Importance of Probability Distributions on Volumetric Characterizations in Resource Assessment

David M. Advocate[1] and Kenneth C. Hood[2]

Search and Discovery Article #42577 (2022)**
Posted October 18, 2022

*Adapted from extended abstract based on oral presentation given at 2022 AAPG Rocky Mountain Section Meeting, Denver Colorado, July 24-27, 2022
 DOI:10.1306/42577Advocate2022

[1]Geologist in Fort Collins, Colorado, USA. (dmadvoc@gmail.com)
[2]Geologist in Monument, Colorado, USA. (kchood@msn.com)

Abstract

Before making investment decisions for hydrocarbon opportunities, explorationists need to realistically evaluate chance of geologic and commercial success. Proper characterization of volume uncertainty, typically using statistical methods to evaluate parameters required for a hydrocarbon accumulation (porosity, net-to-gross, etc.), is essential. Use of past venture analysis has documented the tendency of industry to systematically overestimate the expected mean accumulation size and underestimate the volume range for undrilled prospects. Often such optimistic expectations result from improper characterization of the range and variance of volumetric input parameters.

The preferred choice of probability types used has long been a topic of debate and typically include Lognormal, Normal, Triangular, Beta, Uniform, and Gamma. When assigning distribution type and range, one must consider what each distribution represents. For parameters such as reservoir thickness or porosity, which have spatial or stratigraphic variations in measured values, the distribution represents uncertainty in the mean value for the evaluation unit. Generally, the distribution should be narrower than the range of individual measurements, but if biased, the range could be wider than or offset to measurements. For parameters such as closure area and height that will ultimately be a single measured value, the input distribution represents the range and probability of potential values. The granularity of the volumetric equation can vary, such as Gross Rock Volume as a single aggregate parameter or as multiple input components. Use of multiple components is preferable to enable better control over the uncertainty distribution. Fluid contacts are complex and may be poorly represented by a simple distribution.

In this presentation, we discuss strengths and weaknesses of various options and argue that the Beta distribution is well suited for most symmetrical and asymmetrical volumetric inputs. By using modified inputs, the Beta distribution can be defined using minimum, maximum, mode, and dispersion (λ) parameters. Increasing parameter variance within the defined range is crucial, as the range may be limited by low-end cutoffs (e.g., minimum porosity cutoff) and high-end physical limits (e.g., net-to-gross less than 1). Because naturally unbounded distributions (e.g., Lognormal) must be truncated, the bounded Beta distribution is more intuitive and can reasonably represent the appropriate level of skewness. We discourage using intermediate input values (e.g., P90 and P10 rather than min and max) and allowing software to extend the

range, as this approach can result in unforced errors (extending ranges outside of allowable range) and obscure the ability to learn as prospects are drilled. It is preferable to follow a well-defined workflow to ensure that the range is sufficient rather than depending on software to correct for user underestimation of the range.

Conclusions

1. Probabilistic volumetrics are preferable, especially in areas with significant exploration uncertainty

2. Increasing the distribution range and dispersion of volumetric parameters compensates for the natural human tendency to be overconfident

3. *Interpretation bias* can have a large impact on volumetric estimate

4. Using a simple Beta distribution for most volumetric parameters avoids the complexity of truncating unbounded distributions and the occurrence of unforced errors
    - Spend your time evaluating the range and dispersion rather than arguing about which distribution type to use
    - Simple distribution types are often not suitable for modeling column heights or fluid contacts

5. Take care when using analog data – use the mean values of prospects, not individual measurements from wells

6. Higher granularity of GRV allows easier integration of map, analog and local data

7. Even modest variations in the range and dispersion of parameter distributions can have significant impact on economic analyses

Acknowledgements

- Lisa Stright – drawing our attention to online social interpretations and some seminal geological bias publications

- Jeff Brown, Bill Hood, Michelle Lund, and many others for helpful discussions and for reviewing earlier presentation drafts

- AAPG and RMGS organizers for their preparation for this meeting

- And, most importantly, this audience for your attention!

References:

Citron, G., Bond, M. and Carragher, P., 2017, Exploration assurance team best practices, AAPG Search and Discovery Article #70264, *Adapted from oral presentation given at AAPG 2017 Annual Convention and Exhibition, Houston, Texas, United States, April 2-5, 2017), 34 p.

Edmundson, I.S., R. Davies, L.U. Frette, S. Mackie, E.A. Kavli, A. Rotevatn, G. Yielding, and A. Dunbar. 2021, An empirical approach to estimating hydrocarbon column heights for improved predrill volume prediction in hydrocarbon exploration. AAPG Bulletin, v. 105, no. 12 (December 2021), pp. 2381–2403.

Hood, K. C. 2019. Hydrocarbon Column Height. Risk Coordinator Workshop, Houston. DOI: 10.13140/RG.2.2.27257.75362

Hood, K. C. 2022. Hydrocarbon Column Heights (Part 1). R&A Assurance Blog. https://www.linkedin.com/feed/update/urn:li:activity:6907007940658626560/

Hood, K. C. 2022. Hydrocarbon Column Heights (Part 2). R&A Assurance Blog. https://www.linkedin.com/feed/update/urn:li:activity:6895109553487253504/

James, B., Grundy, A.T., and Sykes, M.A., 2013, The depth-area-thickness (DAT) method for calculating gross rock volume: A better way to model hydrocarbon contact uncertainty: AAPG International Conference & Exhibition, Cartagena, Colombia, September 2013, Poster Session

Murtha, J., and J. Ross. 2009. Uncertainty and the Volumetric Equation. Journal of Petroleum Technology, 61 (09), pp. 20-22.

Ofstad, K., Kullerud, L., and Helliksen, D., 2000, Evaluation of Norwegian Wildcat Wells (Article 1).in *Improving the Exploration Process by Learning from the Past* edited by K. Ofstad, J.E. Kittilsen and P. Alexander-Marrack. NPF Special Publication 9, pp. 23-31, Published by Elsevier Science B.V., Amsterdam. © Norwegian Petroleum Society (NPF), 2000

Sinding-Larsen, R., Lillestol, J., and K. C. Hood, K.C., 2012, Distributions with Beta properties as a model for the parent size distribution of grown oil and gas pools. Ninth International Geostatistics Congress, Oslo, Norway, 2012, 4 p.

Sykes M.A., Welsh, M.B., and Begg, S.H., 2011, Don’t drop the anchor: recognizing and mitigating human factors when making assessment judgments under uncertainty, SPE 146230, p. 1-15 https://www.researchgate.net/publication/241789883_Don't_Drop_the_Anchor_Recognizing_and_Mitigating_Human_Factors_When_Making_Assessment_Judgments_Under_Uncertainty

Importance of Probability Distributions on Volumetric Characterizations in Resource Assessment

David M. Advocate
Kenneth C. Hood

July 27, 2022

Rocky Mountain Section AAPG
Denver, CO
July 24-27, 2022

Outline

1. The importance of probability distributions in volumetric characterization
2. How opportunities are sized – Deterministic Vs. Probabilistic
3. Examine common distribution types
4. Volumetric Workflows for Frontier and Early- to Late-Exploration trends
   - Importance of making the distribution range wider and avoiding unforced errors
5. Modeling Gross Rock Volume (GRV) with different granularity
   - Key uncertainties in the GRV term
6. Where simple distribution are inadequate (Column Height / fluid contacts)
7. Implications of using probabilistic analysis for business decisions
8. Conclusions

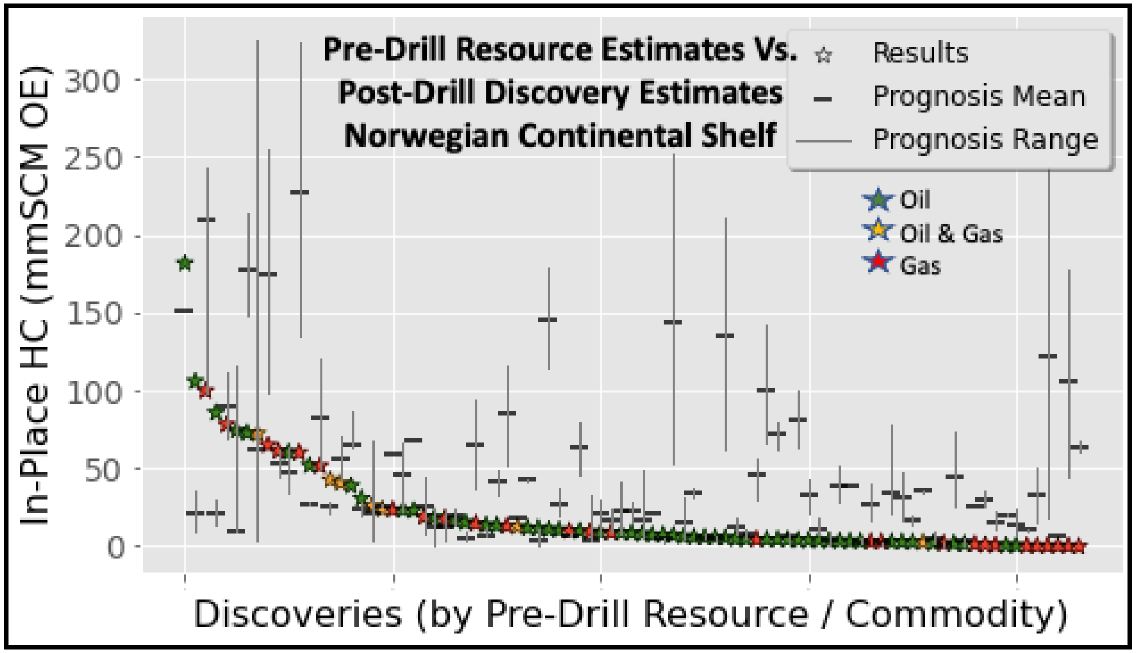
Why is this Important?

Meaningful investment decisions require realistic evaluation of the Geologic Chance of Success, Success-Case Volumes and the Volumetric uncertainty

Past venture analyses have documented systematic overestimates of the expected mean size and underestimation of the volume range

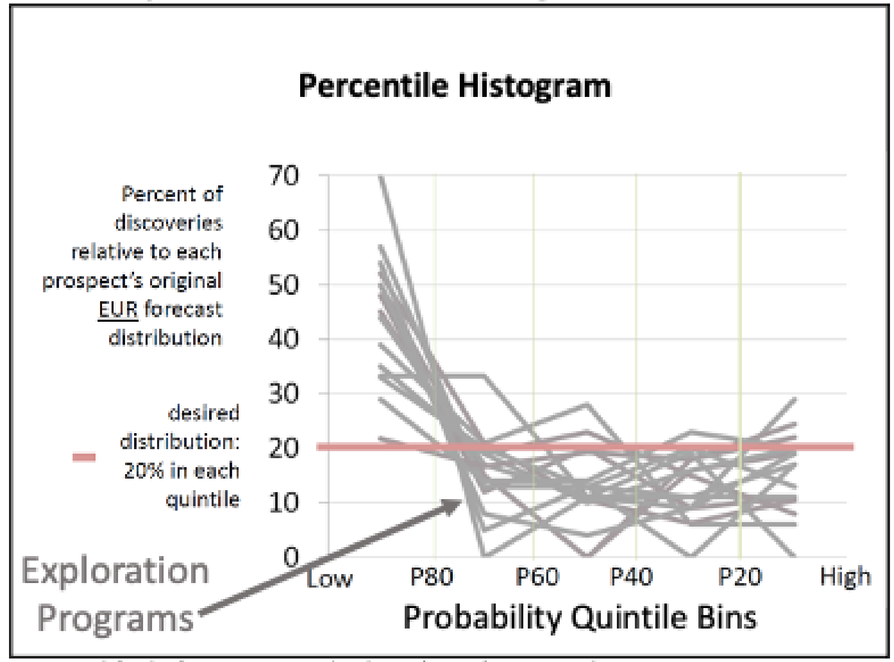
In this talk we will investigate how modeling volumetric parameters impact ability to evaluate prospects

Likely Cause: One or more parameter estimates too high



Data extracted from Ofstad, Kullerud and Helliksen (2000, Fig. 5) -195 wildcat wells, 1990 - 1997

Likely Cause: Parameter ranges too narrow



Modified after Citron and others (2017) – 13 Exploration programs

How are Opportunities Sized

Typically Sized using Volumetric Equation:

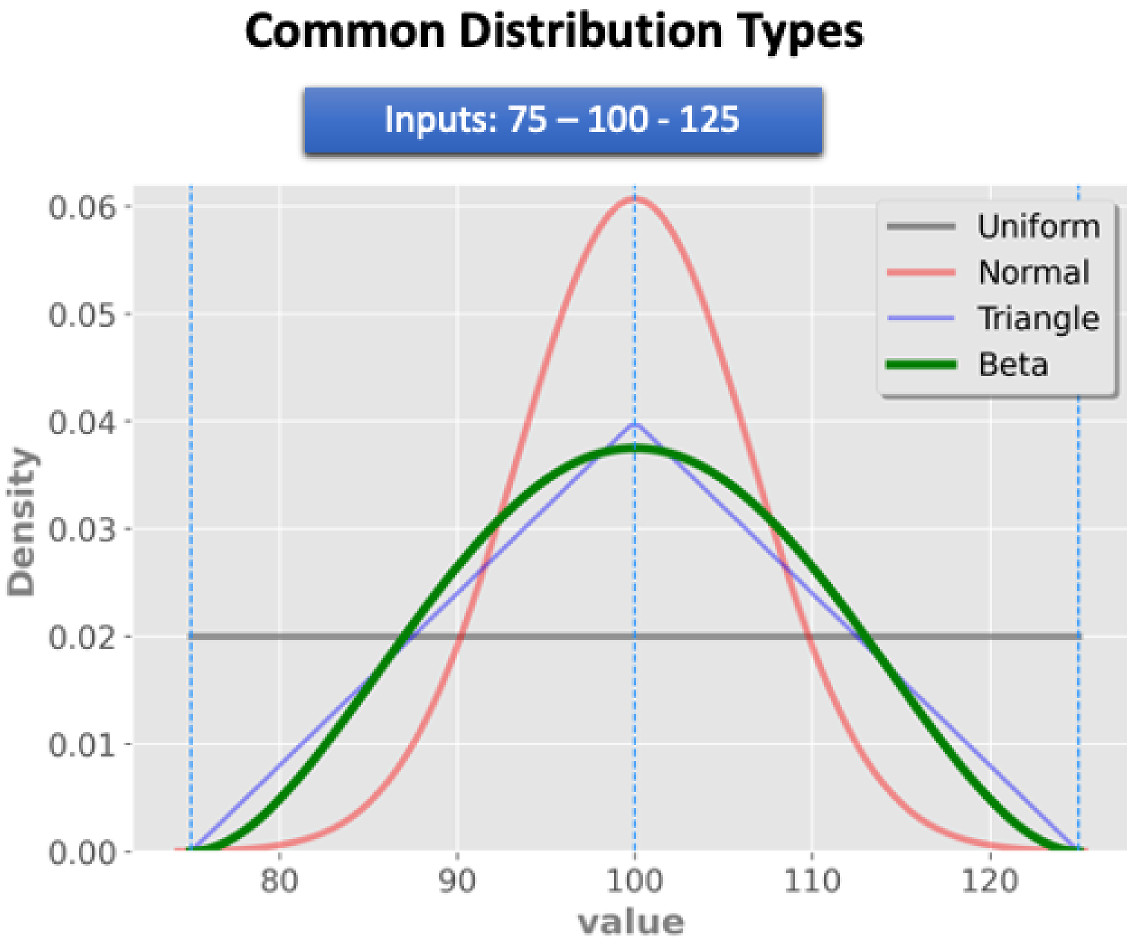
$$EUR = GRV * NTG * \Phi * Sh / FVF * RF * c$$

GRV can take many forms, with variable linkages to key geologic controls and allows prospect specific information from maps and analog data

EUR	Estimated Ultimate Recovery
GRV	Gross Rock Volume
NTG	Net-to-Gross
Φ	Porosity
Sh	Hydrocarbon Saturation
FVF	Formation Volume Factor
RF	Recovery Factor
c	Constant

- A *Deterministic* estimate will only incorporate the single “best estimate” of EUR and says nothing about the potential range of outcomes
- In a *Probabilistic* estimate, each parameter can have an input distribution representing uncertainty in the estimated value: typically convolved using Monte Carlo simulation
  - The distribution type can have a substantial impact on the results, hence a focus of this talk
- Parameter distributions can represent two different families of properties
  1. Uncertainty in the mean of parameters with multiple measurements and spatial or stratigraphic variability, e.g., porosity
  2. Range and probabilities of properties with single measured value, e.g., closure area or fluid contacts

Six Common Distribution Types used in Probabilistic Estimation

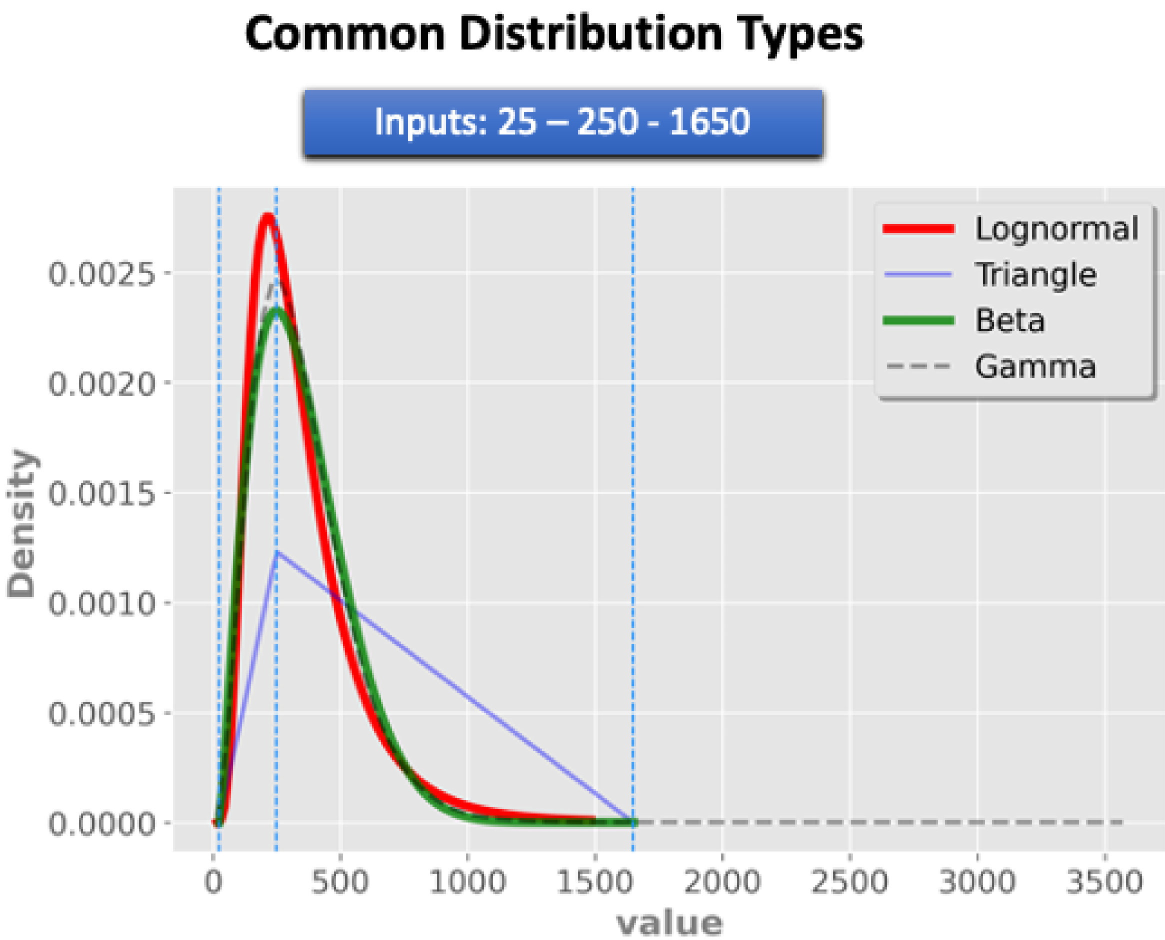


- Distributions are commonly modeled using three points: Minimum, Most-Likely (Mode) and Maximum value
- Three key properties to Consider:
  - **Bounded / Unbounded**: Bounded honors the Min and Max inputs and eliminates the complexity of truncating **Unbound** PDFs
  - **Symmetry / Skew**: flexibility in the shape the PDF
  - **Dispersion**: ability to spread out the distribution
- Commonly used distribution types used or offered in commercial assessment software include:
  1. **Uniform** – Bounded, No mode, broad dispersion
  2. **Normal** – Unbounded, symmetrical, variable dispersion
  3. **Triangle** – Bounded, pronounced mode, moderate dispersion, variable skewness
  4. **Beta** – Bounded, supports variable dispersion and skewness

  <u>For highly right skewed distributions</u>

  5. **Gamma** – Unbound with extended tail
  6. **Lognormal** – Unbounded and variably dispersed

Six Common Distribution Types used in Probabilistic Estimation



- Distributions are commonly modeled using three points: Minimum, Most-Likely (Mode) and Maximum value
- Three key properties to Consider:
  - **Bounded / Unbounded**: Bounded honors the Min and Max inputs and eliminates the complexity of truncating **Unbound** PDFs
  - **Symmetry / Skew**: flexibility in the shape the PDF
  - **Dispersion**: ability to spread out the distribution
- Commonly used distribution types used or offered in commercial assessment software include:

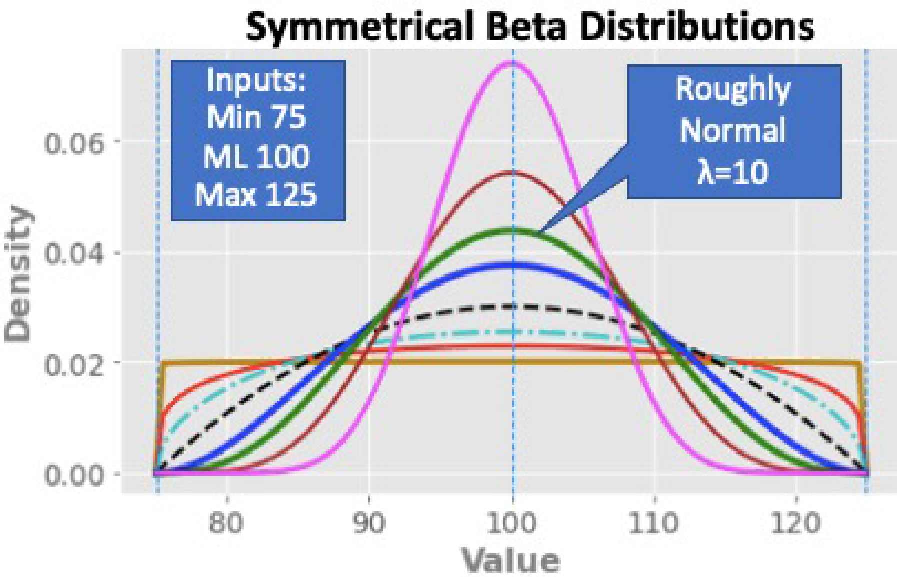
1. **Uniform** – Bounded, No mode, broad dispersion
2. **Normal** – Unbounded, symmetrical, variable dispersion
3. **Triangle** – Bounded, pronounced mode, moderate dispersion, variable skewness
4. **Beta** – Bounded, supports variable dispersion and skewness

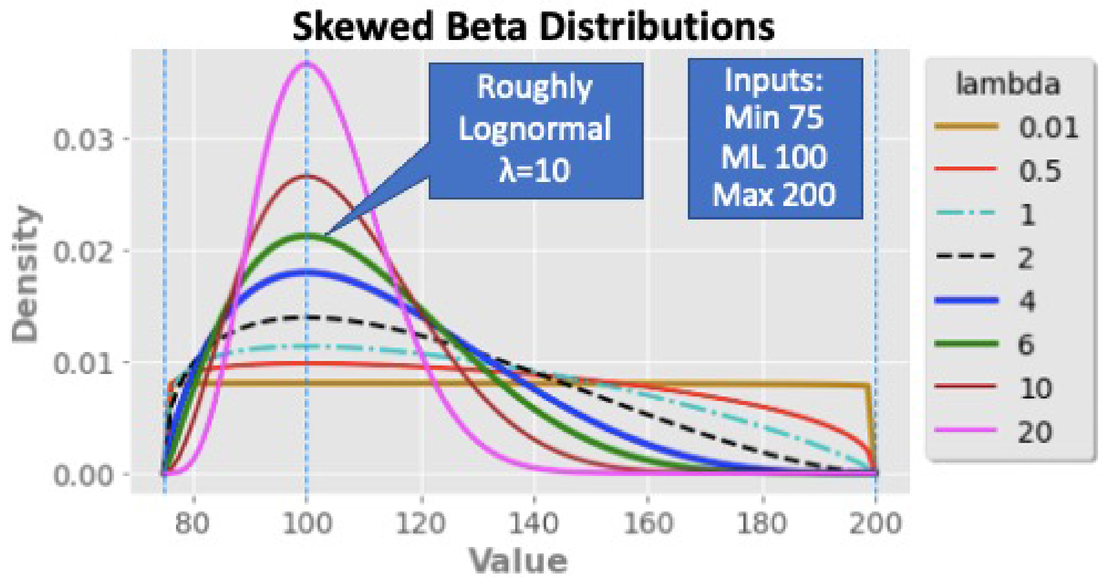
For highly right skewed distributions

5. **Gamma** – Unbound with extended tail
6. **Lognormal** – Unbounded and variably dispersed

Comparison of Beta Distributions with Variable Dispersion & Skew

We recommend the Beta distribution owing to the bounded nature, shape flexibility and adjustable dispersion





- Beta has two shape parameters (α, β). Calculating (α, β) using the dispersion factor (λ) gives this PDF great flexibility

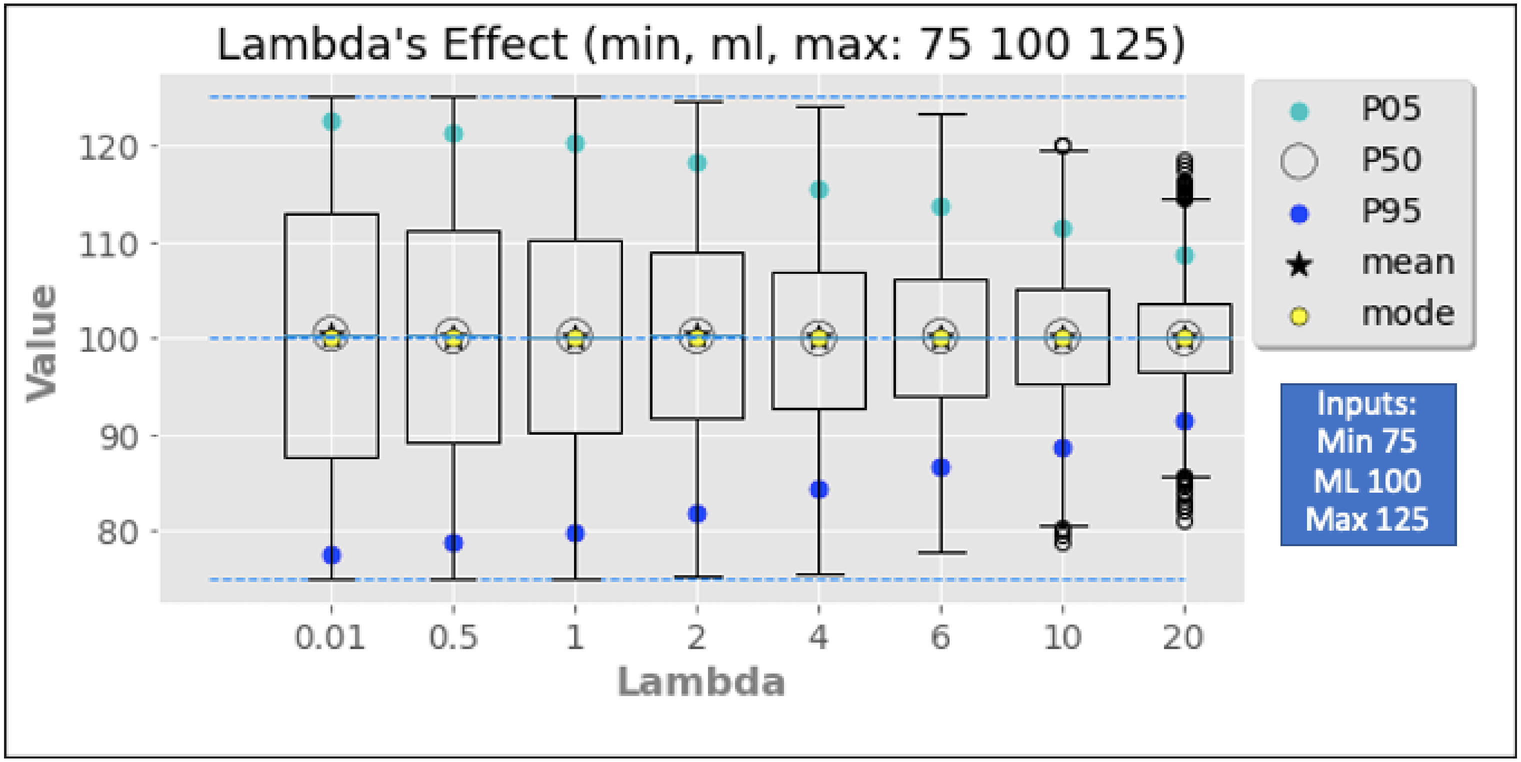
  $\alpha = 1 + \lambda$ ** (mode - min) / (max - min)* $\beta = 1 + \lambda$ ** (max - mode) / (max - min)*

- Lambda <0.5 can approximate a Uniform distribution
- Lambda ~ 10 approximates the Normal and Lognormal
- Lambda of 1 - 4 are ideal for broadening the dispersion yet leaving a discernable most-likely peak

Lambda's Quantitative Effect on Dispersion - Symmetrical Distributions

Symmetric Beta Distribution

Box and Whisker Plot with with P05, P95, Mean and Mode Overlain



Effect of Lambda on a symmetrical distribution with 10,000 random samples:

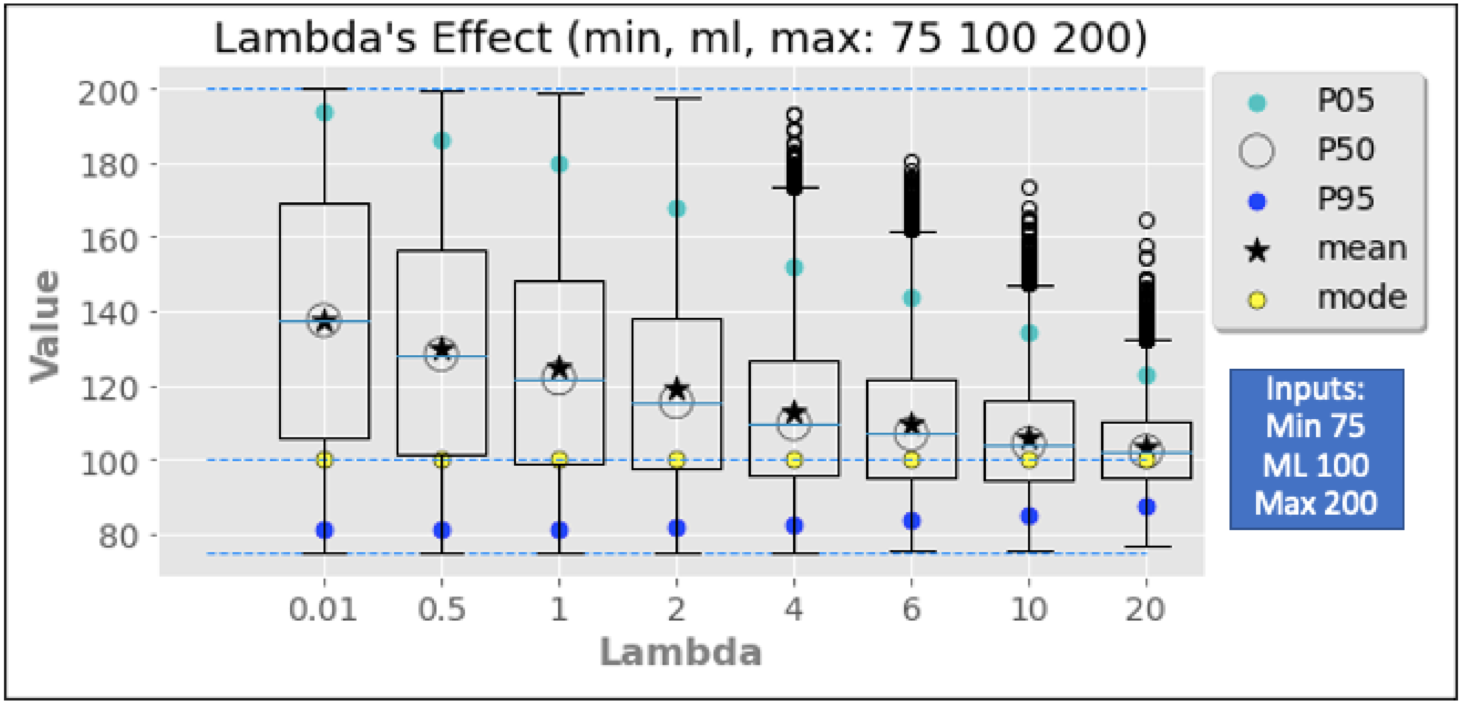
- Lower Lambda values honor the specified input range
- Larger Lambda values (> 4) dramatically reduce the interquartile and and P95 – P05 range, and fail to extend to the specified Min and Max
- In the symmetrical case, the *Mean*, *P50* and *Mode* are aligned

We recommend λ values <= 4

Lambda's Quantitative Effect on Dispersion - Skewed Distributions

Asymmetric Beta Distribution

Box and Whisker Plot with P05, P95, Mean and Mode Overlain



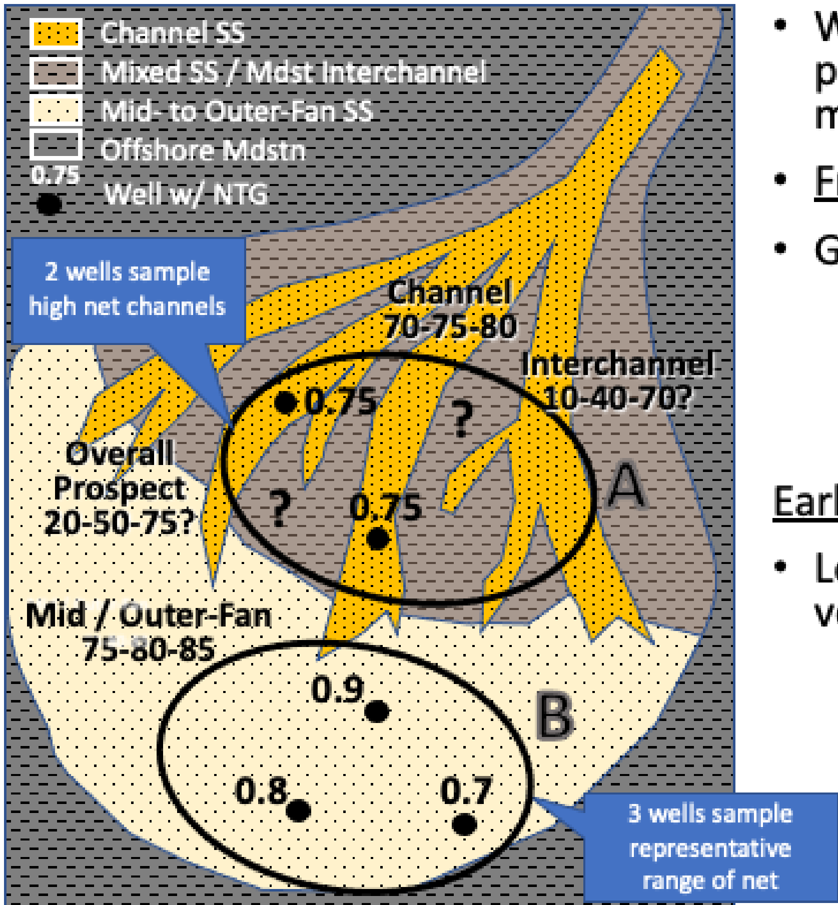
Effect to Lambda on a skewed distribution with 10,000 random samples:

- Higher Lambdas don't sample the high end of the distribution, particularly with λ > 4
- The mode across the full range of Lambdas is honored
- With increasing Lambda, the mean and P50 values approach the mode.
- **Recommend using Lambda values < 4 if you want to preserve the input ranges**

We recommend λ values < 4

Different Workflows Considerations

Early Mature Exploration Play NTG



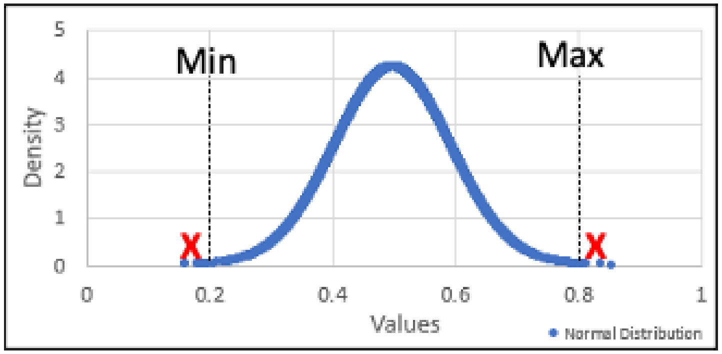
- Identifying the full range and mean of the distribution is perhaps more important that the distribution type
- Workflows for estimating the mean and range of volumetric parameters are considerably different from frontier to early- to mid- and late-mature exploration areas:
- <u>Frontier Exploration Areas</u> (Analog Well Data Only)
- Geophysical data is integrated with regional observations
  - Ensure that the data represent ***Mean*** values from analog discoveries
  - Uncertainty will be high as analogs may not capture full range of variability
  - Assigned range and dispersion volumetric parameters should be high

<u>Early- to Late-Mature Exploration Areas</u> (Analog and Local Data)

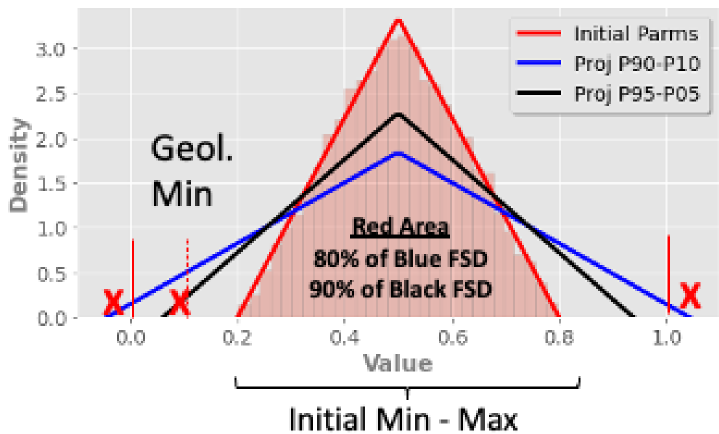
- Local well data are added to the data mix in addition to seismic velocity data, lithology, HC column, fluid properties, etc.
  - With limited well control, **data bias needs to be considered**, e.g., the wells located in the "sweet spot" only
  - May need to supplement local data with analog data
  - With sufficient well control, statistical approaches can be applied to estimate the parameter ranges, e.g., Murtha and Ross (2009) and Sykes et al. (2011)

"Just Make the Range Wider" – Potential Unforced Errors

Unbound Normal Distribution



NTG Initial and Extrapolated Triangle



Initial NTG Inputs

Min	.2
Mode	.5
Max	.8

NTG of 0.2 is the geologic minimum

Extrapolate NTG Inputs

Min	-.04
Mode	.5
Max	1.04

X Unforced Errors

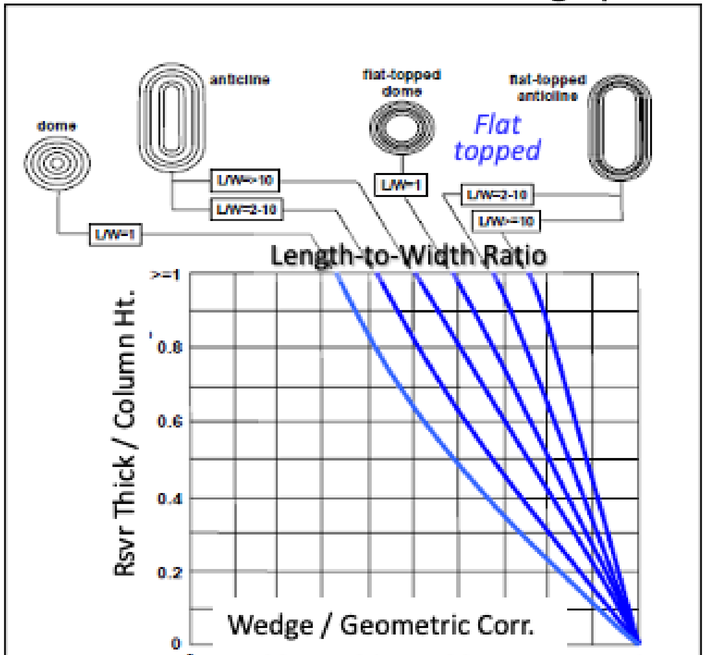
(when min and max values extend beyond practical limits or geologic minima)

- The range must be limited by physical and practical limits, for example
  - NTG < 1
  - Upper limit or Geol. Minimum for Reservoir thickness
- The range can be violated in two ways

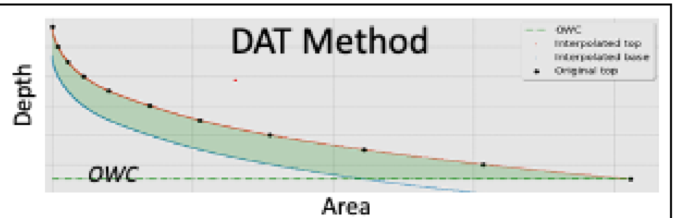
1) Unbounded distribution can generate values beyond specified inputs (**X**)
   - Unbound distributions must be truncated. This adds complexity
2) Extrapolated Min and Max values
   - To compensate for the natural tendency to underestimate the range, some geologists treat the *initial* min and max values as p95/P05 or P90/P10 values and extrapolate to "*true*" min and max values
     - This can lead to extrapolating beyond physical limits and geologic minima
     - We discourage allowing software to extrapolate min and max values blindly. This could result in unforced errors
     - Also, allowing software to extrapolate the Min and Max values could obscure's the geologists' judgements as exploration and production programs are executed

Gross Rock Volume – Granularity

Geometric Correction Nomograph



From James et.al, 2013



- GRV is an aggregation of multiple volumetric elements, each having a distribution
  - Closure area, closure heigh, L/W ratio, flat top fraction, reservoir thickness, column height, geometric correction factor
- Granularity refers to the way these elements are considered

Increasing Granularity

1. **GRV as a single input distribution,** e.g., Mean GRV ± 25%
2. **Closure area, reservoir thickness, and geometric correction** as direct input distributions (*closure height is implicit in the geometric correction*)
3. ***Simple Geometry** – Trap approximated by simple shapes (e.g., cone, prism, trapezoid)
   - Geometric correction is calculated on a trial-by-trial basis
4. ***Complex Geometry** – Uses Depth-Area-Thickness (DAT) triplets or Depth-Area-Pairs (DAP) with reservoir thickness and L/W ratio
   - Structural contour maps are used giving better description of geometric shape
   - DAT allows spatial variation of thickness, DAP method doesn't
5. **Cellular Model** – Grid-based integration of cells that provides an extremely high precision GRV calculation and shows spatial variability
   - Good for deterministic models, but computationally complex for probabilistic modeling – may require many maps

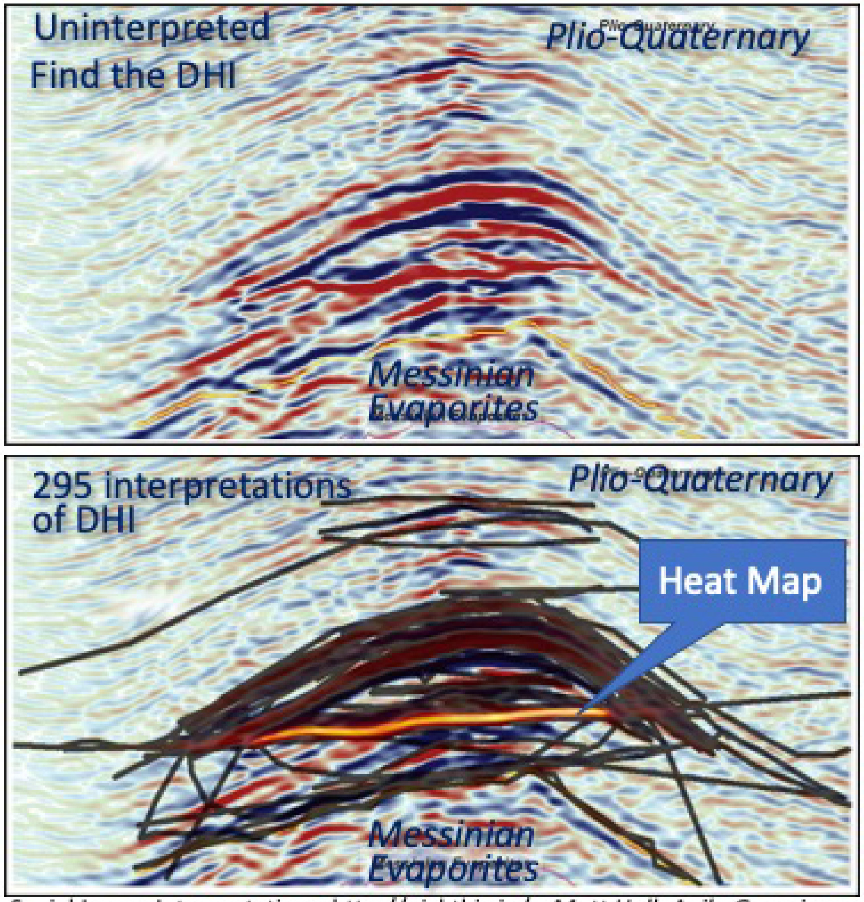
**Fluid contacts, can be represented as volume fill fraction, column height or contact depth distributions*

- **Higher granularity allows easier integration of map, analog and local data**

Key Uncertainties in Volumetric Inputs

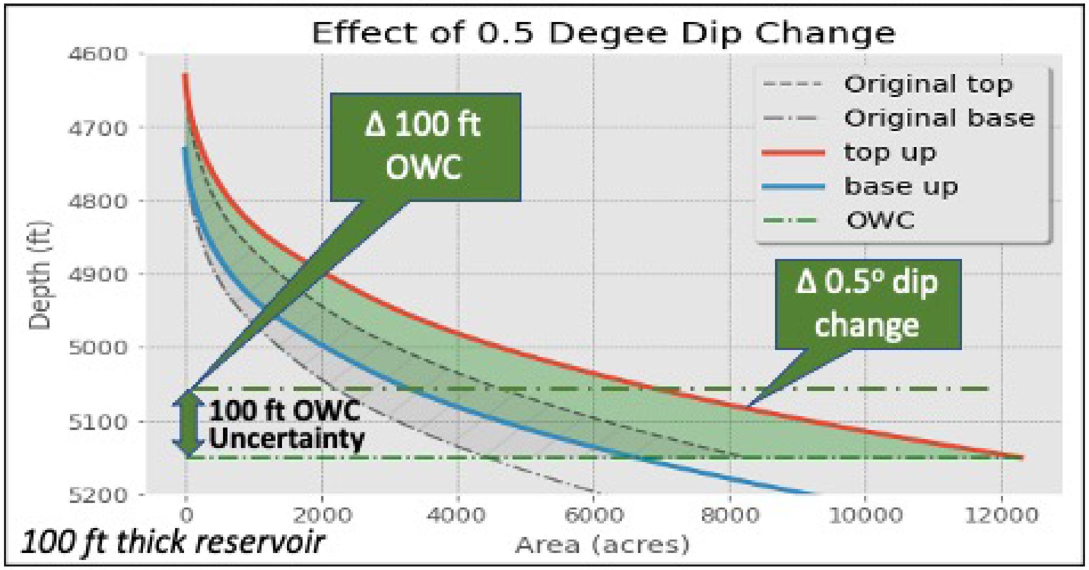
Interpreter Bias Example

Mediterranean Sea Evaporites 2D Seismic



Social Image Interpretation - http://pickthis.io/ *Matt Hall, Agile Geoscience*

- GRV is particularly sensitive to structural dip and fluid contacts
- Each factor can suffer from interpretation bias, particularly in the absence of quality seismic data and without well control
- Fluid contacts may be difficult to interpret pre-drill, especially in over-pressured systems with leaky seals and structural/stratigraphic complexity
- Modest changes in fluid contacts or structural dip can have profound effect on estimated GRV (especially for shallow dipping reservoirs)

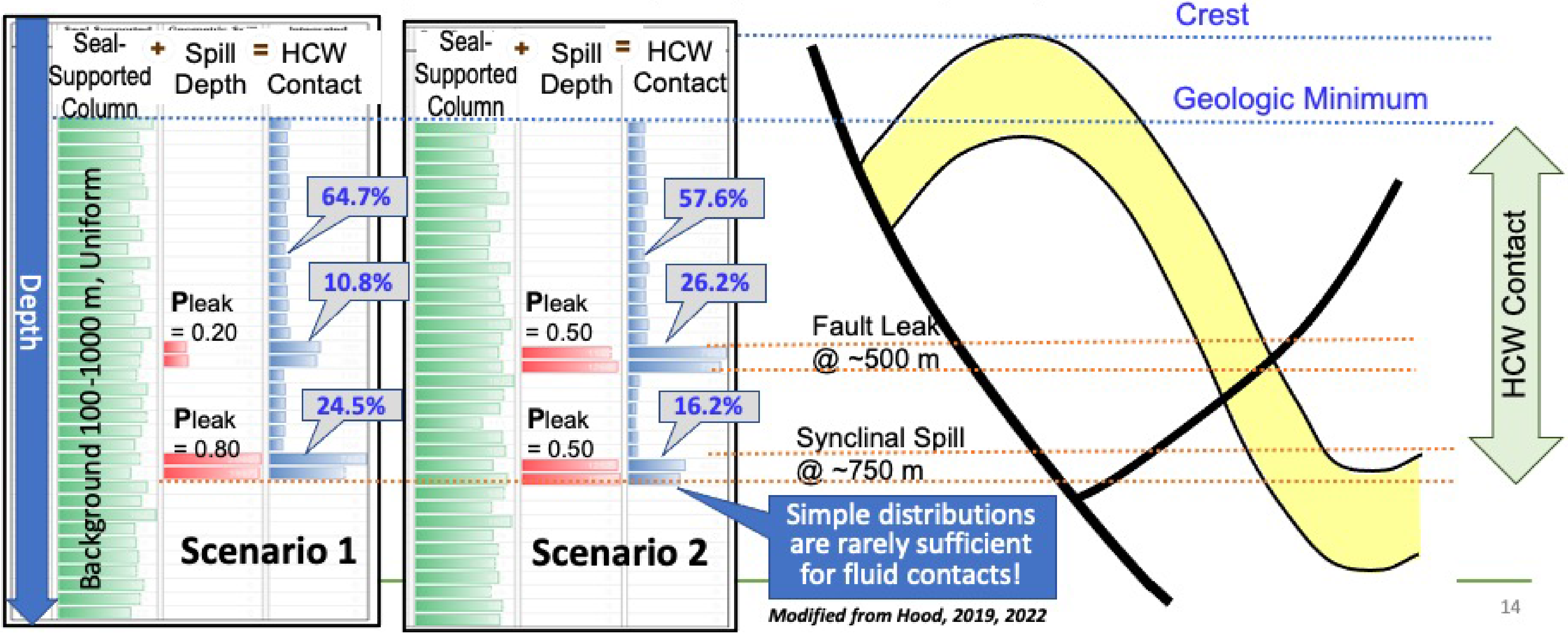


Δ 0.5° (2° to 1.5°) dip = 49% change in GRV Δ100 ft OWC = 49% change in GRV

Column Height or Hydrocarbon-Water Contact

Column Height (Fluid Contract) is an **exception** to guidance to use simple Beta distributions

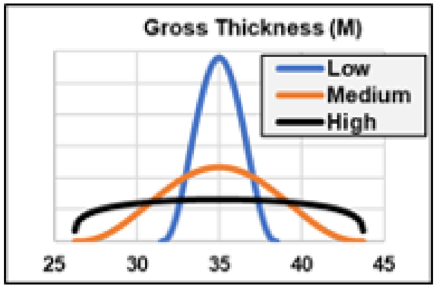
- Column Height is a complex interaction of bed seal capacity and potential geometric leaks (synclinal spill, fault juxtapositions, etc.)
- Preferred implementation is column height distribution representing seal capacity interacting with spill distribution
- In cases with alternative probabilities of leak, multiple scenarios may be required

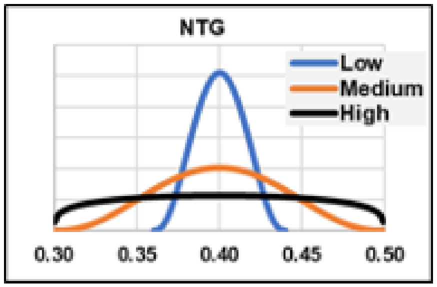


Modified from Hood, 2019, 2022

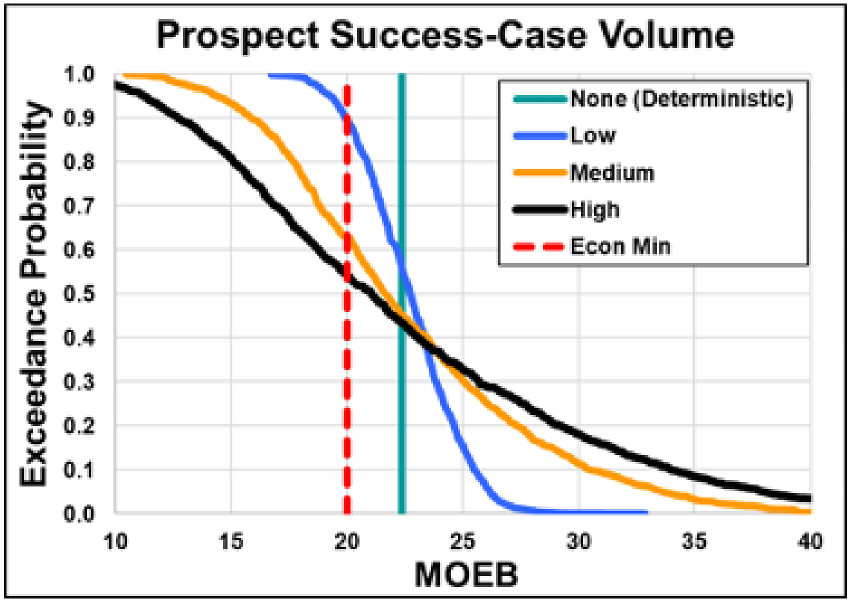
Economic Impacts

Parameter Uncertainty





Prospect Uncertainty Characterization



Prospect Uncertainty Characterization	GCOS	Pct Economic	ECOS	Success Mean (MOEB)		P10:P90	Parameter Uncertainty
				Geologic	Economic		
Uncertainty: None	0.40	100%	0.40	22.6	22.6	n/a	**Determistic**
Uncertainty: Low	0.40	89%	0.36	22.6	23.2	1.3	**+/- 10%; λ = 5**
Uncertainty: Medium	0.40	63%	0.25	22.6	25.9	1.9	**+/- 25%; λ = 5**
Uncertainty: High	0.40	54%	0.22	22.6	28.4	2.7	**+/- 25%; λ = 0.5**

- Example to illustrate significant business impact resulting from range and dispersion of volumetric parameters
- The prospect has a GCOS=0.40 and Economic minimum of 20 MOEB
- Three scenarios were tested, each with increasing uncertainty for all input parameters
- Simulation of each scenario, shown as exceedance probability, illustrates the significant impact on ECOS and success case mean volume
- Business analysis of a prospect with a 40% ECOS (deterministic case) is much different than one with a 22% ECOS (high uncertainty case)

Conclusions

1. Probabilistic volumetrics are preferable, especially in areas with significant exploration uncertainty
2. Increasing the distribution range and dispersion of volumetric parameters compensates for the natural human tendency to be overconfident
3. *Interpretation bias* can have a large impact on volumetric estimate
4. Using a simple Beta distribution for most volumetric parameters avoids the complexity of truncating unbounded distributions and the occurrence of unforced errors
   - Spend your time evaluating the range and dispersion rather than arguing about which distribution type to use
   - Simple distribution types are often not suitable for modeling column heights or fluid contacts
5. Take care when using analog data – use the mean values of prospects, not individual measurements from wells
6. Higher granularity of GRV allows easier integration of map, analog and local data
7. Even modest variations in the range and dispersion of parameter distributions can have significant impact on economic analyses

Acknowledgements

- Lisa Stright – drawing our attention to online social interpretations and some seminal geological bias publications
- Jeff Brown, Bill Hood, Michelle Lund, and many others for helpful discussions and for reviewing earlier presentation drafts
- AAPG and RMGS organizers for their preparation for this meeting
- And, most importantly, this audience for your attention!

References

Citron, G., Bond, M. and Carragher, P., 2017, Exploration assurance team best practices, AAPG Search and Discovery Article #70627, *Adapted from oral presentation given at AAPG 2017 Annual Convention and Exhibition, Houston, Texas, United States, April 2-5, 2017), 34 p.

Edmundson, I.S., R. Davies, L.U. Frette, S. Mackie, E.A. Kavli, A. Rotevatn, G. Yielding, and A. Dunbar. 2021, An empirical approach to estimating hydrocarbon column heights for improved predrill volume prediction in hydrocarbon exploration. AAPG Bulletin, v. 105, no. 12 (December 2021), pp. 2381–2403.

Hood, K. C. 2019. Hydrocarbon Column Height. Risk Coordinator Workshop, Houston. DOI: 10.13140/RG.2.2.27257.75362

Hood, K. C. 2022. Hydrocarbon Column Heights (Part 1). R&A Assurance Blog. https://www.linkedin.com/feed/update/urn:li:activity:6907007940658626560/

Hood, K. C. 2022. Hydrocarbon Column Heights (Part 2). R&A Assurance Blog. https://www.linkedin.com/feed/update/urn:li:activity:6895109553487253504/

James, B., Grundy, A.T., and Sykes, M.A., 2013, The depth-area-thickness (DAT) method for calculating gross rock volume: A better way to model hydrocarbon contact uncertainty: AAPG International Conference & Exhibition, Cartagena, Colombia, September 2013, Poster Session

Murtha, J., and J. Ross. 2009. Uncertainty and the Volumetric Equation. Journal of Petroleum Technology, 61 (09), pp. 20-22.

Ofstad, K., Kullerud, L., and Helliksen, D., 2000, Evaluation of Norwegian Wildcat Wells (Article 1).in *Improving the Exploration Process by Learning from the Past* edited by K. Ofstad, J.E. Kittilsen and P. Alexander-Marrack. NPF Special Publication 9, pp. 23-31, Published by Elsevier Science B.V., Amsterdam. © Norwegian Petroleum Society (NPF), 2000

Sinding-Larsen, R., Lillestøl, J., and K. C. Hood, K.C., 2012, Distributions with Beta properties as a model for the parent size distribution of grown oil and gas pools. Ninth International Geostatistics Congress, Oslo, Norway, 2012, 4 p.

Sykes M.A., Welsh, M.B,, and Begg, S.H., 2011, Don't drop the anchor: recognizing and mitigating human factors when making assessment judgments under uncertainty, SPE 146230, p. 1-15 https://www.researchgate.net/publication/241789883_Don't_Drop_the_Anchor_Recognizing_and_Mitigating_Human_Factors_When_Making_Assessment_Judgments_Under_Uncertainty

Thank You