Building and Using Public Data Sets to Assess Oil and Gas Well Integrity

NETL Well Integrity Workshop:
Identifying well integrity research needs for subsurface energy infrastructure

June 1, 2021

Bureau of Oil and Gas Planning and Program Management
1. Public “Well Integrity” Data Sets
2. Data Set Integration – Plugging Case Studies
   a) Orphan Well Plugging Costs
   b) Geology
   c) Plugging Efficiency
   d) Plugging Integrity
   e) Risks to Plugging Efficiency and Integrity
3. Predicting Failed Plug Integrity Controls
4. Summary
Public “Well Integrity” Data Sets

• Why “Well Integrity” as opposed to Well Integrity?
  
  ▪ Advances in statistical modeling/machine learning have allowed data scientists and analysts to rapidly test many hypotheses in both regression and classification model contexts.
  
  ▪ In many cases, the best predictor variable may not be digitized or otherwise available, but a surrogate variable may be readily accessible.
  
  ▪ Well integrity models may be complex and are likely multivariate in nature...predictors are found in “unexpected” places...hence we might keep an open mind about factors that influence “well integrity!”
Pennsylvania agencies such as the Department of Conservation and Natural Resources (DCNR) and the Department of Environmental Protection (DEP) have amassed many data sets that can be used to evaluate well integrity.

Some resources to consider:

- DCNR’s Geologic Publications and Data
- DCNR’s EDWIN
- DEP’s Plugging Contract Data (available upon request)
- DEP’s Oil and Gas Reports
Orphan Well Plugging Costs

- What predictors drive plugging costs?
  - Historical plugging contract data were paired with geologic data
  - Three machine learning algorithms were used to develop multivariate regression models
    1. Multiple Linear Regression (MLR)
    2. Generalized Additive Model (GAM)
    3. Support Vector Machine Regression (SVMR)
Orphan Well Plugging Costs

• What predictors drive plugging costs?

Data Set Integration – Plugging Case Studies
Orphan Well Plugging Costs

- What predictors drive plugging costs?

<table>
<thead>
<tr>
<th>Model</th>
<th>Predictors</th>
<th>Basis Functions</th>
<th>Tuning Parameters</th>
<th>RMSE</th>
<th>( R^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>MLR</td>
<td>mean contract days per well, mean depth, mean center longitude</td>
<td>linear</td>
<td>not applicable</td>
<td>$23,200</td>
<td>0.7060</td>
</tr>
<tr>
<td>GAM</td>
<td>mean contract days per well, mean depth, mean center longitude</td>
<td>spline</td>
<td>df = 3</td>
<td>$19,200</td>
<td>0.7998</td>
</tr>
<tr>
<td>SVMR</td>
<td>mean contract days per well, mean depth, mean center longitude</td>
<td>radial kernel</td>
<td>cost = 10, gamma = 1</td>
<td>$21,500</td>
<td>0.7551</td>
</tr>
</tbody>
</table>

Data Set Integration – Plugging Case Studies
Orphan Well Plugging Costs

- What predictors drive plugging costs?

Is plugging indeed more difficult moving to the southeast and are there implications for integrity?
Data Set Integration – Plugging Case Studies

Geology

- Appalachian basin evolution

Deep structure, antecedent topography, thin-skin tectonics – translational movement along Salina detachment and fold-belt development
Plugging Efficiency

- Plugging efficiency can be determined by comparing the percentage of the well bore plugged to the time spent on the well, i.e., rig time.
- The most efficient plug jobs are those where it takes the least amount of time to effectively plug the greatest percentage of the well bore.
**Data Set Integration – Plugging Case Studies**

**Plugging Efficiency**

27% of Cluster 1 “Inefficiency Envelope”

18% of Cluster 2 “Inefficiency Envelope”

---

**Solid Cement Plug:** only cement is run from bottom plug to top plug

**Cement & Gel Plug:** only cement and gel are run from bottom plug to top plug
Plugging Efficiency

Inefficiency Ratio = Efficiency “Outlier” per cell / Total Wells per cell
Plugging Integrity

*Integrity Ratio = Failed Plugs per cell / Total Wells per cell*
Data Set Integration – Plugging Case Studies

Risks to Plugging Efficiency and Integrity
Failed Plugs vs. Competent Plugs

• What predictors control plug integrity?
  
  o Plugging certificate information was integrated with geology and production information
  
  o Three machine learning algorithms were used to develop multivariate classification models
    1. Mixture Discriminant Analysis (MDA)
    2. Random Forests with Boosting (TBM)
    3. Support Vector Machine (SVM)
Predicting Failed Plug Integrity Controls

Confusion Matrix - SVM

<table>
<thead>
<tr>
<th></th>
<th>No Leak</th>
<th>Leak</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Leak</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>Leak</td>
<td>1</td>
<td>12</td>
</tr>
</tbody>
</table>

Many Type I errors (false positives)
One Type II errors (false negatives)

Confusion Matrix - MDA

<table>
<thead>
<tr>
<th></th>
<th>No Leak</th>
<th>Leak</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Leak</td>
<td>14</td>
<td>0</td>
</tr>
<tr>
<td>Leak</td>
<td>6</td>
<td>7</td>
</tr>
</tbody>
</table>

No Type I errors (false positives)
Many Type II errors (false negatives)

Confusion Matrix - TBM with Boosting

<table>
<thead>
<tr>
<th></th>
<th>No Leak</th>
<th>Leak</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Leak</td>
<td>14</td>
<td>0</td>
</tr>
<tr>
<td>Leak</td>
<td>9</td>
<td>4</td>
</tr>
</tbody>
</table>

SVM Model (AUC = 0.75)

Random
Predicting Failed Plug Integrity Controls

Failed Integrity

Integrity
Predicting Failed Plug Integrity Controls

LEAKING: NEWER/DEEPER/MORE PRODUCTION CASING ABANDONED IN PLACE/DEEPER TOP PLUG/SLIGHTLY LESS CEMENT

NOT LEAKING: OLDER/SHALLOWER/LESS PRODUCTION CASING ABANDONED IN PLACE/SHALLOWER TOP PLUG/SLIGHTLY MORE CEMENT

Conceptual Models
Summary

- Public data sets are rich sources of well integrity information
- Higher-cost orphan well plugging appears to be most influenced by depth, contract duration, and location
- Geologic factors may be influencing orphan well plugging costs, overall plugging efficiency/degree of difficulty, and plug integrity failure susceptibility
- From a well characteristics standpoint, failed plug integrity appears to be primarily associated with newer and deeper wells with more production casing abandoned in place
Thanks!

Questions?

Seth Pelepkko, P.G.
Program Manager, Subsurface
Bureau of Oil & Gas Planning & Program Management
717.772.2199
(mipelepko@pa.gov)

Elizabeth Cushman, P.G.
Geologist, Subsurface
Bureau of Oil & Gas Planning & Program Management
717.772.2199
(ecushman@pa.gov)

Many thanks to other DEP O&G Team Members: Serena Oldhousier and Harry Wise

Rewriting Pennsylvania’s Legacy (dep.pa.gov/legacywells)