OUTLINE

- CPARM Overview
- Value of Information
- VOI in Concept Selection
- Improvements in Capacitance-Resistive Modeling and Optimization of Hydrocarbon Reservoirs
- Reservoir Classification
- Real Options
CPARM Overview
CPARM Objectives

• To perform research into ways that decisions regarding hydrocarbon exploration and production can be improved

• Improve profitability of E&P operations by research in
  – Methods
  – Processes
  – Culture
  – Tools
DRA Through the Life of a Field

- We repeat the Bayesian update every time new information is acquired and plot the NPV distributions together.
- The figure represents a new type of DRA display that can easily communicate how uncertainty evolves through the life of a field.
The CPARM Team

CPARM Faculty
Larry W. Lake, PGE
Bob Gilbert, CE
Leon Lasdon, IROM
Sanjay Srinivasan, PGE
Jim Dyer, IROM
Chris Jablonowski, EER/PGE
Tom Edgar, ChE
Eric Bickel, OR(ME)/PGE
Emilio Núñez, CPARM

CPARM Students
Tianyang Wang, IROM
Wei Chen, IROM
Namhong Min, CE*
Panos Andonyadis, CE
Morteza Sayarpour, PGE*
Aviral Sharma, PGE*
Yonghoon Lee, CE*
Dan Weber, ChE
Hariharan Ramachandran, PGE
Suryansh Purwar*, PGE
Amin Ettehadtavakkol, PGE
Babafemi Ogunyomi, PGE
Ankesh Anupam, PGE

* Graduated
CPARM Publications – 2007/2008

**Theses**
- Robert MacAskie – “The Value of Oil Price Forecasts”
- Azeez Lawal – “Applications of Sensitivity Analysis in Petroleum Engineering”
- Aviral Sharma – “Classification of Hydrocarbon Recovery Factor Based on Reservoir Databases”

**Dissertation**
- Min Chen – “Inevitable Disappointment and Decision Making Based on Forecasts”
- Namhong Min – “A Method to Establish Non-Informative Prior Probabilities for Risk-Based Decision Analysis”

**Publications**
- Jablonowski and MacAskie – The Value of Oil Price Forecasts (SPE 107570)
- Hultzsch, Lake and Gilbert – Decision and Risk Analysis through the Life of the Field (SPE 107704)
- Chen and Dyer – Inevitable Disappointment in Projects Selected Based on Forecasts (SPE 107710)
- Lasdon, Faya, Lake, Dyer and Chen – Constructing Oil Exploration and Development Project Portfolios Using Several Risk Measures – A Realistic Example (SPE 107708)
- Faya, Lake and Lasdon – Beyond Portfolio Optimization (SPE 107709)
- Liang, Weber, Edgar, Lake, Sayarpour and Al-Yousef – Optimization of Oil Production Based on a Capacitance Model of production and Injection Rates (SPE 107713)
- Hahn, Dyer and Brandao - Using Decision Analysis to Solve Real Option Valuation Problems: Building a Generalized Approach (SPE 108066)
- Gilbert, Lake, Jablonowski, Jennings and Nunez – A Procedure for Assessing the Value of Oilfield Sensors (SPE 109628)
- Sayarpour, Zuluaga, Kabir and Lake – The Use of Capacitance-Resistive Models for Rapid Estimation of Waterflood Performance and Optimization (SPE 110081)
- Jablonowski, Wiboonkij-Arphakul and Neuhold – Estimating the Cost of Errors in Estimates Used During Concept Selection (SPE 110191)
Addressing Risks Throughout the E&P Asset Lifecycle

Appraisal and Conceptual Analysis
Evaluate Alternatives
Define Selected Alternative
Execute
Operate

Portfolio Management and Project Selection

Concept Selection & Development Optimization

Cost and Schedule Estimating; Execution Risk Management

HSE Risk Management

Real-Time Optimization and Risk Management

Inevitable Disappointment
VOI; Impact of Estimates & Methods
Contracting Strategies (lump sum v cost plus?)
MPD & Blowouts; Drilling Safety; Offshore Spills
Life Cycle Assessments
Likelihood Functions from Data Analysis
Real Time Optimization Under Uncertainty

Portfolio Optimization
Uncertainty Updating
Real Options
Dry Gas Model; Compare MC & Decision Trees
UT Tank Model Enhance & Sensitivity Analysis
Cost Modeling of Wells and Facilities

Tool Development:
Value of Information in Exploration and Production
Probability Assessment - Non-Informative Prior Probabilities
Namhong Min, Ph.D.
Supervised by Profs. Robert B. Gilbert and Larry W. Lake

The goal of this research effort was to develop a practical and rational starting point for assessing probabilities. In Bayesian decision analysis, uncertainty and risk are accounted for with probabilities for the possible states, or states of nature, that affect the outcome of a decision. Application of Bayes’ theorem requires non-informative prior probabilities, which represent the probabilities of states of nature for a decision maker under complete ignorance. These prior probabilities are then subsequently updated with any and all available information in assessing probabilities for making decisions. The conventional approach for the non-informative probability distribution is based on Bernoulli’s principle of insufficient reason. This principle assigns a uniform distribution to uncertain states when a decision maker has no information about the states of nature. The principle of insufficient reason has three difficulties: it may inadvertently provide a biased starting point for decision making, it does not provide a consistent set of probabilities, and it violates reasonable axioms of decision theory.

The first objective of this study was to propose and describe a new method to establish non-informative prior probabilities for decision making under uncertainty. The proposed decision-based method focuses on decision outcomes that include preference in decision alternatives and consequences. Specifically, the initial or prior sample space for states of nature is constructed in terms of the possible decision preferences and consequences, and these possibilities are assigned equal probability.

The second objective was to evaluate the logic and rationality basis of the proposed decision-based method. The decision-based method overcomes the three weaknesses associated with the principle of insufficient reason, and provides an unbiased starting point for decision making. It also produces consistent non-informative probabilities. Finally, the decision-based method satisfies axioms of decision theory that characterize the case of no information (or complete ignorance).

The third and final objective was to demonstrate the application of the decision-based method to practical decision making problems in oil and gas production. The examples were based on case histories and touched on topics including production system design, reservoir heterogeneity, and spatial variability in reservoir permeability. Four practical implications are illustrated and discussed with these examples. First, the method is practical because it is feasible in decisions with a large number of decision alternatives and states of nature and it is applicable to both continuous and discrete random variables of finite and infinite ranges. Second, the method provides an objective way to establish non-informative prior probabilities that capture a highly non-linear relationship between states of nature. Third, the method can include any available information through Bayes’ theorem by updating the non-informative probabilities without the need to assume more than is actually contained in the information. Lastly, two different decision making problems with the same states of nature may have different starting point or non-informative probabilities.
Value of Information in E&P

Namhong Min, Bob Gilbert, and Larry Lake

November 2008
Background

Technical Uncertainty

Exploration Data
Spatial Variability
Upscaling
Reservoir Simulation

Value of Information

Economic Decision Making

Development Options

Net Present Value
Volatility
Real Option Valuation
Goals

1. Develop practical methods for value of information analyses
2. Integrate technical models with economic decision making
3. Identify where best to expend research resources in this area
Outline

1. Introduction
2. Theoretical Development
3. Practical Implementation
4. Summary
1. Introduction
Decision Tree

Phase I Development

100 acre Well Spacing

Start Phase 2 (Infill Wells)
in First Year

10th Year

Never
Decision Criteria

Reservoir Scenario → Reservoir Performance

- Expected Production Rate, $q_{osci}$
- Decay Constant, $\lambda$
- Limiting Production Rate, $q_{LIM}$

Expected Economic Performance

- Development Cost
- Operating Cost
- Oil Value
- Discount Rate
- Economic Limit

Discounted Cash Flow Analysis

Production Rate (STB/day)

Time (year)

Benefit
Total Cost
Net Profit

Present Value ($ MM)

Time (year)
Modeling Uncertainty
Average Reservoir Properties
Modeling Uncertainty

Well Variability
Initial Decision

Start Phase 2 (Infill Wells) in First Year

10th Year

Never

Expected Net Profit (NPV/100 acres)

$ - 892,133

$ - 5,561

$ 152,138
Value of Information

Likelihood Function

Production Rate

Time

Porosity (%)

LN(k)

PMF

0

0.05

0.1

0.15

0.2

0.0

-0.1

-0.2

-0.3

LN(k)

Porosity (%)

PMF
Value of Information

Oil Value ($/STB)

VI ($ NPV/100 acres)

Prior Decision

No Go

2nd Well at Year 2

Options somewhere between No Go & Year 2
2. Theoretical Development

- Evaluate the Principle of Insufficient Reason
- Propose a New Method
Bayes’ Theorem
The Basis for VOI Analyses

“Updated Probability” (what we want)

\[ P\left( \text{State of Nature } i \mid \text{Information} \right) = \]
\[ \frac{P\left( \text{Information} \mid \text{State of Nature } i \right) P\left( \text{State of Nature } i \right)}{\sum_{\text{all } i} \left[ P\left( \text{Information} \mid \text{State of Nature } i \right) P\left( \text{State of Nature } i \right) \right]} \]

“Likelihood Function” (what does the available information say)

“Prior Probability” (what do we know with no information)
Bayes’ Theorem
The Basis for VOI Analyses

“Prior Probability”

<table>
<thead>
<tr>
<th>State</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

“Updated Probability”

<table>
<thead>
<tr>
<th>State</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>0.4</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

“Likelihood Function”

<table>
<thead>
<tr>
<th>State</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Basis for Probabilities

\[ P(\text{State } j) = P(\text{State } j | \text{Set of All Possibilities}) \]

- Sample Space for Non-Informative Prior Should Include Set of All Possibilities
- It is Difficult to Define a Sample Space
Objective: Prior probabilities should be unbiased (that is, include “no” information). If they are unintentionally biased, then everything else that goes into the decision is also biased.
Conventional Approach
Bernoulli’s Principle of Insufficient Reason

If a decision maker is completely ignorant as to which state of nature will occur, then the decision maker should behave as if the states are random (equally likely).
The Principle of Insufficient Reason Is Pervasive

1. Used implicitly in all conventional statistical methods (such as maximum likelihood)

2. Used implicitly or explicitly in assigning probabilities in all formal decision analyses

3. Used explicitly in the Principle of Maximum Entropy for Information
Example
The Principle of Insufficient Reason for Priors

<table>
<thead>
<tr>
<th>State</th>
<th>Option A</th>
<th>Option B</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>State 1</td>
<td>State 1</td>
</tr>
<tr>
<td>2</td>
<td>State 2</td>
<td>State 2</td>
</tr>
<tr>
<td>3</td>
<td>State 3</td>
<td>State 3</td>
</tr>
</tbody>
</table>

Utility

| State 1 | 0 |
| State 2 | 1 |
| State 3 | 2 |
| State 1 | 3 |
| State 2 | 2 |
| State 3 | 0 |
Example
Conventional Approach for Priors

<table>
<thead>
<tr>
<th>State</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>State 1</td>
<td>1/3</td>
</tr>
<tr>
<td>State 2</td>
<td>1/3</td>
</tr>
<tr>
<td>State 3</td>
<td>1/3</td>
</tr>
</tbody>
</table>

Utility

<table>
<thead>
<tr>
<th>State</th>
<th>Utility</th>
</tr>
</thead>
<tbody>
<tr>
<td>State 1</td>
<td>0</td>
</tr>
<tr>
<td>State 2</td>
<td>1</td>
</tr>
<tr>
<td>State 3</td>
<td>2</td>
</tr>
</tbody>
</table>

Option A
Option B
Example
Conventional Approach for Priors

We Unintentionally Insert Bias in Decision.
CPARM Conclusion

The Principle of Insufficient Reason for non-informative priors is wrong.
Difficulties with Applying the Principle of Insufficient Reason

- It is Not Possible to Apply it Consistently
- It Does Not Produce Rational Results
- It is Not Theoretically Sound
Intuitive Support for CPARM

Conclusion

1. It is not possible to apply the principle of insufficient reason consistently.
Inconsistency in the Principle of Insufficient Reason

Which distribution is non-informative?

OR

Permeability

Log of Permeability
Inconsistency in the Principle of Insufficient Reason
To which input parameters do we apply it?

\[ k_{\text{eff}} = \frac{1}{\left(\frac{1}{k_1} + \frac{1}{k_2}\right)} \]
Inconsistency in the Principle of Insufficient Reason
To which input parameters do we apply it?

Flow

\[ \frac{1}{k_1} + \frac{1}{k_2} \]

\[ k_{\text{eff}} = \frac{2}{\left( \frac{1}{k_1} + \frac{1}{k_2} \right)} \]

OR

Probability

\[ \frac{1}{3}, \frac{1}{3}, \frac{1}{3} \]

\[ 1, 1.33, 2 \]

\[ k_{\text{eff}} (\text{md}) \]

\[ \frac{1}{3}, \frac{1}{2}, \frac{1}{4} \]

\[ 1, 1.33, 2 \]

\[ k_{\text{eff}} (\text{md}) \]
Practical Support for CPARM

Conclusion

Applying Principle of Insufficient Reason to input does necessarily not produce “insufficient reason” in the output.
### Journal and Deutsch Example

<table>
<thead>
<tr>
<th>Input Permeability Field</th>
<th>Cumulative Production</th>
</tr>
</thead>
<tbody>
<tr>
<td>High Entropy</td>
<td>Low Entropy</td>
</tr>
<tr>
<td>Low Entropy</td>
<td>High Entropy</td>
</tr>
</tbody>
</table>
Journal and Deutsch Example

Gaussian Model

Mosaic Model

Indicator RF Model

Fraction = 0.45

Fraction = 0.78

Fraction = 1.3
Theoretical Support for CPARM Conclusion

Decision theorists (such as Luce and Raiffa) show that Bernoulli’s Principle is not consistent with the fundamental axioms of decision theory.
Theoretical Difficulty with the Principle of Insufficient Reason
Luce and Raiffa Example

<table>
<thead>
<tr>
<th>Decision Alternatives</th>
<th>States of Nature</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$S_1$</td>
</tr>
<tr>
<td>$A_1$</td>
<td>11</td>
</tr>
<tr>
<td>$A_2$</td>
<td>0</td>
</tr>
</tbody>
</table>

Consequences (Utilities)
Theoretical Difficulty with the Principle of Insufficient Reason
Luce and Raiffa Example

<table>
<thead>
<tr>
<th>Decision Alternatives</th>
<th>States of Nature</th>
<th>Expected Utility</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>S₁</td>
<td>S₂</td>
</tr>
<tr>
<td>A₁</td>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td>A₂</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>Probabilities</td>
<td>1/2</td>
<td>1/2</td>
</tr>
</tbody>
</table>
Theoretical Difficulty with the Principle of Insufficient Reason
Luce and Raiffa Example

<table>
<thead>
<tr>
<th></th>
<th>S₁</th>
<th>S₂</th>
<th>Expected Utility</th>
</tr>
</thead>
<tbody>
<tr>
<td>A₁</td>
<td>11</td>
<td>0</td>
<td>5.5</td>
</tr>
<tr>
<td>A₂</td>
<td>0</td>
<td>10</td>
<td>5.0</td>
</tr>
<tr>
<td>Probabilities</td>
<td>1/2</td>
<td>1/2</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>S₁</th>
<th>S₂a</th>
<th>S₂b</th>
<th>Expected Utility</th>
</tr>
</thead>
<tbody>
<tr>
<td>A₁</td>
<td>11</td>
<td>0</td>
<td>0</td>
<td>3.7</td>
</tr>
<tr>
<td>A₂</td>
<td>0</td>
<td>10</td>
<td>10</td>
<td>6.7</td>
</tr>
<tr>
<td>Probabilities</td>
<td>1/3</td>
<td>1/3</td>
<td>1/3</td>
<td></td>
</tr>
</tbody>
</table>
Theoretical Difficulty with the Principle of Insufficient Reason

Luce and Raiffa Example

<table>
<thead>
<tr>
<th></th>
<th>S₁</th>
<th>S₂</th>
<th>Expected Utility</th>
</tr>
</thead>
<tbody>
<tr>
<td>A₁</td>
<td>11</td>
<td>0</td>
<td>5.5</td>
</tr>
<tr>
<td>A₂</td>
<td>0</td>
<td>10</td>
<td>5.0</td>
</tr>
</tbody>
</table>

Adding Duplicate States of Nature Should Not Change Decision

<table>
<thead>
<tr>
<th></th>
<th>S₁</th>
<th>S₂a</th>
<th>S₂b</th>
<th>Expected Utility</th>
</tr>
</thead>
<tbody>
<tr>
<td>A₁</td>
<td>11</td>
<td>0</td>
<td>0</td>
<td>3.7</td>
</tr>
<tr>
<td>A₂</td>
<td>0</td>
<td>10</td>
<td>10</td>
<td>6.7</td>
</tr>
</tbody>
</table>

Labeling States of Nature Should not Change Decision

Probabilities

<table>
<thead>
<tr>
<th></th>
<th>S₁</th>
<th>S₂</th>
<th>Probabilities</th>
</tr>
</thead>
<tbody>
<tr>
<td>A₁</td>
<td>1/2</td>
<td>1/2</td>
<td></td>
</tr>
<tr>
<td>A₂</td>
<td>1/3</td>
<td>1/3</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>S₁</th>
<th>S₂a</th>
<th>S₂b</th>
<th>Probabilities</th>
</tr>
</thead>
<tbody>
<tr>
<td>A₁</td>
<td>1/3</td>
<td>1/3</td>
<td>1/3</td>
<td></td>
</tr>
</tbody>
</table>
CPARM Approach

Principle of Insufficient Reason: If a decision maker is completely ignorant as to which state of nature will occur, then the decision maker should behave as if the states are random (equally likely).

CPARM: If a decision maker is completely ignorant as to which state of nature will occur, then the decision maker should behave as if the preferred decision alternative is random.
Implementation of CPARM Approach

• Decision Outcomes
  1. Preference in Decision Alternatives
  2. Decision Consequence

<table>
<thead>
<tr>
<th></th>
<th>S_1</th>
<th>S_2</th>
<th>S_3</th>
<th>S_4</th>
<th>S_5</th>
</tr>
</thead>
<tbody>
<tr>
<td>A_1</td>
<td>11</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>A_2</td>
<td>0</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>9</td>
</tr>
</tbody>
</table>

A_1 > A_2  A_1 ~ A_2  A_1 < A_2

Decision Outcomes Provide a **Mutually Exclusive and Collectively Exhaustive** Sample Space.
Implementation of CPARM Approach

<table>
<thead>
<tr>
<th></th>
<th>S₁</th>
<th>S₂</th>
<th>S₃</th>
<th>S₄</th>
<th>S₅</th>
</tr>
</thead>
<tbody>
<tr>
<td>A₁</td>
<td>11</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>A₂</td>
<td>0</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>9</td>
</tr>
</tbody>
</table>

- **A₁ > A₂**
- **A₁ ~ A₂**
- **A₁ < A₂**

Preference Outcome:

- A₁ > A₂
- A₁ ~ A₂
- A₁ < A₂

States of Nature:

- S₁
- S₂
- S₃
- S₄
- S₅
CPARM Approach is Based on an Unbiased Starting Point

<table>
<thead>
<tr>
<th>State</th>
<th>Option A</th>
<th>Option B</th>
</tr>
</thead>
<tbody>
<tr>
<td>State 1</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>State 2</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>State 3</td>
<td>2</td>
<td>0</td>
</tr>
</tbody>
</table>

Utility:

<table>
<thead>
<tr>
<th>State</th>
<th>Option A</th>
<th>Option B</th>
</tr>
</thead>
<tbody>
<tr>
<td>State 1</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>State 2</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>State 3</td>
<td>2</td>
<td>0</td>
</tr>
</tbody>
</table>

Probability:

- State 1: 1/2 for Option A, 1/2 for Option B
- State 2: 1/4 for State 1 and 1/4 for State 3, 1/2 for State 2
- State 3: 1/2 for State 3

Preferred Option:

- 1/2 for Option A, 1/2 for Option B
CPARM Approach Provides Consistent Results

Flow

\[ k_1 \quad k_2 \]

\[ k_{\text{eff}} = \frac{2}{1/k_1 + 1/k_2} \]

<table>
<thead>
<tr>
<th>([k_1,k_2] = [1,1])</th>
<th>([1,2])</th>
<th>([2,1])</th>
<th>([2,2])</th>
</tr>
</thead>
<tbody>
<tr>
<td>(k_{\text{eff}} = 1)</td>
<td>1.33</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>(A_1)</td>
<td>5</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>(A_2)</td>
<td>1</td>
<td>5</td>
<td>5</td>
</tr>
</tbody>
</table>

\[ \begin{align*}
\text{A1} & = 57_{\text{0}}^{10} \\
\text{A2} & = 15_{\text{0}}^{20}
\end{align*} \]
CPARM Approach Satisfies Axioms in Decision Theory

<table>
<thead>
<tr>
<th></th>
<th>S₁</th>
<th>S₂</th>
<th>Expected Utility</th>
</tr>
</thead>
<tbody>
<tr>
<td>A₁</td>
<td>11</td>
<td>0</td>
<td>5.5</td>
</tr>
<tr>
<td>A₂</td>
<td>0</td>
<td>10</td>
<td>5.0</td>
</tr>
</tbody>
</table>

Probabilities | 1/2 | 1/2 |

<table>
<thead>
<tr>
<th></th>
<th>S₁</th>
<th>S₂a</th>
<th>S₂b</th>
<th>Expected Utility</th>
</tr>
</thead>
<tbody>
<tr>
<td>A₁</td>
<td>11</td>
<td>0</td>
<td>0</td>
<td>5.5</td>
</tr>
<tr>
<td>A₂</td>
<td>0</td>
<td>10</td>
<td>10</td>
<td>5.0</td>
</tr>
</tbody>
</table>

Probabilities | 1/2 | 1/4 | 1/4 |
3. Practical Application
Practical Application

• Non-Informative Priors from CPARM Approach are Not Uniform

Transmissibility Example

<table>
<thead>
<tr>
<th>Year</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unit 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unit 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3</td>
</tr>
</tbody>
</table>

*Alternative 6819*
Practical Application

• Non-Informative Priors from CPARM Approach are Not Uniform

Transmissibility is Important

Holstein Oil Field
Practical Application

• Non-Informative Priors from CPARM Approach are Not Uniform

Transmissibility Example

Decision-Based Non-Informative Prior Probabilities

Transmissibility (RB/psi/day)
Practical Application

• Add Information Through Bayes’ Theorem

Transmissibility Example

Available Information

P(T* | T < 100) = 0.1
P(T* | T > 100) = 0.9
Practical Application

• Different Decision Frameworks May Provide Different Non-Informative Priors

Transmissibility Example

<table>
<thead>
<tr>
<th>Well Cost for Unit 1</th>
<th>Decision</th>
<th>VPI</th>
</tr>
</thead>
<tbody>
<tr>
<td>$3 MM</td>
<td>$1.8 MM</td>
<td>$6 MM</td>
</tr>
<tr>
<td>$6 MM</td>
<td>$3.1 MM</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Year</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unit 1</td>
<td>3 Wells</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unit 2</td>
<td>2 Wells</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Transmissibility (RB/psi/day)
Different Decision Frameworks May Provide Different Non-Informative Priors

Transmissibility Example

Practical Application

VPI ($ MM)

Well cost for Unit 1 ($ MM)
Practical Application

- **Non-Informative Prior Captures the Complex Relationship between Input Variables**

Spatial Variability Example

- Alternative 1: 1 Production Well
- Alternative 2: 2 Production Wells
- Alternative 3: Abandon
Practical Application

- Non-Informative Prior Captures the Complex Relationship between Input Variables

Spatial Variability Example
Practical Application

• CPARM Approach is Applicable to Practical Decision Making

Production Example

- Gathering Center
- Export Line
- Unit 1
- Unit 2

50625 States of Nature

Alternative 1
- Alternative 6819*
- Alternative 12241

<table>
<thead>
<tr>
<th>* Alternative 6819</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Year</strong></td>
</tr>
<tr>
<td><strong>Unit 1</strong></td>
</tr>
<tr>
<td><strong>Unit 2</strong></td>
</tr>
</tbody>
</table>
Practical Application

- CPARM Approach is Applicable to Continuous Random Variables

Alternative 1 > Alternative 2

Utility

Random variable, x

Alternative 1 < Alternative 2
Practical Application

• CPARM Approach is Applicable to Continuous Random Variables
4. Summary
Objective

To Propose a Method to Establish a Non-Informative Prior Probability Distribution for Decision Analysis

To Evaluate the hypothesis:

*Proposed Method for Decision-Based Prior is Rational, and Practical and it is an Improvement over the Principle of Insufficient Reason*
## Summary

<table>
<thead>
<tr>
<th></th>
<th>Rational</th>
<th>Consistent</th>
<th>Theoretically Sound</th>
</tr>
</thead>
<tbody>
<tr>
<td>Principle of Insufficient Reason</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>CPARM Approach</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

The CPARM Approach is Practical

- Applicable to a Large Number of Decision Alternatives and States of Nature
- Applicable to Both Discrete and Continuous Sample Spaces
- Straightforward and Deterministic
Summary

• Practical Application

1. CPARM Approach: Decision-Based Non-Informative Priors
   ✓ May Not be Uniform
   ✓ Captures Significance of Extremes and Complex Relationships between Input Variables

2. Additional Information about Input Variables Can be Included Through Bayes’ Theorem

3. For the Same Variable, the Decision-Based Non-Informative Prior Probability Distribution that Applies to One Decision May be Different than the Non-Informative Prior Probability Distribution that Applies to Another Decision


VOI in Concept Selection: Assessing the Relative Value of Accuracy in Subsurface, Wells, and Facilities Estimates

Chris Jablonowski

During the concept comparison and selection phase of exploration and production (E&P) capital projects, decision makers estimate the value of competing development concepts. These estimates are used to rank options and to select one option to carry forward to the next project phase. The importance of these estimates cannot be overstated; they determine which concept is selected, and have a strong influence on field architecture, initial capacity of facilities, well counts, production rates, and project schedule. Decisions in concept selection have a large impact on the value ultimately derived from the asset (Evans 2005; Walkup and Ligon 2006). These estimates are also used for other important analyses and decisions during concept selection such as value of information (VOI) analysis.

A variety of input variables are required to estimate the value of competing development concepts. These input variables include estimates for the subsurface (e.g., reserves, flow rates, decline rates), estimates for the surface facilities (e.g., CAPEX, OPEX, schedule, reliability), and estimates for exogenous factors such as commodity price. The true values of these input variables are almost always unknown, and estimates are developed based on the current information set available to the decision maker.

The objective of this research initiative is to examine and compare the loss in value incurred when concept selection decisions are based on erroneous estimates of input variables. Errors can occur in estimates of expected values and in estimates of variance. The conclusion that erroneous estimates of input variables can destroy project value is common sense. What this study attempts to provide are original estimates of the potential magnitude of such losses, and an analysis of which input variable estimates matter more than others.

In practice, one does not know if a current estimate for an input variable is erroneous, but one can estimate the impact of an alternate hypothesis being true, and this is the framework to be adopted here. A procedure for concept selection is defined to model the decision-making process and is used in conjunction with a simplified asset development optimization model to estimate project values. The analysis compares project values resulting from concept selection decisions based on erroneous estimates and decisions based on an alternate hypothesis; in both cases, the alternate hypothesis is taken to be true. The difference in value observed, if any, is caused by sub-optimal initial facility decisions (note that the difference in value can also be interpreted as the maximum willingness to pay to confirm the alternate hypothesis). The approach is similar in form to standard VOI analyses (Coopersmith et al. 2006; Bickel et al. 2006; Prange et al. 2006; Gilbert et al. 2007).
VOI in Concept Selection

“Assessing the Relative Value of Accuracy in Subsurface, Wells, and Facilities Estimates”

Chris Jablonowski

The Center for Petroleum Asset Risk Management (CPARM)
Asset Development Phases

- What is the most cost effective development concept?
- What technology should be used?
- What is the initial configuration and sizing of facilities?
- What pre-investments should be made to accommodate for potential changes in facilities?
- Which uncertainties drive initial decisions?
- Should any uncertainties be mitigated, and how much should be paid to do it?
Concept Selection Drives Value Creation

- The value derived from an asset is largely determined in *concept selection*, driven by:
  - The degree of project team integration
  - The scope of integrated risk assessment
  - The quality of quantitative analysis and optimization

- Shortfalls in these areas in concept selection also contribute to deviation between estimated and actual (value) for the concept selected
  - That is, plans may not be robust to revelations of uncertain variables
Research Objective

• This multi-project initiative develops workflows and supporting models to assess the relative value of accuracy in estimates used in concept selection.

• In summary, we examine the loss in project value caused by planning with inaccurate estimates (for the subsurface, wells, and facilities).
Progress Report

- Workflows have been developed that can help the asset manager and development team with:
  - Concept selection (general)
  - Initial configuration and sizing of facilities

- The workflows address:
  - Reservoir uncertainty
  - Reservoir engineering deliverables
  - Facility decisions
  - VOI computations

- Support tools have been developed to facilitate implementation of the workflows
Workflow for Reservoir Engineering Deliverables

- Define uncertain reservoir variables
- Use Design of Experiments (DoE) to specify $n$ cases
- Define initial reservoir conditions AND Initial facility configuration ($k$ cases)
- Use reservoir simulator to compute cumulative recovery for each case $n$ (at times of interest—e.g. real option points)
- Use regression analysis on $n$ observations to estimate response surfaces for cumulative recovery as a function of uncertain variables (at times of interest)
- Use regression analysis on $n$ observations to estimate a response surface for Auxiliary variables that will be used to transfer reservoir condition(s) from initial to new regime (at times of interest)

All $k$ Cases Run?

START

no

yes

STOP
### Example: DoE and Response Surfaces

Comparison of cumulative oil production for different parameters (Case 1)

<table>
<thead>
<tr>
<th>Case</th>
<th>$K'_{ro}$</th>
<th>$S_{or}$</th>
<th>$K'_{rg}$</th>
<th>$S_{w}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>0.1</td>
<td>0.9</td>
<td>0.325</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>0.2</td>
<td>0.9</td>
<td>0.325</td>
</tr>
<tr>
<td>3</td>
<td>0.8</td>
<td>0.2</td>
<td>0.9</td>
<td>0.325</td>
</tr>
<tr>
<td>4</td>
<td>0.8</td>
<td>0.1</td>
<td>0.9</td>
<td>0.325</td>
</tr>
<tr>
<td>5</td>
<td>0.9</td>
<td>0.2</td>
<td>0.8</td>
<td>0.325</td>
</tr>
<tr>
<td>6</td>
<td>0.9</td>
<td>0.2</td>
<td>1</td>
<td>0.325</td>
</tr>
<tr>
<td>7</td>
<td>0.9</td>
<td>0.1</td>
<td>1</td>
<td>0.325</td>
</tr>
<tr>
<td>8</td>
<td>0.9</td>
<td>0.1</td>
<td>0.8</td>
<td>0.325</td>
</tr>
<tr>
<td>9</td>
<td>0.9</td>
<td>0.15</td>
<td>1</td>
<td>0.3</td>
</tr>
<tr>
<td>10</td>
<td>0.9</td>
<td>0.15</td>
<td>1</td>
<td>0.35</td>
</tr>
<tr>
<td>11</td>
<td>0.9</td>
<td>0.15</td>
<td>0.8</td>
<td>0.35</td>
</tr>
<tr>
<td>12</td>
<td>0.9</td>
<td>0.15</td>
<td>0.8</td>
<td>0.3</td>
</tr>
<tr>
<td>13</td>
<td>1</td>
<td>0.15</td>
<td>0.9</td>
<td>0.3</td>
</tr>
<tr>
<td>14</td>
<td>1</td>
<td>0.15</td>
<td>0.9</td>
<td>0.35</td>
</tr>
<tr>
<td>15</td>
<td>0.8</td>
<td>0.15</td>
<td>0.9</td>
<td>0.35</td>
</tr>
<tr>
<td>16</td>
<td>0.8</td>
<td>0.15</td>
<td>0.9</td>
<td>0.3</td>
</tr>
<tr>
<td>17</td>
<td>0.9</td>
<td>0.15</td>
<td>0.9</td>
<td>0.325</td>
</tr>
</tbody>
</table>
Development Optimization Details: 

_Incorporating Real Options_

*Current reservoir condition(s) at \( t_{\text{switch}} \) become initial conditions for new regime*
Development Optimization Details: 

*Incorporating Real Options*

- How is the optimal new regime selected at $t_{\text{switch}}$?
- Example (CO$_2$ project):
  - Case $k$ (pattern, injection rate) defines the initial configuration
  - For each $k$, there is a table that defines capital expansion costs at the real option point, $\text{CAPEX}_{\text{pat,inj}}$
  - A binary choice variable $DV_{\text{pat,inj}}$ indicates which configuration is chosen for the new regime
  - This choice determines the incremental capital cost:

$$\text{NewCAPEX} = \sum_{\text{pat}} \sum_{\text{inj}} DV_{\text{pat,inj}} \times \text{CAPEX}_{\text{pat,inj}}$$
Development Optimization Details: 
*Incorporating Real Options*

- Example, continued…
- The cumulative production for the new regime is determined by the following equation:

\[
Q_{\text{new}} = \sum_{\text{pat}} \sum_{\text{inj}} DV_{\text{pat,inj}} \times Q^2_{\text{pat,inj}}
\]

- \(Q^2_{\text{pat,inj}}\) is a *Conditional* response surface defined for each potential configuration for the new regime using the *current* reservoir conditions as inputs
- The current reservoir conditions (at the real option time) are computed using the *Auxiliary* response surface(s)
Support Tool: 
*Development Optimization Model*

- Simulate concept selection and initial facility choices as they are done in practice

- Simulate the investment and production process within an optimal control framework
  - Robust treatment of wells and facilities investment, including options to expand
  - Risk-neutral NPV maximizer

- Use Monte Carlo approach to incorporate uncertainty in reservoir properties and other variables
Development Optimization Workflow

- Fix initial facilities ($FAC_0$)
- Reveal reservoir uncertainties from user-defined PDF’s
- Exercise real options
- Compute NPV
- Create CDF for $FAC_0$
- Run additional cases ($FAC_0^k$)
- Select optimal initial facilities ($FAC_0^*$)

This is how projects evolve in practice

Cumulative Distribution Function of NPV
@ Initial Facility Capacity = 10,000 mbopy

- P10 = 635.88
- P50 = 797.60
- P90 = 945.01
- Expected Value = 794.55

Optimal Initial Platform Capacity

- $FAC_0 = 11000$ mbopy
Value of Information

• A procedure is developed to assess the value of accuracy in estimates used in concept selection.

• In summary, the approach examines the loss in project value caused by inaccurate estimates when they are used to make initial facility decisions.

• The result is an estimate of the maximum willingness to pay to mitigate the uncertainty.
If the alternate hypothesis is true, the value of knowing this during planning is equal to C-B.
Current Projects

• Project #1: CO₂ floods (w/ Nguyen) (Completed, writing paper))
  – Examine impacts of uncertain reservoir attributes on CO₂ development decisions (well patterns and injection capacities)
  – **Student:** Suryansh Purwar (now with Landmark)

• Project #2: Reservoir continuity in DW (w/ Lasdon) (Ending Fall 08)
  – Examine the value of information relative to reservoir continuity, with a focus on large deepwater assets
  – **Student:** Hariharan Ramachandran (PGE)

• Project #3: Gas storage (w/ Lake) (Just Starting)
  – Additional external funding provided by Gas Storage Technology Consortium ($60k)
  – Examine impacts of uncertain reservoir attributes and price assumptions on optimal gas storage asset configuration
  – **Students:** Amin Ettehadtavakkol, Babafemi Ogunyomi (PGE)
**Improvements in Capacitance-Resistive Modeling and Optimization of Hydrocarbon Reservoirs**

A simple capacitance-resistive model that characterizes the connectivity between injection and production wells can be used to identify an injection scheme that maximizes the value of the reservoir asset. Model parameters are identified using nonlinear regression. The model is then used together with an optimization algorithm to predict future production rates from an optimal set of injection rates. Research previously conducted has shown that this model, while simple, provides an excellent match to historic data. The optimal injection schemes yield a predicted increase in hydrocarbon recovery of up to 35% over a base case.

The CR model has had extensive development in CPARM over the past three years. These are well documented in previous reports and in the literature. This report summarizes progress in the past year. 

**Using Bottom Hole Pressure.** A simple extension of CR allows incorporation of flowing bottom hole pressure into the model. Such an addition significantly improves total rate prediction and allow an independent evaluation of time-dependent productivity indices. Such a plot should be useful in identifying long-term deterioration of well performance.

**Uncertainty in Parameter Estimation.** The dissertation by Sayarpour shows that multiple fit to the oil production function can give the uncertainty in the input parameters to the model. This uncertainty can be the starting point for large-scale modeling with conventional simulation.

**Large System CR Modeling.** An advantage of using a simple model is the ability to describe large scale systems without incurring a long computation time. However, applying the model to large reservoirs with many wells presents several new challenges. Reservoirs with hundreds of wells have longer production histories that often represent a variety of different reservoir conditions. New wells are created, wells are shut in for a varying periods of time and production wells are converted to injection wells. Additionally, history matching large reservoirs by nonlinear regression is more likely to produce parameters that are statistically insignificant, resulting in a model that is both parameter dense and may be a less accurate reflection of the physical properties of the reservoir. In 2007 we presented an algorithm that eliminated statistically insignificant parameters from the model. We have continued to improve modeling techniques by introducing an algorithm that models shut in periods and a warm-start algorithm that reduces computation time for model fitting.
Optimal Well Location. The CR model can also be used to predict the optimal location of a new well in a reservoir. A finite difference simulator is used to generate production data while varying well location. The capacitance model is then fit to this data and model parameters for well locations that were not explicitly simulated are interpolated from the parameters for the existing models. Well location and allocation is then optimized using a mixed integer programming algorithm. Research is currently concentrated on determining the viability of this approach for increasingly complex reservoir systems.
Reservoir Characterization
From Production and Injection Fluctuations

Center for Petroleum Asset Risk Management
November 2008
Outline

• Background
• Additional Field Cases
• Large Field Application
• User's Manual
• Future Work
The CRM Team

**Faculty**
Larry W. Lake, PGE
Tom Edgar, ChE
Leon Lasdon, IROM
Bob Flake, ECE
Jerry Jensen, UC and TAMU
Emilio Núñez, CPARM

**CRM Students**
Morteza Sayarpour*, PGE
Daniel Kalvani, TAMU
Daniel Weber, ChE
Sami Kaswar, ChE
Alireza Mollaei, PGE
Ahn Phoung Nguyen, ChE
Sirajum Munira, ECE

* Graduated
Hypothesis

• Characteristics of a reservoir can be inferred from analyzing production and injection data only
Boundary Conditions

• Must be injection project
• Rates are most abundant data type
• Rates must vary
• No geologic model required
• Everything done in a spreadsheet
CRM Continuity Equation

**Continuity:**
\[
c_t V_p \frac{dp}{dt} = i(t) - q(t)
\]

**Ordinary Differential Equation:**
\[
\frac{dq(t)}{dt} + \frac{1}{\tau} q(t) = \frac{1}{\tau} i(t) - J \frac{dp_{wf}}{dt}
\]
\[
\tau = \frac{c_t V_p}{J}
\]

**Solution:**
- **Primary**
  \[
  q(t) = q(t_0) e^{-\frac{t-t_0}{\tau}} + I(t) \left( 1 - e^{-\frac{t-t_0}{\tau}} \right)
  \]
- **Injection**
  \[
  \left( c_t V_p \right) \left[ \frac{p_{wf,t} - p_{wf,0}}{t-t_0} \right] \left( 1 - e^{-\frac{t-t_0}{\tau}} \right)
  \]
- **BHP**
  \[
  \left( c_t V_p \right) \left[ \frac{p_{wf,t} - p_{wf,0}}{t-t_0} \right] \left( 1 - e^{-\frac{t-t_0}{\tau}} \right)
  \]
CRMT: Total Production

Continuity Equation

\[ I_{\text{Field}}(t) = q(t_0) e^{-\frac{t-t_0}{\tau_F}} + \left(1 - e^{-\frac{t-t_0}{\tau_F}}\right) I_{\text{Field}}(t) \]

\[ \tau_F = \left(\frac{c_t V_p}{J}\right)_F \]
CRMP: One Producer

Continuity Equation

\[ \dot{q}_j(t) = q(t_0) e^{-\frac{(t-t_0)}{\tau_j}} + 1 - e^{-\frac{(t-t_0)}{\tau_j}} \sum_{\text{injectors}} f_{ij} I_i \]

\[ \tau_j = \left( \frac{c_t V_p}{J} \right)_j \]

CRMT: \( I^* = I_{\text{field}} \)

CRMP: \( I^* = \sum_{i=1}^{N_{\text{inj}}} f_{ij} I_i \)

\( f_{ij} = \text{fraction going} \)

at steady state

(gain or connectivity)
CRMIP: Total Production

Continuity Equation

\[ \dot{q}_{ij}(t) = q(t_0)e^{-(t-t_0)/\tau_{ij}} + q_j(t) = \sum_{i\in\text{injectors}} q_{ij} \]

\[ \tau_{ij} = \left( \frac{c_t V_p}{J} \right)_{ij} \]
CRM: Oil Fractional-Flow Equations

\[ q_0(t) = f_o(t)q(t) \]

\[ f_o(t) = \frac{q_o}{q_o + q_w} = \frac{1}{1 + \text{WOR}(t)} \]

\[ f_o(t) = \frac{1}{1 + \alpha(CWI(t))^{\beta}} \]

\[ \log \left( \frac{1}{f_o(t)} - 1 \right) = \log \alpha + \beta \log(CWI(t)) \]
Outline

- Background
- Additional Field Cases
  - South Wasson Clearfork
  - McElroy San Andres (CO2)
  - Up Ford (40 years data)
  - North Sea (with BHP)
  - North Buck Draw (with tracer)
  - MESL Field (optimization)
  - Okume (daily rate, BHP)
  - Slaughter Levelland
South Wasson Clearfork Injector Results

Representation of gains

Representation of $\tau$
UP Ford Field Injection Rates
Up Ford CRMP Individual Producer
Match with GAMS

![Graphs showing production rates over time for different producers.]

- **CRMP_P002**: Time, Month: 06 01 2018, Total Rate, bbl/M: 50,000
- **P002 D310**: Time, Month: 06 01 2018, Total Rate, bbl/M: 25,000
- **CRMP_P003**: Time, Month: 06 01 2018, Total Rate, bbl/M: 15,000
- **P003 D108**: Time, Month: 06 01 2018, Total Rate, bbl/M: 10,000
- **CRMP_P004**: Time, Month: 06 01 2018, Total Rate, bbl/M: 20,000
- **P004 D623**: Time, Month: 06 01 2018, Total Rate, bbl/M: 15,000
- **CRMP_P011**: Time, Month: 06 01 2018, Total Rate, bbl/M: 10,000
- **P011 D216**: Time, Month: 06 01 2018, Total Rate, bbl/M: 5,000
UP Ford CRMP Individual Producer Oil Match with GAMS
North Sea Case
Production and BHP Data
North Sea CRMP wo BHP

\[ y = 0.7818x + 4920.8 \]

\[ R^2 = 0.7413 \]
North Sea CRMP with BHP

Production Data

$y = 0.9637x + 718.24$

$R^2 = 0.9321$
North Sea CRMP Productivity Indices

[Graph showing productivity indices over time for different layers (J1, J2, J3, J4)]
North Sea CRMIP with BHP

Production Rates

Time, Days

Production Data

y = 0.999x - 3.8297
R² = 0.9584
Adding BHP...
North Buck Draw Comparison

- CM $\tau$ correlates better than Spearman with tracer breakthrough time

![Graph showing the comparison between CM and Spearman methods for tracer breakthrough time. The CM method is represented by a linear equation $\tau_{CM} = 0.26x - 0.03$ with $R^2 = 0.63$.](image)
MESL Field (SPE 110081)

**Permeability**

**Match Oil Rate**

**Total Production**

**Shut in I1; Max I2, I3; Maint I4**

6% increase
Outline

• Background
• Additional Field Cases
• Large Field Application
• User's Manual
• Future Work
Future Work

- More Field Cases
- Graphical IO
- Algorithm for Well Placement
- Determination of Flooded Volume
- Application to EOR
Capacitance-Resistive Modeling and Optimization of Hydrocarbon Reservoirs

Daniel Weber

Supervisors:
Thomas F. Edgar
Larry W. Lake
History Matching and Optimization

• Fit capacitance-resistive model (CRM) to total production data using nonlinear regression
  – Gains are non-negative and must sum to one for each injector
  – Using single objective function – all producers are fit at the same time
• Fit oil fractional flow model to oil production data using nonlinear regression
• Use CRM and oil fractional flow model to optimize future injection to maximize net present value (NPV) of the reservoir
Capacitance-Resistive Model

\[ q_j = q_{0j} e^{-\frac{\Delta t}{\tau_j}} + \left( 1 - e^{-\frac{\Delta t}{\tau_j}} \right) \sum_{i=1}^{n} f_{ij} I_i \]

Where

- \( q_j \) = total production rate of producer \( j \)
- \( q_{0j} \) = initial production rate of producer \( j \)
- \( I_i \) = injection rate of injector \( i \)
- \( f_{ij} \) = weight (gain) between injector \( i \) and producer \( j \)
- \( \tau_j \) = time constant for producer \( j \)
- \( \Delta t \) = time step size
Oil Fractional Flow Model

\[ q_{oj} = f_{oj} q_j \]

\[ f_{oj} = \frac{1}{1 + a_j CWI^{b_j}} \]

Where

- \( q_{oj} \) = oil production rate of producer j
- \( f_{oj} \) = oil fraction of producer j
- \( CWI \) = cumulative water injected in all injectors
- \( a_j, b_j \) = model parameters for producer j
Optimization formulation

Maximize net present value

\[ NPV = \sum_{j=1}^{n_p} \sum_{k=1}^{n_t} \frac{p_o}{(1+ir)^k} q_{oj}(t_k) \Delta t - \sum_{i=1}^{n_i} \sum_{k=1}^{n_k} \frac{p_w}{(1+ir)^k} I_i(t_k) \Delta t \]

Subject to
- CRM
- Fractional flow model
- Upper limit on total injection rate
- Upper limits on rate of each injector

\[ \sum_{i=1}^{n_i} I_i(t_k) \leq I_{TOT} \]

\[ l_i \leq I_i(t_k) \leq u_i \]
Problems with Large Scale Reservoirs

• Statistically insignificant parameters
• Wells shut in for short periods of time
• Long computation time for model fitting
Solutions

• Semi-continuous gains
  – Removes statistically insignificant parameters
  – Marginal increase in model error

• Shut-in logic
  – Removes well shut-in periods from model fit
  – Automatically sets production rate to zero for shut-in periods

• Warm start algorithm
  – Fits each producer separately with relaxed constraints on the model parameters
  – Shown to not only reduce computation time, but to improve model fit
Semi-continuous Gains Model Fit Comparison in Homogeneous Symfield

\[ \sum_{q=1}^{n_t} \text{abs}\left( \frac{q_{k,\text{obs}} - q_{k,\text{est}}}{q_{k,\text{obs}}} \right) \]

\[ err = \frac{1}{n_t - 1} \sum_{k=1}^{n_t} \text{abs}\left( \frac{q_{k,\text{obs}} - q_{k,\text{est}}}{q_{k,\text{obs}}} \right) \]
Cold Start Model Fit

• Total production model for all producers fit simultaneously
• Objective function for model fit

$$\min z = \sum_j \sum_k \left( q_{\text{actual},j,k} - q_{\text{est},j,k} \right)^2$$

• Material balance constraint on gains across injection wells

$$\sum_i f_{i,j} \leq 1$$
Warm Start Model Fit

- Total production model fit separately for each producer
- Objective function for each model fit

\[ \min z_j = \sum_k \left( q_{\text{actual},j,k} - q_{\text{est},j,k} \right)^2 \]

- Gains not constrained across injection wells (a relaxation of the material balance constraint)
- All producers then refit simultaneously (as before) using initial values provided by the warm start fits
Cold Start vs. Warm Start in Southeast Levelland Unit

- Cold Start
  - Objective function value: 13819.6115
  - Computation Time: 1697 sec
- Warm Start
  - Objective function value: 14358.7162
  - Computation Time: 777 sec
- Improvement of 3.75% in objective function value
- Indication of multiple local optima in total production model fit objective function
Proposed Well Location Optimization Algorithm

- Run a small number (<10) of finite difference simulations with varying well locations
- Fit capacitance model to the simulation results
- Interpolate model parameters as a function of well location
- Optimize well location and allocation using a mixed integer programming algorithm
Reservoir Map of Symfield

113 simulations varying injector location – every other grid block in white square

Two producers

P1

P2
Parameter Maps for Interpolation

• Maps show how parameters change with injector location
• Parameter values for un-simulated grid blocks were estimated as an arithmetic average of the values for surrounding grid blocks
• 8 parameters total: 2 gains, 2 time constants, 4 oil model parameters
Proposal for Gains

• Contours are relatively straight (become more curved as distance between wells decreases)

• Propose modeling parameter changes linearly – can run 3 or 4 simulations (preferably at the corners) and fit a plane to the parameter values
Parameter Map for tau1
Parameter Map for \( \tau_2 \)
Proposal for Time Constants

- Contours have no recognizable pattern
- Values are fairly uniform across the reservoir (scale of contour plot is fairly small)
- Propose modeling time constants as constant with respect to well location (can take the average of 3 or 4 simulations)
Parameter Map for b2
Proposal for Oil Model Parameters

• Parameters are constant until they near the well, then they change rapidly to a different relatively constant value

• Appears to be some threshold for when breakthrough occurs – if wells are significant distance apart, can use total production as estimate of oil production

• These parameters may require more sophisticated modeling
Summary

• Work on large scale systems is concentrated on case studies on real data
• Well location optimization research is concerned with determining the viability of the proposed algorithm
Classification of hydrocarbon recovery based on reservoir databases

A statistical procedure for analyzing oil and gas reservoir data sets has been developed in order to come up with deterministic and probabilistic values of ultimate recovery factor for both oil and gas reservoirs. This could be of interest for exploration because the amount of knowledge regarding a newly discovered reservoir is limited and it would be helpful to know some proxy value of recovery factor that could guide later flow simulations. This would also be helpful in projecting the possible revenues that could be generated from the reservoirs. The deterministic models are based on multivariate linear regression. The probability models include the calibration of the likelihood of recovery factor using naïve Bayesian classification. For oil reservoirs, classification accuracies of recovery factor obtained using geological and engineering parameters were compared. For gas reservoirs, the Bayesian classifier model was implemented by fitting a multivariate Gaussian distribution to the predicting variables.

In case of the oil reservoirs, 19 predictor variables were used to estimate the recovery factor, and these included both engineering and geological parameters. First, multivariate regression was performed using 24 reservoirs. Although, a high correlation was achieved between the actual and predicted recovery factor for the test data, the slope of the regression fit was nowhere close to one. These 24 reservoirs showed grouping of the data points in the principal score space, which motivated a cluster analysis. A new data subset consisting of 95 reservoirs was prepared and projected on the principal components, to introduce orthogonality in the data. The k-mean cluster analysis followed by the linear regression analysis provided a reliable model for predicting the recovery factor. For the calibration of the likelihood of recovery factor, principal scores of the predicting variables were used as they are orthogonal and, hence, a naïve Bayesian approach could be easily implemented. The likelihood function of the recovery factor, for the oil reservoirs, is multimodal and non-Gaussian.

The linear regression model performed well compared to the empirical correlations given by Arps et al. (1967) and Guthrie et al. (1995). In gas reservoirs, good prediction was achieved by using the recovery instead of recovery factor as a response function in the regression. The likelihood functions of the recovery factor for both gas and oil reservoirs are multimodal and non-Gaussian. For the oil reservoirs, both geological and engineering parameters played important role for the prediction of recovery factor, which eventually lead to the conclusion that the engineering and geological parameters are not independent.
A Multivariate Statistical Approach to Reservoir Classification

Aviral Sharma
Sanjay Srinivasan
Larry Lake
The University of Texas at Austin
Objective

• Explore if data analysis can yield:

\[
\text{lhd}\{\text{recovery factor} \mid \text{geology, wells, \ldots}\}
\]

a fast proxy for assessing reservoir performance

• Explore if analysis can yield:

\[
P\{\text{geological analog} \mid \text{depositional model, porosity}
\]

\[
\text{permeability, prod. char., \ldots}\}
\]

Precursor to mp statistics based reservoir modeling
GASIS Data

• GASIS (Gas Information System) is a national database of geological, engineering, production, and ultimate recovery data for U.S. oil and gas reservoirs.

• This database is designed to be used as input for modeling and forecasting.

• GASIS has two components:
  1. A "Reservoir Data System" of geologic and engineering data.
GASIS Data

• The reservoir data system contains more than 8000 reservoir records with 185 variables per record for Gulf of Mexico (GOM).

• These reservoirs represent most of the historical gas production in the areas covered.

• In general, GASIS reservoirs are those with at least 10 Bcf of cumulative gas production. (5 Bcf in Rocky Mountain region)
<p>| | | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>17520</td>
<td>21.0</td>
<td>15.00</td>
<td>361</td>
<td>15487.6800</td>
<td>50.000</td>
<td>0.6200</td>
<td>650</td>
<td>1969</td>
<td></td>
</tr>
<tr>
<td>16721</td>
<td>63.0</td>
<td>18.00</td>
<td>348</td>
<td>14938.5414</td>
<td>27.000</td>
<td>0.6200</td>
<td>31231</td>
<td>121188</td>
<td></td>
</tr>
<tr>
<td>15050</td>
<td>70.0</td>
<td>17.00</td>
<td>340</td>
<td>13814.3950</td>
<td>52.000</td>
<td>0.6300</td>
<td>560</td>
<td>1987</td>
<td></td>
</tr>
<tr>
<td>16177</td>
<td>35.0</td>
<td>17.00</td>
<td>340</td>
<td>14415.3247</td>
<td>53.000</td>
<td>0.6500</td>
<td>994</td>
<td>3978</td>
<td></td>
</tr>
<tr>
<td>15020</td>
<td>8.0</td>
<td>19.00</td>
<td>340</td>
<td>15015.4940</td>
<td>39.000</td>
<td>0.6200</td>
<td>505</td>
<td>2020</td>
<td></td>
</tr>
<tr>
<td>13860</td>
<td>7.0</td>
<td>30.00</td>
<td>332</td>
<td>9883.5660</td>
<td>30.000</td>
<td>0.6000</td>
<td>747</td>
<td>2987</td>
<td></td>
</tr>
<tr>
<td>16246</td>
<td>38.0</td>
<td>18.00</td>
<td>330</td>
<td>14327.3474</td>
<td>38.000</td>
<td>0.6200</td>
<td>2902</td>
<td>56203</td>
<td></td>
</tr>
<tr>
<td>14128</td>
<td>19.0</td>
<td>28.00</td>
<td>316</td>
<td>12799.9680</td>
<td>50.000</td>
<td>0.6600</td>
<td>1912</td>
<td>7648</td>
<td></td>
</tr>
<tr>
<td>14292</td>
<td>33.0</td>
<td>18.00</td>
<td>309</td>
<td>11760.8868</td>
<td>31.000</td>
<td>0.6000</td>
<td>19946</td>
<td>80235</td>
<td></td>
</tr>
<tr>
<td>18410</td>
<td>10.0</td>
<td>21.00</td>
<td>306</td>
<td>16270.7580</td>
<td>56.000</td>
<td>0.6000</td>
<td>53</td>
<td>214</td>
<td></td>
</tr>
<tr>
<td>14600</td>
<td>16.0</td>
<td>26.00</td>
<td>301</td>
<td>12951.6600</td>
<td>20.000</td>
<td>0.6500</td>
<td>4083</td>
<td>19342</td>
<td></td>
</tr>
<tr>
<td>18050</td>
<td>23.0</td>
<td>14.00</td>
<td>300</td>
<td>15342.5000</td>
<td>31.000</td>
<td>0.6300</td>
<td>0</td>
<td>44288</td>
<td></td>
</tr>
<tr>
<td>15525</td>
<td>18.0</td>
<td>23.00</td>
<td>299</td>
<td>13497.4350</td>
<td>25.000</td>
<td>0.6200</td>
<td>7760</td>
<td>31038</td>
<td></td>
</tr>
<tr>
<td>14050</td>
<td>14.0</td>
<td>22.00</td>
<td>299</td>
<td>12347.1400</td>
<td>28.000</td>
<td>0.6500</td>
<td>11473</td>
<td>52875</td>
<td></td>
</tr>
<tr>
<td>14488</td>
<td>16.0</td>
<td>21.00</td>
<td>298</td>
<td>9759.7920</td>
<td>30.000</td>
<td>0.6000</td>
<td>340</td>
<td>1359</td>
<td></td>
</tr>
<tr>
<td>17349</td>
<td>11.0</td>
<td>20.00</td>
<td>296</td>
<td>8081.1642</td>
<td>30.000</td>
<td>0.6000</td>
<td>25</td>
<td>98</td>
<td></td>
</tr>
<tr>
<td>15200</td>
<td>31.0</td>
<td>26.00</td>
<td>295</td>
<td>13786.6400</td>
<td>34.000</td>
<td>0.6800</td>
<td>2176</td>
<td>8706</td>
<td></td>
</tr>
<tr>
<td>17316</td>
<td>9.0</td>
<td>26.00</td>
<td>295</td>
<td>8067.5244</td>
<td>28.000</td>
<td>0.6300</td>
<td>105</td>
<td>421</td>
<td></td>
</tr>
<tr>
<td>17500</td>
<td>18.0</td>
<td>24.00</td>
<td>294</td>
<td>14453.2500</td>
<td>34.000</td>
<td>0.6000</td>
<td>1732</td>
<td>6929</td>
<td></td>
</tr>
<tr>
<td>16904</td>
<td>10.0</td>
<td>27.00</td>
<td>290</td>
<td>7791.0536</td>
<td>28.000</td>
<td>0.6200</td>
<td>59</td>
<td>236</td>
<td></td>
</tr>
<tr>
<td>14308</td>
<td>21.0</td>
<td>27.00</td>
<td>289</td>
<td>11915.7024</td>
<td>22.000</td>
<td>0.6500</td>
<td>4983</td>
<td>19933</td>
<td></td>
</tr>
<tr>
<td>15050</td>
<td>9.0</td>
<td>26.00</td>
<td>289</td>
<td>14115.3950</td>
<td>28.000</td>
<td>0.6400</td>
<td>16028</td>
<td>59363</td>
<td></td>
</tr>
<tr>
<td>15450</td>
<td>58.0</td>
<td>19.00</td>
<td>288</td>
<td>12360.0000</td>
<td>27.000</td>
<td>0.6200</td>
<td>3006</td>
<td>12031</td>
<td></td>
</tr>
<tr>
<td>13137</td>
<td>45.0</td>
<td>15.00</td>
<td>288</td>
<td>8977.8258</td>
<td>29.000</td>
<td>0.6100</td>
<td>6965</td>
<td>20485</td>
<td></td>
</tr>
<tr>
<td>15000</td>
<td>6.0</td>
<td>25.00</td>
<td>288</td>
<td>13140.0000</td>
<td>36.000</td>
<td>0.6000</td>
<td>2750</td>
<td>11002</td>
<td></td>
</tr>
<tr>
<td>16400</td>
<td>9.0</td>
<td>17.00</td>
<td>287</td>
<td>8255.7600</td>
<td>40.000</td>
<td>0.6200</td>
<td>148</td>
<td>592</td>
<td></td>
</tr>
<tr>
<td>16709</td>
<td>8.0</td>
<td>27.00</td>
<td>287</td>
<td>7784.7231</td>
<td>29.000</td>
<td>0.6200</td>
<td>95</td>
<td>380</td>
<td></td>
</tr>
<tr>
<td>14750</td>
<td>26.0</td>
<td>18.00</td>
<td>286</td>
<td>10932.7000</td>
<td>37.000</td>
<td>0.5900</td>
<td>9888</td>
<td>39550</td>
<td></td>
</tr>
<tr>
<td>13135</td>
<td>22.0</td>
<td>21.00</td>
<td>286</td>
<td>12240.5065</td>
<td>39.000</td>
<td>0.7000</td>
<td>7133</td>
<td>19280</td>
<td></td>
</tr>
<tr>
<td>16550</td>
<td>11.0</td>
<td>28.00</td>
<td>285</td>
<td>7710.6450</td>
<td>27.000</td>
<td>0.6000</td>
<td>966</td>
<td>3871</td>
<td></td>
</tr>
<tr>
<td>14100</td>
<td>11.0</td>
<td>27.00</td>
<td>285</td>
<td>11018.6200</td>
<td>22.000</td>
<td>0.6500</td>
<td>131</td>
<td>523</td>
<td></td>
</tr>
<tr>
<td>15700</td>
<td>26.0</td>
<td>28.00</td>
<td>284</td>
<td>8014.8500</td>
<td>36.000</td>
<td>0.8200</td>
<td>2880</td>
<td>11521</td>
<td></td>
</tr>
<tr>
<td>16350</td>
<td>45.0</td>
<td>30.00</td>
<td>283</td>
<td>12710.4900</td>
<td>30.000</td>
<td>0.6200</td>
<td>2667</td>
<td>10660</td>
<td></td>
</tr>
<tr>
<td>15133</td>
<td>5.0</td>
<td>29.00</td>
<td>282</td>
<td>7051.9780</td>
<td>22.000</td>
<td>0.6200</td>
<td>657</td>
<td>2628</td>
<td></td>
</tr>
<tr>
<td>12523</td>
<td>8.0</td>
<td>24.00</td>
<td>281</td>
<td>11330.6104</td>
<td>39.000</td>
<td>0.6500</td>
<td>619</td>
<td>17043</td>
<td></td>
</tr>
<tr>
<td>13100</td>
<td>26.0</td>
<td>16.00</td>
<td>280</td>
<td>8871.3200</td>
<td>35.000</td>
<td>0.6100</td>
<td>1021</td>
<td>4084</td>
<td></td>
</tr>
<tr>
<td>15353</td>
<td>22.0</td>
<td>25.00</td>
<td>280</td>
<td>12220.9800</td>
<td>32.000</td>
<td>0.7300</td>
<td>8245</td>
<td>32992</td>
<td></td>
</tr>
<tr>
<td>14548</td>
<td>15.0</td>
<td>27.00</td>
<td>280</td>
<td>12796.4208</td>
<td>40.000</td>
<td>0.6000</td>
<td>1024</td>
<td>4094</td>
<td></td>
</tr>
<tr>
<td>14000</td>
<td>9.0</td>
<td>17.00</td>
<td>280</td>
<td>10515.6000</td>
<td>40.000</td>
<td>0.6500</td>
<td>161</td>
<td>584</td>
<td></td>
</tr>
</tbody>
</table>
Preprocessing

Out of 185 features following 23 features were selected (most of which are continuous variables):

- Reservoir Type
- Tight Gas
- Field Elevation
- Offshore Water Depth
- Specific Lithology
- General Trap Type
- Acreage
- Spacing/Well
- Depth Top
- Net-Pay
- Porosity
- Permeability
- Temperature
- Initial Pressure
- Water Saturation
- Drive Mechanism
- Gas Gravity
- Liquid Gravity
- Stimulation
- Gas-in-Place
- Gas Contents
- Cumulative Liquid Production
- Cumulative Gas Production
Preprocessing Contd.

• Some reservoirs removed.
  ✔ No recovery mentioned.
  ✔ Associated gas reservoirs are removed because recovery is for non-associated gas reservoirs.
  ✔ Recovery factor was greater than 1.0.

\[
\text{Rec. Factor} = \frac{\text{Recovery}}{\text{Gas-in-Place}}
\]
Preprocessing Contd.

• Missing Initial Pressure was calculated using gradient and depth.

• Elevation variable was removed because most records have zero elevation.

• If records have missing values for Sw, Phi, Pay thickness, record removed.

• Percentage of gas content variable present for only 115 reservoirs.
Box Plots of Features

- Depth (ft)
- H2O Depth (ft)
- Thickness (ft)
- Porosity (%)

These box plots illustrate the distribution of various features, including depth, H2O depth, thickness, and porosity, measured in feet and percentage, respectively.
A Multivariate Statistical Approach to Reservoir Classification
• Only the extreme points of Box-plots were removed.
• Most of the data was kept for the modeling.
Data for Analysis

Finally we brought it down to 7331 GOM reservoirs and following 15 variables for the analysis:

- Offshore Water Depth
- Depth Top
- Net-Pay
- Porosity
- Temperature
- Initial Pressure
- Water Saturation
- Gas Gravity
- Gas-in-Place
- Cumulative Liquid Production
- Cumulative Gas Production
Distribution of RF in dataset
Kmean Cluster Analysis of ultimate recovery factor has been performed to determine number of classes for recovery factor.
Silhouette Plot

\[ s(i) = \begin{cases} 
1 - a(i)/b(i) & \text{if } a(i) < b(i) \ (A) \\
0 & \text{if } a(i) = b(i) \ (B) \\
b(i)/a(i) - 1 & \text{if } a(i) > b(i) \ (C) 
\end{cases} \]
Average Silhouette Values (Smeans)

- Smeans for the clusters are:
  - 3 clusters = 0.77
  - 4 clusters = 0.78
  - 5 clusters = 0.76
Classes Definition

• On the basis of pdf of ultimate recovery factor (URF) and cluster analysis, it can be seen that four cluster (classes) are the optimum.

From the cluster analysis:
- Class 1: URF ∈ (0,0.37]
- Class 2: URF ∈ [0.38,0.56]
- Class 3: URF ∈ [0.57,0.73]
- Class 4: URF ∈ [0.74,1)
Missing Values

• Dealing with missing values for certain features within a class:
  ✓ Cluster analysis in each class.
  ✓ Fill the missing values using mean of the attribute in each cluster.
Bayesian Classifier

\[
P(C_i / x) = \frac{P(x / C_i)P(C_i)}{P(x)}
\]

\[
P(x / C_i) = \frac{1}{(2\pi)^{d/2} \left| \Sigma \right|^{1/2}} \exp\left\{-0.5(x - \mu)^T \Sigma^{-1} (x - \mu)\right\}
\]

Objective is to fit the means using a training procedure
Training/Test set results

• Training set comprised of 6165 reservoirs.
• Testing was done over 1166 reservoirs.

Observations:
 – The classification accuracy : 75.04%
 – Many of the Class 2 reservoirs misclassified.

Conclusions:
 – Divide into 3 classes instead of 4.
3 Classes Problem

Class 1: URF ∈ (0,0.45]
Class 2: URF ∈ [0.46,0.7]
Class 3: URF ∈ [0.71,1]
Linear Regression
on GASIS
Predictive Modeling-Linear Regression

- Linear regression is a method for representing observed outputs (Target $y$) as linear combinations of functions of the inputs (Predictors $x$)
- Express as a linear combination of basis functions
  \[ y(x, w) = \sum_{i=1}^{M} w_i \phi_i(x) = w^T \phi(x) \]
- Minimize sum of squares error function
  \[ E = \frac{1}{2} \sum_{n=1}^{N} \{w^T \phi(x_n) - y_n\}^2 \]
Measure of goodness of fit

- $R^2$ measure

$$R^2 = 1 - \frac{SSE}{SST}$$

$$SSE = \sum_i (y_i - \hat{y})^2$$

$$SST = \sum_i (y_i - \bar{y})^2$$
Predictors Selected

1. Offshore water depth (ft)
2. Depth of Reservoir (ft)
3. Thickness (ft)
4. Porosity (%)  
5. Reservoir Temperature (deg F)
6. Initial Pressure (Psia)
7. Water Saturation (%)  
8. Gas Gravity  
9. Cumulative Gas Production (mmcf)  
10. Gas-in-Place (mmcf)
Linear regression on each Class

- Most data points in each class of recovery are grouped into one cluster.

- Regression model for each class of recovery.

- For the four class problem, data set in each class has been divided into train and test data.
Distribution of Recovery Factor in Each Class

- Ultimate Recovery Factor vs. Frequency
- Frequency distribution for different ranges of Ultimate Recovery Factor.
Regression Results

Class 1

\[ y = 0.0399x + 0.2559 \]
\[ R^2 = 0.0384 \]

Class 2

\[ y = 0.054x + 0.4652 \]
\[ R^2 = 0.06 \]

Class 3

\[ y = 0.0395x + 0.5629 \]
\[ R^2 = 0.0536 \]

Class 4

\[ y = 0.0378x + 0.7353 \]
\[ R^2 = 0.0852 \]
Regression Results

• From the linear regression results it can be seen that predicted recovery factor for each class is very poor.

• Another approach: Predicting recovery instead of the recovery factor.

\[
\text{Rec. Factor} = \frac{\text{Recovery}}{\text{Gas – in – Place}}
\]
Distribution of Recovery

Over-all data Points

Class 1

Class 2

Class 3

Class 4
Regression Results

Class 1
\[ y = 0.8196x + 318.88 \quad R^2 = 0.9404 \]

Class 2
\[ y = 0.9620x + 385.57 \quad R^2 = 0.9882 \]

Class 3
\[ y = 0.9425x + 55.688 \quad R^2 = 0.9932 \]

Class 4
\[ y = 1.015x - 534.54 \quad R^2 = 0.9966 \]
Validation of Model on Tight Gas Reservoirs

• Use the existing composite model (one big cluster) for validating tight gas reservoirs.
• Tight gas reservoirs were not used in training data-set
• Weights from belongs to one cluster have been used.
• 86 tight gas reservoirs have been used for validation.
RESULTS

\[ y = 0.8793x + 1976.7 \]

\[ R^2 = 0.995 \]

Correlation coefficient between actual and predicted recovery = 0.97
Likelihood Function of Recovery Factor

\[ P(C_i \mid X) = \frac{P(X \mid C_i)P(C_i)}{P(X)} \]

\[ X = \{x_1, x_2, \ldots, x_d\} \]

\[ P(C_i \mid X) = \frac{P(x_1 \mid C_i)P(x_2 \mid C_i)\ldots P(x_d \mid C_i)P(C_i)}{P(x_1)P(x_2)\ldots P(x_d)} \]

If the attributes are conditionally independent, we can apply Naïve Bayesian method.
Principal Scores

• The Likelihood function was calculated by introducing orthogonality in the data.
• This is done by first calculating the principal components of the data.
• Principal score = Data matrix × Principal Components
• Each column of principal scores are orthogonal to each other
• This helps in using naïve Bayesian to calibrate likelihood of recovery factor.
## Principal Components

<table>
<thead>
<tr>
<th></th>
<th>Water Depth</th>
<th>-2.1E-03</th>
<th>-7.8E-03</th>
<th>-1.8E-02</th>
<th>-1.0E+00</th>
<th>1.2E-02</th>
<th>1.3E-02</th>
<th>1.4E-03</th>
<th>2.3E-03</th>
<th>7.5E-06</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Depth</strong></td>
<td>4.2E-03</td>
<td>-2.4E-02</td>
<td>8.4E-01</td>
<td>5.5E-01</td>
<td>-1.6E-02</td>
<td>1.9E-04</td>
<td>-9.7E-03</td>
<td>8.8E-04</td>
<td>-2.5E-04</td>
<td>2.4E-06</td>
</tr>
<tr>
<td><strong>Thickness</strong></td>
<td>1.0E-04</td>
<td>-8.2E-05</td>
<td>5.3E-04</td>
<td>-6.3E-04</td>
<td>-1.1E-02</td>
<td>-1.0E+00</td>
<td>1.8E-02</td>
<td>8.7E-02</td>
<td>-6.4E-03</td>
<td>1.7E-05</td>
</tr>
<tr>
<td><strong>Φ</strong></td>
<td>-1.2E-06</td>
<td>1.6E-05</td>
<td>-5.4E-04</td>
<td>-2.0E-04</td>
<td>-2.8E-03</td>
<td>-1.2E-03</td>
<td>-3.5E-02</td>
<td>-8.0E-02</td>
<td>-1.0E+00</td>
<td>9.6E-05</td>
</tr>
<tr>
<td><strong>Temp.</strong></td>
<td>6.6E-05</td>
<td>-3.0E-04</td>
<td>1.1E-02</td>
<td>1.6E-03</td>
<td>1.2E-02</td>
<td>1.0E-02</td>
<td>1.0E+00</td>
<td>-8.8E-02</td>
<td>-2.8E-02</td>
<td>7.6E-05</td>
</tr>
<tr>
<td><strong>P</strong></td>
<td>3.6E-03</td>
<td>-2.1E-02</td>
<td>5.5E-01</td>
<td>-8.4E-01</td>
<td>1.0E-02</td>
<td>6.1E-04</td>
<td>-4.7E-03</td>
<td>-8.6E-05</td>
<td>2.0E-05</td>
<td>-3.9E-07</td>
</tr>
<tr>
<td><strong>Sw</strong></td>
<td>-2.1E-05</td>
<td>-9.8E-06</td>
<td>1.8E-04</td>
<td>-3.5E-04</td>
<td>3.3E-03</td>
<td>8.9E-02</td>
<td>8.4E-02</td>
<td>9.9E-01</td>
<td>-8.2E-02</td>
<td>-5.1E-05</td>
</tr>
<tr>
<td><strong>Gas Gravity</strong></td>
<td>4.5E-09</td>
<td>-7.6E-08</td>
<td>2.5E-06</td>
<td>1.6E-06</td>
<td>-7.2E-06</td>
<td>-2.1E-05</td>
<td>6.8E-05</td>
<td>-6.4E-05</td>
<td>-9.3E-05</td>
<td>-1.0E+00</td>
</tr>
<tr>
<td><strong>Cum. Gas Prod.</strong></td>
<td>5.1E-01</td>
<td>8.6E-01</td>
<td>2.4E-02</td>
<td>-3.5E-03</td>
<td>-1.8E-03</td>
<td>1.8E-05</td>
<td>2.2E-06</td>
<td>1.9E-05</td>
<td>4.5E-06</td>
<td>1.9E-09</td>
</tr>
<tr>
<td><strong>Gas-in-Place</strong></td>
<td>8.6E-01</td>
<td>-5.1E-01</td>
<td>-2.1E-02</td>
<td>2.9E-03</td>
<td>1.3E-03</td>
<td>1.0E-04</td>
<td>-1.0E-05</td>
<td>5.1E-06</td>
<td>-2.4E-06</td>
<td>-1.6E-08</td>
</tr>
</tbody>
</table>
Likelihood Function of Recovery Factor

Likelihood function is plotted for 14 principal score sets
Conclusions

• For the new reservoir data-set, ensemble model can be used: Bayesian classifier, Linear regression on classes, Composite Linear regression.

• An ensemble of models would help determine the uncertainty in the recovery factor and hence in the reservoir model.

• Likelihood function of recovery factor is multimodal and non-Gaussian.
An Implied Binomial Tree Approach to Value Multi-factor Real Options

Tianyang Wang, James S. Dyer

McCombs School of Business, The University of Texas at Austin, Austin, Texas 78712

{tianyang.wang@phd.mccombs.utexas.edu, jim.dyer@mccombs.utexas.edu}

This paper proposes an algorithm for solving a multi-factor real options problem by approximating it with an implied binomial tree. The implied binomial tree is constructed to be consistent with simulated market information. By simulating European option prices for real options on projects as artificial market information, we extend the finance-based Implied Binomial Tree (IBT) method for real options valuation when no tradable options on the market are available. Compared to the current literature, this algorithm offers a more flexible probability distribution assumption for project values than the ones suggested previously in the literature, and provides a simple, computationally efficient and accurate way to price high dimensional real options. For risk managers, it serves as an easy capital budgeting method for projects with managerial flexibility.

As opposed to the current approaches to multi-factor real options valuation with a binomial tree, this paper relaxes the limiting assumption that the project cash flows follow a specific distribution. The implied binomial tree approach is nonparametric in nature and it allows for skewness and non-constant volatilities in the project cash flow process. As a result, we extend the literature by allowing a generalized diffusion process for the project cash flows in real options valuation.
Project Description

- **Title:** An Implied Binomial Tree Approach to Value Multi-factor Real Options

- **Author:** Tianyang Wang, Jim Dyer

- **Abstract:** This paper proposes an algorithm for solving a multi-factor real options problem by approximating it with a single factor implied binomial tree. The implied binomial tree is constructed to be consistent with the simulated market information. By simulating European option prices as artificial market information, we extend the finance-based Implied Binomial Tree (IBT) method for real options valuation when no tradable options on the market are available. Compared to the current literature, this algorithm offers a more flexible distribution assumption for project value than the ones suggested previously in the literature, and provides a simple, computationally efficient and accurate way to price high dimensional real options problems with general smooth distributions. For risk managers, it serves as an easy capital budgeting method for projects with managerial flexibility.

- **Key words:** real options; implied binomial tree; multi-factor; simulation
Application in Oil and Gas Industry

- Real options are common in oil/gas exploration and development, and are widely used in Oil and Gas industry.
- Practical complex oil/gas projects contain multiple sources of uncertainty.
- Other solution methods for problems with options are subject to the “curse of dimensionality” for large, realistic problems, or need to assume a particular distribution for the project value.
- This research provides a general approach to approximate the stochastic process of the oil/gas project value with an implied binomial tree by simulating market information necessary for its implementation.
- This approach offers a more flexible distribution assumption for the project value.
- This approach is accurate and computationally efficient.
- The implied binomial tree/lattice has a simple, intuitive interface which lends itself to practical applications in oil/gas industry.
## A Case Example with Three Uncertainties

<table>
<thead>
<tr>
<th>Year</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Remaining Reserves</td>
<td>90.0</td>
<td>81.0</td>
<td>72.0</td>
<td>63.0</td>
<td>54.0</td>
<td>45.0</td>
<td>36.0</td>
<td>27.0</td>
<td>18.0</td>
<td>9.0</td>
<td></td>
</tr>
<tr>
<td>Production Level</td>
<td>9.00</td>
<td>9.00</td>
<td>9.00</td>
<td>9.00</td>
<td>9.00</td>
<td>9.00</td>
<td>9.00</td>
<td>9.00</td>
<td>9.00</td>
<td>9.00</td>
<td>9.00</td>
</tr>
<tr>
<td>Variable Op Cost Rate</td>
<td>10.2</td>
<td>10.4</td>
<td>10.6</td>
<td>10.8</td>
<td>11.0</td>
<td>11.3</td>
<td>11.5</td>
<td>11.7</td>
<td>12.0</td>
<td>12.2</td>
<td></td>
</tr>
<tr>
<td>Oil Price</td>
<td>25.0</td>
<td>25.0</td>
<td>25.0</td>
<td>25.0</td>
<td>25.0</td>
<td>25.0</td>
<td>25.0</td>
<td>25.0</td>
<td>25.0</td>
<td>25.0</td>
<td>25.0</td>
</tr>
<tr>
<td>Revenues</td>
<td>225.0</td>
<td>225.0</td>
<td>225.0</td>
<td>225.0</td>
<td>225.0</td>
<td>225.0</td>
<td>225.0</td>
<td>225.0</td>
<td>225.0</td>
<td>225.0</td>
<td>225.0</td>
</tr>
<tr>
<td>Production Cost</td>
<td>(96.8)</td>
<td>(98.6)</td>
<td>(100.5)</td>
<td>(102.4)</td>
<td>(104.4)</td>
<td>(106.4)</td>
<td>(108.4)</td>
<td>(110.4)</td>
<td>(112.6)</td>
<td>(114.7)</td>
<td></td>
</tr>
<tr>
<td>Cash Flow</td>
<td>128.2</td>
<td>126.4</td>
<td>124.5</td>
<td>122.6</td>
<td>120.6</td>
<td>118.6</td>
<td>116.6</td>
<td>114.6</td>
<td>112.4</td>
<td>110.3</td>
<td></td>
</tr>
<tr>
<td>Profit Sharing</td>
<td>(32.1)</td>
<td>(31.6)</td>
<td>(31.1)</td>
<td>(30.6)</td>
<td>(30.2)</td>
<td>(29.7)</td>
<td>(29.2)</td>
<td>(28.6)</td>
<td>(28.1)</td>
<td>(27.6)</td>
<td></td>
</tr>
<tr>
<td><strong>Net Cash Flows</strong></td>
<td>96.2</td>
<td>94.8</td>
<td>93.4</td>
<td>91.9</td>
<td>90.5</td>
<td>89.0</td>
<td>87.5</td>
<td>85.9</td>
<td>84.3</td>
<td>82.7</td>
<td></td>
</tr>
<tr>
<td><strong>PV of Cash Flows</strong></td>
<td><strong>696.6</strong></td>
<td>731.4</td>
<td>667.0</td>
<td>600.8</td>
<td>532.8</td>
<td>463.0</td>
<td>391.1</td>
<td>317.2</td>
<td>241.3</td>
<td>163.1</td>
<td>82.7</td>
</tr>
<tr>
<td><strong>Cash Flow Ratio</strong></td>
<td>13.1%</td>
<td>14.2%</td>
<td>15.5%</td>
<td>17.3%</td>
<td>19.5%</td>
<td>22.8%</td>
<td>27.6%</td>
<td>35.6%</td>
<td>51.7%</td>
<td>100.0%</td>
<td></td>
</tr>
</tbody>
</table>

2008 CPARM Report
Three Uncertain Factors

☐ Production level
  ■ Triangle Distribution
  ■ Auto-correlated

☐ Mineral price
  ■ GBM process-this is essentially a process that assumes prices are modeled by a lognormal distribution in each period, and the prices are auto-correlated.

☐ Variable operating costs
  ■ GBM process
Managerial Flexibility (Real Options)

- Three alternatives at the end of year 5:
  - continuing the project
  - buying out the partner’s 25% share for $40 million
  - selling your share for $100 million (divesting)
Available Approaches and Drawbacks

- Multi-factor Decision Tree—one chance node for each uncertainty in each period
  - The Tree size grows exponentially
  - Hard to incorporate correlations among factors

- Simulation
  - Hard to implement for complicated real options (e.g., early exercise options, compound options)

- Other multi-factor approaches based on estimating the uncertain value of the project over time
  - Need specific presumption about the project value distribution (e.g. GBM, mean reverting)
A Snapshot of the 3 factor Decision Tree

- 30 chance nodes
- Three factors in each period
- Ten periods
- The tree is too large to solve using commercial software (DPL)
The Invalidity of the GBM Assumption

Percentile mismatch of the original model and the model under GBM assumption means that the solution will be a poor approximation.

Figure 1: PVt Percentile Comparison

![Figure 1: PVt Percentile Comparison](image_url)
## Implied Binomial Tree

<table>
<thead>
<tr>
<th>Step</th>
<th>State</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>778.48</td>
</tr>
<tr>
<td>1</td>
<td>1213.51</td>
</tr>
<tr>
<td>2</td>
<td>1473.68</td>
</tr>
<tr>
<td>3</td>
<td>1748.36</td>
</tr>
<tr>
<td>4</td>
<td>2114.03</td>
</tr>
<tr>
<td>5</td>
<td>2654.54</td>
</tr>
<tr>
<td>6</td>
<td>3074.07</td>
</tr>
<tr>
<td>7</td>
<td>3045.09</td>
</tr>
<tr>
<td>8</td>
<td>2274.41</td>
</tr>
<tr>
<td>9</td>
<td>1184.90</td>
</tr>
<tr>
<td>10</td>
<td>10</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Step</th>
<th>State</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>492.12</td>
</tr>
<tr>
<td>5</td>
<td>745.11</td>
</tr>
<tr>
<td>6</td>
<td>906.03</td>
</tr>
<tr>
<td>7</td>
<td>1012.21</td>
</tr>
<tr>
<td>8</td>
<td>1094.26</td>
</tr>
<tr>
<td>9</td>
<td>904.39</td>
</tr>
<tr>
<td>10</td>
<td>888.11</td>
</tr>
<tr>
<td>State</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>853.67</td>
</tr>
<tr>
<td>6</td>
<td>848.35</td>
</tr>
<tr>
<td>7</td>
<td>631.08</td>
</tr>
<tr>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>9</td>
<td>8</td>
</tr>
<tr>
<td>10</td>
<td>8</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Step</th>
<th>State</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>268.50</td>
</tr>
<tr>
<td>4</td>
<td>453.57</td>
</tr>
<tr>
<td>5</td>
<td>599.77</td>
</tr>
<tr>
<td>6</td>
<td>692.82</td>
</tr>
<tr>
<td>7</td>
<td>599.45</td>
</tr>
<tr>
<td>8</td>
<td>593.26</td>
</tr>
<tr>
<td>9</td>
<td>548.42</td>
</tr>
<tr>
<td>10</td>
<td>451.07</td>
</tr>
<tr>
<td>State</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>336.12</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>9</td>
<td>8</td>
</tr>
<tr>
<td>10</td>
<td>8</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Step</th>
<th>State</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>151.99</td>
</tr>
<tr>
<td>3</td>
<td>239.64</td>
</tr>
<tr>
<td>4</td>
<td>347.23</td>
</tr>
<tr>
<td>5</td>
<td>355.40</td>
</tr>
<tr>
<td>6</td>
<td>341.49</td>
</tr>
<tr>
<td>7</td>
<td>294.77</td>
</tr>
<tr>
<td>8</td>
<td>179.02</td>
</tr>
<tr>
<td>9</td>
<td>117.12</td>
</tr>
<tr>
<td>10</td>
<td>105.27</td>
</tr>
<tr>
<td>State</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>5</td>
<td>8</td>
</tr>
<tr>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>9</td>
<td>8</td>
</tr>
<tr>
<td>10</td>
<td>8</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Step</th>
<th>State</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>7.87</td>
</tr>
<tr>
<td>2</td>
<td>9.28</td>
</tr>
<tr>
<td>3</td>
<td>7.67</td>
</tr>
<tr>
<td>4</td>
<td>4.75</td>
</tr>
<tr>
<td>5</td>
<td>4.09</td>
</tr>
<tr>
<td>6</td>
<td>2.18</td>
</tr>
<tr>
<td>7</td>
<td>2.18</td>
</tr>
<tr>
<td>8</td>
<td>2.18</td>
</tr>
<tr>
<td>9</td>
<td>2.18</td>
</tr>
<tr>
<td>10</td>
<td>2.18</td>
</tr>
</tbody>
</table>

The problem is solved easily in an Excel spreadsheet and the solution is accurate.
Advantage of The Implied Binomial Tree

- Relaxes the limiting assumption that the project cash flows follow a specific distribution, and allows general smooth probability distributions for the project cash flows in real options valuation.
- Allows for skewness and non-constant volatilities in the project cash flow process.
- Reduces the multi-factor problem into a simple implied binomial tree of a manageable size.