Reserves Appreciation in the Gulf of Mexico

by

Ted D. Tupper

Overview

It is a well known that oil and gas reserves estimates in tend to increase through time, hence the term reserves appreciation. In the 1990's the MMS Gulf of Mexico Resource Evaluation conducted a statistical analysis of field reserves and found that Annual Growth Factor (AGF) could be found using a negative exponential function of the age of the field. MMS 2000-087 (2000 Assessment, p. 47) presents a description of the prior work. This paper is a reexamination of the topic.

Reserves appreciation is an important contributor to remaining reserves. For example the most recent reserves report MMS 2009-064 (p. 44) shows that:

Remaining proved reserves	Oil (billion bbl)	Gas(TCF)
Previous estimates, as of 12/31/2005	5.19	17.90
Discoveries	0.29	0.7
Revisions (net appreciation)	0.21	1.2
Production during 2006	-0.47	-2.9
Net Change	0.03	-1.0
Estimate as of 12/31/2006	5.22	16.9

The data used in this study shows that over 40% of the proven reserves in the Gulf of Mexico is due to reserves appreciation.

There have been various claims that we are running out of oil. This is the "peak oil" hypothesis. The most recent reserves reports show that the 2006 remaining proved oil reserves are 5.22 billion bbl as compared to 1975 estimate of 2.79 billion bbl. Reserves appreciation is key part of the story why the projected shortage of oil keeps moving into the future.

Data

The data source used for this analysis the Reserve History for Proven Fields (http://www.gomr.homepg/offshore/fldresv/2006-HIST.pdf). This data set has reserves estimates from 1975 through 2006. All these estimates we created by MMS Gulf of Mexico Resource Evaluation in a consistent manner in accordance with industry standards. The following table is exact of the data for the first field in the history, AC024.

Year	Disc Year	Original Oil	Original Gas	Original BOE	Remaining Oil	Remaining Gas	Remaining BOE
2000	1998	4967694	7451541	6293591	4967694	7451541	6293591
2001	1998	7125984	5700787	8140359	7125984	5700787	8140359
2002	1998	7125984	5700787	8140359	5525036	4395428	6307141
2003	1998	7125984	5344488	8076961	3021610	2214021	3415564
2004	1998	9639114	7036553	10891170	3223156	2297121	3631897
2005	1998	9639114	7036553	10891170	1769751	1280835	1997658
2006	1998	10442374	7936204	11854510	1570634	1462584	1830880

The Year Column is estimate for December 31st of that year. Disc Year is the year the field was discovered. The next three columns are the estimated Original Oil, Gas, and BOE (Barrel Oil Equivalent) for the field. Oil and BOE are in Barrels. Gas is in MCF (thousand cubic feet). The last three columns show the remaining Oil, Gas and BOE. The data set also has cumulative production. Cumulative value can be computed by taking the original value and subtracting the remaining value. Cumulative values were not used in this analysis.

The data for AC024 shows that reserves estimate are volatile. The Original Oil estimate in 2000 was almost 5 million barrels. In 2001 the estimate increases to over 7.1 million barrels. In 2004 the estimate is over 9.6 million barrels and in 2006 the estimate is now up to 10.4 million barrels. In six year the original oil estimate doubles. On the gas reserves is a different pattern. The 2000 gas estimate is about 7.5 BCF (billion cubic feet). In 2001 gas estimate decreases to 5.7 BCF. In 2003 another downward revision to 5.3 BCF. In 2004 the trend is reversed with a gas estimate of 7 BCF. In 2006 the gas estimate is almost 8 BCF. That is a little more than the 2000 estimate.

The data for AC024 is volatile. There are examples where the estimates are more dramatic as in field MC397:

Field	Year	Original BOE
MC397	1984	460966
MC397	1985	460966
MC397	1986	10520458
MC397	1987	10060365
MC397	1991	27091755
MC397	1992	28410392
MC397	1993	80336090
MC397	1994	86069198
MC397	1995	72045770
MC397	1996	80330690
MC397	1997	88014613
MC397	1998	88919605
MC397	1999	87272531
MC397	2000	88567375
MC397	2001	78959296
MC397	2002	77936960
MC397	2003	77481223
MC397	2004	76946708
MC397	2005	77358066
MC397	2006	77875374

Estimating individual field reserves appreciation will be high uncertain. There are well over thousand proved fields in Gulf of Mexico which in aggregate may provide insight into the reserves appreciation process.

The data used in this study starts with 2006 History table. Those rows with Year - Disc Year < 0 are dropped. Fields with records in 1975 were dropped. Field "EI0127" was also dropped. These filters insure that the initial reserve estimates were made in 1976 or later.

Initial Approach

Three levels of estimates were developed: Age basis, Age & Year basis, Age & Year & Field. The regression models will compute CGF (Cumulative Growth Factor). The Age basis is comparable to the previous analysis. Here the data is examined at the highest level of aggregation. The Age & Year basis approach is looking a fields in cohorts. Fields that have their initial reserve estimates in the same year are examined as a group. Age & Year & Field attempts to make an estimate of reserves appreciation at the field level.

The potential predictive variables are computed from the 2006 History. For the Age & Year & Field estimates attributes of the field such as water depth were considered. Some of the variables were created by making transformation of variables such as log of the age. The residuals (actual - prediction) of preliminary estimates were examined to develop new variables.

For the Age basis the variables considered are:

Age	= Year - Min{Year for Field in History Data}
REM_PCT	= Remaining BOE / Original BOE
OIL_PCT	= Original Oil / Original BOE
CGF	= Original BOE / (Original BOE for Min{Year for Field})

All the variables are initially tested in a linear regression model. The variable the least contribution (i.e. smallest absolute t statistic) is dropped. The process is repeated until the remaining variables had t statistic which exceed 3.0 in absolute value in the final model. Residuals were are examined for patterns which may suggest additions variables to add to model. If new variables are found the process is repeated.

The Age and REM_PCT variables are strong predictors, but also highly negatively correlated. In an attempt create an even better estimator a hybrid estimator was considered: Age * (1 - REM_PCT).

The analysis yielded this the Age Basis Model

CGF = -(4.009266 * REM_PCT) + (4.726084 * OIL_PCT) + (-0.04372 * (Age * (1.0 - REM_PCT)))
+ 2.538217

This model has an adjusted R-squared 0.9481 with a residual standard error = 0.0915.

For Age & Year model the following additional variables were considered:

Recip_YR	= 1.0 / Year
LogYR	= Log(Year)
From_1982	= | Year - 1982 |
From_885	= | Year - 1988.5 |
LogAge	= Log(Age)
NegExpAge	= Exp(-Age)
Recip_From_885	= 1.0 / From_885

CGF = (0.48500 * Age) + (1.33964 *OIL_PCT) + (-0.41715 * (Age * (1.0 - REM_PCT)))

This model has an adjusted R-squared of 0.8483 and a residual error of 0.9755.

The reserves appreciation data is highly volatile, but it often strongly autocorrelated. That is last year's reserve estimate is often near this year's estimate. The use of previous year's data is called lagging in econometrics.

If the prior year CGF (CGF_1) is considered then model becomes:

CGF = (0.212345 * REM_PCT) + (0.999524 * CGF_1)

This model has an adjusted R-squared of 0.9904 and a residual error of 0.2456.

That is a much more precise and simpler estimate.

For Age & Year & Field model also considered the following additional variables:

WD	= Water Depth in feet
IType	= 1 is Oil, 0 is Gas
WD_1000	= WD > 1000, then 1, else 0
HiVar	= AC in {MP, GI, EW}
NegR	= AC in {SS, SP, SM}
PosR	= AC in {MP, DC, GI}

AC is the Area Code. The Area Code is a geographic region in the Gulf of Mexico. It use brings in location a variable into reserves growth. HiVar (high variance), NegR (negative residual), and PosR (positive residual) categories came from examining residual of preliminary models. Water Depth and Field Type is from Table 4 of the reserves report.

The model found was:

CGF = (1.1094866 * Age) + (-1.9613164 * OIL_PCT) + (0.0005418 * WD) +
(0.1815710 * From_1982) + (0.2092554 * From_855) + (4.2352739 * NegR) +
(-6.4299825 * PosR) + (6.7861847 * HiVar) + (-0.9299767 * (Age * (1 - REM_PCT)))

This model has an adjusted R-squared of 0.18 and residual error of 10.98.

If the prior YR CGF (CGF_1) is considered then model becomes:

CGF = (0.060356 * Age) + (1.0569904 * REM_PCT) + (0.981186 * CGF_1) + (0.262146 * NegR) +
(-0.047031 * (Age * (1 - REM_PCT))) -0.336683

This model has an adjusted R-square of 0.9272 and residual error of 3.055.

Follow On Approach

The results from the initial approach for the Age & Year & Field models were disappointing with high residual errors.

Next various subsets of the fields were examined in attempts to find partitions of the fields that would provide a basis of better models. The data exploration revealed that fields with smaller BOE sizes have higher variation in reserve appreciation. That is fortunate. Large fields contain most of the total reserves and also have comparatively smaller variations reserve appreciation, which enables more precise modeling. I spilt the fields into three groups (Big, Middle, Small) based on the Initial Estimate of Original BOE (First_BOE). Where Big is over 20 million BOE and Small is under 2 million BOE. The 20 to 2 million BOE compose the Middle. The values of 20 and 2 million BOE are modeling judgements. Here are results of modeling Age & Year & Fields:

Big Fields

The model found was:

CGF = (-0.04183 * YR) + (0.04238 * DISC) + (-0.000000001589 * First_BOE) + (0.08401 * Age) + (0.6536 * OIL_PCT) + (0.0001193 * WD) + (-0.02632 * From_855) + (0.9802 *NegR)

This model has an adjusted R-square of 0.7023 and residual error of 0.9684.

If the prior YR CGF (CGF_1) is considered then the model becomes:

CGF = (0.085369 * REM_PCT) + (0.997614 * CGF_1)

This model has an adjusted R-square of 0.9717 and residual error of 0.2986.

Middle Size Fields

The model found was:

CGF = (0.0002776 * DISC) + (0.5701 * Age) + (4.344 * OIL_PCT) + (0.001058 * WD) + (-1.537 * IType) + (-0.1163 * From_855) + (6.162 * HiVar) + (0.9979 * NegR) + (-6.246 * PosR) + (-0.3573 * (Age * (1 - REM_PCT)))

This model has an adjusted R-square of 0.3823 and residual error of 4.63.

If the prior YR CGF (CGF_1) is considered then the model becomes:

CGF = -0.259685 + (0.042568 * Age) + (0.560218 * OIL_PCT) + (0.570902 * REM_PCT) + (-0.285569 * IType) + (-0.033426 * (Age * (1 - REM_PCT))) + (0.995156 * CGF_1)

This model has an adjusted R-square of 0.9522 and residual error of 1.1.

Small Fields

The model found was:

CGF = 0.03856 + (-0.1950 * DISC) + (2.325 * Age) + (0.01122 * WD) + (0.1103 * HiVar) + (8.444 * NegR) + (-0.1048 * PosR) + (-1.985 * (1 - REM_PCT)))

This model has an adjusted R-square of 0.1632 and residual error of 17.11.

If the prior YR CGF (CGF_1) is considered then the model becomes:

CGF = (2.362104 * REM_PCT) + (0.974024 * CGF_1)

This model has an adjusted R-square of 0.9334 and residual error of 0.5259.

Partitioning the fields into size groups enables a dramatic improvement into accuracy of the models at the Age & Year & Fields level.

Summary

This study is a major expansion of the previous work. Three levels of aggregation were considered. Many more predictive variables were considered. Some of which are better predictors than found in the prior work.

Reserves Appreciation is major contributor to total reserves, that is over 40% to the total proven original reserves in the Gulf of Mexico.

Last year's reserves estimate (CGF_1) is the best estimator of this year's reserves.

Remaining Reserves Percentage (REM_PCT) is a better estimator of reserves appreciation than Age.

Other variables that influence reserve appreciation are Oil Percentage (OIL_PCT), Water Depth (WD), Years from 1982 (From_1982), Years from 1985.5 (From_855), and fields located in groupings of area codes.

Grouping Fields by the size enables a better estimate of large field size reserves appreciation for a specific field.