[PS] The Limitations of Lognormal Distributions: Using Subsurface Data to Make More Accurate Resource Estimations*

David Liner[1]

Search and Discovery Article #41848 (2016)**
Posted August 8, 2016

*Adapted from poster presentation given at AAPG Annual Convention and Exhibition, Calgary, Alberta, Canada, June 19-22, 2016

[1]Corporate Reserves, Marathon Oil Corporation, Houston, Texas, United States (dhliner@marathonoil.com)

Abstract

The outcomes produced by multiplying two independent variables are lognormally distributed (e.g. the product of two die or area). Exploration risk and uncertainty pioneer, E. C. Capen, first advocated using lognormal distributions to estimate petroleum reserves in an AAPG short course titled Evaluating and Managing Petroleum Risk in 1984. Subsequently, working together with R. E. Megill and P. R. Rose, Capen offered this course more than 50 times in the succeeding years, and the use of lognormal distributions became the industry standard when describing a range of potential outcomes for everything from EUR to gross rock volume. Lognormal distributions give more accurate reserve estimates but have one inherent flaw—they start at zero and extend to infinity. Capen addressed this issue in 1992 by explaining that one has to "sense check" the high side outputs and truncate appropriately. However, this upper truncation is affected by its own uncertainty. How big is "too big"? Building on this previous work, a new workflow has been designed that reduces the uncertainty in predrill resource estimates and constrains high-side estimates to geologically reasonable values. "Full field" uncertainty analysis allows for stochastic Monte Carlo simulation, while accounting for potential variance in the mapped horizon.

References Cited

Capen E.C., 1976, The Difficultly of Assessing Uncertainty: Journal of Petroleum Technology, v. 28/8, p. 843-850.

Capen, E.C., R.E. Megill, and P.R. Rose, 1984, Evaluating and Managing Petroleum Risk: Tulsa, OK, Am. Assoc. Pet. Geol.

Capen, E.C., 1992, Dealing with Exploration Uncertainties: in Steinmetz, R., ed., The Business of Petroleum Exploration, Tulsa, OK, Am. Assoc. Pet. Geol., p. 29-61.

The Limitations of Lognormal Distributions: Using Subsurface Data to Make More Accurate Resource Estimations

David Liner[1]

[1]*Corporate Reserves, Marathon Oil Corporation, Houston, TX, United States*

Abstract

The outcomes produced by multiplying two independent variables are lognormally distributed (e.g. the product of two die or area). Exploration risk and uncertainty pioneer, E. C. Capen, first advocated using lognormal distributions to estimate petroleum reserves in an AAPG short course titled *Evaluating and Managing Petroleum Risk* in 1984. Subsequently, working together with R. E. Megill and P. R. Rose, Capen offered this course more than 50 times in the succeeding years, and the use of lognormal distributions became the industry standard when describing a range of potential outcomes for everything from EUR to gross rock volume. Lognormal distributions give more accurate reserve estimates but have one inherent flaw—they start at zero and extend to infinity. Capen addressed this issue in 1992 by explaining that one has to "sense check" the high side outputs and truncate appropriately. However, this upper truncation is affected by its own uncertainty. How big is "too big"? Building on this previous work, a new workflow has been designed that reduces the uncertainty in predrill resource estimates and constrains high-side estimates to geologically reasonable values. "Full field" uncertainty analysis allows for stochastic Monte Carlo simulation, while accounting for potential variance in the mapped horizon.

Background

As early as 1976, Ed Capen was addressing a major challenge facing the oil and gas industry, characterizing uncertainty. To illustrate the issue, Capen asked industry professionals around the country to answer ten questions by giving a "90-percent range" around each answer (Capen, 1976). The results conclusively showed people were more likely to guess an accurate range if they knew very little about the topic, and more likely to have an inaccurate range when the topic was familiar to them. While the concept seems counterintuitive, it showed that oil and gas professionals are overconfident in topics that were related to their industry. To combat this bias Capen suggested using lognormal distributions to show a range of prospective resources (Capen, 1984). Given that lognormal distributions are found when independent distributions are multiplied together, it is expected that EUR's, field size distributions, and even annual rainfall will be lognormal.

$$\text{Recoverable Hydrocarbon} = \left(\frac{7758 * GRV * N{:}G * \phi * (1 - Sw)}{FVF}\right) * RF$$

Equation 1: Ample multiplication in the Recoverable Hydrocarbon equation to create a lognormal distribution

Potential Issues

While lognormal distributions work well for large data sets they can also be deceiving when used on the individual prospect level. The Gross Rock Volume (GRV) and Net Rock Volume (NRV) of a prospect will have a lognormal shape (see Fig. 7), but the high end of the distribution will rarely be geologically possible.

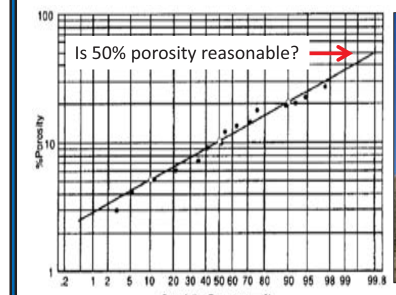


Fig. 1: Average Canadian field porosities are lognormal (Capen, 1992)

Fig. 2: Is the volume of El Capitan lognormal?

Why Does it Matter?

The tornado plot below shows when Gross Rock Volume and Net to Gross are combined (Net Rock Volume) they affect greater than 50% of the overall resource distribution. **Determining the accuracy of these numbers ensures the overall resource range is reasonable.**

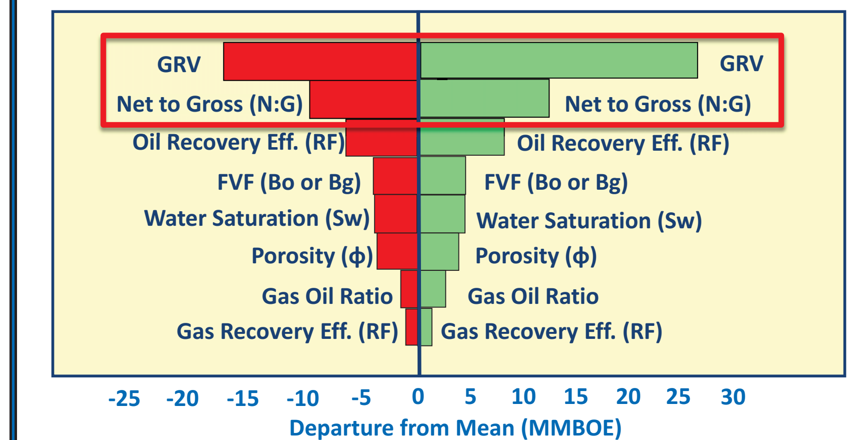


Fig. 3: Each variable's effect on the total resource distribution

Prospect Overview

Grand Slam Background	
Location	Earth (offshore)
Type of structure	Large, high relief 4-way
Image quality	Medium (sub-salt)
Mapping confidence	Medium
Reservoir	Thin bedded sands
Depositional system	Channel levee margin

Table 1: Prospect background

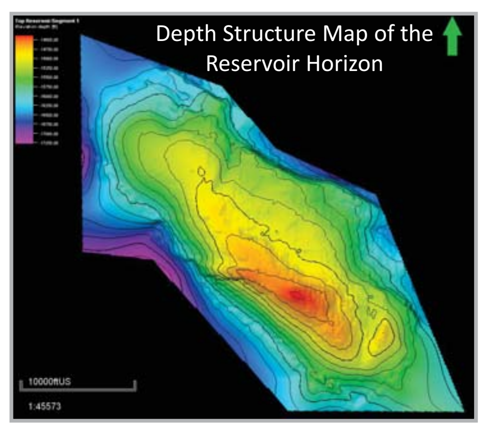


Fig. 4: Depth structure map of the Grand Slam prospect

Why is There Uncertainty?

Geoscientists can never be 100% certain of a mapped surface, no matter how much data is available.

Horizons mapped on 3-D seismic
- Velocity Model (especially sub-salt)
- Different versions of seismic
- New surveys/processing
- Alternate interpretations/interpreters

Horizons mapped on 2-D seismic
- Line density
- Image quality

Horizons mapped from well data
- Data density
- Gridding algorithm (IDW, spline, kriging, etc.)

Traditional Stochastic Approach

Monte Carlo simulators (e.g. Crystal Ball, MMRA, etc.) begin with a user input of a low & high side GRV range for the prospect, which are used to create a lognormal distribution.

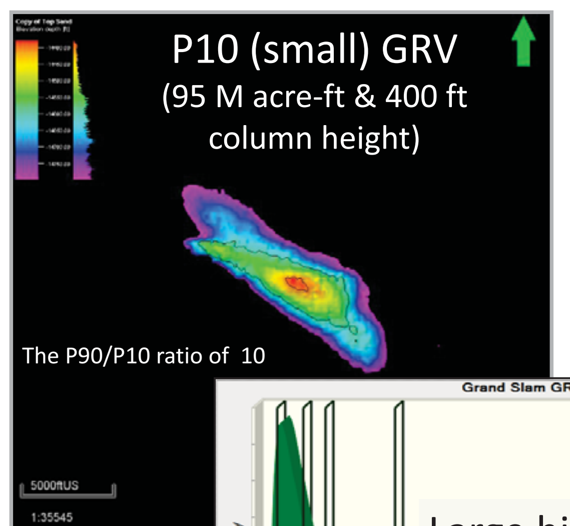


Fig. 5: P10 GRV

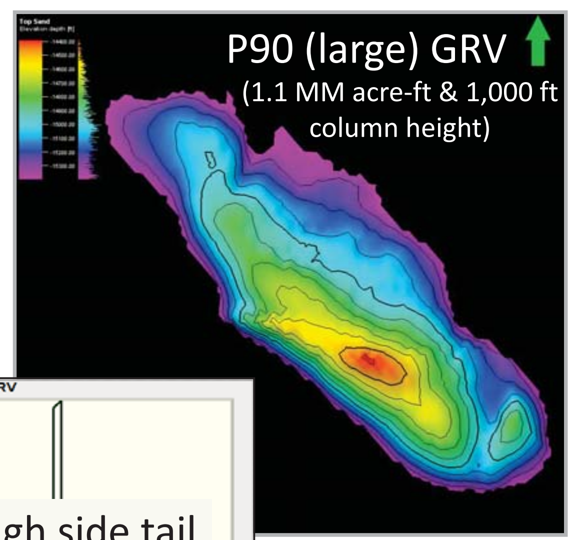


Fig. 6: P90 GRV

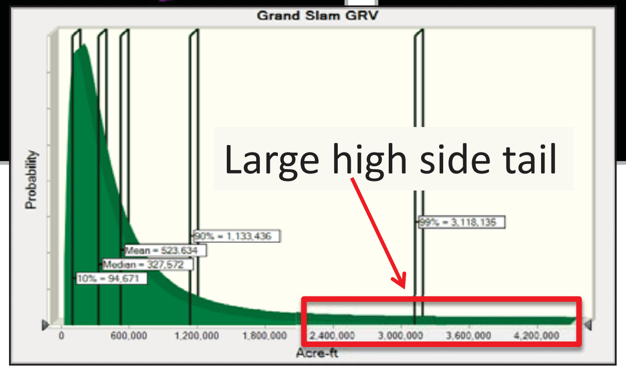


Fig. 7: Lognormal distribution of GRV

The Problem with Truncation

The high end of the GRV distribution is larger than geologically possible (it is several hundred feet below the spill point). When truncating the high end of a lognormal distribution you have two options: resample or spike.

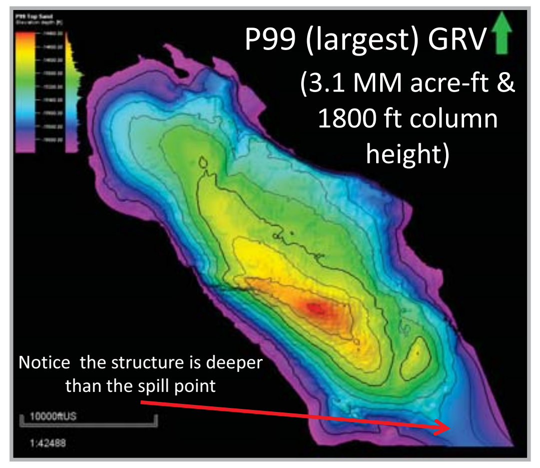


Fig. 8: The P99 GRV is geologically impossible

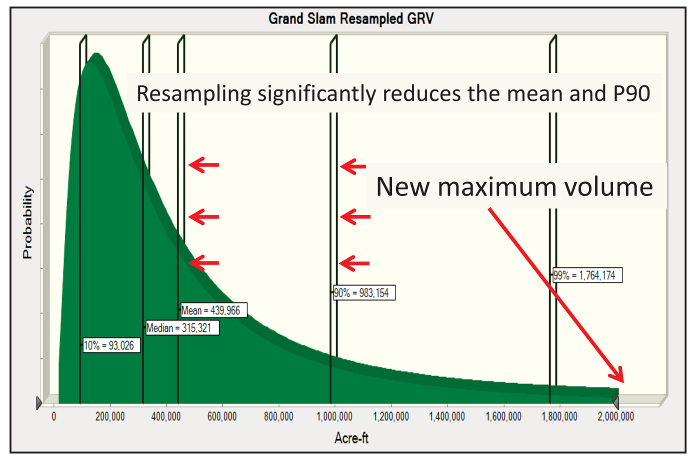


Fig. 9: The resampled GRV takes values greater than the truncation limit and resamples them

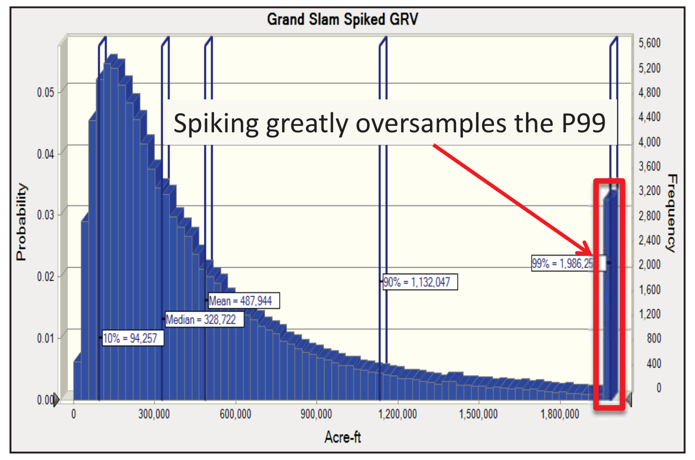


Fig. 10: The spiked GRV takes all values over the truncation limit and makes them the truncated value

Gross Rock Volume Distributions

Distribution Type	P10 (MAcre-ft)	P50 (MAcre-ft)	Mean (MAcre-ft)	P90 (MAcre-ft)	P99 (MAcre-ft)
Untrucated Lognormal	95	328	524	1,133	3,118
Resampled Lognormal	93	315	440	983	1,764
Spiked Lognormal	94	328	488	1,133	1,986

Table 2: GRV values from resampled and spiked truncations

Input Parameters for Recoverable Oil Calculation

Resource Parameter	Distribution Type	P10	Mean	P90
Net to Gross	Normal	25%	38%	50%
Porosity	Normal	22%	25%	28%
Sw	Normal	40%	27%	15%
FVF	Triangular	1.27	1.23	1.20
Rec. Factor	Normal	21%	25%	29%

Table 3: Stochastic ranges used to calculate recoverable oil

Recoverable Oil Distributions

Distribution Type	P10 (MMBO)	P50 (MMBO)	Mean (MMBO)	P90 (MMBO)	P99 (MMBO)
Untruncated Lognormal	8.7	33.1	56.3	124.6	362
Resampled Lognormal	8.6	31.8	47.4	107.7	222.8
Spiked Lognormal	8.7	33.1	52.5	121.9	267.4

Table 4: Recoverable oil from resampled and spiked truncations

Full-field Uncertainty Modeling

The traditional stochastic approach does not address uncertainty in the mapped horizon and causes unintended consequences when truncating lognormal distributions. Because this is a sub-salt prospect with medium mapping confidence it can be assumed the mapped horizon is not perfect, so ± 250 ft was used for mapping uncertainty.

Structural Uncertainty

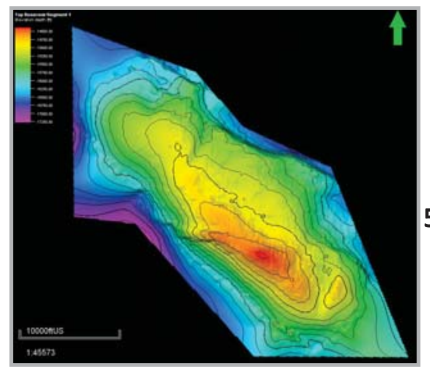


Fig. 11: The prospect's original horizon

X

501 Runs

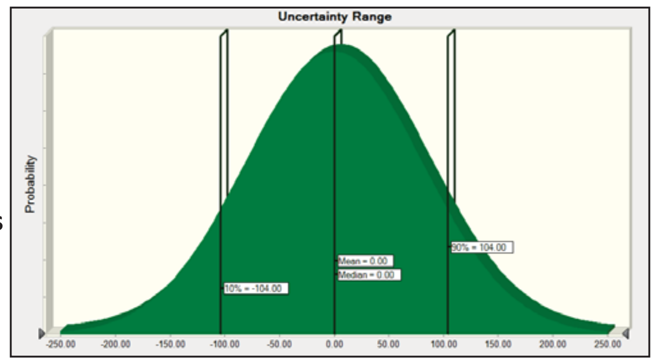


Fig. 12: The distribution will flex the horizon ± 250 ft

=

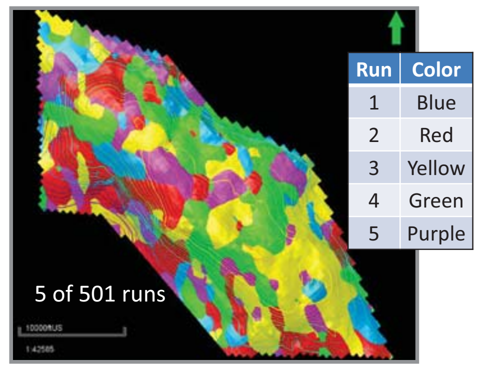


Fig. 13: Each color is a different structural horizon

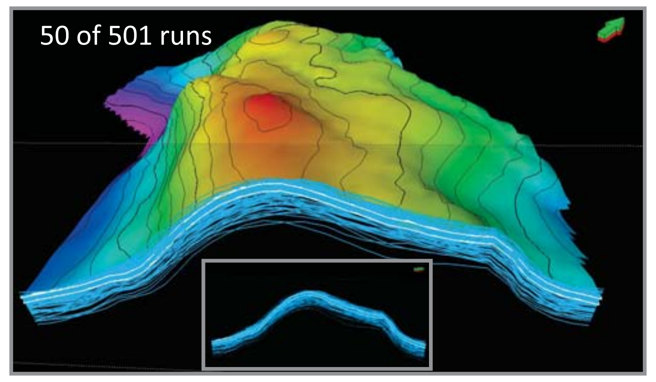


Fig. 14: 50 of the 501 horizons in cross section view

Results

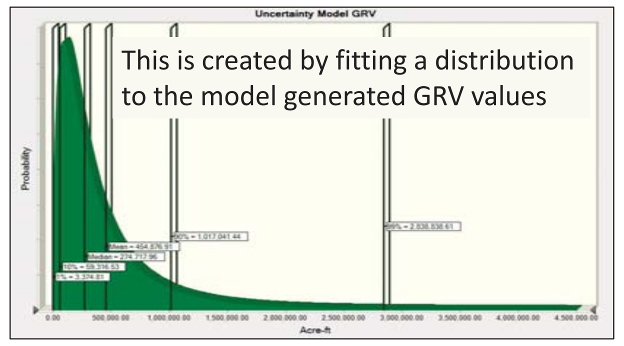


Fig. 15: The GRV range derived from the full-field uncertainty model

Model Derived GRV Values and Variance from Lognormal Distributions

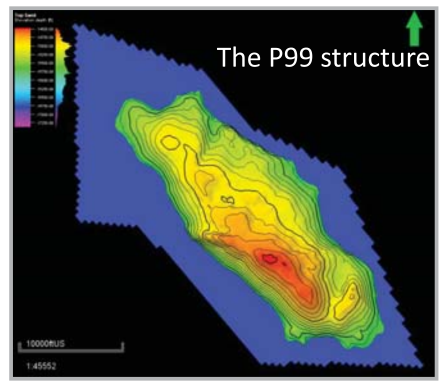
Distribution Type	P10 (MAcre-ft)	P50 (MAcre-ft)	Mean (MAcre-ft)	P90 (MAcre-ft)	P99 (MAcre-ft)
Uncertainty Model	59	275	455	1,017	2,839
Variance from Untruncated	-38%	-16%	-13%	-10%	-9%
Variance from Resampled	-36%	-13%	+3%	+3%	+38%
Variance from Spiked	-37%	-16%	-7%	-10%	+30%

Table 5: The uncertainty model gives geologically feasible results without arbitrarily truncating the higher values

Recoverable Oil Values and Variance from Lognormal Distributions

Distribution Type	P10 (MMBO)	P50 (MMBO)	Mean (MMBO)	P90 (MMBO)	P99 (MMBO)
Uncertainty Model	6.9	27.8	50.3	112.3	350.2
Variance from Untruncated	-21%	-16%	-11%	-10%	-3%
Variance from Resampled	-20%	-13%	+6%	+4%	+36%
Variance from Spiked	-21%	-16%	-4%	-7%	+24%

Table 6: The recoverable oil results mimic the GRV results, but the P99 volume can now be validated



Input Parameters for P99 Volume

GRV (MMAcre-ft)	2.4
Net to Gross	47%
Porosity	27%
Sw	17%
FVF	1.24
Rec. Factor	29%

Fig. 16: The input parameters that create the P99 volume

Conclusions

- **Full-field uncertainty gives an alternative methodology that allows the P1 – P99 volumes to be constrained by the geology and "fall out" of the data**
- Lognormal distributions work better on large data sets, not an individual prospect's area or volume
- Lognormal distributions need to be checked for geologic feasibility, especially on the high end (P99)
- Truncating lognormal distributions causes unintended statistical errors
- **There is no "one size fits all" solution in volumetric estimation, every estimator needs to carefully consider the inputs and validate outputs**

References & Acknowledgements

I would like to thank Marathon Oil Corporation for allowing me to publish this work and my colleagues who contributed to the development of this concept.

Capen E.C., 1976, The Difficultly of Assessing Uncertainty: Journal of Petroleum Technology, v. 28, no. 8, p. 843-850.

Capen, E.C., Megill, R.E., Rose, P.R. 1984, Evaluating and Managing Petroleum Risk, Tulsa, OK, Am. Assoc. Pet. Geol.

Capen, E.C., 1992, Dealing with Exploration Uncertainties *in* Steinmetz, R., ed., The Business of Petroleum Exploration, Tulsa, OK, Am. Assoc. Pet. Geol., p. 29-61.