# Does the Hubbert Linearization Ever Work?

Posted by Robert Rapier on March 22, 2007 - 11:15am
Topic: Supply/Production
Tags: hubbert linearization, hubbert peak, peak oil, saudi arabia, texas [list all tags]



## Part II: The Titanic, Oilsville, and Saudi Arabia

In Part I, we examined the evolving Hubbert Linearization (HL) of Texas oil production, and found that the predictive precision of the technique was quite poor. The error range was on the order of 3 decades. However, some have suggested that the trends just need to stabilize, and then we can be more confident in the predictions. Others offered rationalizations for why Texas behaved as it did, and suggested that the HL is still a useful predictive tool provided we somehow filter the data. Still others suggested that it is futile to attempt to linearize non-linear data. In this essay, we will push this issue further. I will examine more cases that cast grave doubts in my mind that the HL can accurately predict anything.

But first, does this debate even matter?

---

## The Sinking of the Titanic

A number of people have latched on to the sinking of the Titanic as an appropriate analogy for the debate over the methodology and timing of peak oil forecasts. I see those analogies as missing a number of key points. If you wish to apply an appropriate analogy, try the following.

After the Titanic hit the iceberg, a debate quickly ensued on how fast it would sink. One train of thought was that it didn’t matter, because everyone was doomed anyway after the precious few lifeboats were utilized. Most of these people sat idly by biting their fingernails while others made decisions that would affect them. One train of thought suggested that based on the alignment of the stars – which had correctly forecast prior disasters - sinking in 2 hours was inevitable. These people proceeded to warn everyone to expect the ship to go down in 2 hours. They demonstrated unwavering faith in their forecast.

But another group employed a more scientific approach. First, they challenged the faith-based group to show a case in which their methodology had ever been demonstrated to work. Challenges to their arguments were met with repetitious assertions of the original claims. Yet the scientific group continued to challenge the argument – not because they didn’t think the ship was sinking – but because they saw some inherent dangers in faith-based arguments and they were not satisfied with the integrity of the conclusions. In response to their challenges, the scientific group was accused of mounting frantic, even hysterical attacks on the faith-based arguments. They were accused of having agendas and of being non-believers. The faith-based proponents were praised by their followers for maintaining faith in the face of these challenges.

However, the challenges persisted, and some in the scientific group pointed to data suggesting that the ship would actually sink in 4 hours. (To be fair, some scientifically minded investigators concluded that the ship would sink in 2 hours). But two hours after the iceberg was struck, the faith-based group began to cry “We are all doomed!” A number of people, unable to bear the tension any longer, decided to end it all before they were plunged into the icy water.

But the 2-hour mark came and went, and while it was clear that the ship was sinking, the voices of the faith-based group had been silenced. They had either jumped overboard as the 2-hour mark bore down upon them, or they were ignored because their assertions now rang hollow. They still sought influence, but their previous tactics now undermined the possibility of sounding a warning.

At the 3-hour mark, a few small rescue boats made it to the area. While there were not enough boats to save everyone, some of those who had taken a cautious approach toward the faith-based group were saved. However, there was a very large group of people who did not differentiate between the arguments of the faith-based group and those of the scientific group. They decided to discount both messages. “After all,” they argued, “these sinkers have already shown that they can’t forecast very well.” So this group called into question the credibility of the scientific group based on the missed forecast of the faith-based group. They convinced great numbers that the Titanic wouldn’t sink at all, and that they were safer taking their chances there than by climbing into a rescue boat on the open sea.

Of course we know how the story turned out. The cry of “wolf” at the 2-hour mark resulted in great loss of life when the Titanic actually did go down at the 4-hour mark. Those who survived were the ones who valued the scientific approach of cautiously evaluating and challenging the data. Many doomers were wiped out at the 2-hour mark, and many cornucopians – along with the vast majority of the passengers - were wiped out at the 4-hour mark. So the debate over the timing of the Titanic’s sinking actually had great significance for all 3 groups. And while we don’t know if any rescue boats will actually arrive or just how long we have until the ship sinks, the debate does matter to many people.

## Oilsville – The Constant Production Case

Consider the case of a fictional society called Oilsville. Oilsville discovered a great oil reserve in 1955 and began full production on January 1, 1960. The size of the recoverable reserve is at least 5 trillion barrels. Oilsville’s policy makers decided that they would maintain their oil production constant at 10 billion barrels per year. It was estimated that the new find would provide for the needs of Oilsville for at least 500 years, provided they maintained their longstanding policy of keeping their population relatively constant.

Note that we have made no assumptions about how they produce their oil. We don’t know if there is a water-cut, we don’t know if they are using horizontal wells, and we don’t know how hard they are pumping their fields. All we know is that they are keeping production constant. So, let’s do a series of HLs for Oilsville to better understand the HL behaviour under these circumstances. First, fix in your mind what you think the HL for Oilsville will look like after 20 years of steady production. Form your hypothesis. So, what does the HL actually look like?

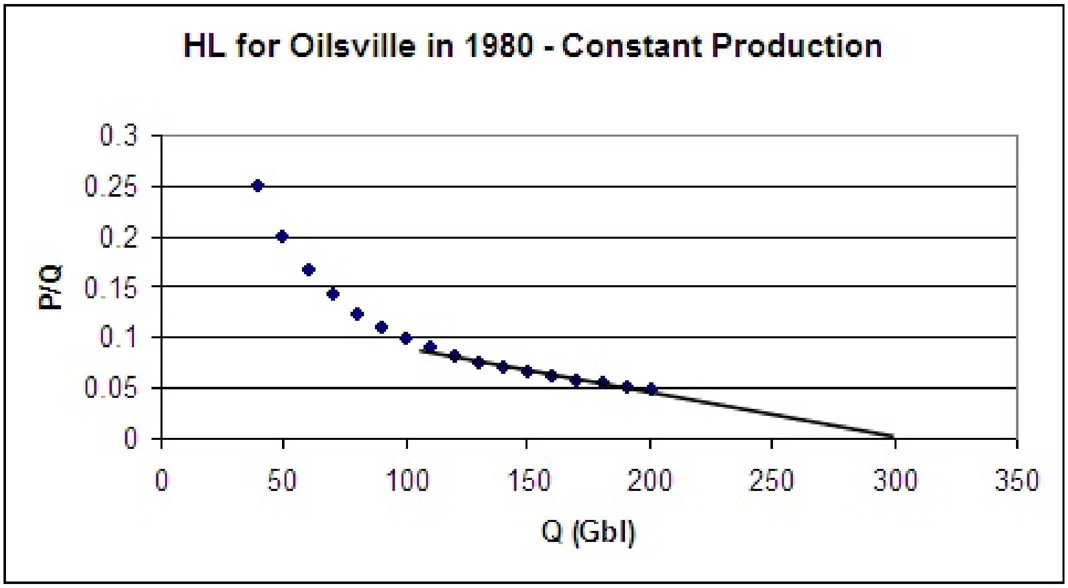


**Figure 1. Hubbert Linearization of Flat Production Case after 20 Years.**

At the end of 1980, they have produced 200 billion barrels, or 4% of their total URR of 5 trillion barrels. The last 10 points on the graph have stabilized such that we can make an estimate of URR based on the HL. It appears from the HL that the predicted URR is 300 billion barrels. However, this is only 6% of the actual URR. The % Qt indicated by the HL is (200/300) or 66.7%, indicating a fairly advanced stage of depletion and likely well past peak. Yet the actual depletion is 4%. Furthermore, recall that even though the production rate is constant, the slope of the line is negative. This may confuse some into thinking that production is declining.

Maybe that one was a fluke. Let's look at 1990.

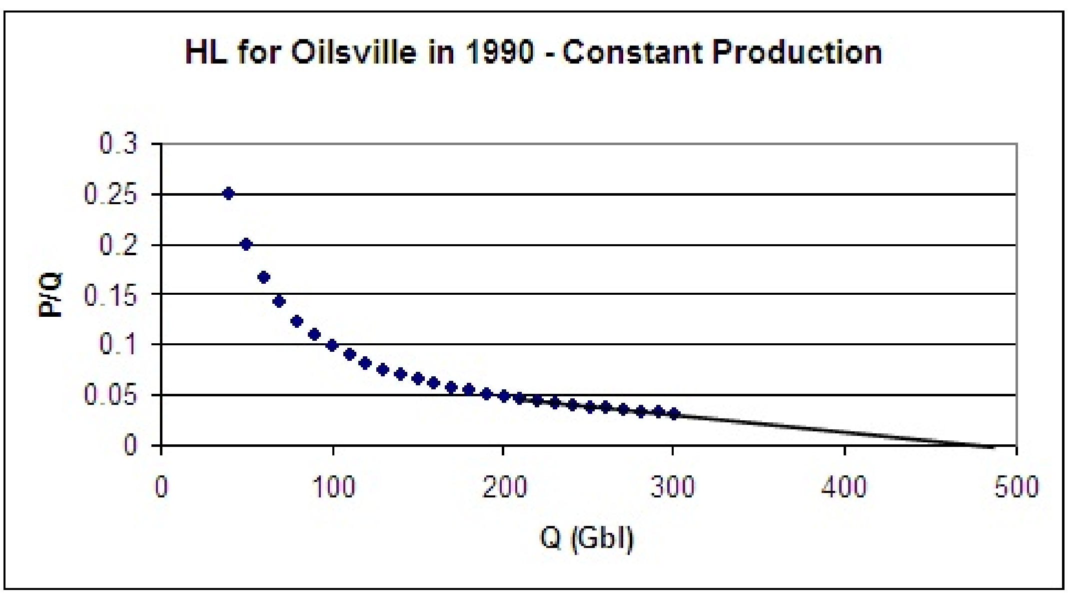


**Figure 2. Hubbert Linearization of Flat Production Case after 30 Years.**

Hmm. The URR according to the HL prediction has now increased to 460 billion barrels, still only 9.2% of the actual URR. The % Qt implied by the HL is now (300/460), or 65.2%, while the true % Qt is (300/5,000), or 6%. Also, by 1990 they had produced 300 billion barrels, which only 10 years earlier was predicted as the URR by the HL. These HL predictions are not remotely in the ballpark. Let's jump to 2000.

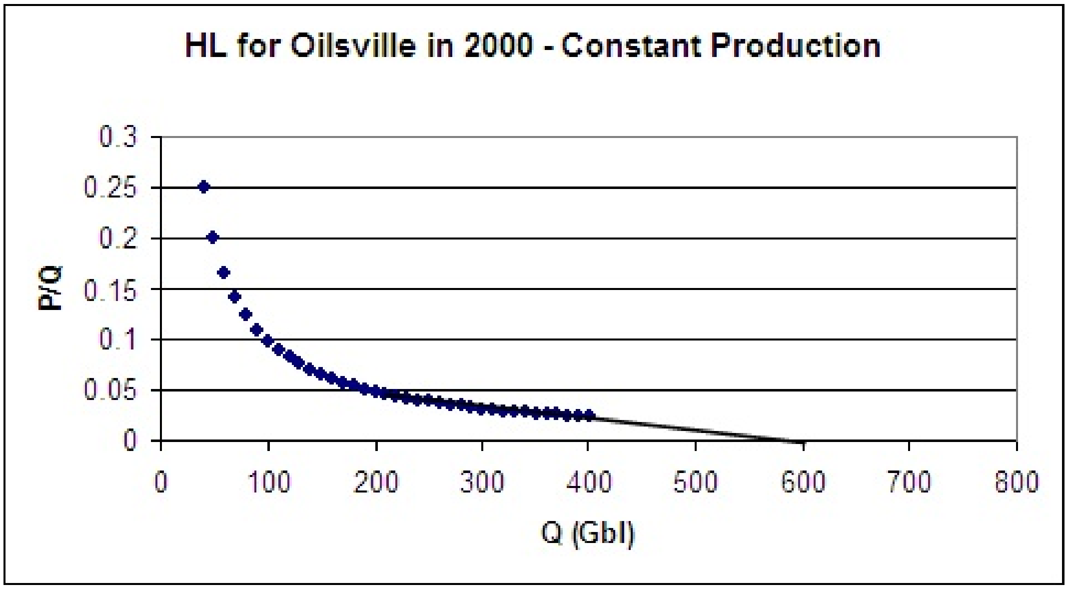


**Figure 3. Hubbert Linearization of Flat Production Case after 40 Years.**

It's more of the same. The URR is trending to a higher value. The HL is now predicting that the URR is 580 billion barrels, or 11.6% of the true value. Cumulative production in 2000 is 400 billion barrels, giving an implied Qt of (400/580), or 69%. The actual % Qt at this point is (400/5000), or 8%. But let's look at one more in 2010:

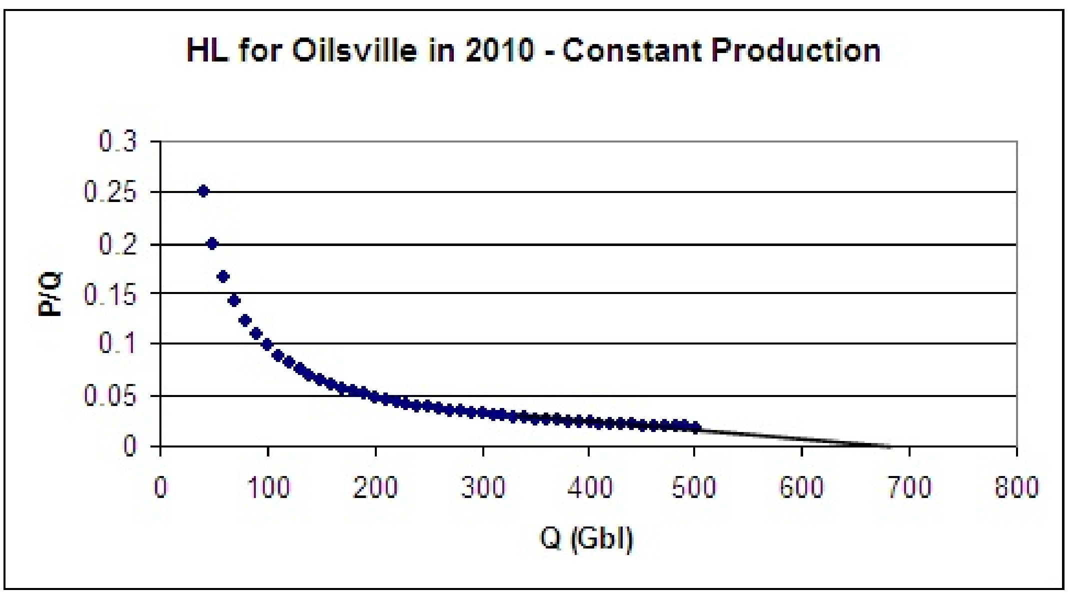


**Figure 4. Hubbert Linearization of Flat Production Case after 50 Years.**

Again, the pattern is the same. URR as indicated by the HL continues to grow, and now stands at 630 billion barrels, or 12.6% of the true URR. Yet cumulative production is now at 500 billion barrels, giving a predicted (500/630), or 79% of Qt. That's right; the HL is telling us that they are 79% depleted, when they are actually only (500/5,000), or 10% depleted.

The HL has failed in this case, and furthermore it will show the same tendency for any area having relatively flat production over an extended period of time. I will leave that one for the reader to do as an exercise. But the problem is – as someone mentioned following my previous essay – linearization of non-linear data. That is why we so often see "doglegs" when we stop and plot the graphs. Often, the dogleg is a mere manifestation of the fact that we are drawing a straight line through a curve, and the points at the end don't fit.

## The Rising Production Case

Now, if that doesn't convince you that there is a problem, perhaps this will. Instead of a constant production rate, let's assume they increase production each year by 0.1 billion barrels per year. Let's also say we no longer know anything about reserves. You could assume any number you want. Let's revisit the HL from 1960 with this new assumption:

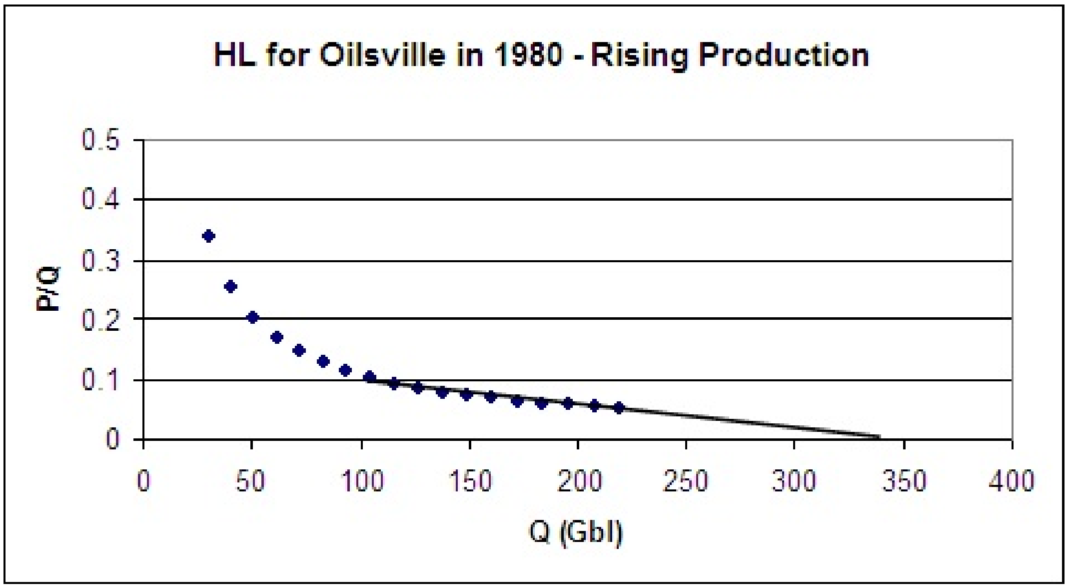


**Figure 5. Hubbert Linearization of Flat Production Case after 20 Years.**

At the end of 1980, they have produced 219 billion barrels. Despite the fact that production has increased each year, and despite the fact that the reserves could for all practical purposes be infinite, the HL predicts a URR of 348 billion barrels. So, 20 years into production, the HL says they are at (219/348), or 62.9% of Qt. Furthermore, the R-squared of the last 10 points is a respectable 0.969, indicating a very good fit for the trend line. To me this is indicative of a major malfunction in the HL, given that all we know is that production has increased every year.

Once again, the HL has failed badly in Oilsville. And while you may dismiss this case as an unrealistic example, consider the implications. This example suggests that it is not possible to forecast a production peak prior to the production actually turning downward. Given this, I conclude that it is highly unlikely that Hubbert actually used this sort of linearization technique in making his forecasts.

## The Ideal Production Case

"No fair," you say. "You are feeding the HL contrived and unrealistic cases. You can't expect it to behave in these instances." So, let's examine an ideal case. Consider the case in Oilsville where the production rate increases by 5% each year for 20 years, and then starts to decline at 5% a year indefinitely. What we see in this case is an eventual convergence toward a URR of 861.3 Gbl. This should be an ideal case for the HL. Peak production will occur in 1980, and we will cross the 50%

depletion mark in 1984. We know all of the important parameters. But can the HL predict the results? Let's look first at the HL for 1980:

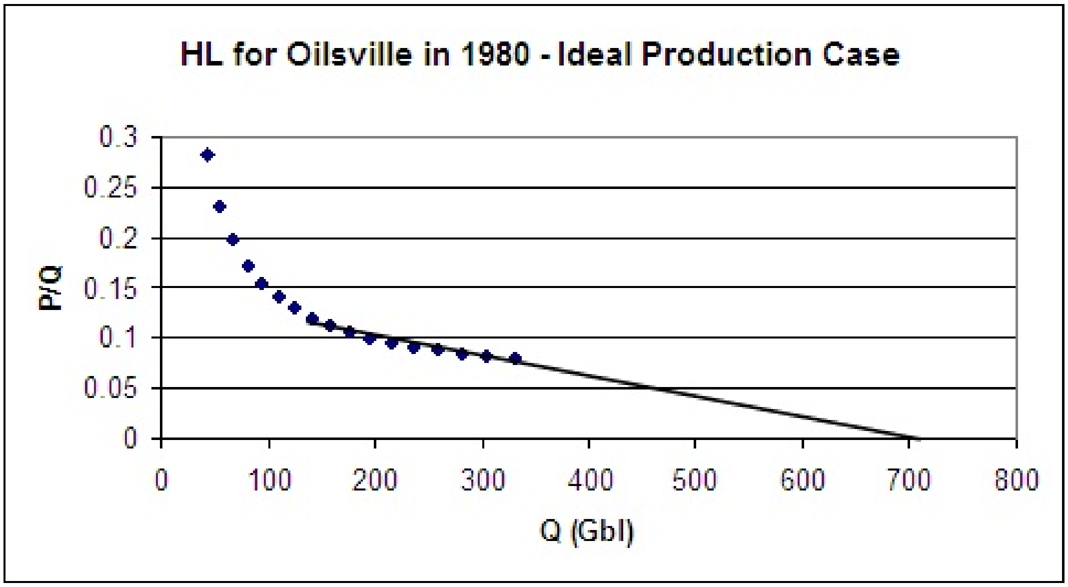


**Figure 6. Hubbert Linearization of Ideal Production Case after 20 Years.**

Even though production has risen each year to 1980, the HL is still under-predicting the URR. The R-squared is indicating a good fit at 0.945. The trend line is predicting a URR of 702 Gbl, an under-prediction of 159 Gbl. Cumulative production in 1980 is 330.7 Gbl, which means that they are at (330.7/702), or 47% of predicted Qt. But the actual position is (330.7/861.3), or 38% of Qt. Remember, that last point represents the year of peak production.

Let's jump forward 10 years.

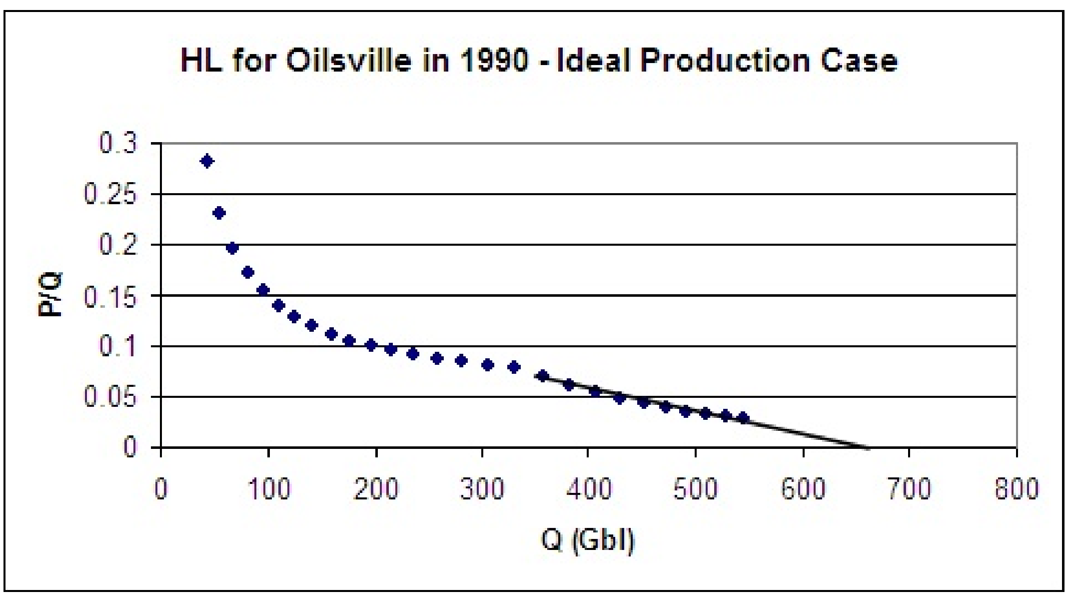


**Figure 7. Hubbert Linearization of Ideal Production Case after 30 Years.**

The 50% depletion mark was crossed in 1984. The HL is now predicting a Qt of 667 Gbl, which is a greater error than the previous prediction. The R-squared in this case implies a better fit than in the previous graph (0.986 now versus 0.945 for the previous case). Yet the actual URR is being under-predicted by 194 Gbl. Cumulative production in 1990 is 543.6 Gbl, putting the actual % Qt at (543.6/861.3), or 63%. The predicted % Qt is (543.6/667), or 81.5%.

The HL is lying to us even in the ideal case where we have specified all of the important parameters. In none of these cases has the HL come close to providing accurate predictions.

## Saudi Arabia

If the HL is not even reliable when all the important parameters are defined, then it is pointless to use it in a case in which we are trying to predict the parameters. The HL has failed all validation tests. As someone accurately said in another thread, my motive here is that I really want to know "what the heck is going on." Given the enormous errors in the HL predictions where URR is persistently under-predicted, I don't have any confidence in the ability of this model to do that. We saw this in the case of Texas, where the URR grew by 50% in 20 years.

The HL has failed us in every single test case. But let's go ahead and look at the case of Saudi Arabia. So, what does the actual HL look like for Saudi? First, we need production statistics from Saudi. After 1960, those numbers are available from the EIA. (1, 2) Prior to that, I had to cobble data together from several sources. According to Matt Simmons' Twilight in the Desert, (3) the first major oil discovery in Saudi Arabia occurred in 1938. Simmons further indicates that production was slow during World War II, and that at the end of the war Saudi was producing 60,000 barrels per day (bpd). Figure 3.3 in Twilight shows that in 1950, Saudi was producing about 0.5 million bpd, and by 1960, production had risen to just over 1 million bpd.

To calculate cumulative production prior to 1960, I assumed an average of 30,000 bpd from 1938 through 1945, and then an average of 250,000 bpd until 1950, and finally an average of 750,000 bpd from 1950 to 1960. There may be an actual cumulative number available, but I haven't found it. Regardless, later production rates are high enough that even if we assumed no production prior to 1950, it would have minimal effect on the results. But based on my assumptions, I calculated a total of 3.2 billion barrels produced in Saudi Arabia before 1960. Using those assumptions, the entire HL graph for Saudi Arabia looks like this:

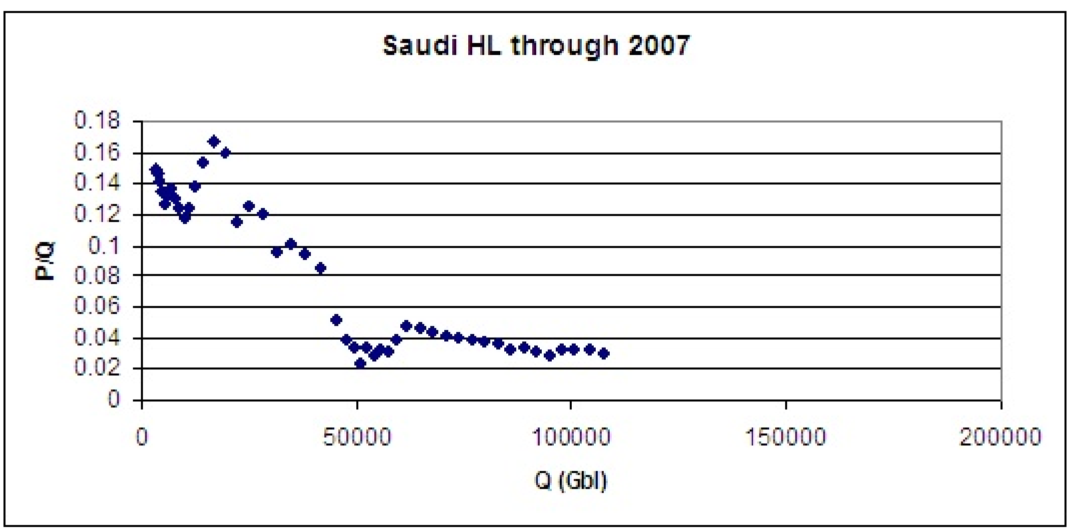


**Figure 8. Hubbert Linearization of Saudi Arabia in 2007.**

This curve is quite consistent with previous HLs of Saudi Arabia. The HL was very erratic until 1991, at which point the slope became relatively stable. By 2000, there were 10 points that looked to have stabilized:

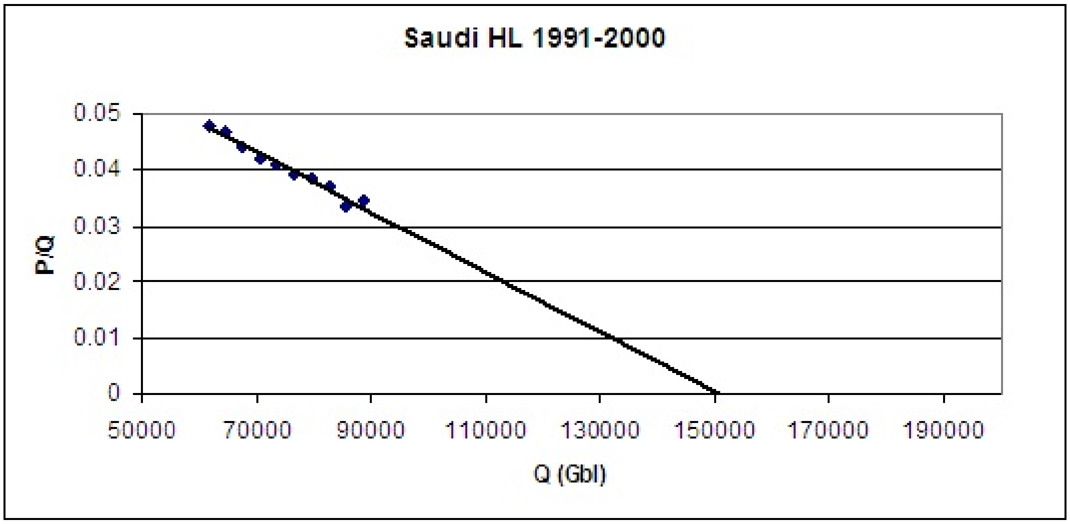


**Figure 9. Hubbert Linearization of Saudi Arabia 1991-2000.**

The HL performed in 2000 predicted a URR of 151 Gbl. That would put cumulative production through 2000 at (88.7/151), or 58.7% of Qt. The R-squared is 0.969. Now let's include production numbers through 2006:

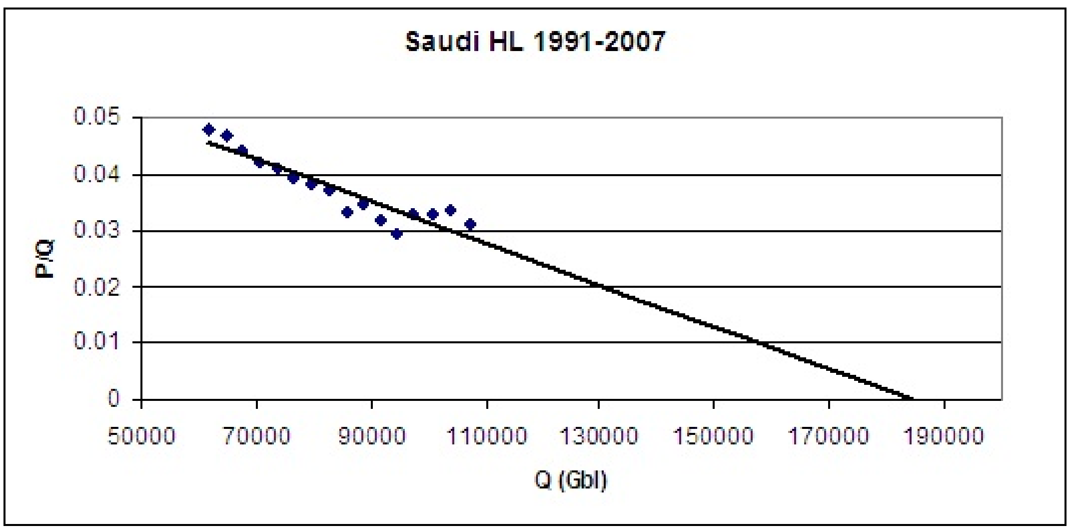


**Figure 10. Hubbert Linearization of Saudi Arabia 1991-2007.**

The addition of a mere 6 years of production data has shifted the URR prediction to 185 Gbl. Between 2000 and the end of 2006 Saudi Arabia produced 19 Gbl, but the URR prediction increased by 34 Gbl. Because of this, the 2006 HL is at a slightly lower % of Qt than the 2000 HL. In 2000, the graph showed that Saudi was at 58.7% of Qt, but in 2006 the graph indicates that Saudi is at (107.5/185), or 58.1% of Qt. This implies that one could have used the HL to make a stronger case for an imminent Saudi peak in 2000 than in 2007. Furthermore, the R-squared in 2006 – 0.856 - is worse than in 2000, indicating that the stability of the line has degraded. Yet this is the data that is being used by some to forecast the peak of Saudi oil production. Perhaps you can now understand my extreme scepticism that the HL has the ability to accurately make such a prediction.

## Conclusion

Based on the case studies I have carried out – both hypothetical and real – I conclude that the Hubbert Linearization technique does not have the capability of predicting a production peak in real time. Furthermore, I believe it is the responsibility of those who are using this technique as a basis of their peak predictions to validate the model. Each case I have investigated showed a failure of the HL technique and an incorrect prediction. This was true for Texas, it was true for the idealized cases, and we are now seeing evolution in the HL for Saudi Arabia. But it is not my responsibility to investigate every possible variation of when the HL "might" work. I could spend the rest of my life evaluating case after case. I think it is a much more reasonable request to ask proponents to provide details on some cases in which it did work - or would have worked – to predict a peak in something resembling real time.

If we are to have any confidence in the HL model, I think it is fair to request the following of HL proponents:

1). Please validate your model. Demonstrate any case – real or hypothetical – in which the HL would have predicted a peak in real time. Details are preferred over assertions.

2). Please define the conditions and parameters of the HL (e.g., % Qt, P/Q intercept) that indicate a production peak.

3). Please identify cases in which the HL would be expected not to work.

4). Please identify any other criteria (filtering, etc.) that need to be fulfilled before the method should be used.

I know that some have a difference of opinion on whether this debate even matters. I am of the belief that it does, and that we should rigorously challenge all of our assumptions. Furthermore, I know that there are some who believe that peak is imminent, but who also agree that continually testing and evaluating conclusions is necessary. I want to make it clear that there is a big difference – especially in the ability to persuade others – in the faith-based approach and the more rigorous approaches. This is true regardless of the ultimate conclusion of the approach.

## References

1 . World Oil Production 1960-2005, Energy Information Administration, Accessed March 18, 2007.

2. Saudi Production 2006, Energy Information Administration, Accessed March 18, 2007.

3. Simmons, Matthew R., Twilight in the Desert, John Wiley and Sons, Inc., 2005.

**Note:** There were a lot of calculations involved in this one, and it is quite possible that there may be an error or two. If you spot one, I would be grateful if you would call it to my attention and I will fix it. Thanks.