Petroleum Assessment via Hierarchical Modeling*

Gordon M. Kaufman[1] and Jack Schuenemeyer[2]

Search and Discovery Article #40558 (2010)
Posted June 30, 2010

* Adapted from an oral presentation at AAPG Annual Convention and Exhibition, New Orleans, Louisiana, USA, April 11-14, 2010

[1]Sloan School of Management, MIT, Cambridge, MA. (jackswsc@q.com)
[2]Southwest Statistical Consulting, Cortez, CO.

Abstract

How to elicit expert judgment about probabilistic dependencies among distinct but geologically related petroleum units in a basin is a long-standing question. Assessments are often based on the assumption of probabilistic independence among assessment units while others assume total fractile dependency. In most cases, the most appropriate set of assumptions lies in between these extremes. However, capturing dependencies by specification of all pairwise correlations among a large set of assessment units is often difficult.

We describe a framework--hierarchical modeling--for probabilistic assessment of Alaskan North Slope gas hydrates that takes into account dependencies among assessment units without burdening experts with assessment overload and that guarantees logical coherence of assessments.

Website

USGS Circum-Arctic Resource Appraisal: Web accessed 28 June 2010, http://energy.usgs.gov/arctic/

Petroleum Assessment via Hierarchical Modeling

AAPG April 19, 2010

G.M. Kaufman, Massachusetts Institute of Technology &

J. Schuenemeyer, Southwest Statistical Consulting, LLC

Acknowledgements

- The authors appreciate the support of U.S. Minerals Management Service (MMS) and especially Matt Frye.
- We also appreciate the assistance of Tim Collett & Ron Charpentier, U.S. Geological Survey.
- We also wish to thank Ray Faith for much computational assistance.
- We anticipate that this methodology will be applied to an MMS assessment of gas hydrate resources in the Alaskan off-shore Federal waters.

Outline

- The Problem
- A Solution
- Application to Alaskan North Slope onshore Hydrates
- How Hierarchical Modeling works

The Problem

- How to aggregate probabilistically dependent assessment units in a logically coherent manner.
- Why is this a problem?
  - Because there are logical constraints on pairwise correlations within and among assessment units that must be obeyed.

A simple example

- The following set of pairwise correlations between assessment units (AUs) is **not a correlation matrix!**

	AU1	AU2	AU3
AU1	1	0.6	0.0
AU2	0.6	1	0.9
AU3	0.0	0.9	1
Numbers are correlations			

A more complex example: The USGS Circum Arctic Assessment

- Assessors specified pairwise correlations among 48 assessment units

- Often leads (and did lead) to an improper probability distribution for the joint distribution of in place oil/gas
  - Correlation matrix was not positive definite symmetric
  - Solution required successive iterations and statistical based modification of the correlation matrix

Circum Arctic

AU codes	AU name	LM1	LM3	EB1	EB3	TPB2	TPB3	EBB1	EBB2	EBB3
LM1	Makarov Basin Margin	1.00								
LM3	Siberian Passive Margin	0.70	1.00							
EB1	Lena Prodelta	0.20	0.27	1.00						
EB3	Nansen Basin Margin	0.20	0.20	0.30	1.00					
TPB2	Main Basin Platform	0.20	0.20	0.20	0.20	1.00				
TPB3	Foredeep Basins	0.20	0.20	0.20	0.20	0.80	1.00			
EBB1	Kolguyev Terrace	0.20	0.20	0.20	0.20	0.80	0.80	1.00		
EBB2	South Barents Basin and Ludlov Saddle	0.20	0.20	0.20	0.20	0.60	0.60	0.90	1.00	
EBB3	North Barents Basin	0.20	0.20	0.20	0.20	0.50	0.50	0.80	0.90	1.00

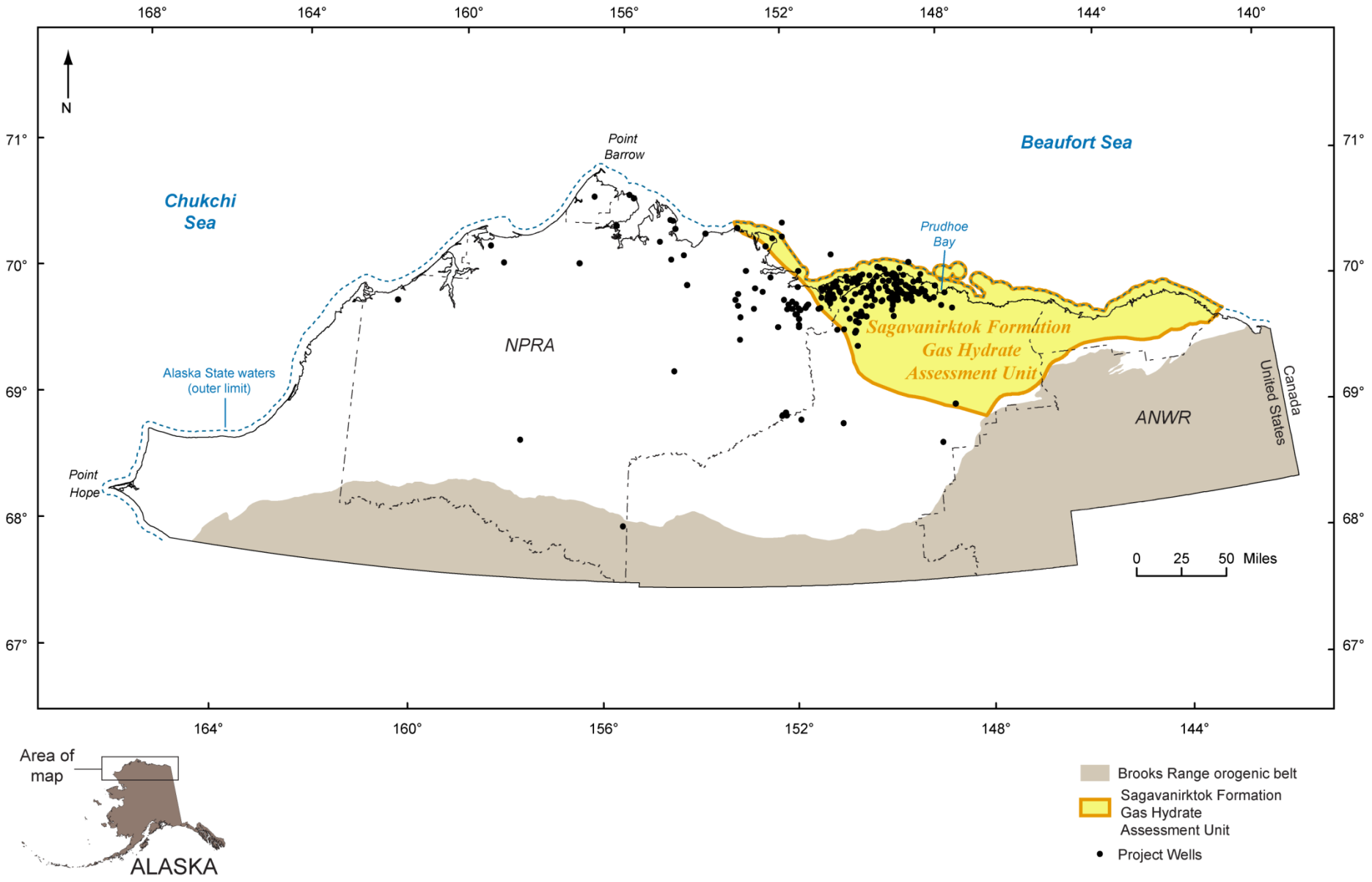
Hierarchical modeling guarantees a logically consistent correlation matrix

- Hierarchical modeling is a device for encoding subjective probability judgments about:
  - Size distributions of undiscovered accumulations
  - Numbers of accumulations
  - Correlations
    - Between accumulations within an assessment unit
    - Between assessment units

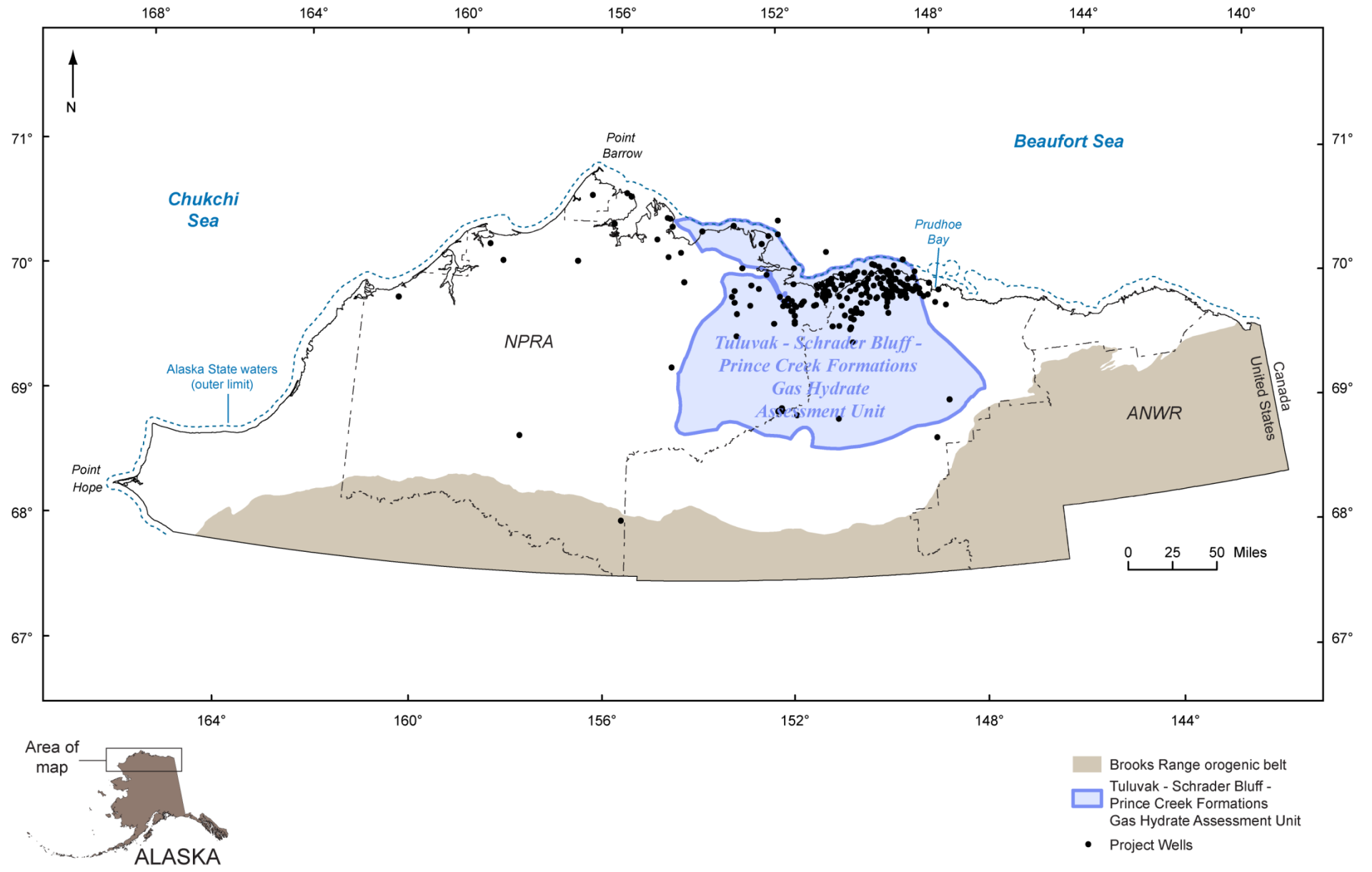
USGS Northern Alaska Probabilistic assessment

- Gas Hydrate Regions (maps and summary statistics courtesy Tim Collett, USGS)
  - Sagavanirktok Formation
  - Tuluvak - Schrader Bluff – Prince Creek Formations
  - Nanushuk Formation

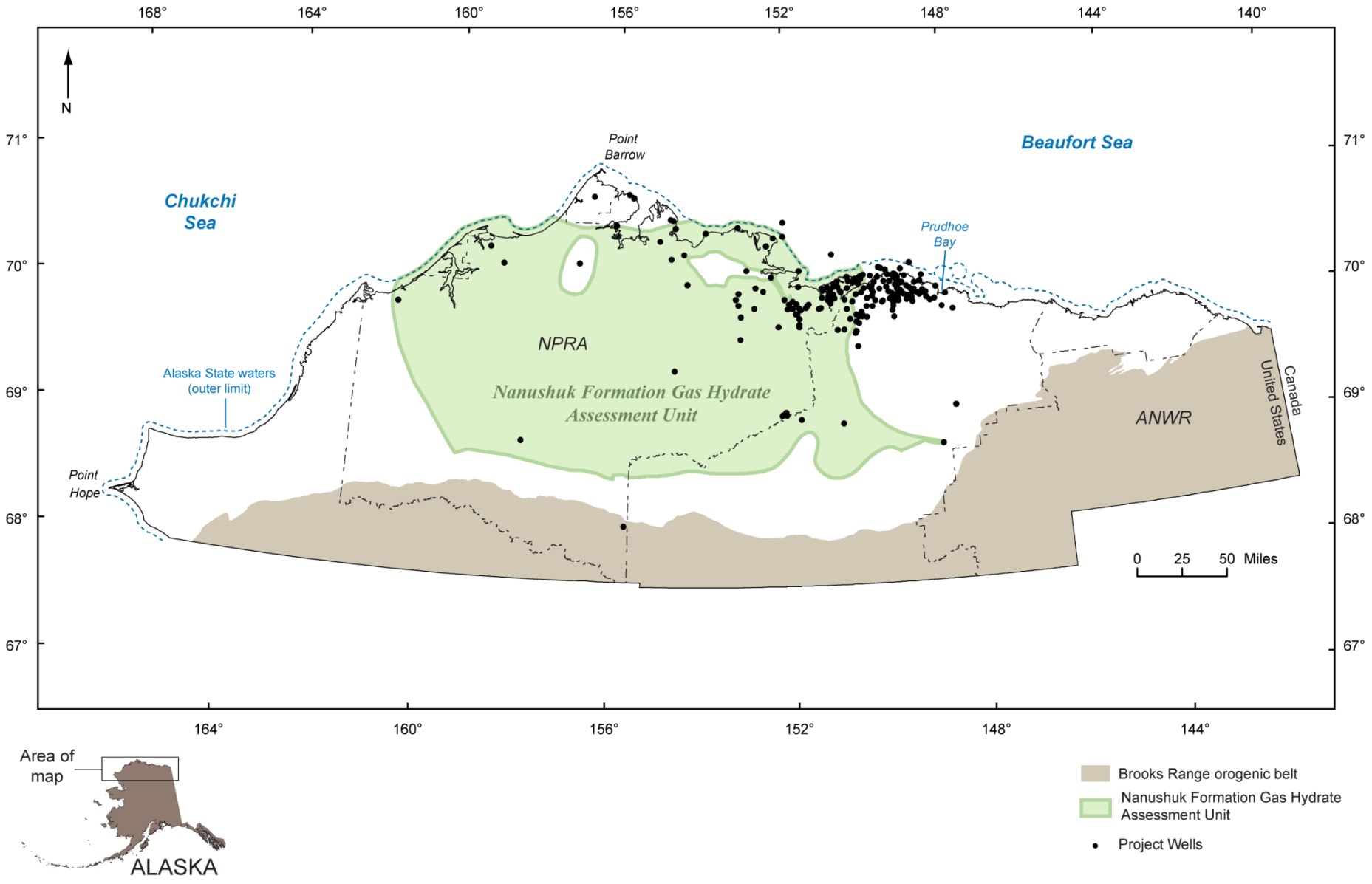
Sagavanirktok Formation Gas Hydrate AU



Tuluvak - Schrader Bluff – Prince Creek Formations Gas Hydrate AU



Nanushuk Formation Gas Hydrate AU



USGS Alaska Hydrate Assessment--Hierarchical Model Comparison

- Three AUs
- Within AU pairwise correlations

$$r=r_1=r_2=r_3$$

- Between AUs correlations

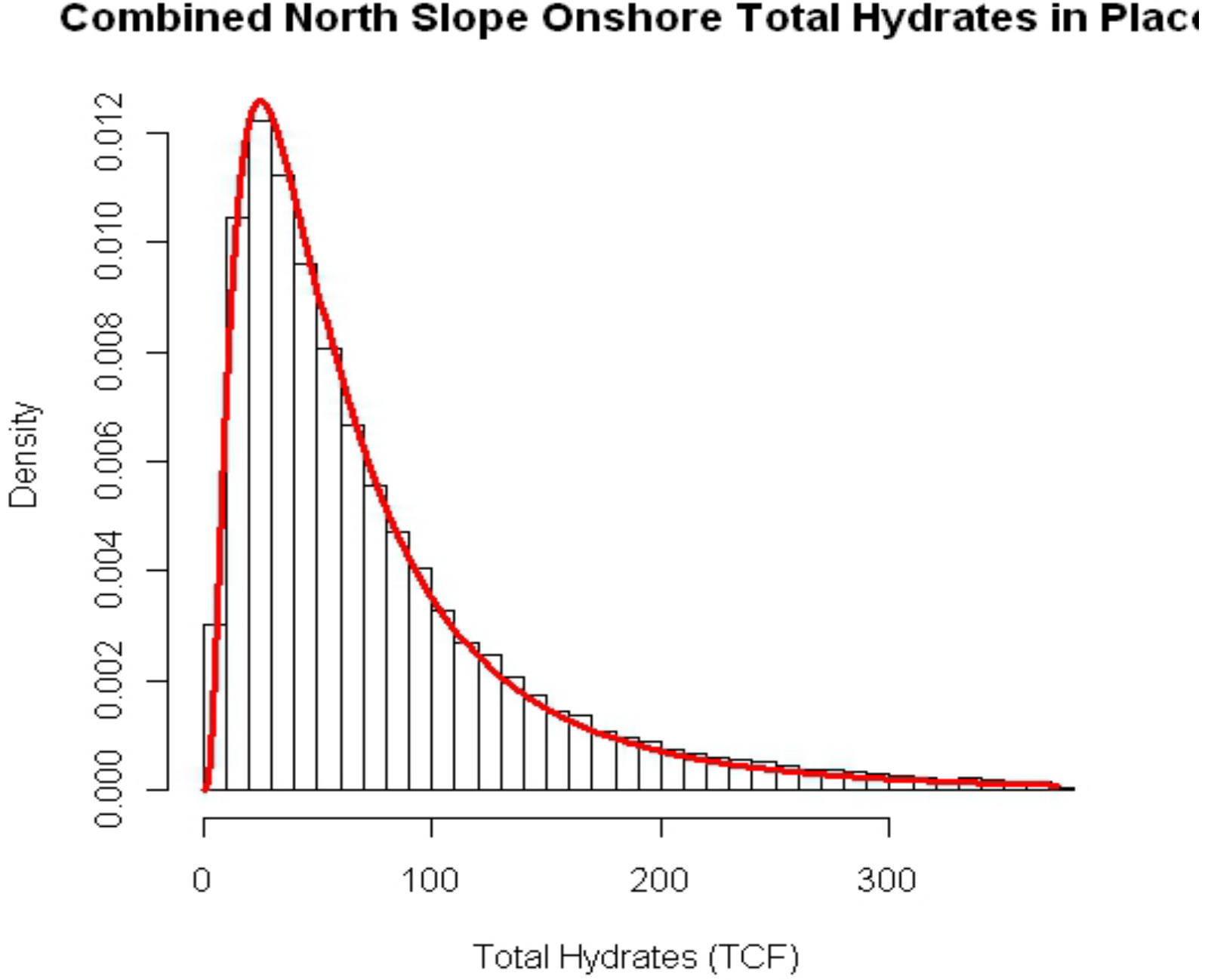
$$\text{rho}=\rho_{12}=\rho_{13}=\rho_{23}$$

Comparison, Alaska Hydrates

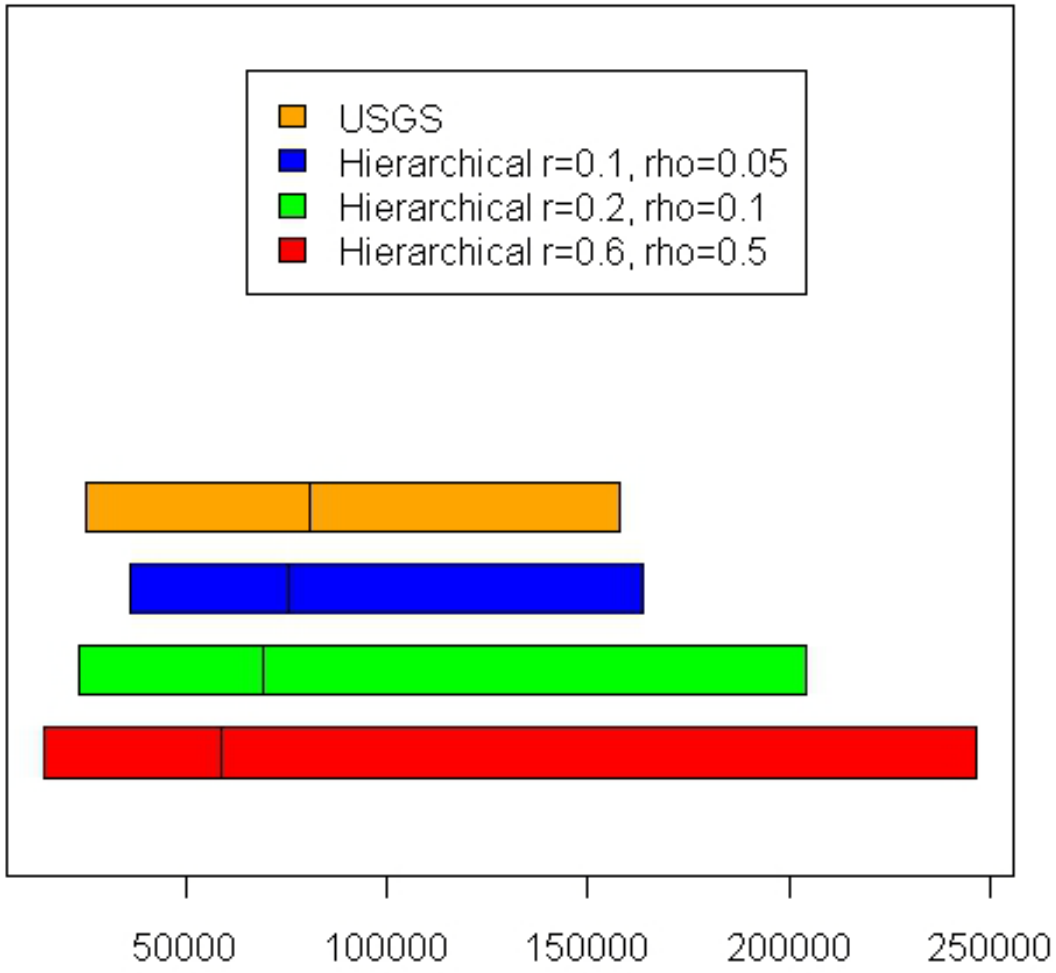
Within AU correlation $r_1=r_2=r_3=0.9$

Between AUs correlation $\rho_{12}= \rho_{13}= \rho_{23} =0.6$

	Total Gas (BCF), 3 AUs			
	Fractile			
	F95	F50	F5	Mean
USGS Probabilistic	25,233	81,030	157,831	85,427
Hierarchical	14,895	59,058	246,662	85,471


Combined North Slope Onshore Total Hydrates in Place
Density
0.000
0.002
0.004
0.006
0.008
0.010
0.012
0
100
200
300
Total Hydrates (TCF)

USGS and Hierarchical Model Comparison



Hierarchical Model Consists of 3 levels

Level I Data Generating Process: Regression Equation for k th Assessment Unit

$$Y_{jk} = \ln(X_{jk}) = \boldsymbol{\beta}_k \mathbf{z}_k + \varepsilon_{jk},\ j = 1, \ldots, n_k$$

- The $\mathbf{z}_k$ s are a combination of explanatory geologic variables and variables that control means and variances of accumulation sizes
- The $\boldsymbol{\beta}_k$ are parameters
  - The $\boldsymbol{\beta}_k$ *s are uncertain quantities*

Level II: Super-population Process For Regression Parameters of AU_k

- Parameters $\boldsymbol{\beta}_k$ are Multivariate Normal with mean **b** and variance matrix

Level III: Prior distribution of Super-population Parameter b

- Super-population mean **b** is an uncertain quantity
- **b** is Multivariate Normal with mean **b'** and variance matrix $\mathbf{V}_b$

Generating Accumulation Sizes

- Simulate accumulation counts $N_1,\ldots,N_K$

- Successively Monte Carlo
  - A value of **b**
  - Regression parameters $\boldsymbol{\beta}_k \quad k = 1,..,K$ given **b**
  - Accumulation sizes in each unit given $\boldsymbol{\beta}_k s$

Association Metrics

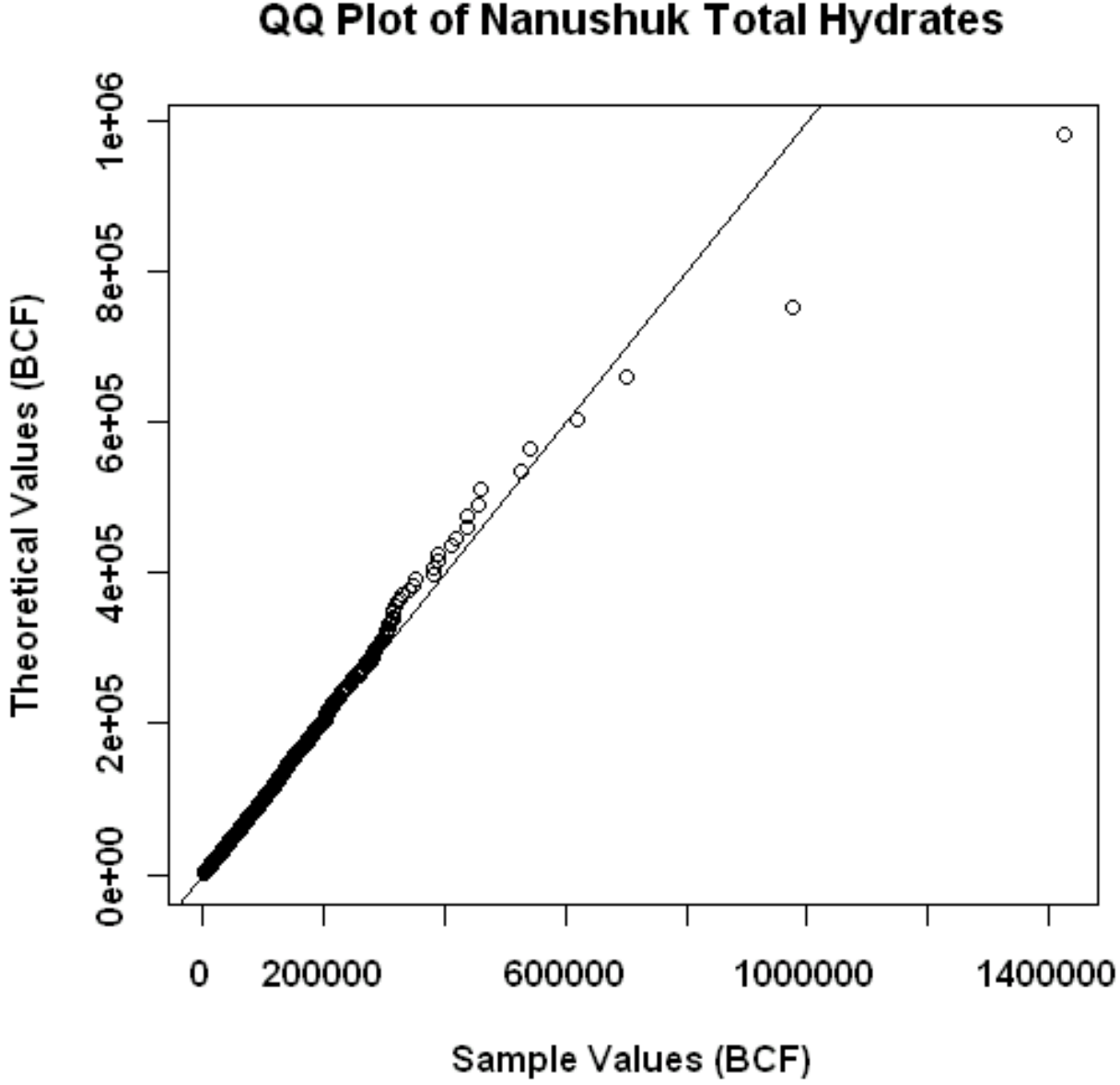
- **Conditional Dependence**
  - Suppose AU 1 median accumulation size turns out to be twice that originally assessed
  - How does this fact influence judgments about AU 2 accumulation sizes?

A Surprise!

- TOTAL Hydrates in a unit is a sum of a random number of Lognormal accumulation sizes

- A Lognormal distribution is an excellent fit to the empirical relative frequency distribution of TOTAL HYDRATES in a unit

- Also true for Alaska North Slope TOTAL HYDRATES


QQ Plot of Nanushuk Total Hydrates
Theoretical Values (BCF)
0e+00
2e+05
4e+05
6e+05
8e+05
1e+06
0
200000
600000
1000000
1400000
Sample Values (BCF)

In Summary, Hierarchical Modeling

- Incorporates probabilistic dependencies among and within assessment units in a transparent fashion
- Dovetails with an assessment protocol designed to elicit judgments about probabilistic dependencies
- Produces reasonable predictive probability distributions for total in place hydrates

How Hierarchical Modeling Works

- **Level I**: Given all model parameters
  - Accumulation sizes in an assessment unit obey a regression like data process generating model
- **Level II**: Level I Regression parameters are uncertain
  - Assigned a probability distribution dependent on a (super-population) parameter
- **Level III**: Level II Super-population parameter is uncertain
  - Assigned a probability distribution

USGS & Hierarchical Model Simulation

Within AU$_k$

- A random number n_k of accumulations of random size X_k
- For the k^{th} unit, Monte Carlo simulation yields:

$$\mathbf{X}_k = (X_{k1}, \ldots, X_{kn_k})$$

- For a given trial total hydrates in unit k is

$$S_k = \sum_{j=1}^{n_k} X_{kj}$$

Over all *K* AUs

- For a given trial total hydrates in all K units is

$$S = S_1 + \ldots + S_K$$

Note: Correlation may be present both *within* and *between* AUs

Within Unit Correlations

- Residual error variance and total accumulation size variance determine within unit pair-wise correlation

$$\text{Within Unit Correlation } r_k = 1 - \frac{v_k}{V_{kk}}$$

- Regression model error variance $v_k = Var(\varepsilon_{jk})$

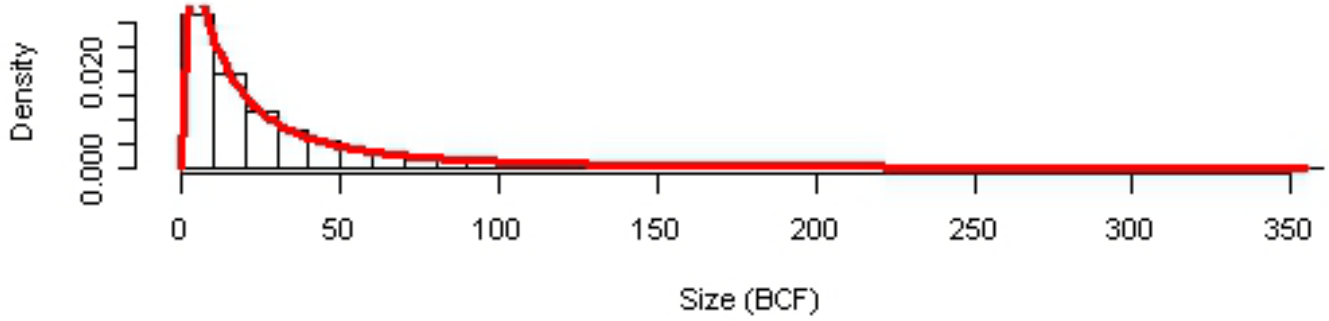
- V_{kk} = unconditional variance of the log of a generic accumulation size in unit *k*.

Between Assessment Unit Correlations

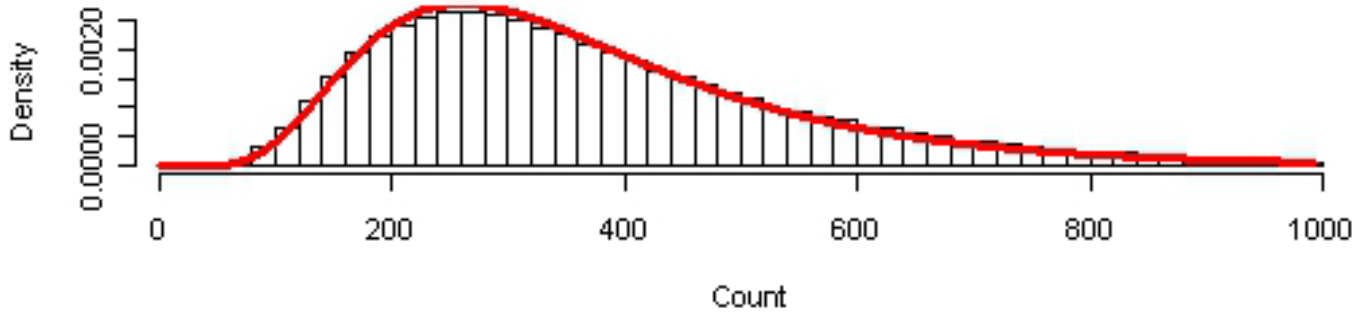
Pair-wise correlation between Units *k* and *m* is less than or equal to the geometric mean of Within Unit correlations

$$\rho_{km} \leq \sqrt{r_k r_m}$$

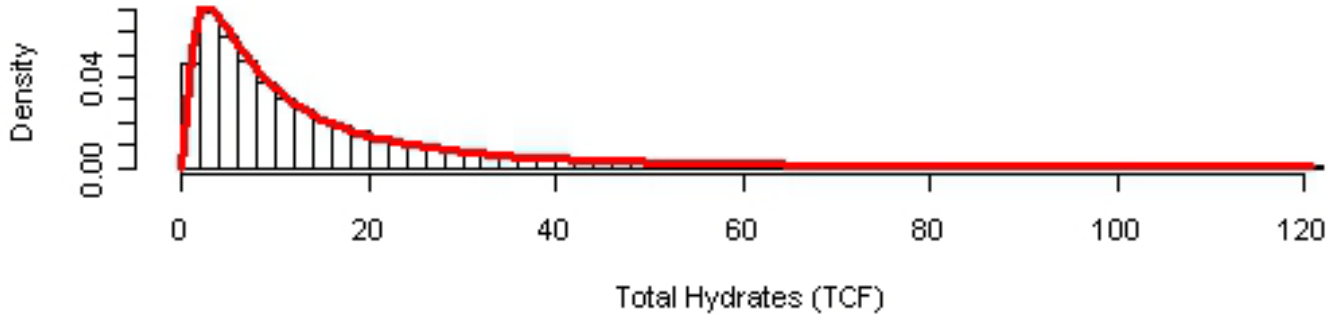
Sagavanirkitok Size

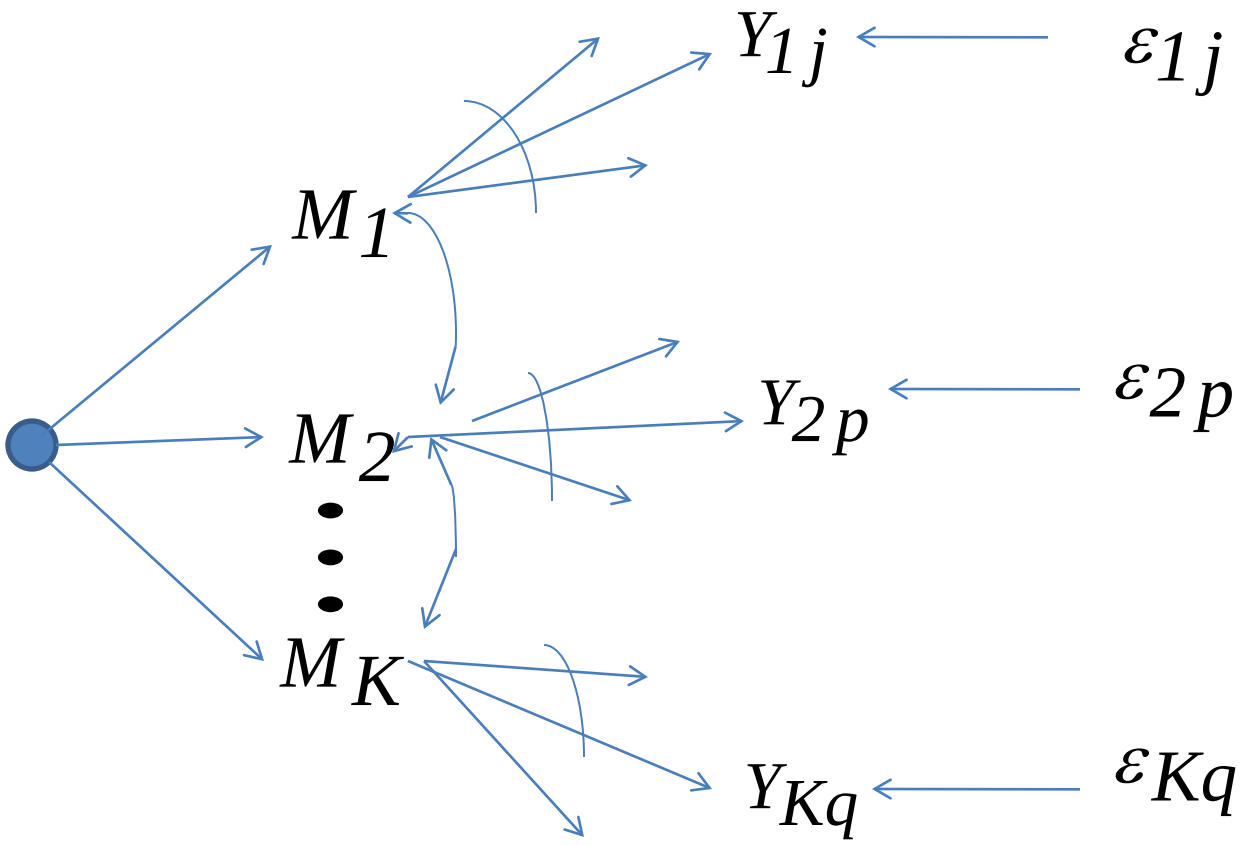


Sagavanirkitok Count



Sagavanirkitok Total Hydrates




Y_{1j}
ε_{1j}
M_1
M_2
Y_{2p}
ε_{2p}
M_K
Y_{Kq}
ε_{Kq}

$$M_k \equiv \boldsymbol{\beta}_k \mathbf{z}_k$$

$$\mathbf{M} = (M_1, \ldots, M_K)^t$$

$$\mathbf{M} \sim MVN(\overline{\mathbf{M}}, Var(\mathbf{M}))$$

$$\overline{\mathbf{M}} = \mathbf{b}' \mathbf{Z}$$

$$Var(\mathbf{M}) = V_\beta \times \mathbf{Z}^* \mathbf{Z}^{*t} + V' \times \mathbf{Z}\mathbf{Z}^t$$

$$\varepsilon_{kj} \sim iid\ N(0, v_k)$$

USGS Northern Alaska Probabilistic assessment

Alaska North Slope Gas Hydrate Assessment Results				
		Gas (BCF)		
		Fractiles		
Petroleum System	**F95**	**F50**	**F5**	**Mean**
Sagavanirktok Formation	6,285	19,400	37,791	**20,567**
Tuluvak Schrader Bluff-Prince Creek Formation	8,173	26,532	51,814	**28,003**
Nanushuk Formation	10,775	35,008	68,226	**36,857**
Total Undiscovered Resources	**25,233**	**81,030**	**157,831**	**85,427**

Thanks for your attention

- Questions?