Resource Analysis to Improve Recovery of Unconventional Oil and Gas

Don Remson
Energy Systems Analysis Team

U.S. Department of Energy
National Energy Technology Laboratory
Mastering the Subsurface Through Technology Innovation, Partnerships and Collaboration: Carbon Storage and Oil and Natural Gas Technologies Review Meeting
August 13-16, 2018
Discussion Agenda

• Scope and Overview

• Pilot Project
  – Study Area
  – Data
  – Results and Findings

• Next Steps/Ongoing Work

• Questions/Discussions

Scope and Objectives
Detailed Analysis of Industry Performance in Marcellus Shale

• Evaluate region-specific industry performance data with the goal of identifying R&D needs conducive to improving the recovery of oil and gas in unconventional reservoirs.

  – Apply regression-style techniques to develop a model capable of predicting EUR based on available data parameters.
  – Test several machine learning regression algorithms and assess relevance in O&G applications.
  – Use sensitivity analysis or other means to quantify the relative contribution of each input parameter on productivity.
  – Identify most critical research needs and pass that information to fundamental researchers.
Pilot Evaluation – Western Marcellus
Western Marcellus Shale – Wet Gas Region; 2007 Through 2016 1st Production Year Wells

<table>
<thead>
<tr>
<th>Type</th>
<th>Name</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Geology</td>
<td>Surface Hole Latitude</td>
<td>DrillingInfo (DI)</td>
</tr>
<tr>
<td></td>
<td>Surface Hole Longitude</td>
<td>DI</td>
</tr>
<tr>
<td></td>
<td>GR</td>
<td>Well Logs, DI</td>
</tr>
<tr>
<td></td>
<td>Thickness</td>
<td>Well Logs, DI, Lit. Review</td>
</tr>
<tr>
<td></td>
<td>$R_0$ (VR)</td>
<td>Core Data, DI, Lit. Review</td>
</tr>
<tr>
<td></td>
<td>True Vertical Depth</td>
<td>DI</td>
</tr>
<tr>
<td></td>
<td>First 12m GOR</td>
<td>Calculated</td>
</tr>
<tr>
<td>Technology</td>
<td>Perf. Interval Length</td>
<td>DI</td>
</tr>
<tr>
<td></td>
<td>Total Additive Per ft</td>
<td>DI</td>
</tr>
<tr>
<td></td>
<td>Total Fluid Per ft</td>
<td>DI</td>
</tr>
<tr>
<td></td>
<td>Total Proppant Per ft</td>
<td>DI</td>
</tr>
<tr>
<td></td>
<td>Azimuth</td>
<td>DI</td>
</tr>
<tr>
<td></td>
<td>Spacing</td>
<td>DI</td>
</tr>
<tr>
<td></td>
<td>Pad Drilled (Y/N)</td>
<td>DI</td>
</tr>
<tr>
<td>Prod.</td>
<td>First 12m Production</td>
<td>DI</td>
</tr>
</tbody>
</table>

![Map of Western Marcellus showing study extent, surface gas, and liquid gas regions.](image)
Justification for Use of 1st Year Production

- Not a predicted value.
- Explicitly measured.
- Strongly correlated to predicted EUR.
- Better parameter for pilot-testing machine learning.
Machine Learning Framework
To Evaluate the Impact of Technology and Geology Parameters on Well Productivity
Modeling Training Results Overview

Western Marcellus Predictive Model

- Nine algorithms with various parameter combinations (up to a total of 14) were tested in this study to compare model performance.
- Non-linear algorithms performed better, indicating complexity in predicting production.

Model performance (Kernel Ridge)
Assessing Parameter Impact on Accuracy
R² Loss Evaluation on Down-Selected Parameter Set

Initial Model Parameter Set
(All 14 Parameters Included)

- Perf. Length
- Location
- Additive
- Gamma Ray
- %Ro
- 12m GOR
- Thickness
- Pad Drill
- TVD
- Azimuth
- Water
- Proppant
- Spacing

Impact to baseline accuracy (all parameters included)
Initial R² = 0.85

Finalized Model Parameter Set
(10 Parameters Included)

- Perf. Length
- %Ro
- Water
- Thickness
- Gamma Ray
- TVD
- Additive
- Pad Drilled
- 12m GOR
- Spacing

Impact to baseline accuracy (10 parameters included)
Initial R² = 0.83
Pilot Study Conclusions

- Publicly available data can be used to develop reasonably performing regression models that can predict well productivity.
- Geology and technology parameters are needed in combination, in order to fully explain variance in well productivity.
- There is a need for expanded data sets, both in number of samples and in number of parameters in each sample.
- Early sensitivity analysis shows that there is room for optimization in all wells analyzed.
Next Steps/Ongoing Work
Expanded Study Area and Well Counts

The data was downloaded from Drilling Info at Marcellus formation on May 16, 2018.

<table>
<thead>
<tr>
<th>Type of Data</th>
<th>2007</th>
<th>2008</th>
<th>2009</th>
<th>2010</th>
<th>2011</th>
<th>2012</th>
<th>2013</th>
<th>2014</th>
<th>2015</th>
<th>2016</th>
<th>2017</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Newly Added Wells</td>
<td>2</td>
<td>15</td>
<td>121</td>
<td>451</td>
<td>845</td>
<td>1,107</td>
<td>1,035</td>
<td>854</td>
<td>550</td>
<td>406</td>
<td>937</td>
<td>6,323</td>
</tr>
<tr>
<td>Western Marcellus Study</td>
<td>8</td>
<td>26</td>
<td>117</td>
<td>212</td>
<td>365</td>
<td>488</td>
<td>641</td>
<td>694</td>
<td>624</td>
<td>393</td>
<td>3,568</td>
<td></td>
</tr>
</tbody>
</table>

9,891

More data could reduce training vs. validation gap
### Expanded Evaluation – Marcellus Shale

Marcellus Shale – 2007 Through 2017 1\textsuperscript{st} Production Year Wells

<table>
<thead>
<tr>
<th>Type</th>
<th>Name</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Geology</td>
<td>Surface Hole Latitude</td>
<td>DrillingInfo (DI)</td>
</tr>
<tr>
<td></td>
<td>Surface Hole Longitude</td>
<td>DI</td>
</tr>
<tr>
<td></td>
<td>GR, Porosity, Res, Den</td>
<td>Well Logs, DI</td>
</tr>
<tr>
<td></td>
<td>Thickness</td>
<td>Well Logs, DI, Lit. Review</td>
</tr>
<tr>
<td></td>
<td>$R_0$ (VR)</td>
<td>Core Data, Lit. Review</td>
</tr>
<tr>
<td></td>
<td>True Vertical Depth</td>
<td>DI</td>
</tr>
<tr>
<td></td>
<td>Upper, Lower, All Marcellus</td>
<td>Well Logs</td>
</tr>
<tr>
<td></td>
<td>First 12m GOR</td>
<td>Calculated</td>
</tr>
<tr>
<td>Technology</td>
<td>Perf. Interval Length</td>
<td>DI</td>
</tr>
<tr>
<td></td>
<td>Total Additive Per ft</td>
<td>DI</td>
</tr>
<tr>
<td></td>
<td>Total Fluid Per ft</td>
<td>DI</td>
</tr>
<tr>
<td></td>
<td>Total Proppant Per ft</td>
<td>DI</td>
</tr>
<tr>
<td></td>
<td>Azimuth</td>
<td>DI</td>
</tr>
<tr>
<td></td>
<td>Spacing</td>
<td>DI</td>
</tr>
<tr>
<td></td>
<td>Pad Drilled (Y/N)</td>
<td>DI</td>
</tr>
<tr>
<td>Prod.</td>
<td>First 12m Production</td>
<td>DI</td>
</tr>
</tbody>
</table>

![Map showing study extent with symbols for Wet Gas and Dry Gas](image-url)
Expanding the Geologic Dataset

Data Acquisition and Interpretation

Stratigraphic Property Assessment

Geologic Property Contouring

- Well log correlation
- Core data
- Lit. data

- Parameter values
- Thickness

- Depth
- Zonal segregation

Ave. Gamma Ray

Bulk Density
Preliminary – Geologic Assessment

Isopach and Thermal Maturity

- Well log interpretation completed to assess geologic factors across play.
Recovery Factor (RF) Assessment
Marcellus Shale – West Virginia

- RF is the ratio of the EUR of a specific entity (i.e., well, lease area, or play) divided by the total in-place resource.

- Acquire OGIP data.
- Evaluate RF for areas totally developed or nearly developed.
- Use info to inform the regression analysis if possible.
- Analyze the data parameters to determine their individual impact on well productivity (EUR) and RF.
- Collaboration with the West Virginia Geologic Survey.

Boswell, R. 2017 - Recovery Efficiency in UOG Development
Desired data sets

- Only partial understanding can be attained from publicly-available data/information alone.
  - State reporting requirements strongly influence data availability and quality across plays

- Expanded datasets would enable for refined models, and enable better determination of parameters influencing production.

- Desired datasets:
  - Well logs (i.e. .las files)
  - Completion-related information (i.e. stage count, total perforations, and pressures)
  - Additive type, proppant size and type
  - Well orientation (toe-up vs. toe down; % in zone)
  - Well spacing
  - Pre-stimulation pay-zone pore pressures
  - Geochemical and geophysical data
  - Natural fracture extent
  - Others...
Acknowledgements

NETL Research & Innovation Center
Kristin Gerdes – Associate Director Systems Engineering & Analysis (SEA) Division
Peter Balash – Energy Systems Analysis Team (ESAT) Supervisor
Ale Hakala – Onshore Unconventional Resources Field Work Proposal Portfolio Lead

Mission Execution and Strategic Analysis (contractors)
Derek Vikara
Chung Yan Shih
Anna Wendt
ShangMin Lin
Aranya Venkatesh
Questions ?
Backup Slides
Impact of Correlated Parameters on Accuracy

Water and Proppant Correlation

• Volumes of water and proppant injected were found to be strongly correlated.
• Should either of the two parameters be excluded in model training, the other compensates, suggesting that neither parameter has importance.
• But, when both parameters are removed, the test scores drop considerably.
Variation in Parameter Impact on Accuracy
Comparison of Different Studies Predicting Production

Fuzzy Pattern Recognition

Wolfcamp shale – Delaware Basin, Texas
Battelle / Baker Hughes - 2015

Focused Marcellus Region – Appalachian Basin
WVU / Intelligent Solutions - 2017

Western Marcellus – Appalachian Basin
NETL - 2018
Optimization of Well Design
Modifying Additive, Fluids, and Proppant per Perforated Interval Length

- Most wells in the preliminary test showed that the design can be improved.
- Increasing parameter values does always yield best results.
- Additive was decreased to optimize well design for a case-study well.
Production Performance Summary
Marcellus Shale – All Wells (2007 – 2017)

EUR Per Well

First 6-month Cumulative Production

Peak Rate

EUR per 1,000-foot perforated interval

First 12-month Cumulative Production

Initial Decline Rate (Di)

Newly Added Wells
Western Marcellus Study
Well Completion/Design Summary
Marcellus Shale – All Wells (2007 – 2017)

- **Total Additive Injected**
- **Proppant per 1,000-foot perforated interval**
- **Water per 1,000-foot perforated interval**

Graphs showing trends in perforated interval length, total proppant, total additive injected, and total water for the years 2007 to 2017.
Recovery Factor (RF) Assessment

- RF is a concept not readily applied to UOG.
- EUR is a function of the marriage of technology and geology.
  - Technology changes with time (future >>> past).
  - Geology changes with location (core >>> margins).
  - Assessments can get EUR very wrong for either (both) reasons.
- In-place volumes subject to great uncertainty.
- RF is better with gas. Also better with depth/pressure.
- RF is likely better than we think in core areas and worse than we think at the margins.
- Minor improvements in RF can be directly translated into immense and tangible economic and national security benefits.

<table>
<thead>
<tr>
<th>Table 1: Major Shale Gas Plays: Reserves, Recovery Factors, Production Potential, Well-Productivities – 2013</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plays</td>
</tr>
<tr>
<td>Gas-in-place, bcf/sq mile</td>
</tr>
<tr>
<td>Year-end Output, bcf/d</td>
</tr>
<tr>
<td>Cumulative Production, tcf</td>
</tr>
<tr>
<td>Reserve/EUR, tcf</td>
</tr>
<tr>
<td>Recovery Factor, %</td>
</tr>
<tr>
<td>Production Potential, bcf/d</td>
</tr>
<tr>
<td>Present Well-Productivity, Mcf/d/well</td>
</tr>
<tr>
<td>Year-end Producing Wells</td>
</tr>
<tr>
<td>Current 180-day Well IPs, MMcf/d</td>
</tr>
<tr>
<td>Well-Productivity Decline Rate, %/year</td>
</tr>
<tr>
<td>Well EUR, bcf/dwell</td>
</tr>
<tr>
<td>Well-Productivity by 2020, Mcf/d/well</td>
</tr>
</tbody>
</table>

Sandrea and Sandrea, OGJ, 2014
Shale Well Production Economic Model

- Well spacing/design typically based on spacing patterns that yield the highest NPV.
- Coupling data-driven predictive model with cash flow model enables economic evaluation of well/pad/lease optimization.
- Enables comparison of improving recovery (DOE mission) vs. maximizing profitability/NPV (Industry mission).

Influencing Factors
- Well spacing and SRV.
- Well interference.
- Over-capitalized field development.
- Economic vs. technically recoverable.
## Parameter Overview by Well Vintage
### Average Values

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<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>Perforated lateral length</td>
<td>foot</td>
<td>2,712</td>
<td>2,258</td>
<td>2,821</td>
<td>3,313</td>
<td>3,441</td>
<td>4,090</td>
<td>4,712</td>
<td>5,555</td>
<td>6,098</td>
<td>6,612</td>
<td>193%</td>
</tr>
<tr>
<td>Water used for hydraulic fracturing</td>
<td>bbl/1,000 foot perforated</td>
<td>NA</td>
<td>NA</td>
<td>27,824</td>
<td>34,573</td>
<td>33,317</td>
<td>29,529</td>
<td>35,939</td>
<td>41,853</td>
<td>39,685</td>
<td>42,983</td>
<td>54%</td>
</tr>
<tr>
<td>Proppant used for hydraulic fracturing</td>
<td>pound/foot perforated</td>
<td>NA</td>
<td>NA</td>
<td>1,251</td>
<td>444</td>
<td>672</td>
<td>1,127</td>
<td>1,521</td>
<td>1,733</td>
<td>1,711</td>
<td>1,975</td>
<td>345%</td>
</tr>
<tr>
<td>Additive used for hydraulic fracturing</td>
<td>pound/foot perforated</td>
<td>15</td>
<td>25</td>
<td>72</td>
<td>58</td>
<td>63</td>
<td>81</td>
<td>61</td>
<td>99</td>
<td>66</td>
<td>143</td>
<td>850%</td>
</tr>
<tr>
<td>Well azimuth trajectory*</td>
<td>degree</td>
<td>139</td>
<td>139</td>
<td>128</td>
<td>132</td>
<td>131</td>
<td>137</td>
<td>139</td>
<td>140</td>
<td>138</td>
<td>140</td>
<td>10%</td>
</tr>
<tr>
<td>Well spacing</td>
<td>foot</td>
<td>601</td>
<td>2,709</td>
<td>1,617</td>
<td>1,360</td>
<td>1,083</td>
<td>1,251</td>
<td>1,283</td>
<td>1,167</td>
<td>1,313</td>
<td>1,328</td>
<td>351%</td>
</tr>
<tr>
<td>GOR cumulative at 12 months</td>
<td>mcf/bbl</td>
<td>4,445</td>
<td>2,401</td>
<td>2,870</td>
<td>3,793</td>
<td>4,397</td>
<td>3,773</td>
<td>3,763</td>
<td>3,650</td>
<td>4,729</td>
<td>7,272</td>
<td>203%</td>
</tr>
<tr>
<td>True vertical depth</td>
<td>foot</td>
<td>7,088</td>
<td>6,799</td>
<td>7,528</td>
<td>7,868</td>
<td>7,768</td>
<td>7,435</td>
<td>7,469</td>
<td>7,494</td>
<td>7,588</td>
<td>7,804</td>
<td>16%</td>
</tr>
<tr>
<td>Thickness</td>
<td>foot</td>
<td>30</td>
<td>29</td>
<td>28</td>
<td>28</td>
<td>29</td>
<td>28</td>
<td>29</td>
<td>29</td>
<td>29</td>
<td>29</td>
<td>7%</td>
</tr>
<tr>
<td>Gamma ray</td>
<td>API</td>
<td>261</td>
<td>259</td>
<td>268</td>
<td>268</td>
<td>271</td>
<td>265</td>
<td>270</td>
<td>276</td>
<td>271</td>
<td>273</td>
<td>7%</td>
</tr>
<tr>
<td>Thermal maturity</td>
<td>% R₀</td>
<td>1.5305</td>
<td>1.5304</td>
<td>1.6007</td>
<td>1.6373</td>
<td>1.6305</td>
<td>1.5708</td>
<td>1.5894</td>
<td>1.5964</td>
<td>1.6238</td>
<td>1.6555</td>
<td>8%</td>
</tr>
</tbody>
</table>
# Predictive Models for 12-mo Productivity

## Comparative Analysis

<table>
<thead>
<tr>
<th>Team KeyLogic</th>
<th>MIT [1]</th>
<th>BEG UT Austin [2, 3]</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Technology</strong></td>
<td>Perforated Lateral Length, Proppant, Fluid (Water), Pad Drilled, Well Spacing</td>
<td>Lateral Length, Fluid (Water), Proppant</td>
</tr>
<tr>
<td><strong>Geology</strong></td>
<td>Thickness, VR, Gamma Ray, Depth, Location</td>
<td>Location</td>
</tr>
<tr>
<td><strong>Algorithm</strong></td>
<td>Kernel Ridge</td>
<td>Regression-Kriging</td>
</tr>
<tr>
<td><strong>Prediction</strong></td>
<td>12 Month Cumulative Gas</td>
<td>12 Month Cumulative Gas</td>
</tr>
<tr>
<td><strong>Scores</strong></td>
<td>MASE 0.28 (Lower the Better)</td>
<td>0.62</td>
</tr>
<tr>
<td></td>
<td>R² 0.83 (Higher the Better)</td>
<td>---</td>
</tr>
</tbody>
</table>

## Key Take Aways

- Using a comprehensive geology data set instead of location data (latitude and longitude) will provide more accurate production outlooks
- Initial results suggested that well completion designs can still be optimized to improve the overall production
- RK modelling can be used to develop supply curves for different economic scenarios or optimize design parameters at different well locations
- To prevent overly optimistic potential well production projections, the chosen modeling method must consider the influence of location
- Lateral length does not significantly affect recovery factor
- Completion type and well spacing were revealed to be the most significant factors affecting productivity
- Recovery factor can be increased in the low to mid productivity range

---


Geology
Differences Between Coordinates and Geology (gamma ray, thickness, $R_o$)

• Algorithms trained exclusively with either (1) spatial coordinates, or (2) GR, thickness, and $R_o$.

• Production varies spatially, likely due to changes in geologic quality.
  – Most studies use coordinates (lat/long) as a proxy for geology.
  – For this study, the geologic assessment enabled extrapolation of geologic parameters to entire study area.
  – Extrapolation imposes less certainty than explicit well-specific measurements.

• Results indicate that geologic parameters acquired (despite extrapolation) have similar test score trend as using coordinates.
Effect of Spacing
“Distance to Nearest Well” and “Pad Drill”

- Accuracy remains after removing both spacing related parameters.
  - It is known that wells can interfere when drilling too close to each other.
- Possible conclusions:
  - Noisy data about well spacing (i.e., not accurately reflecting well spacing).
  - Wells in the dataset are at spacings that are not causing interference or “frac hits.”
- R&D Pursuit: Evaluation of optimal spacing in Marcellus to maximize production and improve RF.
  - Parent/Child well impacts.

Note: This baseline does not include coordinates and azimuth
Desired Datasets

- Only partial understanding can be attained from publicly available data/information alone.
  - State reporting requirements strongly influence data availability and quality across plays.
- Expanded datasets would enable for refined models, and enable better determination of parameters influencing production.
- Desired datasets:
  - Well logs (i.e., .las files)
  - Completion-related information (i.e. stage count, total perforations, and pressures)
  - Additive type, proppant size and type
  - Well orientation (toe-up vs. toe down; % in zone)
  - Pre-stimulation pay-zone pore pressures
  - Lateral trajectory data
  - Geochemical and geophysical data
  - Natural fracture extent
  - Others…
Methods to Determining R&D Needs
Parameter Impact Assessment
Requires Various Approaches to Extract Actual Parametric Impact

- Removing Fluid or Proppant alone does not show significant impact to the overall accuracy.
- However, removing both parameters shows the real impact of fracture fluid and proppant.
- This problem is non-linear and certain parameters are likely collinear and/or have high degree interaction.
- Simple one-at-a-time sensitivity tests not suitable for identifying the parameter importance.
  - Monte-Carlo variance-based approach.
  - Sobol total index approach
  - Decision tree analysis.

![Bar chart showing test scores for different parameters]
Decision Tree Analysis

Exploration of Parameters that Contribute to “Extreme” Well Performance

- Dataset with low and high performing wells.
  - <25th percentile (low) and >75th percentile (high).
- Used key features to “classify” wells.
- Preliminary results show that:
  - All left branches at each node = True, all right branches at each node = False.
  - gini is a ‘score’ for each node (zero when all cases in a node are classified into a single category).
  - Value represents number of samples classified into each category [Low, High].
# Literature Review

## Machine Learning for Unconventional Oil and Gas Applications

<table>
<thead>
<tr>
<th>Study</th>
<th>Region</th>
<th>Methods</th>
<th>Data used</th>
<th>Key parameters/ findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zhou et al</td>
<td>West Virginia</td>
<td>Multiple linear regression, principal components analysis and k-means</td>
<td>Fracture fluid, proppant, true vertical depth (TVD), lateral length (LL), stages, treatment rate, thermal maturity (TM), thickness</td>
<td>Stages, lateral length</td>
</tr>
<tr>
<td>Izadi et al</td>
<td>Bakken</td>
<td>Multiple linear regression, boosted tree models</td>
<td>well location, LL, azimuth, stages, fracturing fluid, proppant type and volumes</td>
<td>Well location, proppant quantity</td>
</tr>
<tr>
<td>Schuetter et al</td>
<td>Wolfcamp shale</td>
<td>$R^2$-loss for model selection, decision trees</td>
<td>Latitude and longitude, TVD, LL, proppant quantity and concentration, stages</td>
<td>TVD, proppant quantity, LL</td>
</tr>
<tr>
<td>Montgomery and O'Sullivan</td>
<td>Williston Basin</td>
<td>Multiple linear regression, fixed-effects regression, kriging</td>
<td>Latitude and longitude, LL, water, proppant volumes</td>
<td>Location data,</td>
</tr>
<tr>
<td>Mohaghegh et al</td>
<td>Marcellus</td>
<td>Neural networks, Monte Carlo simulation, optimization</td>
<td>TVD, thickness, porosity, TOC, LL, clusters per stage, clean volume, proppant quantity per ft LL</td>
<td>Net thickness, well spacing, LL</td>
</tr>
<tr>
<td>Mishra et al</td>
<td>Literature review</td>
<td>Decision trees, gradient boosting machine, support vector machine, neural networks, kriging</td>
<td>1) cross-validation typically not been done in O&amp;G studies 2) most studies analyze only a handful of regression models 3) these studies typically ignore records with missing data points 4) they do not typically evaluate relative variable importance.</td>
<td></td>
</tr>
</tbody>
</table>