Making Sense of Seismic Attributes: Attribute-Type Framework for Machine Learning Characterization

MEWBOURNE COLLEGE OF EARTH AND ENERGY
SCHOOL OF GEOSCIENCES
The UNIVERSITY of OKLAHOMA

Hilmi Putra, Dr Heather Bedle, Dr David Lubo-Robles
School of Geosciences, University of Oklahoma

ATTRIBUTE ASSISTED SEISMIC PROCESSING AND INTERPRETATION

Introduction

Seismic attributes are different mathematical operations used to measure seismic traces' various aspects. Analysis of the different attributes enables geophysicists to find **anomalies** within seismic data that could indicate **geological objects of interest**. Machine learning provides a means to summarize information from the multi-dimensional aspects through quantitative pattern recognition—viewing the seismic signal from multiple perspectives at once—and *objectively* point towards features that might be of interest.

While many seismic attributes exist, several **overlap** in what aspect of the signal they describe. Selecting the **appropriate set** of seismic attributes is crucial to optimally capture the diverse signal expressions that characterize subsurface heterogeneity and produce a useful multi-attribute image.

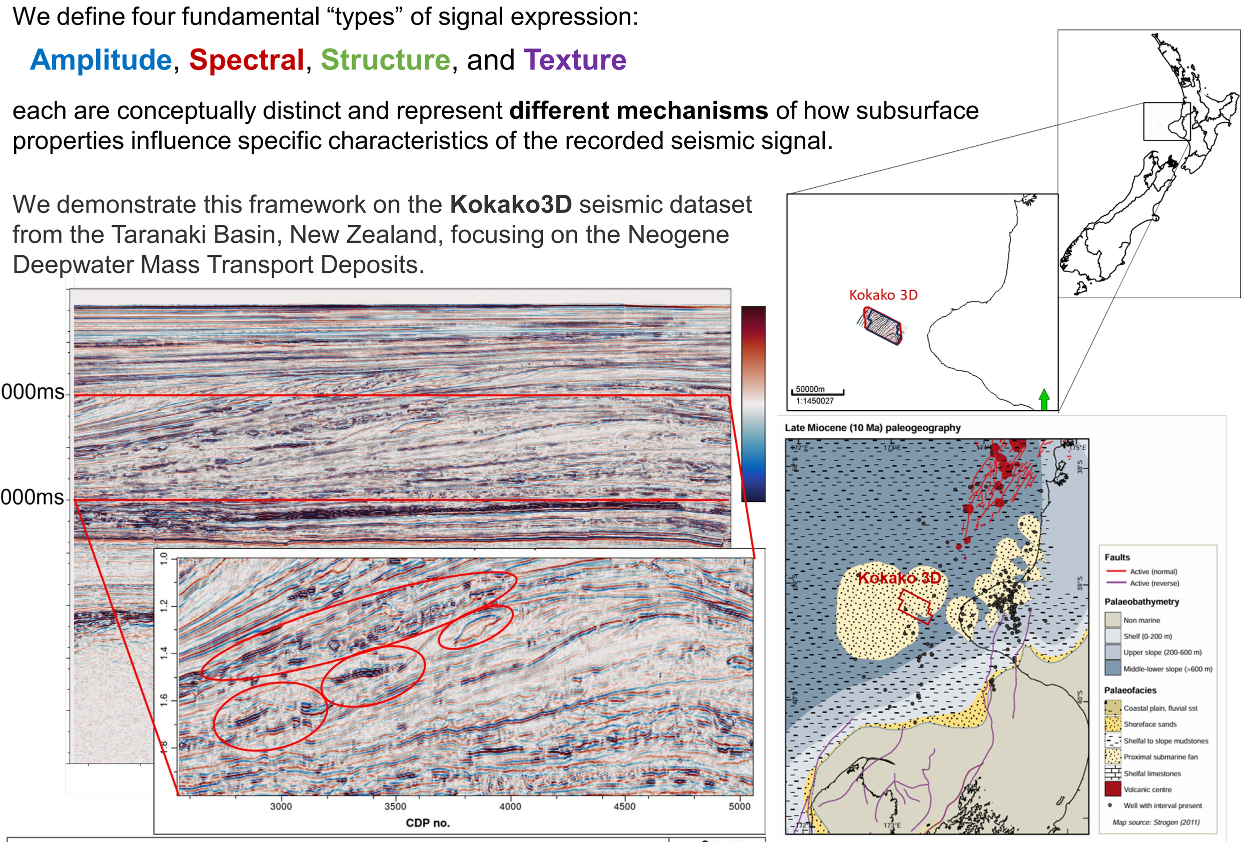
We present a targeted approach to attribute selection for unsupervised machine learning, aimed at quickly visualizing and characterizing subsurface features.

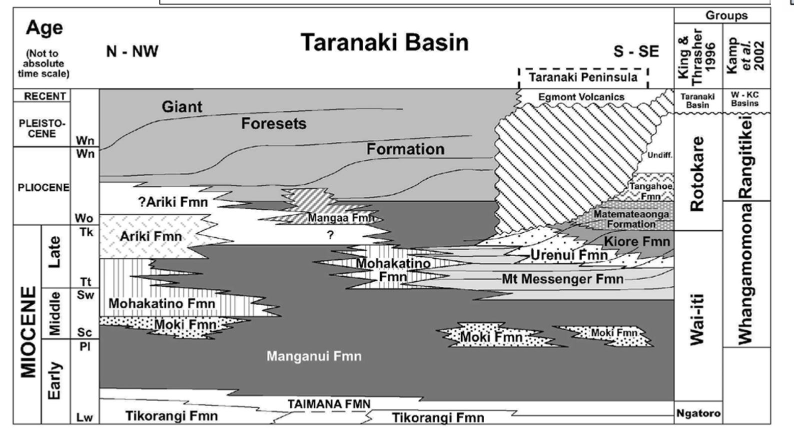
We define four fundamental "types" of signal expression:

Amplitude, **Spectral**, **Structure**, and **Texture**

each are conceptually distinct and represent **different mechanisms** of how subsurface properties influence specific characteristics of the recorded seismic signal.

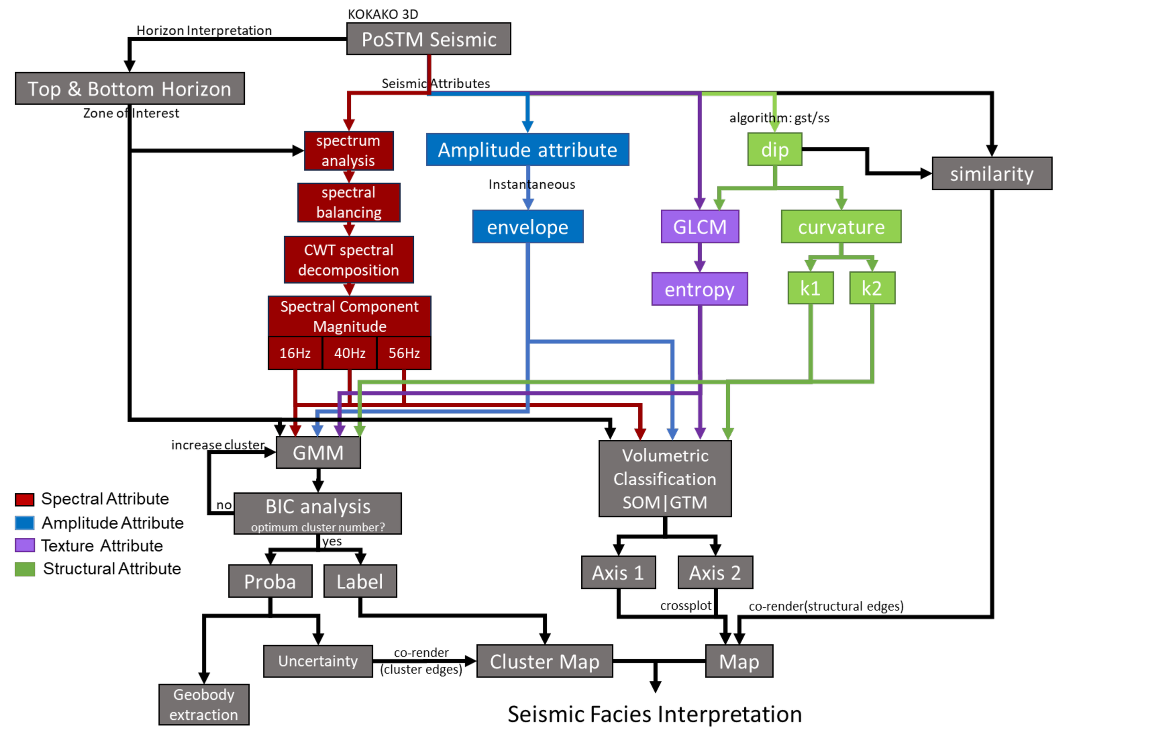
We demonstrate this framework on the **Kokako3D** seismic dataset from the Taranaki Basin, New Zealand, focusing on the Neogene Deepwater Mass Transport Deposits.





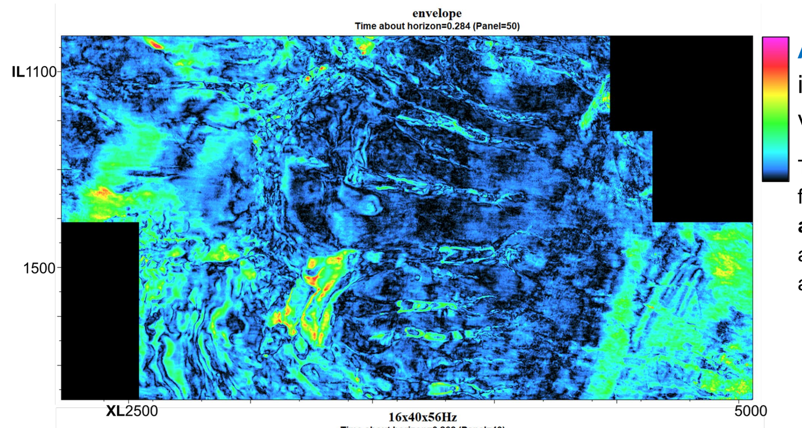
Turbidite systems developed from Late Eocene time (Tangaroa Fm.). It was noted that shelf progradation during the Neogene was mud-dominated and punctuated by discrete **turbidite depositional phases** in the Late Oligocene (Tariki Sandstone), Middle Miocene (Moki Formation), **Late Miocene (Mt Messenger Formation)**, and **Pliocene (Mangaa Formation).**

Methodology



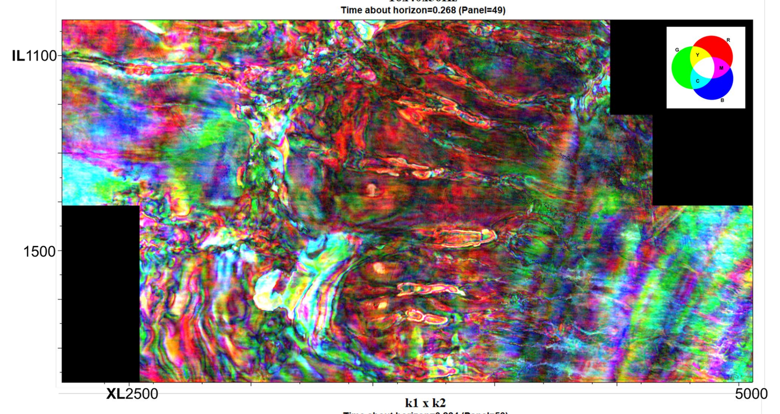
We start with the Post-stack Time Migrated Seismic Data as Input—which in this case is the Kokako 3D dataset. Seismic attributes are computed for each of the attribute types to have all the relevant variations within the seismic signal; making sure all our machine learning input attributes are independent of each other.

Attribute Types



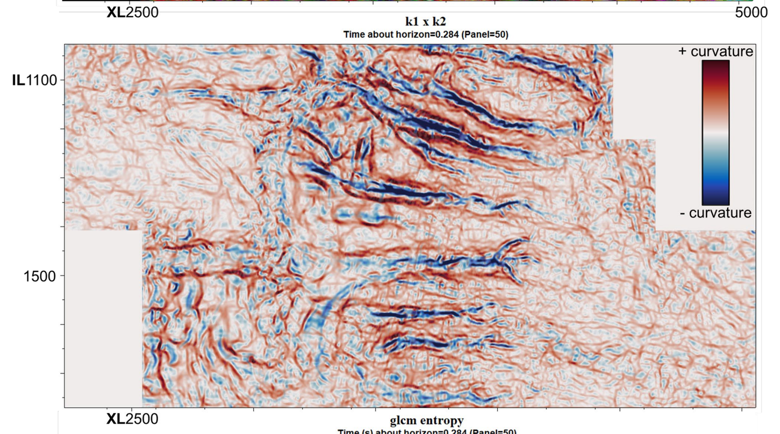
Amplitude attributes represent acoustic impedance contrasts (density and P-wave velocity)

Therefore, are often tied to changes in lithology or fluid content. Attributes such as **envelope** or **RMS amplitude** are direct measures of signal strength and can delineate features like bright spots or amplitude terminations.



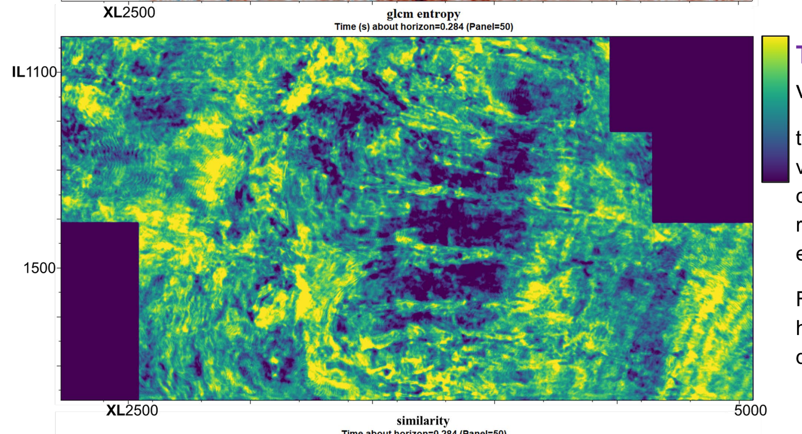
Spectral attributes reflect how the frequency content of the seismic wavelet varies in space.

These variations captured here are typically due to the interaction of the strata with the moving wavelet that could indicate stratigraphic complexities such as bed thickness (e.g., tuning effects) and attenuation due to lithology. For example, the **magnitude of** a **16Hz** or **56Hz frequency component** can help isolate features with specific vertical-time resolution or layering.



Structure attributes evaluate the spatial organization of reflectors.

Attributes such as curvature (**K1 and K2**) or **aberrancy** detect features such as channels, levees, or faults—elements defined not by amplitude but by configuration across traces. These are critical for identifying morphological expressions of mass transport processes such as basal scours, folded locks, or slump bodies.



Texture attributes, are based on local variability or patterning within the data.

these attributes quantify how seismic amplitude values change over a small neighborhood, capturing arrangements that may indicate flow regimes, internal structures, or depositional environments.

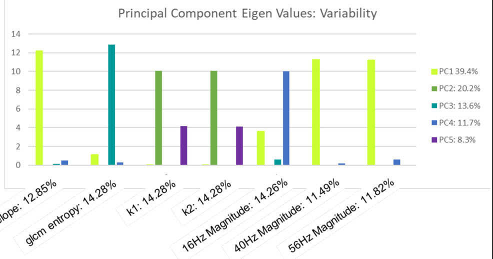
For example, **GLCM entropy** highlights regions of high signal disorder, which often correlate with disrupted bedding or weakly conforming layers.

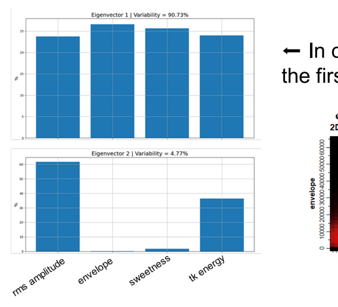
Similarity/Variance highlights structural edges and is useful to make visual boundaries for facies distribution on the maps.

Although, by definition, it conforms alongside the **Structure attribute** type, we did not include this attribute for the **clustering algorithm** as it marks not the seismic facies itself but the **termination/edges** of seismic facies/feature and thus will be useful on its own as the *third axis* on our map color bar (by co-rendering).

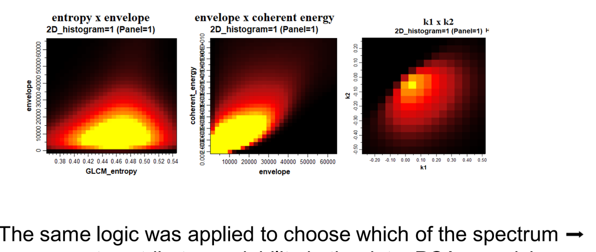
we utilize Principal Component Analysis (PCA) as a validation step to confirm ➞ that the selected attributes contribute unique and uncorrelated information.

The relatively even distribution of eigenvector contributions in our PCA results supports our intent that the chosen attributes reflect multidimensional variability rather than redundancy, confirming their relevance to signal discrimination and avoids biasing our clustering.

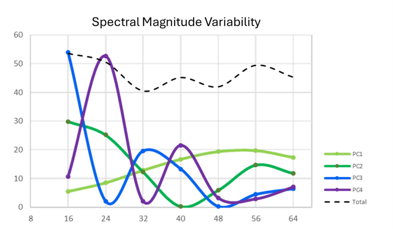




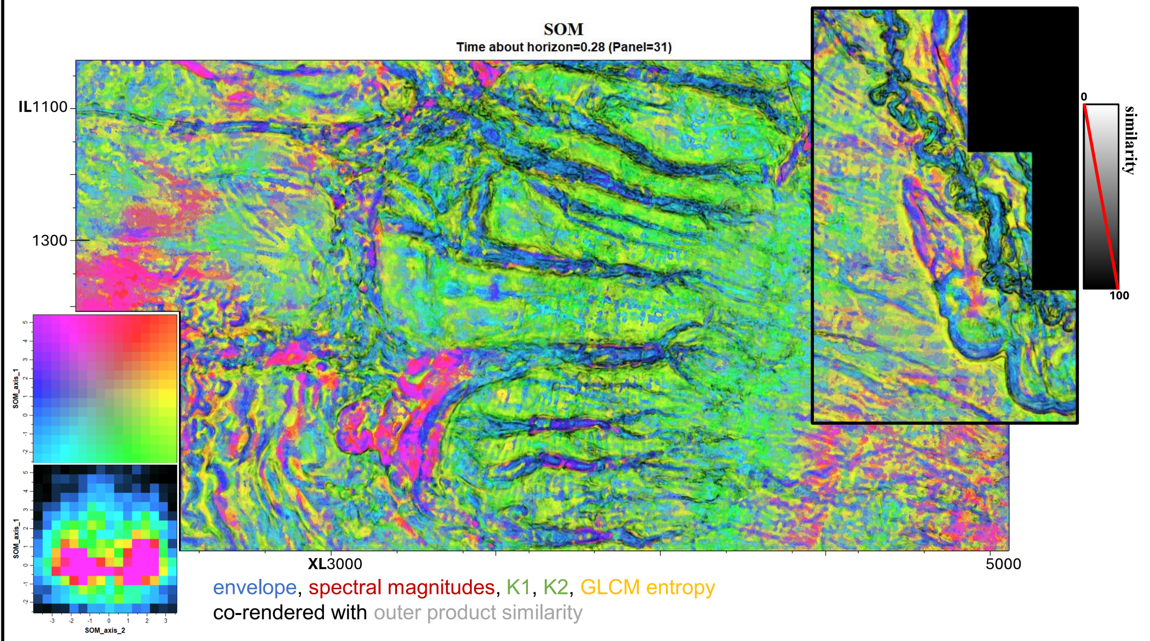
⟵ In contrast, PCA analysis of the *amplitude-type* attributes returns 90% of the variation captured by the first eigen vector, suggesting redundancy.



The same logic was applied to choose which of the spectrum ➞ component contributes variability in the data. PCA can pick up the bands that represents the most spectral variation.

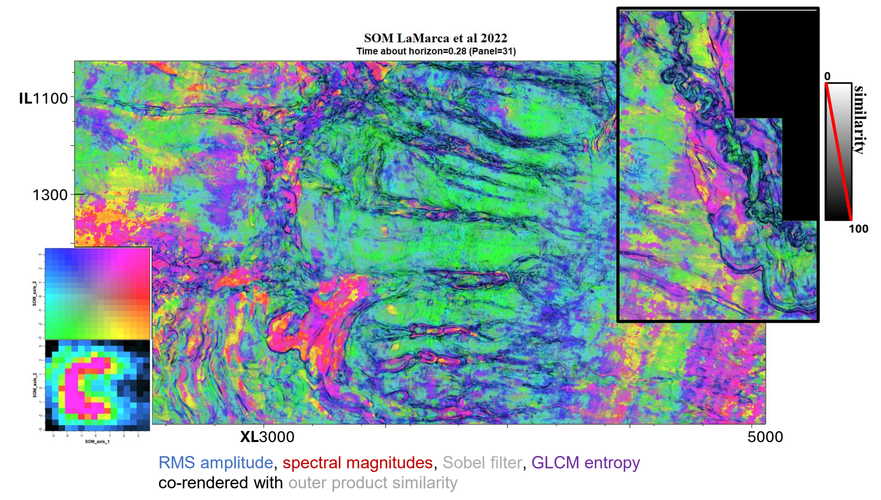


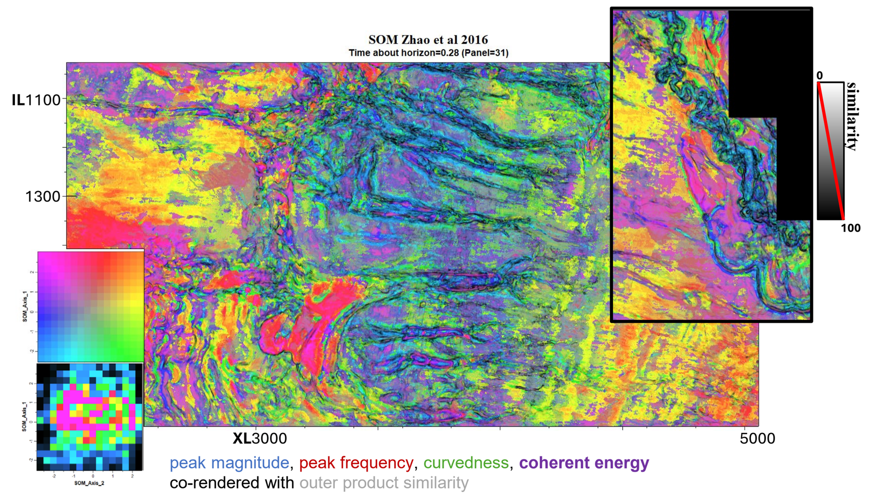
Self Organizing Maps – Volumetric Classification

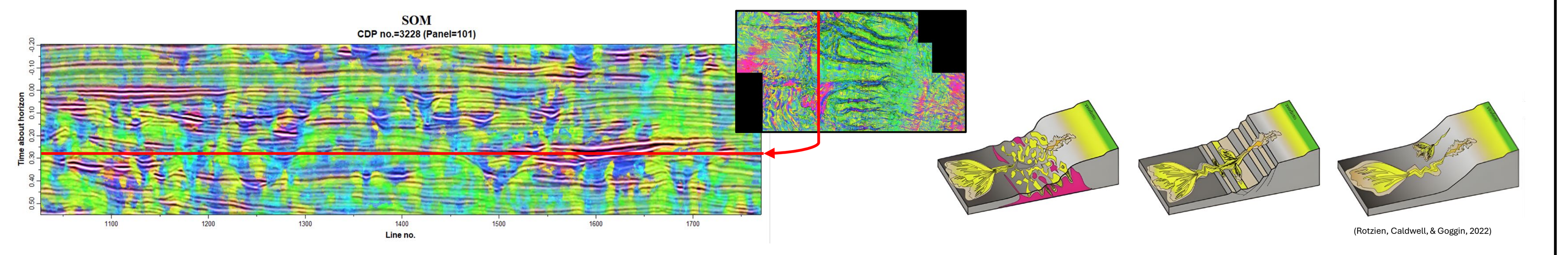


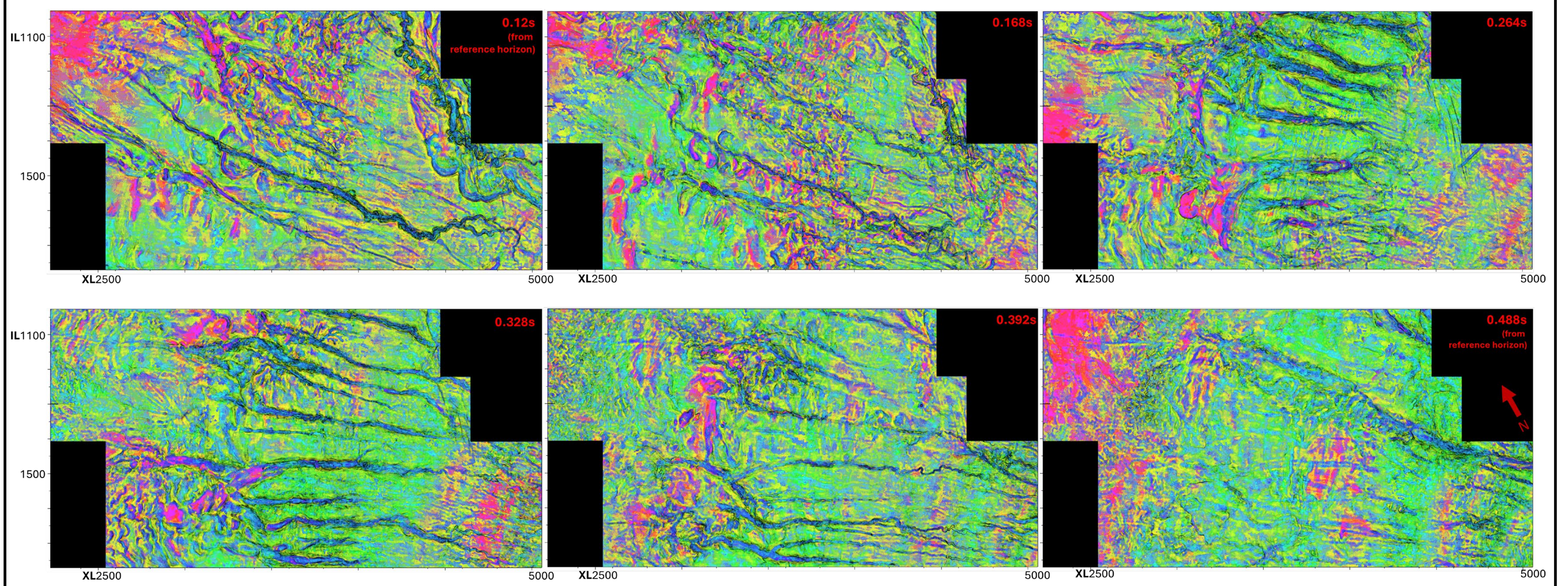
Self-Organizing Maps (SOM) project the high-dimensional seismic attribute space onto a two-dimensional latent space, aggregating data points that has similarity in attribute(s) expression.

It produces a 2D color-coded map that reflects **gradational transitions** between classes, effective in capturing the continuous nature of geological facies, preserving interpretive nuance whilst highlighting regions of similar signal characteristics.









The SOM maps follow a phantom horizon around 0.8-2s time-depth in the dataset. Starting from the bottom right, we can observe the evolution of the mass transport deposit, with the distinct facies marked with the different colors; co-rendered with outer product similarity to visually aid the interpreter with the structural edges. We observe the unconstrained flow 'fan' structures, marked with the magenta facies. It appears to precede straight channel facies which then moves the sediment further west, forming a new slumping fan or undulating structure; building up sediment as the progradation continues.

Summary

The method is not tied to a specific list of attributes but rather maintains representation from each of **the four categories**.
This allows us to focus on **optimizing** attribute algorithms and **parameters** that best capture the intended signal variations, rather than arbitrarily choosing from the hundreds of attributes available.

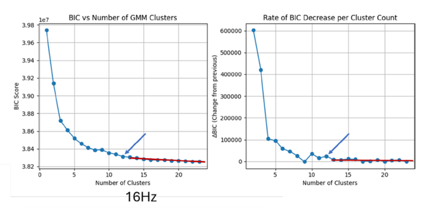
While the current study uses **post-stack attributes**, this workflow is extensible to include **AVO** (amplitude versus offset) or **elastic property** attributes from angle stacks, which would add a **fifth signal expression type** to add into the cluster definition.

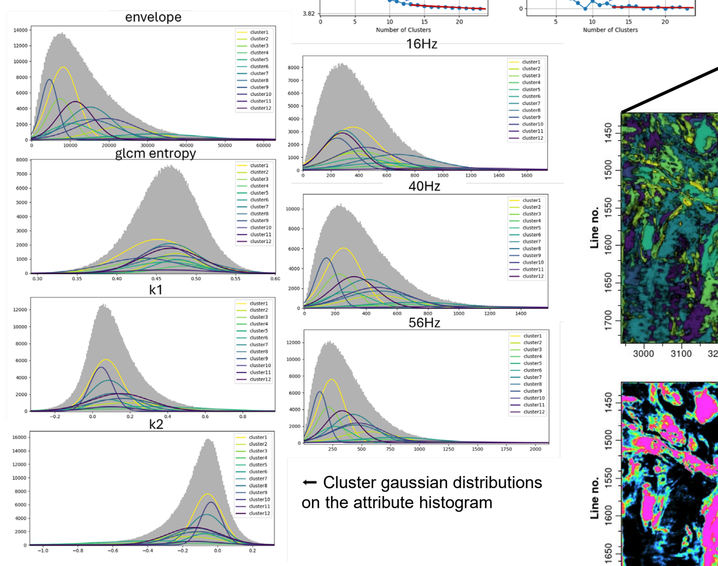
Whether visualization maps with SOM, or quickly extracting geobodies with GMM, the common thread is that the attribute input strategy leads to meaningful and interpretable groupings in the data.

Gaussian Mixture Models – Unsupervised Clustering

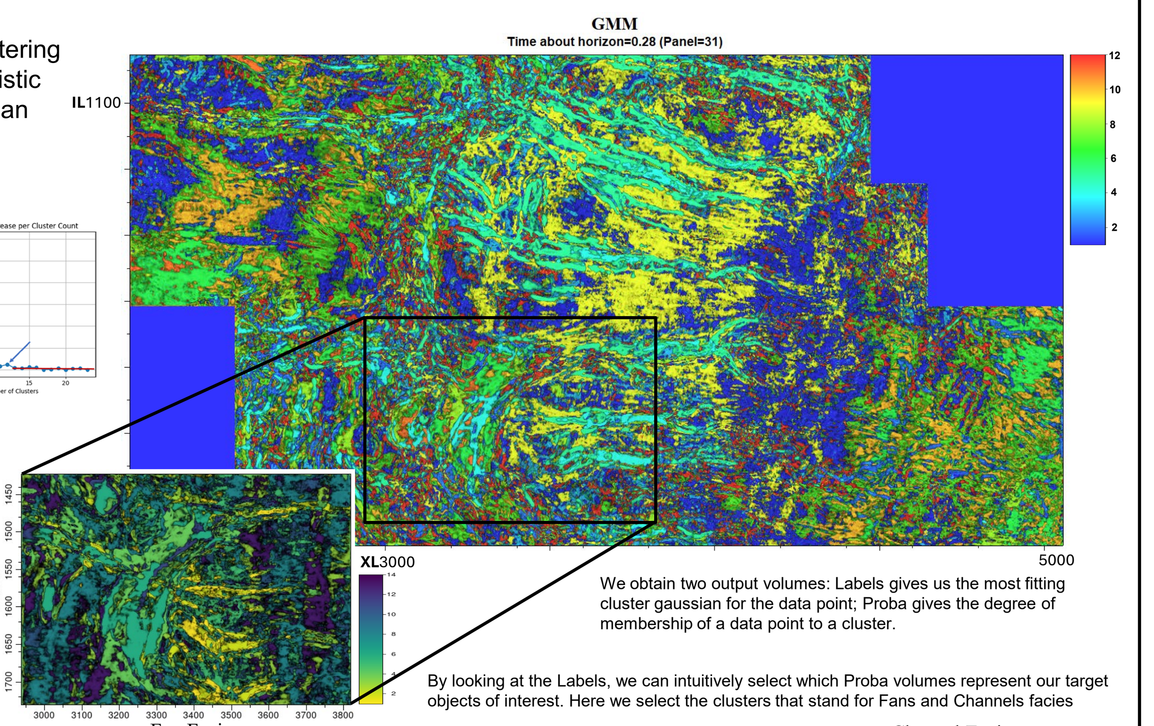
Gaussian Mixture Models (GMM) offer a soft-clustering alternative where data points are assigned probabilistic membership across a predefined number of Gaussian distributions.

BIC (Bayesian Information Criterion) ➔ Analysis is performed to obtain the optimum number of cluster



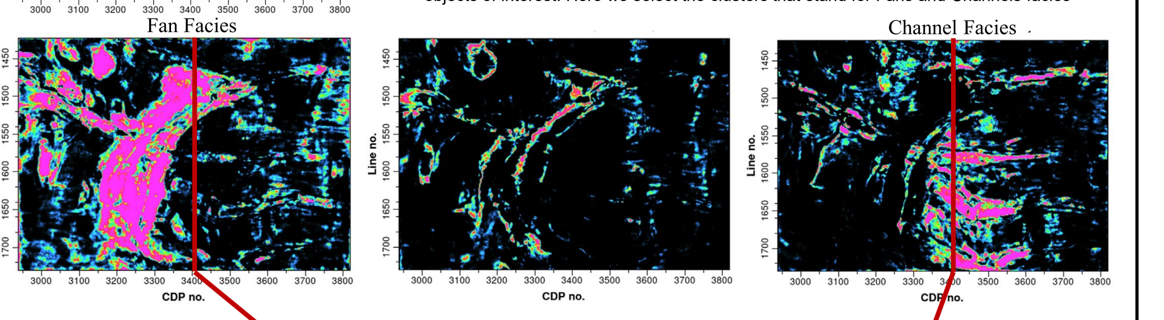


⟵ Cluster gaussian distributions on the attribute histogram



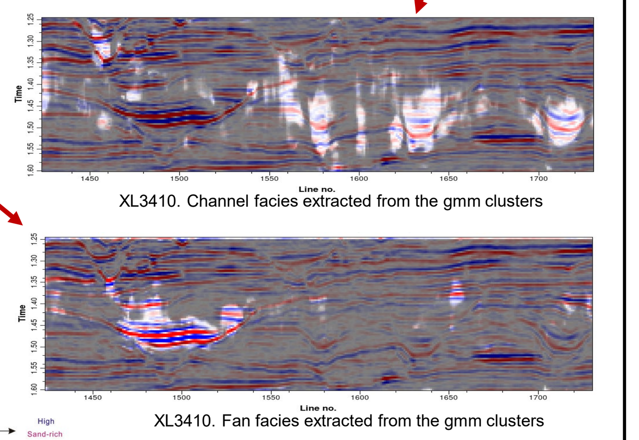
We obtain two output volumes: Labels gives us the most fitting cluster gaussian for the data point; Proba gives the degree of membership of a data point to a cluster.

By looking at the Labels, we can intuitively select which Proba volumes represent our target objects of interest. Here we select the clusters that stand for Fans and Channels facies



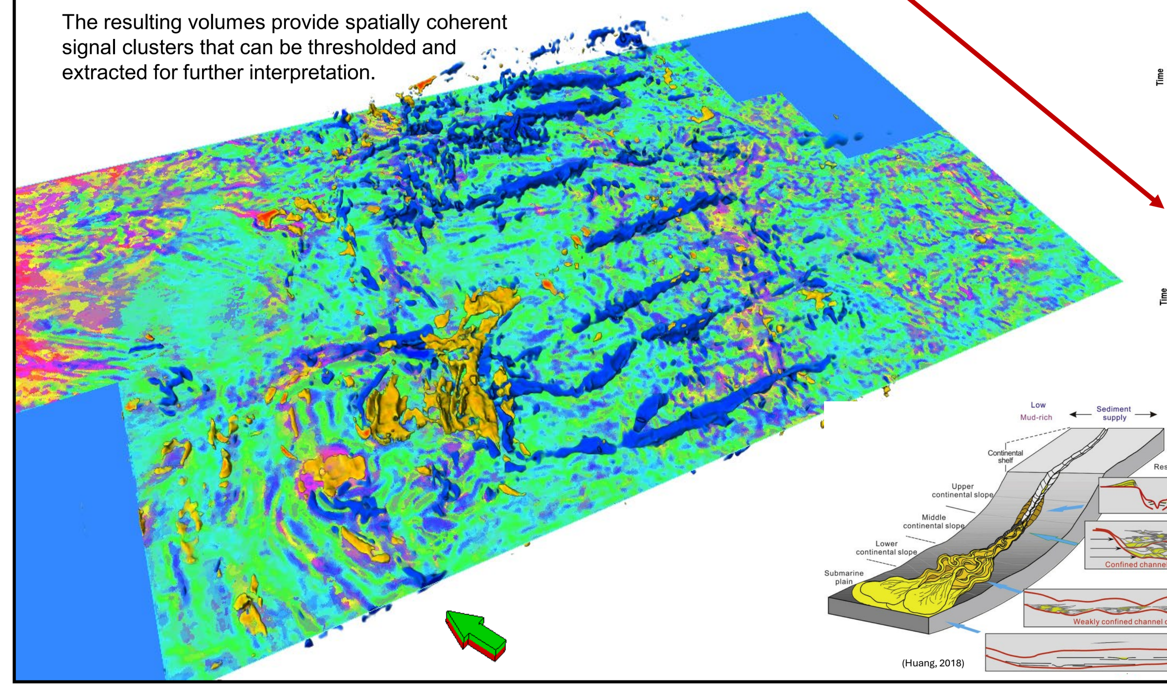
This method is particularly effective for geobody extraction, as it quantifies the degree to which each voxel belongs to a cluster based on its multi-attribute expression.

The resulting volumes provide spatially coherent signal clusters that can be thresholded and extracted for further interpretation.



XL3410. Channel facies extracted from the gmm clusters

XL3410. Fan facies extracted from the gmm clusters

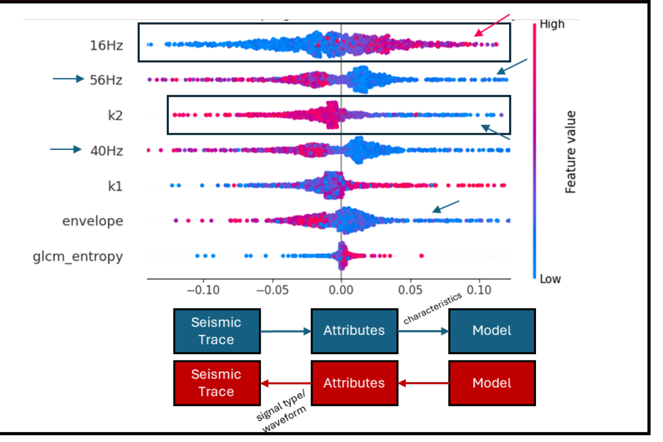


Co-rendering the Proba volume as a mask on top of the seismic trace volume can show us what a certain cluster's waveform looks like for interpretation.

Future Work

Now that we have the signal expressions well represented for machine learning pattern recognition, our next step is to assess from the resulting clusters itself what attributes is the key characteristics of the seismic facies of interest. We will now explore how to gain insight from the clustering methods to help *describe* each of the facies that was able to be picked up by the algorithm, adding meaning to the highlights for interpretation.

Once we have a working workflow on getting interpretation insights from multi-attribute machine learning methods, we'll turn the workflow around and try to simulate the signal expressions (that is analyzed by the seismic attributes); and its key characteristic contributions; from a rock property/physics model to experiment possible geological setups of the subsurface and what the signal anomalies will look like.



Acknowledgments

The author would like to thank Dr Kurt Marfurt for the guidance and discussions in understanding and optimizing the seismic attributes.
We also would like to thank our friends in the AASPI research group; and Mewbourne School of Geosciences students, faculty, and staff,
to New Zealand Petroleum and Minerals for making the Kokako 3D dataset available to us,
and to the AASPI sponsors:

References