The 2019 Permian Basin Geophysical Society’s
60th Annual Exploration Meeting

Platform for Geophysical Knowledge Share in the Permian Basin since 1949

Thursday, May 9th, 2019
Petroleum Club of Midland

imagery courtesy of EAGLEEYE GEOSPATIAL
A Dawson Geophysical Company
Abstract

Over the last few years, because of the increase in low-cost computer power, individuals and companies have stepped up investigations into the use of machine learning in many areas of E&P. For the geosciences, the emphasis has been in reservoir characterization, seismic data processing, and to a lesser extent interpretation. The benefits of using machine learning (whether supervised or unsupervised) have been demonstrated throughout the literature, and yet the technology is still not a standard workflow for most seismic interpreters. This lack of uptake can be attributed to several factors, including a lack of software tools, clear and well-defined case histories and training. Fortunately, all these factors are being mitigated as the technology matures. Rather than looking at machine learning as an adjunct to the traditional interpretation methodology, machine learning techniques should be considered the first step in the interpretation workflow.

By using statistical tools such as Principal Component Analysis (PCA) and Self Organizing Maps (SOM) a multi-attribute 3D seismic volume can be “classified”. The PCA reduces a large set of seismic attributes both instantaneous and geometric, to those that are the most meaningful. The output of the PCA serves as the input to the SOM, a form of unsupervised neural network, which, when combined with a 2D color map facilitates the identification of clustering within the data volume. When the correct “recipe” is selected, the clustered or classified volume allows the interpreter to view and separate geological and geophysical features that are not observable in traditional seismic amplitude volumes. Seismic facies, detailed stratigraphy, direct hydrocarbon indicators, faulting trends, and thin beds are all features that can be enhanced by using a classified volume.

The tuning-bed thickness or vertical resolution of seismic data traditionally is based on the frequency content of the data and the associated wavelet. Seismic interpretation of thin beds routinely involves estimation of tuning thickness and the subsequent scaling of amplitude or inversion information below tuning. These traditional below-tuning-thickness estimation approaches have limitations and require assumptions that limit accuracy. The below tuning effects are a result of the interference of wavelets, which are a function of the geology as it changes vertically and laterally. However, numerous instantaneous attributes exhibit effects at and below tuning, but these are seldom incorporated in thin-bed analyses. A seismic multi-attribute approach employs self-organizing maps to identify natural clusters from combinations of attributes that exhibit below-tuning effects. These results may exhibit changes as thin as a single sample interval in thickness. Self-organizing maps employed in this fashion analyze associated seismic attributes on a sample-by-sample basis and identify the natural patterns or
clusters produced by thin beds. Examples of this approach to improve stratigraphic resolution in both the Eagle Ford play, and the Niobrara reservoir of the Denver-Julesburg Basin will be used to illustrate the workflow.

Introduction

Seismic multi-attribute analysis has always held the promise of improving interpretations via the integration of attributes which respond to subsurface conditions such as stratigraphy, lithology, faulting, fracturing, fluids, pressure, etc. The benefits of using machine learning (whether supervised or unsupervised) has been demonstrated throughout the literature and yet the technology is still not a standard workflow for most seismic interpreters. This lack of uptake can be attributed to several factors, including a lack of software tools, clear and well-defined case histories, and training. This paper focuses on an unsupervised machine learning workflow utilizing Self-Organizing Maps (Kohonen, 2001) in combination with Principal Component Analysis to produce classified seismic volumes from multiple instantaneous attribute volumes. The workflow addresses several significant issues in seismic interpretation: it analyzes large amounts of data simultaneously; it determines relationships between different types of data; it is sample based and produces high-resolution results and, reveals geologic features that are difficult to see in conventional approaches.

Principal Component Analysis (PCA)

Multi-dimensional analysis and multi-attribute analysis go hand in hand. Because individuals are grounded in three-dimensional space, it is difficult to visualize what data in a higher number dimensional space looks like. Fortunately, mathematics doesn’t have this limitation and the results can be easily understood with conventional 2D and 3D viewers.

Working with multiple instantaneous or geometric seismic attributes generates tremendous volumes of data. These volumes contain huge numbers of data points which may be highly continuous, greatly redundant, and/or noisy. (Coleou et al., 2003). Principal Component Analysis (PCA) is a linear technique for data reduction which maintains the variation associated with the larger data sets (Guo and others, 2009; Haykin, 2009; Roden and others, 2015). PCA can separate attribute types by frequency, distribution, and even character. PCA technology is used to determine which attributes may be ignored due to their very low impact on neural network solutions and which attributes are most prominent in the data. Figure 1 illustrates the analysis of a data cluster in two directions, offset by 90 degrees. The first principal component (eigenvector 1) analyses the data cluster along the longest axis. The second principal component (eigenvector 2) analyses the data cluster variations perpendicular to the first principal component. As stated in the diagram, each eigenvector is associated with an eigenvalue which shows how much variance there is in the data.

Figure 1. Two attribute data set illustrating the concept of PCA

The next step in PCA analysis is to review the eigen spectrum to select the most prominent attributes in a data set. The following example is taken from a suite of instantaneous attributes over the Niobrara formation within the Denver-Julesburg Basin. Results for eigenvectors 1 are shown with three attributes: sweetness, envelope, and relative acoustic impedance being the most prominent.
Utilizing a cutoff of 60% in this example, attributes were selected from PCA for input to the neural network classification. For the Niobrara, eight instantaneous attributes from the four of the first six eigenvectors were chosen and are shown in Table 1. The PCA allowed identification of the most significant attributes from an initial group of 19 attributes.

Table 1. Results from PCA for Niobrara Interval shows which instantaneous attributes will be used in a Self-Organizing Map (SOM).

### Selected Attributes

<table>
<thead>
<tr>
<th>ATTRIBUTE NAME</th>
<th>LINE</th>
<th>EIGENVECTOR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inst_Envelope</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Inst_Instantaneous Frequency</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Inst_Instantaneous Q</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Inst_Normalized Amplitude</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Inst_Relative Acoustic Impedance</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Inst_Sweetness</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Inst_Thin Bed</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Niobrara Training Enhanced 1 ms</td>
<td>3</td>
<td></td>
</tr>
</tbody>
</table>

Self-Organizing Maps

Teuvo Kohonen, a Finnish mathematician, invented the concepts of Self-Organizing Maps (SOM) in 1982 (Kohonen, T., 2001). Self-Organizing Maps employ the use of unsupervised neural networks to reduce very high dimensions of data to a classification volume that can be easily visualized (Roden and others, 2015). Another important aspect of SOMs is that every seismic sample is used as input to classification as opposed to wavelet-based classification.

Figure 3 diagrams the SOM concept for 10 attributes derived from a 3D seismic amplitude volume. Within the 3D seismic survey, samples are first organized into attribute points with similar properties called natural clusters in attribute space. Within each cluster new, empty, multi-attribute samples, named neurons, are introduced. The SOM neurons will seek out natural clusters of like characteristics in the seismic data and produce a 2D mesh that can be illustrated with a two-dimensional color map. In other words, the neurons “learn” the characteristics of a data cluster through an iterative process (epochs) of cooperative than competitive training. When the learning is completed each unique cluster is assigned to a neuron number and each seismic sample is now classified (Smith, 2016).
Figures 4 and 5 show a simple example using 2 attributes, amplitude, and Hilbert transform on a synthetic example. Synthetic reflection coefficients are convolved with a simple wavelet, 100 traces created, and noise added. When the attributes are cross plotted, clusters of points can be seen in the cross plot. The colored cross plot shows the attributes after SOM classification into 64 neurons with random colors assigned. In Figure 5, the individual clusters are identified and mapped back to the events on the synthetic. The SOM has correctly distinguished each event in the synthetic.
Results for Niobrara and Eagle Ford

In 2018, Geophysical Insights conducted a proof of concept on 100 square miles of multi-client 3D data jointly owned by Geophysical Pursuit, Inc. (GPI) and Fairfield Geotechnologies (FFG) in the Denver-Julesburg Basin (DJ). The purpose of the study is to evaluate the effectiveness of a machine learning workflow to improve resolution within the reservoir intervals of the Niobrara and Codell formations, the primary targets for development in this portion of the basin. An amplitude volume was resampled from 2 ms to 1 ms and along with horizons, loaded into the Paradise® machine learning application and attributes generated. PCA was used to identify which attributes were most significant in the data, and these were used in a SOM to evaluate the interval Top Niobrara to Greenhorn (Laudon and others, 2019).

Figure 6 shows results of an 8X8 SOM classification of 8 instantaneous attributes over the Niobrara interval along with the original amplitude data. Figure 7 is the same results with a well composite focused on the B chalk, the best section of the reservoir, which is difficult to resolve with individual seismic attributes. The SOM classification has resolved the chalk bench as well as other stratigraphic features within the interval.

Figure 6. North-South Inline showing the original amplitude data (upper) and the 8X8 SOM result (lower) from Top Niobrara through Greenhorn horizons. Seismic data is shown courtesy of GPI and FFG.
Figures 8 and 9 show SOM results for the Eagle Ford from studies conducted in 2016 and 2017 (Santogrossi, 2017, Roden and others, 2017). On the left, a NW-SE seismic section across the location of Well #6 is comprised of a conventional amplitude seismic display with a red/white/black ID color scheme. The figure shows the Austin Chalk to the Buda stratigraphic interval resolved in roughly 3 peak/trough cycles. The graphic on the right displays the results of a SOM of multiple instantaneous attributes colored by neurons of up to 64 classes. A seismic interval from 10ms below the Buda to 100ms above the Buda or near to the top of the Austin Chalk was chosen for the SOM run. Shown clearly is the resolution improvement provided by classification when the interpretation interval is reduced to just the area of interest.

Eagle Ford - Enhanced Resolution via Horizon Delimitation

- Amplitude Data – “boosted”

- SOM Classification

Figure 8. 8X8 SOM results through the Eagle Ford. The primary target, the Lower Eagle Ford shale had 16 neuron classes over 14-29 milliseconds of data. Seismic data shown courtesy of Seitel.
The results shown in Figure 9 reveal non-layer cake facies bands that include details in the Eagle Ford basal clay-rich shale, high resistivity and low resistivity Eagle Ford shale objectives, the Eagle Ford ash, and the upper Eagle Ford marl, which are overlain disconformably by the Austin Chalk.

Convolutional Neural Networks (CNN)

A promising development in machine learning is supervised classification via the applications of convolutional neural networks (CNNs). Supervised methods have, in the past, not been efficient due to the laborious task of training the neural network. CNN is a deep learning seismic classification. We apply CNN to fault detection on seismic data. The examples that follow show CNN fault detection results which did not require any interpreter picked faults for training, rather the network was trained using synthetic data. Two results are shown, one from the North Sea, Figure 10, and one from the Great South Basin, New Zealand, Figure 11.

Figure 9: Eagle Ford SOM classification shown with well results. The SOM resolves a high resistivity interval, overlain by a thin ash layer and finally a low resistivity layer. The SOM also resolves complex 3-dimensional relationships between these facies.
Conclusions

Advances in compute power and algorithms are making the use of machine learning available on the desktop to seismic interpreters to augment their interpretation workflow. Taking advantage of today’s computing technology, visualization techniques, and an understanding of machine learning as applied to seismic data, PCA combined with SOMs efficiently distill multiple seismic attributes into classification volumes. When applied on a multi-attribute seismic sample basis, SOM is a powerful nonlinear cluster analysis and pattern recognition machine learning approach that helps interpreters identify geologic patterns in the data and has been able to reveal stratigraphy well below conventional tuning thickness.

References


Guo, H., K. J. Marfurt, and J. Liu, 2009, Principal component spectral analysis: Geophysics, 74, no. 4, 35–43.


Laudon, C., Stanley, S., and Santogrossi, P., 2019, Machine Learning Applied to 3D Seismic Data from the Denver-Julesburg Basin Improves Stratigraphic Resolution in the Niobrara, URTeC 337, in press

