

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/346345661>

# A Geoscientist-Guided Multiattribute SOM Analysis. AAPG-Explorer, Geophysical corner. November, 2020

Article · November 2020

CITATIONS

0

READS

46

2 authors:



[Karelia La Marca](#)

University of Oklahoma

9 PUBLICATIONS 3 CITATIONS

[SEE PROFILE](#)



[Heather Bedle](#)

University of Oklahoma

60 PUBLICATIONS 240 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Multi- attribute analysis and optimization of machine learning techniques like SOMs to properly characterize seismic facies in deepwater depositional settings [View project](#)



Gas Hydrates and Machine Learning [View project](#)

## Geophysical Corner

## A Geoscientist-Guided Multiattribute SOM Analysis

“Machine learning” has become a common phrase in geophysics. These methods, based on complex algorithms and statistics, allow geoscientists to speed up and improve their interpretations. However, as interpreters, we can feel intimidated and concerned about how much of our expertise can be replaced by machine learning algorithms. To better understand the limitations, we assess the importance of human validation and participation in one machine learning process, highlighting the upsides and downsides of a machine-derived process versus a geoscientist-guided selection of attributes. As Earth scientists, we explored a suite of seismic attributes and selected those that were meaningful for interpreting a deepwater channel system and compared our results with the attributes derived from principal component analysis.

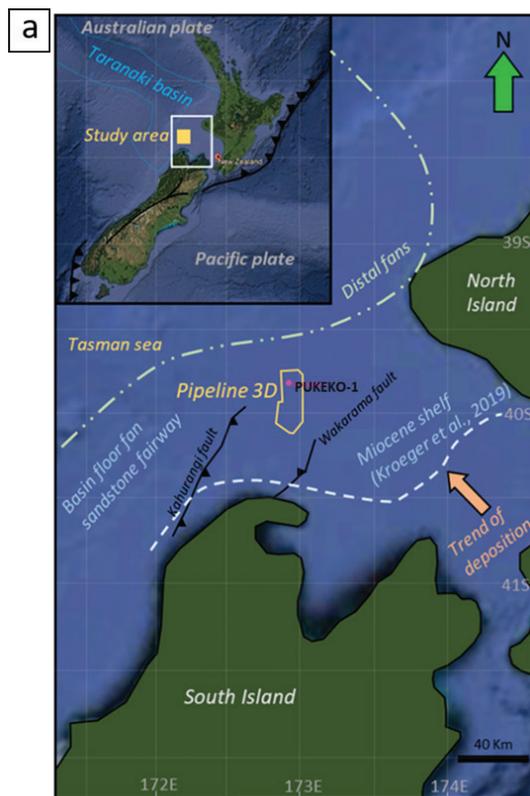
Machine-learning methods, such as PCA, reduce the initial number of attributes down to a smaller subset, which attempt to contain most of the variation within the dataset through eigenvectors. PCA results, or multiple attributes that are user-selected, can then be analyzed using self-organizing maps. SOMs, provide a fast organization of data in groups that aid in geological interpretation. Using the Pipeline 3-D seismic dataset in the southern Taranaki Basin of New Zealand (figure 1), results are presented that improve the understanding of the significance of the interpreter’s presence throughout the application of SOMs.

## Two Study Phases

This analysis comprises two study phases (figure 2).

After inspecting the seismic volume and defining the area of interest, we interpreted a horizon within the middle Miocene Moki Formation. Then, we calculated a suite of instantaneous, geometrical, and spectral attributes based on user selection and experience. The choice of the attributes was done according to the study objective (identify different seismic facies and architectures that correspond to deepwater deposits). Therefore, attributes that could differentiate mud-filled channels from sand-filled ones (RMS amplitude and GLCM entropy, for example) and isolated mass-transport and overbank deposits were selected. These attributes were used as inputs for the SOMs to identify deepwater architectural elements and were cross-correlated with well data (Pukeko-1) to support the proposed interpretations.

The second phase consisted of the

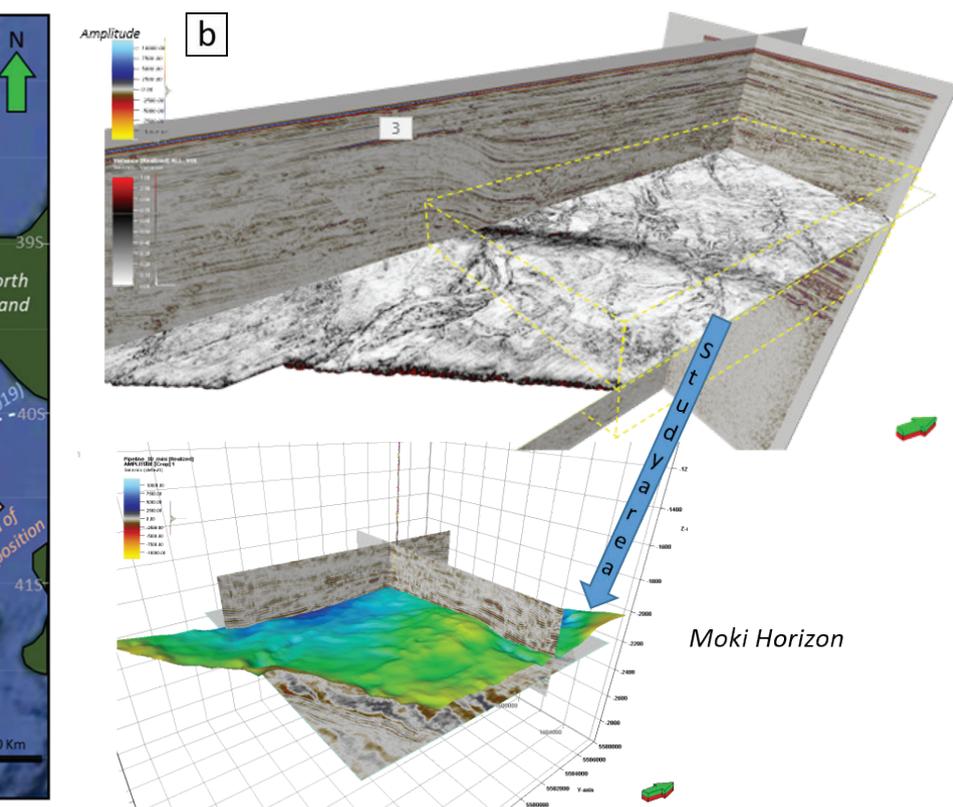


determination of the most meaningful attributes for a machine via PCA. Attributes selected were used as input for SOM and analyzed similarly to the first study. In the end, we compared the results of both phases.

## SOMs Results from User-Selected Attributes Versus Machine-Derived Inputs

Figure 3 depicts the comparison between the SOM derived from attributes selected in an unsupervised fashion, using PCA in figure 3a. The SOM results from meaningful attributes characterized by us (the geologists) in figure 3b and the amplitude expression in figure 3c. Both the geologist and PCA-derived SOM results greatly improve upon seismic amplitude, revealing architectural elements in the deepwater system.

When comparing results from the SOM using the attributes selected by the machine using PCA (figure 3a) and the SOM derived from the user-method (figure 3b), we noticed that the PCA method allows the distinction of predominant architectural elements. However, more details of the channel architecture are revealed with the user-selected attributes. These details, like the geomorphology of the channels and definition of subtler features within architectural elements, allow for a better picture, and therefore interpretation in the case of the user-selected attributes SOM. We included attributes that PCA did



## Upsides of the approaches

- ▶ The user-driven approach allows us to inspect attributes individually and select the most geologically meaningful for their interpretation purpose or goal.
- ▶ Regardless of the goal, the PCA is a dimensionality reduction technique that narrows down the data to a smaller dataset containing most of the variability in the data.
- ▶ PCA offers a quick way to explore a dataset. It is more time-efficient than a multi-attribute analysis and may consider attributes that were not previously inspected by the user. An example of this is the envelope attribute well known to aid in seismic facies interpretation.
- ▶ The user-selected methods allow integration of different data types that can be

not consider (spectral, for example), which provide additional detail in the classification. Spectral and instantaneous attributes are known for unraveling lithological content and for their capability of distinguishing differently-sized features associated with channel complexes: small architectures like levees are usually related to high frequencies, and significant, master channels can be identified with small frequencies.

In the user-defined approach we considered a mixture of GLCM entropy, RMS, sobel filter and spectral attributes for SOM. Nevertheless, the unsupervised PCA method also provided good results, although just not as detailed as when the algorithm is assisted by an expert interpreter. This suggests that the use of PCA preceding a SOM can be a good practice when the interpreter’s experience, or time, is limited.

This work, revealed several insights, summarized below.

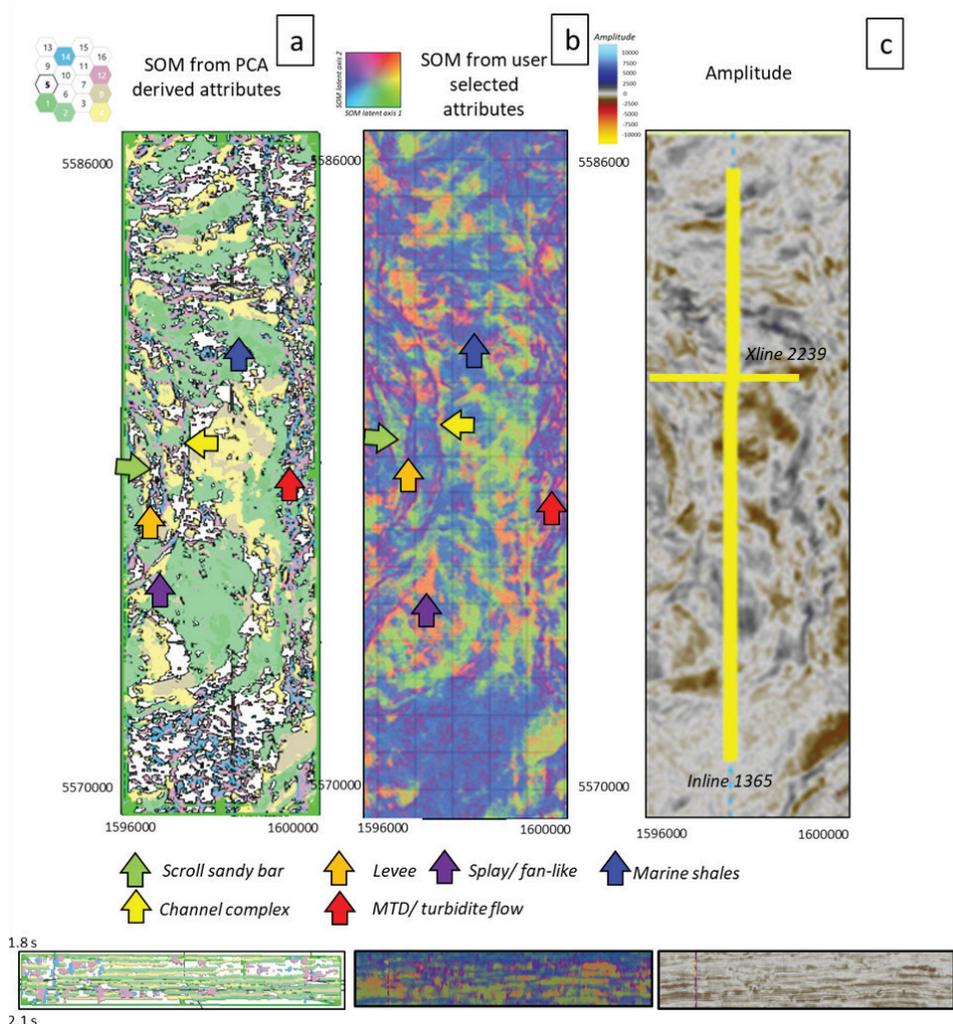
Continued on next page ▶



**Karelia La Marca** received a bachelor’s in geological engineering from Universidad de Los Andes, Venezuela. She started her master’s in geophysics at the University of Oklahoma under the guidance of Heather Bedle after joining Schlumberger, where she worked as a geologist. Currently, she is a research assistant for the Subsurface Detective Agency and Attribute Assisted Seismic Processing and Interpretation consortiums at the University of Oklahoma and pursues a doctorate in geophysics. Her research interests include reservoir characterization, seismic interpretation and application of seismic attributes with machine learning techniques for seismic facies definition, especially in deepwater settings. She received the “Provost certificate in teaching for outstanding graduate assistant.” She placed among the first 25 presentations in the 89th annual Society of Exploration Geophysicists meeting. She is an active member of the AAPG and SEG since 2018, and most recently of the Geophysical Society of Houston, the Society for Sedimentary Geology, the Geological Society of America and GSOC.



**Heather Bedle** received a bachelor’s in physics from Wake Forest University and then worked as a systems engineer in the defense industry. She later received a master’s and then a doctorate from Northwestern University. After nine years of working with Chevron, she instructed at the University of Houston for two years and is now an assistant professor at the University of Oklahoma. Her primary research interests merge seismic interpretation, rock physics and attribute analysis. She is an active member of AAPG and the Society of Exploration Geophysicists.



◀ Continued from previous page

tested or classified to reduce redundancy.

▶ Both methods implemented an unsupervised machine learning technique, SOM, which allows recognition of elements that a human interpreter might not have been able to identify by himself, such as levees and sediment waves.

**Downsides and Limitations**

▶ The optimal selection of attributes relies on the interpreter's experience and understanding of each attribute significance and it might be biased.

▶ The PCA results might show attributes that are not the most suitable for the goal (cosine of phase, for example), and also requires data quality control, because PCA is sensitive to noise. The PCA will reduce the amount of data examined, but results may have little to do with physical features of interest.

▶ The multiattribute approach might be time-consuming if started from zero, but effective if combinations used for every setting are documented and followed to apply in similar geological settings.

▶ The PCA method suggests attributes that are "main contributors", but that may be similar or redundant, such as most positive curvature and maximum curvature.

▶ Both methods, applied to SOMs, require human inspection and approval in pivotal moments. The most straightforward example is the validation and interpretation of the SOMs outputs regardless of the initial approach used. Interpreters must apply seismic geomorphology principles to define the groups of data points with "similar" characteristics that belong to each cluster.

**Conclusions**

In this comparison of user-based (interpreter) and machine-based (PCA) attribute selections to implement in a SOM technique for deepwater seismic facies interpretation, the PCA allowed the reduction of a large dataset of 28 attributes to only nine attributes contained in the principal components. The first, second and third principal components contained

most of the variability within the data. The machine-based attribute selection included amplitude, cosine of instantaneous phase, envelope, most positive and most negative curvature, sweetness and textural attributes such as GLCM, dissimilarity, entropy, and homogeneity to use in the SOMs.

But, we concluded that the machine-based attribute selection is composed of redundant attributes such as GLCM homogeneity and GLCM entropy, or amplitude and envelope, and this redundancy may have contributed to the lesser SOM results, when compared to the user-chosen attributes. We found that the user-based approach offers a better picture and geological significance to the interpreter than the machine-based approach (PCA). The user-based method helped to better illustrate the architecture and geomorphology of the channel elements and define subtle differences within them. Keep in mind that if the geological setting and objective differs from the one in this research, the user should define the most suitable attributes for his geological goal.

The major difference between the compared methods is the time efficiency and accuracy: A user-based approach produces better results but is not as time efficient as the machine-based method. Nonetheless, machine learning methods have to be treated as a tool to work efficiently to extract information from multiple datasets at the same time, to overcome interpreter's limitations. In the future, machine algorithms will improve, and but we underpin the idea that, for now, geoscientist interpreters are irreplaceable.

**Acknowledgements**

We want to thank NZPM and GNS for the dataset and reports provided for the study. We appreciate AASPI consortium for their software, used for seismic attributes, and to Schlumberger and Geophysical insights for their software licenses provided to OU. 

*(Editors Note: The Geophysical Corner is a regular column in the EXPLORER, edited by Satinder Chopra, Founder and President of SamiGeo, Calgary, Canada, and a past AAPG-SEG Joint Distinguished Lecturer.)*

