Attribute assisted seismic facies classification on a turbidite system in Canterbury Basin, offshore New Zealand

Tao Zhao and Kurt J. Marfurt, University of Oklahoma*

Summary

In this study, we show an example of first-hand interpretation of a turbidite system in Canterbury Basin, offshore New Zealand without involving any traditional structural interpretation. By choosing the right seismic attributes, features in a turbidite system can be identified automatically by pattern recognition techniques, which is particularly valuable to accelerate the interpretation process and provide guidance when working on a new resource play where previous study and data are limited.

Introduction

Traditionally, skilled interpreters delineated seismic facies on a suite of 2D lines by visually examining seismic waveforms, frequency, amplitude, phase, and geometric configurations. Facies would then be posted on a map and hand contoured to generate a seismic facies map. With the introduction of 3D seismic data and volumetric attributes, such analysis is both more quantitative and more automated. Nevertheless, the size of 3D seismic data volumes and the number of seismic attributes have increased to the extent that very few interpreters examine every seismic line and time slice.

To address this growing volume of seismic data, several alternative algorithms have been proposed and successfully applied to computer-assisted seismic facies classification, including k-means, self-organizing mapping, generative topographic mapping, support vector machines, Gaussian mixture models, and artificial neural networks. Although well documented in the literature, the terminology and complexity of these algorithms may bewilder the average seismic interpreter, and few papers apply these competing methods to same data volume.

In this study, we apply one unsupervised and one supervised classification techniques to a single 3D seismic data volume acquired in the Canterbury Basin, New Zealand, where the objective is to differentiate deposition features in a turbidite system. Not surprisingly, the most important parameter in this analysis is the choice of the correct input attributes, which in turn depends on careful pattern recognition by the interpreter. By incorporating both unsupervised and supervised learning techniques using the same input seismic attributes, we provide a first-hand interpretation of the turbidite system while illustrating the merits and demerits of each technique, which may be used as a reference for selecting the suitable classification techniques.

Geologic Setting

We perform the classification on the Waka-3D seismic survey acquired over the Canterbury Basin, offshore New Zealand, generously made public by the New Zealand Petroleum and Minerals. Readers can request this data set through their website for research purposes. Figure 1 shows the location of this survey, where the red rectangle corresponds to time slices shown in subsequent figures. The study area lies on the transition zone of continental slope and rise, with abundance of paleocanyons and turbidite deposits of Cretaceous and Tertiary ages. These sediments are deposited in a single, tectonically driven transgressive – regressive cycle (Uruski, 2010). Being a very recent and underexplored prospect, publically available comprehensive studies of the Canterbury Basin are somewhat limited. The modern seafloor canyons shown in Figure 1 are good analogs of the deeper paleocanyons illuminated by the 3D seismic amplitude and attribute data.

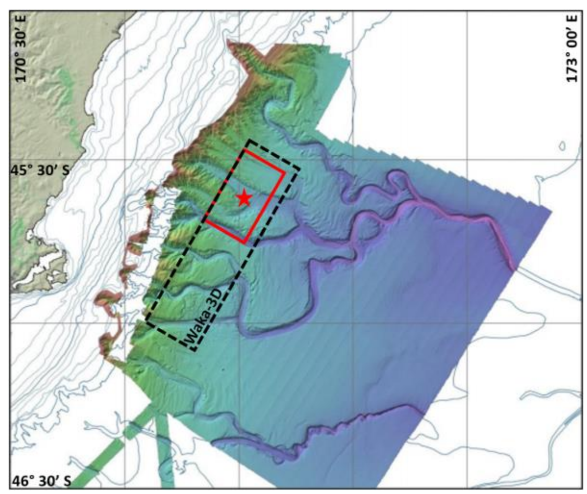


Figure 1: A map showing the location of the 3D seismic survey acquired over the Canterbury Basin, offshore New Zealand. The black rectangle denotes the limits of the Waka-3D survey, while the smaller red rectangle denotes the part of the survey shown in subsequent figures. Colors represent the relative depth of the current seafloor, warm being shallower and cold being deeper. Current seafloor canyons are delineated in this map, which are good analogs for the paleocanyons in Cretaceous and Tertiary ages. (Modified from Mitchell and Neil, 2012)

Attribute selection

The number of attributes should be as small as possible to discriminate the facies of interest. While it may be fairly easy to represent three attributes with a deformed 2D manifold, increasing the dimensionality results in increased deformation, such that our manifold may fold on itself or may not accurately represent the increased data variability. Because the survey is new to us two authors, we have tested numerous attributes that we think may highlight different facies in the turbidite system. Among these attributes, we find the shape index to be good for visual classification but dominates the unsupervised classifications with valley and ridge features across the survey. After such analysis we chose four attributes that are mathematically independent but should be coupled through the underlying geology: peak magnitude, peak frequency, GLCM homogeneity, and curvedness, as the input to our classifiers.

Figure 2 shows a time slice at *t=1.88 s* through the seismic amplitude volume on which we identify channels (black arrows), high amplitude deposits (yellow arrows), and slope fans (red arrows). Figure 3a shows an equivalent time slice through peak spectral frequency co-rendered with peak spectral magnitude that emphasizes the relative thickness and reflectivity of the turbidite system and surrounding slope fan sediments into which it was incised. The edges of the channels are delineated by Sobel filter similarity. We show equivalent time slices through (Figure 3b) GLCM homogeneity, and (Figure 3c) co-rendered shape index and curvedness. In Figure 4 we show a representative vertical slice at line AA' in Figure 2 cutting through the channels through (Figure 4a) seismic amplitude, (Figure 4b) seismic amplitude co-rendered with peak spectral magnitude/peak spectral frequency, (Figure 4c) seismic amplitude co-rendered with GLCM homogeneity, and (Figure 4d) seismic amplitude co-rendered shape index and curvedness. Block arrows indicate several of the key facies. We note a time slice *t=1.88 s* where one can identify some of the incised valleys (black arrows).

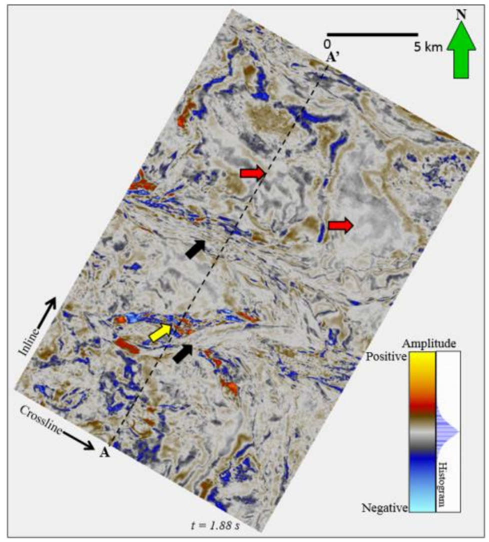
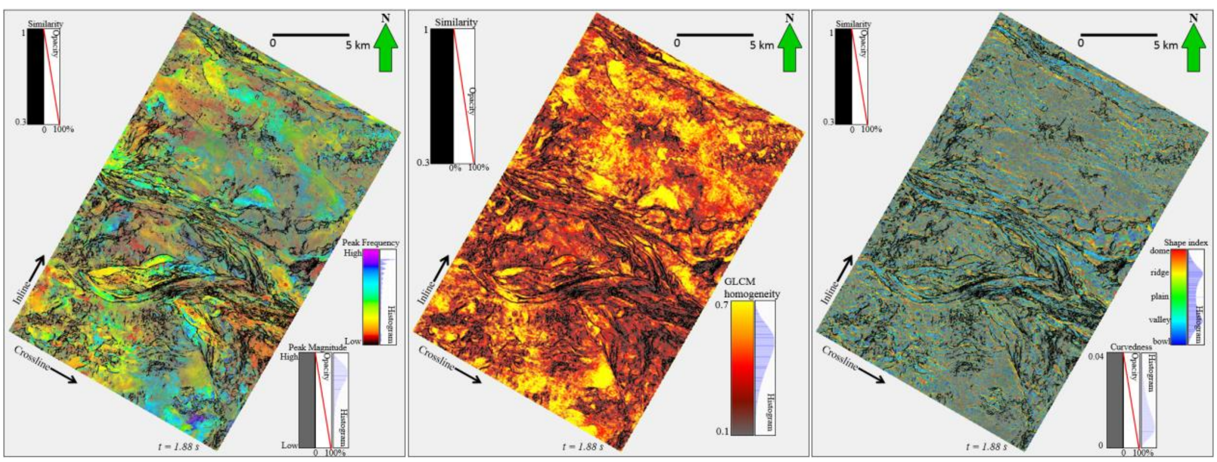


Figure 2. Time slice at *t=1.88 s* through the seismic amplitude volume. Black arrows indicate potential channel/ canyon features. The yellow arrow indicates a high amplitude feature. Red arrows indicate relatively low energy, gently dipping area. AA' denotes a cross section shown in Figure 4.



(a) **(b)** **(c)**

Figure 3. **a**) Time slice at *t=1.88 s* through peak spectral frequency co-rendered with peak spectral magnitude that emphasizes the relative thickness and reflectivity of the turbidite system and surrounding slope fan sediments into which it was incised. The two attributes are computed using a continuous wavelet transform algorithm. The edges of the channels are delineated by Sobel filter similarity. **b**) Time slice at *t=1.88 s* through the GLCM homogeneity attribute co-rendered with Sobel filter similarity. Bright colors highlights areas with potential fan sand deposits. **c**) Time slice at *t=1.88 s* through the co-rendered shape index, curvedness, and Sobel filter similarity. The shape index to be good for visual classification but dominates the unsupervised classifications with valley and ridge features across the survey.

Application

We perform generative topographic mapping (GTM) (Bishop et al., 1998) and proximal support vector machine (PSVM) (Fung and Mangasarian, 2001, 2005) classifications using the four seismic attributes that have been previously discussed. GTM is a nonlinear projection technique that provides a probabilistic representation of the data-vectors in a corresponding lower dimensional latent space, which in a geophysics sense means generating a continuous facies map from multiple seismic attribute inputs. PSVM is a supervised classification technique which can be used to recover the latent relation between seismic attribute inputs and a target property (facies, rock properties, or engineering parameters).

First we apply GTM to the four selected attributes. We compute two "orthogonal" projections of data onto the manifold and thence onto the two dimensions of the latent space. Rather than define explicit clusters, we project mean posterior probability distribution onto the 2D latent space and then cross-plot onto the 2D space where different clusters can be later defined. We also map the latent space against a 2D color bar (Figure 5a). In this slice, we have channels delineated by purple colors (black arrows), point bar and crevasse splays in pinkish colors (yellow arrows), and slope fans in lime colors (red arrows). We can also identify some thin, braided channels at the south end of the survey (blue arrow). We use similarity to separate the incised valleys from the slope fans. However, the geological meaning of the orange colored facies is somehow vague. This is the nature of unsupervised learning techniques that the clusters represent topological differences in the input data vectors, which are not necessarily the facies differences we wish to delineate. We can ameliorate this shortcoming by adding a posteriori supervision to the GTM manifold. The simplest way to add supervision is to compute the average attribute vectors about a given seismic facies and map it to the GTM crossplot. Then, the interpreter can manually define clusters on the 2D histogram by constructing one or more polygons (Figure 5b), where we cluster the data into four facies: channel thalwegs (blue), high-energy point bar and crevasse splay deposits (yellow), slope fans (green), and "everything else" (red). A more quantitative methodology is to mathematically project these average clusters onto the manifold, and then cross multiply the probability distribution of the control vectors against the probability distribution function of each data vector, thereby forming the Bhattacharya distance (Roy et al., 2013, 2014). Such measures then provide a probability ranging between 0 and 100% as to whether the data vector at any voxel is like the data vectors about well control (Roy et al., 2013, 2014) or like the average data vector within a facies picked by the interpreter.

The posterior supervision added to GTM is the critical prior supervision necessary for supervised classification such as PSVM. In this study we use the same four attributes as input for both unsupervised and supervised learning techniques. Our supervision consists of picked seed points for the three main facies previously delineated using the unsupervised classification results, which are channel thalwegs, point bar and crevasse splay deposits, and slope fans, plus one more facies of channel flanks. The seed points are shown in Figure 6a. We then compute averages of the four input attributes within a 7 trace X 7 trace X 24 ms window about each seed point to generate a training table which consists of four-dimensional of input attributes and one dimensional targets (the labeled facies). We cluster our four-dimensional input data using PSVM. The training and cross-validation performance is 92% training correctness and 85% cross-validation correctness. Figure 6b shows the PSVM classification result at time $t = 1.88\ s$. Generally the PSVM map follows the same pattern as we have seen on the posteriori supervised GTM map, but because of the nature of supervised classification, the PSVM map does not provide the smooth transition as we would expect in the GTM map. PSVM picks out the main slope fans (black arrows), but missed some crevasse splays that are previously visible in the GTM map (yellow arrow). We also see a great amount of facies variation within the incised valleys, which is reasonable because of the multiple course changes of a paleochannel that results in multiple channel thalwegs. Finally, we note some red lines following NW-SE direction (red arrows) which correspond to acquisition footprint.

Conclusion

In this study we used both supervised and unsupervised classification techniques assisted by seismic attributes to perform a first-hand interpretation in a turbidite system. We demonstrate that unsupervised classification products can be used to construct not only an initial estimate of the number of classes, but also a validation tool to determine if separate clusters have been incorrectly lumped together. We advise computing unsupervised GTM prior to picking seed points for subsequent supervised learning, to clarify the topological differences mapped by our choice of attributes. Such mapping will greatly improve the picking confidence, because the seed points are now confirmed by both human experience and mathematical statistics.

Acknowledgement

We thank New Zealand Petroleum and Minerals for providing the Waka-3D seismic data to the public. Financial support for this effort and for the development of our own GTM and PSVM algorithms was provided by the industry sponsors of the Attribute-Assisted Seismic Processing and Interpretation (AASPI) consortium in the University of Oklahoma. We thank colleagues for their valuable input and suggestions. All the 3D seismic displays were made using licenses to Petrel, provided to the University of Oklahoma for research and education courtesy of Schlumberger.

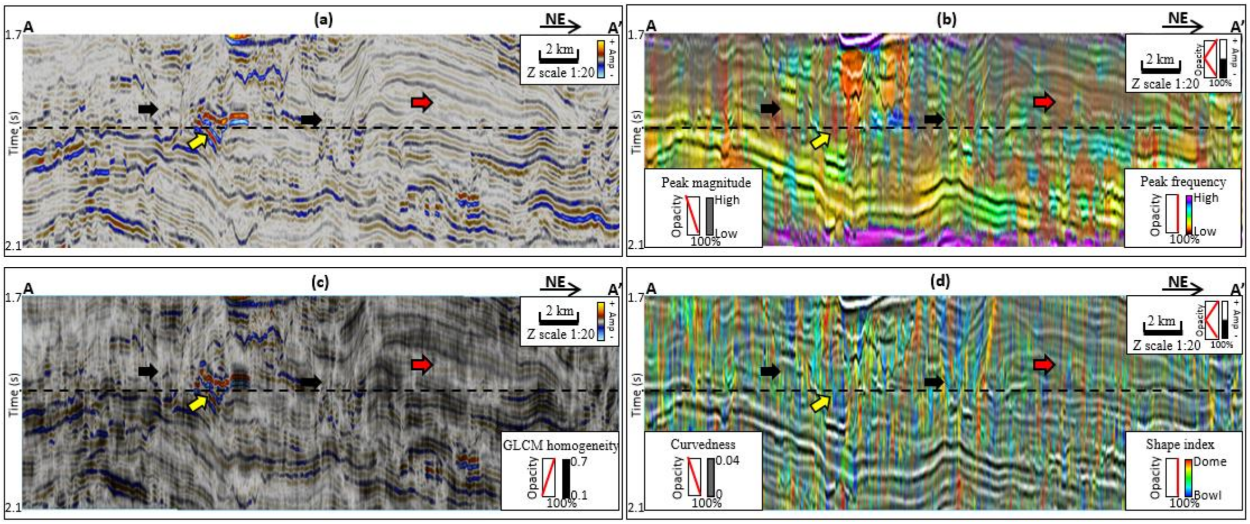


Figure 4. Cross sections along line AA' (location shown in Figure 2). **a**) Seismic amplitude. **b**) Seismic amplitude co-rendered with peak spectral magnitude and peak spectral frequency. **c**) Seismic amplitude co-rendered with GLCM homogeneity. **d**) Seismic amplitude co-rendered with shape index and curvedness. Black arrows indicate incised channel and canyon features. The yellow arrow indicates at a high amplitude reflector. Red arrows indicate relatively low amplitude, gently dipping areas.

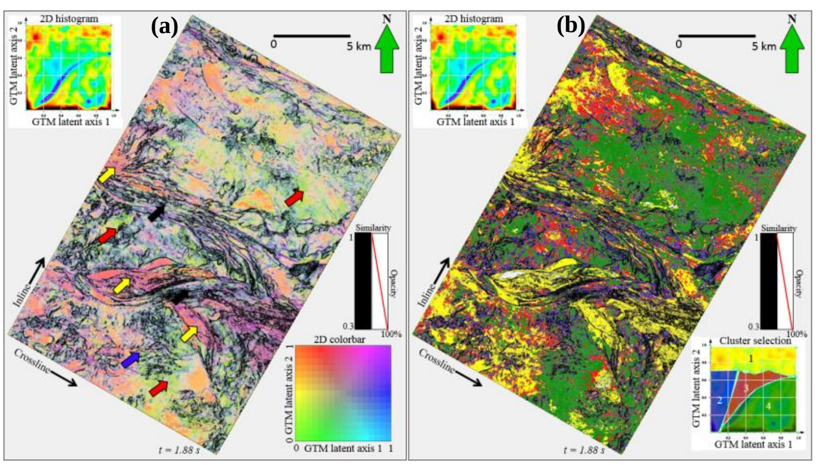


Figure 5. **a**) Time slice at *t=1.88 s* through crossplotting GTM projection 1 and 2 using a 2D colorbar. Black arrows indicate channel-like features. Yellow arrows indicate possible overbank deposits. Red arrows indicate possible slope fan deposits. The blue arrow indicates a braided channel system. The colors indicates the location of the mean probability each data vector represented by the Gaussians on the deformed 2D manifold then projected onto a 2D latent space. **b**) Same time slice but manually picking GTM clusters. Four facies are defined on the posterior probability distribution, with blue being channels /canyons, green being slope fan deposits, yellow being overbank deposits, and red being "everything else".

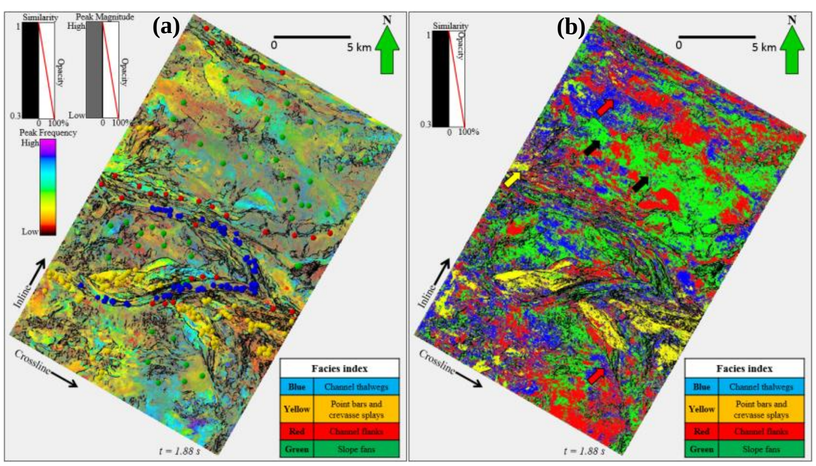


Figure 6. **a**) Time slice at *t=1.88 s* through co-rendered peak spectral frequency, peak spectral magnitude, and Sobel filter similarity volumes. Seed points (training data) are shown with colors for the picked four facies, blue being channel thalwegs, yellow being point bar and crevasse splays, red being channel flanks, and green being slope fans. Attributes at these seed points are used as training data in supervised classification. **b**) Same time slice through PSVM classification result. Black arrows indicate more correctly classified slope fans. Yellow arrow indicates crevasse splays. Red arrows show the misclassifications due to possible acquisition footprint.