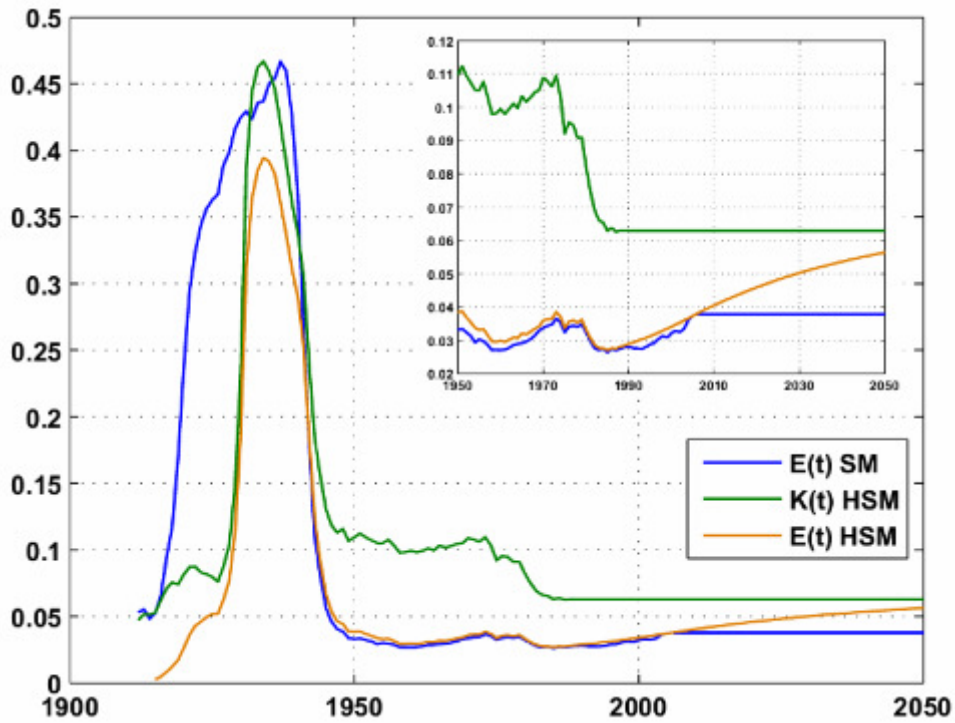


Fig 4. Estimated shock functions for the HSM and SM using the 1900-2006 baseline .

## Modeling Reserve Growth

Reserve growth is probably one of the most contentious and complex aspect of peak oil (see Rembrandt's [posts](#) for a great overview). A quick definition from Verma et al. [2]:

Reserve growth is a term used to refer to estimated increases in the total technically and economically recoverable petroleum reserves of a field that commonly occur through time because (1) additional reservoir and geologic information leads to increases in estimates of hydrocarbons in-place of existing reservoirs or pools; (2) new reservoirs or pools are discovered in existing fields; and (3) improvements take place in the hydrocarbon recovery factor owing to better understanding of reservoir characteristics and behavior through use of 3D/4D seismic interpretation, better geophysical logging tools, and improved reservoir simulation techniques. Additionally, application of horizontal-well drilling technology and enhanced recovery methods improve the hydrocarbon recovery factors significantly, resulting in increased estimates of reserves, particularly in oil reservoirs.

Unfortunately, reserve growth is also very sensitive to the reserve booking and reporting requirements in each country. For the US, reserve growth has been spectacular and is believed to follow a kind of parabolic curve spanning over 95 years. The recent modified Arrington Cumulative Growth Function (CGF) proposed by Verma [2] is forecasting an increase of 0.42% per year for oil fields (onshore) between 1996 and 2030 for the US onshore fields:

$$CGF = \alpha(t)^\beta \quad (5)$$

with  $\alpha=1.75752$  and  $\beta=0.30050$  and  $t$  in the 1-95 years range ( $CGF=1$  for  $t=0$ ). This reserve growth model has also been observed on West Siberia by Vermak and Ulmishek [1] with the parameters  $\alpha=1.56639$  and  $\beta=0.36060$ . I'm proposing to model reserve growth using also a linear filtering derived from the above parabolic curve, the filter is the following:

$$h_{Growth}(t) = \begin{cases} 0 & t = 0 \\ \left(\frac{t+t_{Start}}{t_{Start}}\right)^\beta - \left(\frac{t-1+t_{Start}}{t_{Start}}\right)^\beta & 0 < t \leq t_{max} \end{cases}$$

where  $t_{Start}$  is the initial reference year where we assume growth is taking place (i.e. the initial discovery number is including some reserve growth, up to  $t_{Start}$  years of equivalent growth). The growth filter impulsional response (6) is shown on Figure 5 for  $t_{Start}=20$  years.

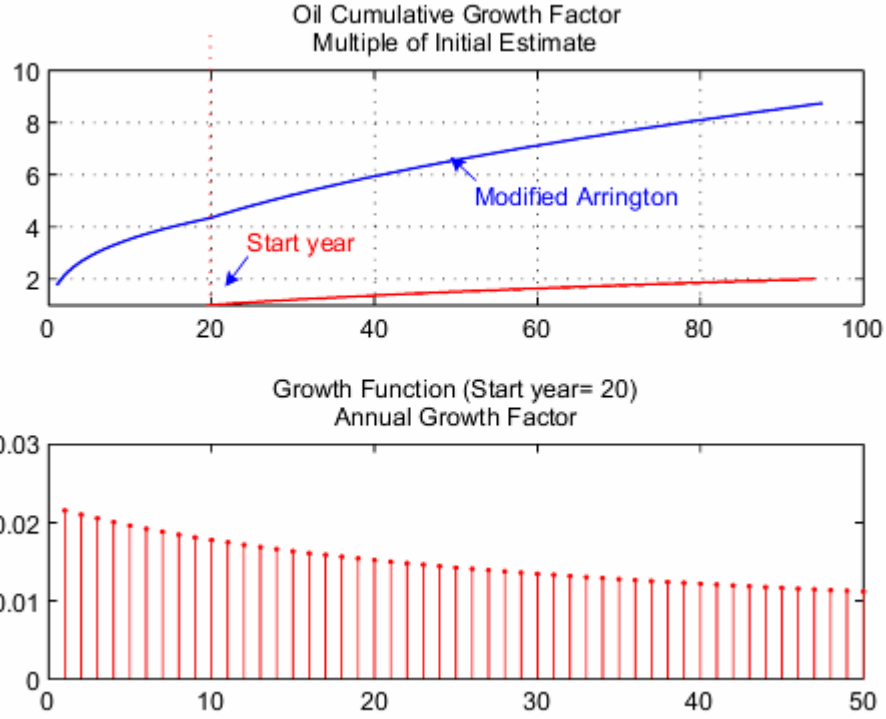


Fig 5. Proposed linear filter response (bottom) for the modeling of reserve growth derived from the empirical CGF function (top).

Because reserve growth is generally occurring as soon as the first producing well is put in place, the reserve growth  $G(t)$  is the result of the convolution of  $h_{Fallow}$ ,  $h_{Build}$  and  $h_{Growth}$ :

$$\begin{aligned}
 R(t) &= (h_{Mature} \otimes h_{Fallow} \otimes h_{Build} \otimes D)(t) \\
 G(t) &= (h_{Growth} \otimes h_{Fallow} \otimes h_{Build} \otimes D)(t) \\
 T(t) &= R(t) + G(t)
 \end{aligned} \quad (7)$$

This linear filtering of initial discoveries simulates the response of the oil production infrastructure to new oil discovery  $D(t)$ . Last time we have shown that this impulsional response was approximated by a Gamma function (in green on Figure 6) therefore simulating lagging effect and spreading reserve additions over time. The fourth convolution by  $h_{Growth}$  makes the impulsional response larger with a heavier tail as shown in blue on Figure 6. The area under the blue curve is greater than one hence simulating reserve growth.

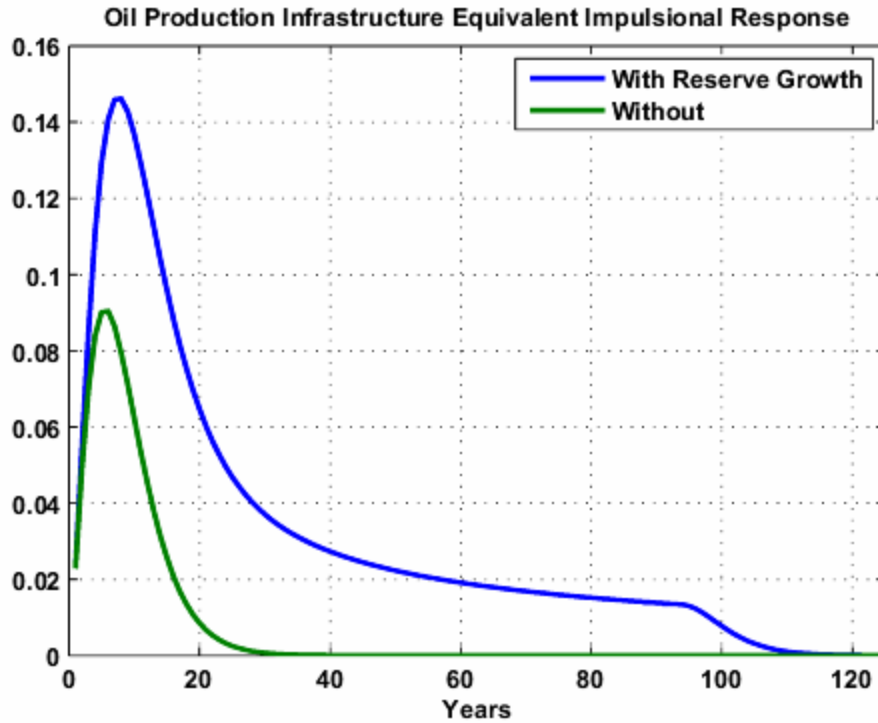


Fig 6. Effect of the reserve growth on the system impulsional response. The Modified Arrington CGF is used here.

The tricky part is to find an appropriate value for  $t_{Start}$ . It seems logical that  $t_{Start}$  should depend on the discovery age because the discovery curve already includes an unknown amount of reserve growth. The chart below is taken from Robelius PhD thesis [3] and is showing how much reserve growth we have experienced in the 1994-2005 period and how it has affected the shape of the discovery curve. It's pretty obvious that reserve growth cannot be neglected and that a static view of oil production will underestimate future production levels. Also, there is no obvious correlation between discovery age and the amount of reserve growth. The amount of reserve growth between 1994 and 2005 is an astonishing 427 billion barrels (Gb). However, only 170-190 Gb seems to be genuine reserve growth (see Rembrandt [post](#)).

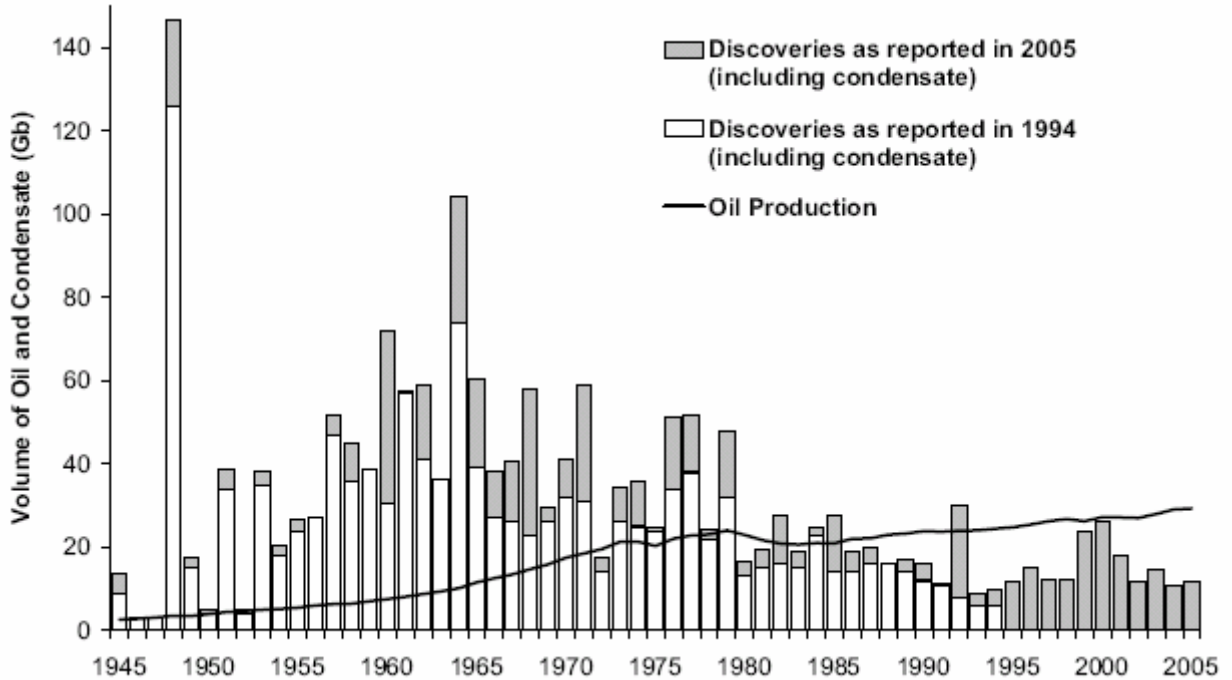


Fig 7. Global annual discoveries of both oil and condensate, as reported in 1994 and 2005, together with oil production in billion of barrels (Gb) The difference reported discoveries is the reserve growth. Source: Based on data from IHS Energy, ASPO and Oil & Gas Journal (from Robelius [3], page 71).

I assume the following model for  $t_{Start}$ :

$$t_{Start} = \begin{cases} t - t_{Ref} + t_{Init} & t \leq t_{Ref} \\ t_{Init} & t > t_{Ref} \end{cases} \quad (8)$$

where  $t_{Ref}$  is the reference year for which the backdated discovery curve has been issued and  $t_{Init}$  is a general offset that gives us a better control on the model. Below are a few [claims](#) that may help us calibrate our algorithm:

1. The USGS is forecasting 612 billion barrels (mean estimate) for conventional oil between 1996 and 2025.
2. Albrandt et al. (USGS) conclude that approximately 28% percent or 171 billion barrels of the forecasted 612 billion barrels for conventional oil had been added to the reserve pool between 1996 and 2003.
3. 2005 resource growth in pre-2005 discoveries was only 8 Gb.

In the simulations below, I made a few assumptions:

1. discovery data is the ASPO backdated (1932-2004) so  $t_{Ref} = 2004$ .
2. pre-1932 total discovery is 30 Gb (mainly US)
3. post-2004 discovery forecast is based on a logistic decline.
4. lambda= 3 years.

	<b>Peak</b>	<b>URR (Gb)</b>	<b>Total Reserve Growth</b>	<b>1996-2025 Reserve Growth</b>	<b>1996-2003 Reserve Growth</b>	<b>2005 Reserve Growth</b>	<b>Pre-2005 Reserve Growth</b>
<b>Logistic</b>	2005 @ 25.2 Gb	2023	0	0	0	0	0
<b>SHM</b>	2005 @ 26.9 Gb	1962	0	0	0	0	0
<b>Modified Arrington</b> $t_{Init}= 1$	2036 @ 36.8 Gb	3937	1976	704	146	21	427
<b>Modified Arrington</b> $t_{Init}= 15$	2017 @ 28.7 Gb	2616	655	230.8	59	7.7	236
<b>Modified Arrington</b> $t_{Init}= 20$	2015 @ 28.1 Gb	2500	538	191	50	6.5	204
<b>West Siberia Growth</b> $t_{Init}= 1$	2018 @ 29.3 Gb	2603	640	293	40	10.6	144

*Table I. Peak estimates for crude oil + condensate derived from various model. The logistic fit was obtained using the Hubbert Linearization technique (1983-2006).*

The result of the SHM + Modified Arrington with  $t_{Init}=1$  year (third row in Table I) is shown on Figure 8 and replicates closely the rosy USGS/DOE/CERA view of future production with a large amount of reserve growth to come. We can see clearly the effect of the relation (8) with a ramp up of reserve growth prior to 2004 and a huge input of reserve growth on new discoveries post 2004. The SHM + Modified Arrington with  $t_{Init}= 15$  years and the SHM + Russian growth ( $t_{Init}=1$ ) give similar results with a peak in 2017-2018 and a URR around 2.6 trillion barrels. For the peak date to be before 2015, reserve growth should be around 6-7 Gb for 2006 and decrease afterward.

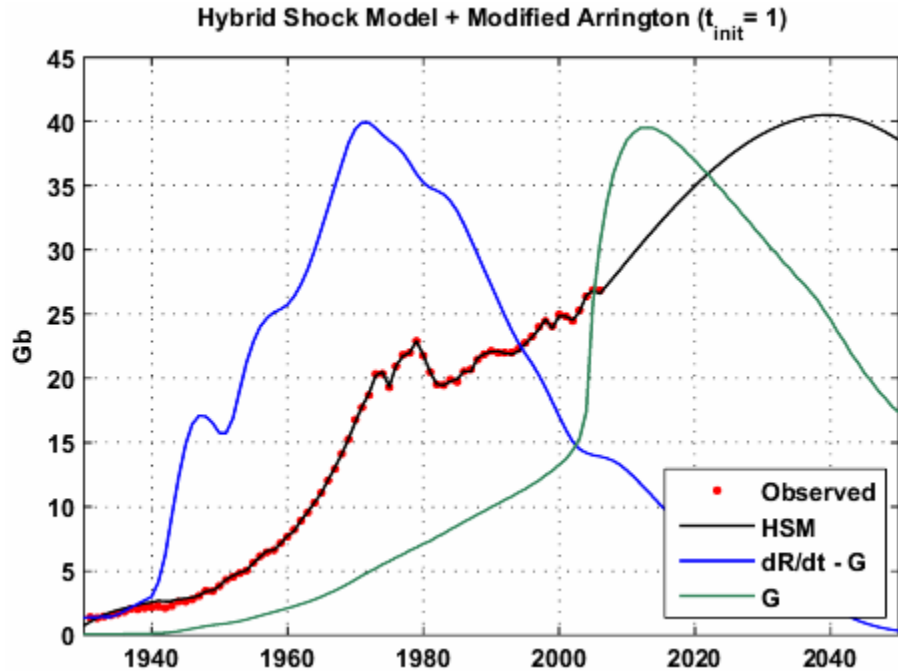


Fig 8. World production forecast (C&C) produced by the HSM assuming the modified Arrington model for the reserve growth .

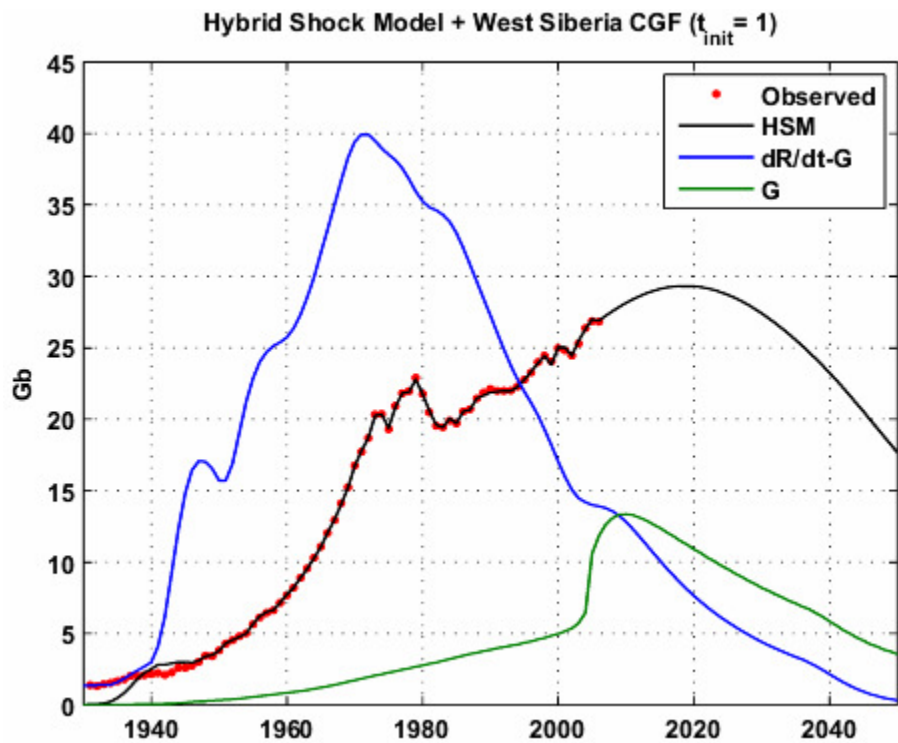


Fig 9. World production forecast (C&C) produced by the HSM assuming Verma's model (West Siberia) for the reserve growth.

On Figure 10, we can see that the HSM + Modified Arrington (in orange) fits the proven reserves from BP but does not fit with the recent record prices. The HSM + West Siberia

CGF (in green) is closed to the corrected BP reserve numbers except since 2001. However the green curve is going down precisely when prices started their climb which makes me think that the proven reserve increases for 2001 and 2002 are probably bogus.

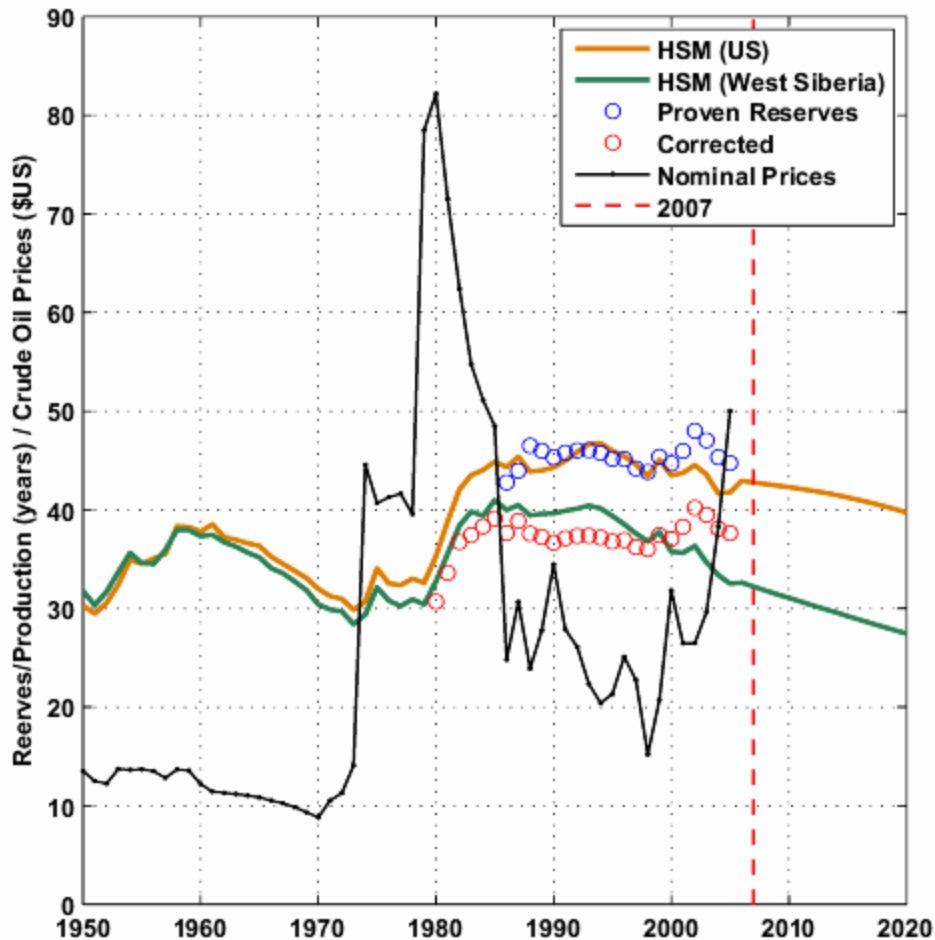


Fig 10. Reserves to production ratio values. Proven reserves are from BP. The corrected reserves account for anomalous Middle-East reserve revisions.

HSM (US) and HSM (West Siberai) are the production curve shown on Figures 8 and 9.

## Conclusions

About the Hybrid Shock Model:

1. The interest of the Shock Model approach resides in its capacity to exploit the discovery data, the production profile and the reserve growth models.
2. The URR is not an output of the model as it is the case for the Hubbert Linearization but results directly from the discovery curves and the application of reserve growth models. The HSM is a nice way to inject prior information about the URR.
3. The method can potentially deal with difficult multi-modal production profiles such as Saudi Arabia.

4. The logistic case can be seen as a particular case of the HSM when the

About reserve growth:

1. Using a fourth convolution function derived from empirical Cumulative Growth Factors, I was able to derive an estimate close to the USGS forecast on reserve growth.
2. Note that we don't know how much true reserve growth is included in the 171 Gb figure. However, the reserve growth in 2005 for pre-2005 discoveries was only 8 Gb, it should have been 21.5 Gb (171/8). If we assume that 8 Gb per year is the true reserve annual addition we get 64 Gb for 1996-2003.
3. In my opinion, peak oil proponents should pay more attention to reserve growth issues. Very often, the argument is only focus on new discoveries but reserve growth is poorly understood and may have a significant contribution especially within a high oil prices environment.
4. Using the West Siberia reserve growth factor and a decreasing number of new discoveries, I estimate the peak to be at most in 2018 for conventional oil. The interesting thing is that it seems to match the 8 Gb in 2005. This result assumes that reserve growth related technologies will be applied aggressively and extensively. Also, the two CGFs that I used are for onshore fields and they are probably very different for offshore fields (new discoveries to come will be increasingly offshore). Therefore, I consider this result as being an upper bound on conventional production.
5. It will be very important to watch reserve growth estimates for the year 2006 in order to confirm (or infirm) a decrease in reserve growth that was observed in 2005 (8 Gb). In particular, a collapse in reserve growth (2-3 Gb) could indicate that the peak for crude oil + condensate is likely to be in 2005-2006.
6. It's important to note that the CGF model (5) is significantly different between large fields and small fields [1]. Because new discoveries are likely to be small fields, reserve growth post-2005 is likely to be smaller.

Of course, this is a work in progress and more tests are needed. The US, Norway and the UK should constitute a nice benchmark for the HSM (maybe in a part III). By the way, if anyone has discovery datasets, please [contact me](#).

## References:

- [1] M. K. Verma and G. F. Ulmishek, Reserve growth in Oil fields of West Siberian Basin, Russia. Ulmishek of the United States Geological Survey. [pdf](#).
- [2] M.K. Verma, Modified Arrington Method for Calculating Reserve Growth—A New Model for United States Oil and Gas Fields, U.S. Geological Survey Bulletin 2172-D, [pdf](#).
- [3] Giant oil fields and their importance for future oil production, Fredrik Robelius, [PhD Thesis](#).

The data for the world production of crude oil + condensate is composed of:

- 1900-1959: API Facts and Figures Centennial edition 1959.
- 1960-2006: [EIA](#) data (includes tar sands production from Canada and Venezuela).

**Code:**

The [code](#) is available in [R](#) language, the windows version of R software is available [here](#). You will need to install the Matlab package (go in "Packages/Install Package(s)" then choose a CRAN site and select Matlab from the package list). To execute the program, open the file (ShockModel\_Part2.txt) and click on "Edit/Run all".