Rapid Basin-scale Geologic Reconnaissance Using Advanced Unsupervised Learning

Didi Ooi
Drew Derenthal, Michael Ashby, Christian Noll
DISCLAIMER 2019

The contents of this presentation are for informational purposes only. Halliburton** makes no representation or warranty about the accuracy or suitability of the information provided in this presentation and any related materials. Nothing in this presentation constitutes professional advice or consulting services. No contractual relationship with Halliburton is established by attending or viewing this presentation. No rights to intellectual property of Halliburton are granted through this presentation. The opinions expressed in this presentation are those of the author and do not necessarily represent the views of Halliburton.

Key Takeaways

1. Unsupervised learning (UL) of well logs can provide data-driven insights into the combined effects of multiple properties that can be interpreted as the environmental of deposition on a basin-scale.

2. UL methods such as clustering can be evaluated using a combination of gap statistics, scree plot, Hartigan-Wong plot, silhouette coefficients and GEOLOGISTS’ input.

3. Clustering of averaged well logs within the Bone Spring formation show depositional model ties with the existing regional study.

4. Clustering all of the Delaware Basin core measurement data using Ward and Spectral algorithms show clear lithologic separation.

5. This set of UL workflows set the stage to aid geologic interpretations (2D and 3D), anomaly detection and as an input to other multivariate analyses.
Clustering: an Unsupervised Learning Method

An exploratory method for discovering inherent patterns or clusters

**Goal:** Model underlying structure and distribution in the data

**Methods:** Cluster analysis, dimensionality reduction

**Uses:** Creates new labels/classes, validate existing interpretation

**Example:** Market segmentation, targeted ads, movie recommendation
Towards a Basin-scale Lithofacies Model

- **Stats**: Synthesize tops and logs QCs and runs full suite of stats analyses
- **LogUL**: Clusters well logs per formation of interest or over set depth intervals
- **CoreUL**: Discover lithologic separation from raw core data
- **FSL**: Facies prediction using wireline log and core data

**Well Clusters**

**Core Clusters**

**Lithofacies Model**
Overview of Workflow

Wells with Logs

Preprocessing

Log Norm. Log QC Stats Calc.

Unsupervised Learning

Clustering

Log Clusters

Restrict training of supervised learning within well clusters

INPUT

OUTPUT

Gamma Ray, Caliper, Resistivity, Neutron Porosity, Bulk Density, DTC, Photoelectric Index etc…
Well Clustering
Why do we care?

Delaware Basin
~6.4 million acres!
100,000+ wells
10,000+ digital logs
1000+ core data

How can we use 10,000s well logs to rapidly understand the geology?
Delaware Basin: Bone Spring paleogeography

Basinal clastics surrounded by carbonate shelves.

Cyclic-reciprocal deposition a result of highly oscillatory sea-level

- Carbonates deposited during highstand cycles
- Siliciclastics deposited during lowstand cycles
Case Study: Bone Spring Cluster Analysis

- 5000 wells analyzed
- **Variables used:** Average and standard deviation of triple combo logs
- **Formation:** Each series/formation i.e. Bone Spring 1, 2, 3 grouped to one
- **Comparison of UL of raw logs with**
  - Existing Regional Study
  - 3rd Bone Spring Carbonate \( V_{\text{lime}} \) map
Comparing with regional interpretation

**Cluster 1:**
Low gamma ray, porosity
High resistivity, bulk density

**Interpretation:**
Carbonate-dominant on the proximal northwest and on the rim

<table>
<thead>
<tr>
<th>Cluster</th>
<th>GR</th>
<th>RESD</th>
<th>NPHI</th>
<th>RHOB</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>48</td>
<td>1740</td>
<td>0.06</td>
<td>2.62</td>
</tr>
<tr>
<td>2</td>
<td>56</td>
<td>726</td>
<td>0.09</td>
<td>2.59</td>
</tr>
<tr>
<td>3</td>
<td>70</td>
<td>326</td>
<td>0.13</td>
<td>2.56</td>
</tr>
<tr>
<td>4</td>
<td>68</td>
<td>53</td>
<td>0.15</td>
<td>2.56</td>
</tr>
</tbody>
</table>
Bone Spring Cluster 2

Comparing with regional interpretation

- **Cluster 2:**
  - ↑ gamma ray and porosity
  - ↓ resistivity and bulk density

**Interpretation:**
Decrease carbonate content, increase in sand on the shelf slope

### Table 1. Average logs per cluster

<table>
<thead>
<tr>
<th>Cluster</th>
<th>GR</th>
<th>RESD</th>
<th>NPHI</th>
<th>RHOB</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>48</td>
<td>1740</td>
<td>0.06</td>
<td>2.62</td>
</tr>
<tr>
<td>2</td>
<td>56</td>
<td>726</td>
<td>0.09</td>
<td>2.59</td>
</tr>
<tr>
<td>3</td>
<td>70</td>
<td>326</td>
<td>0.13</td>
<td>2.56</td>
</tr>
<tr>
<td>4</td>
<td>68</td>
<td>53</td>
<td>0.15</td>
<td>2.56</td>
</tr>
</tbody>
</table>
Comparing with regional interpretation

- **Cluster 3**:
  - ↑ gamma ray and porosity
  - ↓ resistivity and bulk density

**Interpretation:**
Lack of carbonate, more sandy. Clastic-dominant in the distal part of the basin towards the southeast.

**Table 1. Average logs per cluster**

<table>
<thead>
<tr>
<th>Cluster</th>
<th>GR</th>
<th>RESD</th>
<th>NPHI</th>
<th>RHOB</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>48</td>
<td>1740</td>
<td>0.06</td>
<td>2.62</td>
</tr>
<tr>
<td>2</td>
<td>56</td>
<td>726</td>
<td>0.09</td>
<td>2.59</td>
</tr>
<tr>
<td>3</td>
<td>70</td>
<td>326</td>
<td>0.13</td>
<td>2.56</td>
</tr>
<tr>
<td>4</td>
<td>68</td>
<td>53</td>
<td>0.15</td>
<td>2.56</td>
</tr>
</tbody>
</table>
Bone Spring Cluster Distribution

Cluster 1

Cluster 2

Cluster 3

by Mark Boehm, after Montgomery
The Skeptic Geologists: Comparison With Regional Study

Bone Spring Cluster Distribution

Cluster 1
Cluster 2
Cluster 3

Bone Spring Cross-Section

Cluster 1
Cluster 2
Cluster 3

after Aislyn Barclay
Clustering is an iterative process

Cannot be used blindly

Performance can be evaluated using a combination of:

- Silhouette Coefficient
- Elbow Plot - figure below
- Subject matter expertise

Silhouette Score: +0.36
Silhouette Score: -0.25

Goldilocks window!
Cluster analysis can pick out outliers

- Inadequate logged depth coverage OR
- Vintage logs OR
- Geologic interest?

**Takeaway:**
Well clustering reliable front-end workflow to analyze 1000s of logs in significantly reduced time
Facies Classification using unsupervised learning
Delaware Basin Core Data Clustering Analysis

Core Data:
TOC, XRD (Clastic, Carbonate, Clays)

Formation:
For apple-to-apple comparison and to reduce complexity, clustering is done for all formations through Delaware Basin

Method:
Ran 10 cluster algorithms with performance indicators

Results:
5 clusters determined using Ward and Spectral algorithm
Cluster Distribution in the Bone Spring

**Interpretation of clusters:**
1. Mixed Clastic-Carbonate
2. Siliciclastic-dominant
3. Carbonate-dominant
4. Organic-rich siliceous

**Results match field observations:**
Bone Spring is known for highly reciprocal (cyclical) sedimentation – interbedded aeolian sands and carbonate

**Takeaway:**
Cluster analyses of 1000s+ of core data can provide a quick mineralogy profile of the basin

**Table 2. Average core measurements per cluster**

<table>
<thead>
<tr>
<th>Facies</th>
<th>TOC</th>
<th>Clastic</th>
<th>Carbonate</th>
<th>Clay</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>2.6</td>
<td>51</td>
<td>8</td>
<td>36</td>
</tr>
<tr>
<td>1</td>
<td>2.3</td>
<td>44</td>
<td>39</td>
<td>13</td>
</tr>
<tr>
<td>2</td>
<td>6.0</td>
<td>65</td>
<td>6</td>
<td>22</td>
</tr>
<tr>
<td>3</td>
<td>1.1</td>
<td>20</td>
<td>72</td>
<td>5</td>
</tr>
<tr>
<td>4</td>
<td>2.0</td>
<td>69</td>
<td>9</td>
<td>19</td>
</tr>
</tbody>
</table>
Towards a Basin-scale Lithofacies Model

Stats:
Synthesize tops, logs. QC's and runs full suite of stats calculation on logs.

LogUL:
Clusters well logs per formation of interest or over set depth intervals

CoreUL:
Discover lithologic separation from raw core data

FSL:
Facies prediction using wireline log and core data

Well Clusters
Core Clusters
Lithofacies Model
1. Unsupervised learning (UL) of well logs can provide first-pass insights into the environmental of deposition on a basin-scale

2. UL methods such as clustering can be evaluated using a combination of gap statistics, scree plot, Hartigan-Wong plot, silhouette coefficients and GEOLOGISTS’ input

3. Clustering of averaged well logs within the Bone Spring formation show depositional model ties with the existing regional study

4. Clustering all of the Delaware Basin core measurement data using Ward and Spectral algorithms show clear lithologic separation

5. This set of UL workflows set the stage to aid geologic interpretations (2D and 3D), anomaly detection and as an input to other multivariate analyses
What’s Next?

**Workflow augmentation:**
- Sub-interval or member-level clustering
- Interpolated data clustering
- Probabilistic (soft) clustering
- Deep clustering (auto encoder)
- Log Character Clustering

**Feature engineering:**
- Basin-scale log normalization while maintaining geologic variability (GR, NPHI)
- Log coverage
- Pre-processing with dimensionality reduction
- Declustering and debiasing of dataset

*It’s so exciting!!*
Acknowledgments

Drew Derenthal  Michael Ashby  Christian Noll
Nikolaos Mitsakos  Luis Moraes  Ingrid Tobar  Seth Brazell  Dawn Hayes

Advanced Analytics and Emerging Technology
and Delaware Basin Subsurface Team
Geology has always been data-rich...

What’s your next classification project?
Towards a Basin-scale Lithofacies Model
An integrated ML approach using well logs and core in the Permian Basin

Didi Ooi
Drew Derenthal, Michael Ashby, Christian Noll
Your feedback is very important to us. Please open the LIFE2019 app to answer a few short questions on this presentation.