Interpreter-Driven Multiattribute Classification

Sutton, Katy
Paton, Gaynor
ffA
The Northpoint Suite e3
Aberdeen Science & Energy Park
Exploration Drive
Aberdeen
AB23 8HZ, UK

Lowell, James
Purves, Stephen
ffA
Generator Studios
Trafalgar Street
Newcastle Upon Tyne
NE1 2LA, UK

Abstract

Facies classification by means of multiattribute clustering is becoming a more widely used method of identifying and extracting multiple facies within a dataset. Typically these methods are entirely data driven, have minimal input, and therefore do not take into account any *a priori* knowledge from the interpreter. We present a new interpreter-driven technique that provides simple, real time interaction with multidimensional attribute space.

Volumetric attribute analysis is an effective way of isolating individual properties of the seismic response. These properties can be drawn from models of the wavelet, from local geometry and shape, or from statistical behavior. Typically, more than one attribute is required to fully characterize an area of interest. Simultaneous interpretation of multiple attributes has been greatly facilitated by modern color blending and opacity blending schemes. However, objectively the simultaneous analysis of multiple attributes is difficult for the human interpreter. The interpreter-driven technique presented here is based on diverse context information stored in the human brain. By defining facies within the 3D sub-surface space these contexts are converted into class, or cluster, centers. These class centers define seismic facies based on the statistical distribution of the attributes included in the classification.

We apply our methodology to diverse hydrocarbon provinces at both regional and reservoir scale classification.

Introduction to Interpreter-Driven Multiattribute Classification

In order to gain a full understanding of geological features using seismic analysis, multiple attributes need to be considered and analyzed to best represent the different seismic facies. Techniques available to examine multiple attributes simultaneously vary from visually blending the attribute volumes (Leppard *et al.*, 2009; Henderson *et al.*, 2007) to sophisticated multiple attribute classification schemes. These classification methods are typically data-driven and often by their nature do not have the flexibility to be guided by the interpreter, such that the results can be difficult to understand. In this paper we present a new real-time, interactive, interpreter-centered classification technique.

Classification Background

Data classification is the process that allows correlations or patterns to be found in high dimensional data. Classification can be defined by two distinct approaches. In the first supervised learning approach, we are given a set of observations and classes are established to cluster the data. Typically, a volume that has been labeled with multiple examples (or observations) of geological features (predefined classes) such as channel and overbank deposits. The computer system learns from these data representing "past geological experiences" and places new observations into one of these known classes. In the second unsupervised learning approach, , the computer defines geological classes or (or features) without any prior knowledge other than an estimate of the number of classes.

In order to avoid the time-consuming construction of training data required for supervised learning, we have developed an Interactive Facies Classification Tool (IFC) whereby the interpreter defines a subset of the classes found in unsupervised learning (Fig. 1).

One of the most popular unsupervised classification techniques is cluster analysis. Clustering is the processing of partitioning a set of data into meaningful sub-classes, or clusters. A cluster is a collection of objects (in this case voxels) that are similar to one another and can be treated collectively as one group. Using a simplistic example to explain this concept further, if a volume contains three colors, red, green and blue, we could naturally divided the volume into three groups—one for each color, and each of these groups would also have a class center. However, if the interpreter wanted only two groups, the question would be how the volume should be divided. In this example there would be three options red and not red, blue and not blue or green and not green. If the interpreter wanted to insure the volume was sub-divided into red and not red voxels the position of the class centers could be influenced.

In the IFC tool, a composite of attribute volumes are divided into regions with similar characteristics, thus helping us understand the natural grouping or structure of the volume set. Clustering algorithms can be divided into four groups: exclusive, overlapping, hierarchical, and probabilistic. In exclusive clustering, data are grouped in such a manner that if a given datum belongs to one definite cluster it cannot be a member of any other cluster. In overlapping clustering, data are clustered using fuzzy sets, so that each datum may belong to two or more clusters with different degrees of membership. Hierarchical clustering is based on the union between the two nearest clusters. The beginning condition is realized by setting every datum as its own cluster. After a few iterations we reach the final desired number of clusters. Probabilistic clustering is based on a model based, such as Gaussian statistics.

Before any classification method is used it is important to select attributes that represent different seismic characteristics of the dataset. The IFC tool uses exclusive clustering, in which every voxel is assigned to one and only one class. The assignment is achieved using a similarity measure in multidimensional attribute space. One dimension is added for every volume that is included in the analysis. Thus if we have one attribute volume, we have a one dimensional space. If we have two attribute volumes, we have a two dimensional space, If we have five–attribute volumes, we have a five dimensional space. With the latter, each voxel would have five corresponding attribute values, one for each volume.

With a single voxel having representative values from multiple volumes, it is important that each volume have equal importance. Variations in dynamic range of attribute volumes can cause unequal attribute contributions and result in biased classifications. To compensate for these differences in attribute contributions, volume standardization, and normalization is performed.

Although this classification technique is unsupervised, the interpreter can add supervision and thus influence the classification by determining the location of the class centers. The class centers are selected by simply defining a point or polygon on or within any given attribute volume, which then defines a position in the feature space. The classification can be further influenced by altering the relative weights of each component attribute to bias the classification to achieve the desired result (Fig. 1).

Once two or more classes have been created, all voxels are then assigned to the closest class center in n-dimensional attribute space. The fit of the classification for each class can be investigated by viewing the distance metric. This metric assigns a value describing how similar each point is to the picked point by calculating the minimum distance between the voxel and class center (Fig. 2). The distance metric indicates which areas are poorly represented by the classification and therefore alerts the interpreter that additional classes may be needed.

Onshore USA Case Study 1

Geologic setting

Teapot Dome (also known as Naval Petroleum Reserve No. 3) dataset is located 27 miles north of Casper, Wyoming, in the southwest part of the Powder River Basin and is a northwest- southeast trending anticline that is a southward extension of the Salt Creek anticline (Miliken 2006). The anticline has formed due to basement thrusting that can be observed in the Teapot Dome dataset. The Teapot Dome Field has numerous producing strata but the focus in this study is the Pennsylvanian (Upper Carboniferous) Tensleep Formation (Fig. 3). The general stratigraphic sequence consists of 300 ft of Devonian which unconformably underlie the Upper Mississippian Madison Formation carbonates, which in turn underlie unconformably 160 ft of Pennsylvanian Amsden Formation dolomite. The 320 ft thick Tensleep Formation consists of aeolian sandstones interbedded with dolostone shallow marine carbonates. The top of the Tensleep Formation is marked by the Red Shale that is found throughout the region.

Two levels of classification have been computed for this study: the objective of the first classification at the regional level is to separate the dome from the basement and the thrust deformation as well as to identify the different lithology in the dome itself. The second classification focuses on the Tensleep sandstones and has the objective of identifying the reservoir.

Regional geological classification

With interpreter-driven facies classification, large scale regional geological events can be easily delineated, such as the thrust graben and high amplitude reflectors seen in the Teapot Dome data in Figure 4, where the thrust belt and basement appear as red. The antiformal structure consists of large sand bodies separated by dolomites, muds, and silts that are folded against the thrust. The regional geology has been captured by classifying an envelope and chaos attribute (Fig. 5). The envelope attribute is calculated using a Hilbert Transform and identifies high amplitude zones in the reflectivity. The chaos attribute is computed from the eigenvalues of the gradient structural tensor and measures chaotic behavior by examining the local orientation in the data.. In this case, it detects the deformation associated with the thrusting and the chaotic reflectors of the basement.

Figure 4 shows the classification result with the thrust and basement zones defined by the interpreter to be the red class. There are some regions on the apex of the dome which have been included in the thrust /basement class suggesting a similar seismic character is present in these areas. Within the dome, the high amplitude coherent sandstones fall into the yellow class, and the lower amplitude shale or dolomite strata fall into the blue class.

This regional classification is useful for determining gross structures and as a reconnaissance volume to determine the areas for future focus. Table 1 lists the different classes that were computed.

Detailed geological classification

The Teapot Dome Field has numerous producing strata. The focus of this effort is the detailed geological classification of the Pennsylvanian Tensleep Formation that consists of aeolian sandstone dunes interbedded with shallow marine dolostone carbonates. The top of the producing sandstone is marked by the Tensleep horizon in Figure 5. In the Tensleep Formation the reservoir is found in the southwest portion of the Teapot Dome survey and is approximately 5400 ft deep and has a net thickness of 50 ft. The reservoir forms a structural high but is also bounded in the north by a fault (fault 1 in Fig. 3). The main Tensleep production has been from the crest of the structural high; exploratory wells outside of this high have not been commercially viable. The reservoir is water driven, with regeneration from the Big Horn Mountains.

In order to give the Tensleep reservoir its own class to separate it from the surrounding sandstone and to investigate why the productivity is variable across the Tensleep sandstone, dip and envelope attributes have been used in the classification, as dip can indicate structural variations and envelope can indicate pore fill and lithology variations.

The instantaneous dip attribute is calculated using the gradient structural tensor that provides small scale (high resolution) estimates of the dominant orientation at every point in a 3D volume. The dip calculation can also be run on a large scale to show regional structure by using a larger window containing more traces in forming the gradient structural tensor.

Classification based on these two attributes gave rise to the four classes shown in Table 2. Class centers picked on the Tensleep horizon (Fig. 6) consists of the reservoir sandstone (Class 1 in pink), the nonproducing sandstone (Class 2 in turquoise), the dolomites (Class 3 in blue) and the surrounding strata (Class 4 in red).

Calibration

In the full volume, the lateral extent of the classes can be examined and correlated with well information. Class 1 (pink) is the Tensleep reservoir, Figure 6 shows that the class is within the structural high as indicated by the yellow Tensleep horizons. The classification also shows that the reservoir class has not extended past the fault to the north marked as Fault 1 in Figure 3. These results correlate with the known production in the field.

Once the classification has been calibrated at a known reservoir, the resulting volume can allow rapid identification of other zones in the seismic amplitude and attribute volumes that have a similar character.

Onshore USA Case Study 2

Geologic setting

The Stratton dataset is located in South Texas and consists of Oligocene Frio and Vicksburg formations. The reservoirs here are found in thin-bedded compartmentalized fluvial systems. The fluvial reservoirs are narrow and meandering, as thin as 10 ft and as narrow as 200 ft (Hardage 1994). In this paper we examine a channel in the upper Frio that is just at the limit of seismic resolution (Fig. 7). The channel is expressed in the amplitude data as a continuous reflector that shows a slight dip and a drop in amplitude. This channel has proven difficult to extract by conventional means such as single attribute segmentation; however, it has been classified successfully using this multiattribute classification tool.

Channel classification

The selection of attributes is key to any successful classification and this channel shows up clearest in flexure and deformation attributes, although the single attributes failed to define the channel on their own. Flexure is a form of curvature analysis and often is used to pick up faults that are expressed as an inflection (or flexure) in the reflectivity data, as well as characterizing faults and seismic distortion around faults throughout a dataset. However, flexure is also useful for characterizing many other geological formations across a range of scales, including the thin channels seen in this dataset.

Deformation is an orientation-based attribute that analyses the distribution of the 3D orientation vectors within a neighborhood surrounding the current point. The process creates a number of outputs, in which each output describes a different aspect of the vector distribution within the analysis window. One of these outputs is the mean norm which calculates the average length of the component vectors and is effective at highlighting this channel.

The class picks have been made directly on the reflectivity shown in Figure 7 and are defined to denote a channel (yellow), overbank deposit (green), over burden (red) and surrounding strata (blue) facies (Table 3). The result of this classification can be observed in Figure 8 where the discrete channel is classified.

Using the classification

Since the channel is classified as a single seismic facies, it can be extracted as a discrete geobody. Once extracted, the size and thickness of the channel can be measured; the thickness variation can be seen in Figure 9. The upper and lower surface of the channel can be extracted and then embedded back into the reflectivity data, thus allowing the subtle channel to be traced in the data.

Offshore North Sea Case Study

This dataset is from the mid-Norwegian North Sea and exhibits distinctive mud slides in different strata. It is a simple exercise to extract these mud slides as geobodies as they have a distinctive low amplitude, chaotic reflectors. What is more interesting is to determine if the classification can highlight multiple sliding episodes. The input attributes for the classification are eigenvalues of the gradient structural tensor, structurally-orientated semblance, and texture.

Three eigenvalues (and corresponding eigenvectors) are computed from the local gradient structural tensor which in turn is constructed by cross-correlating the partial derivatives of the data within the data analysis window. By construction eigenvalue 1, λ_1, represents the largest component, eigenvalue 2, λ_2, the intermediate component, and eigenvalue 3, λ_3, the smallest component of the gradient strutcture tensor. Constant amplitude planar events are represented by $\lambda_1 >> \lambda_2 > \lambda_3$. Cylindrical consistent amplitude events such as anticlines and synclines are represented by $\lambda_1 \approx \lambda_2 > \lambda_3$. Chaotic events are characterized by $\lambda_1 \approx \lambda_2 \approx \lambda_3$. Eigenvalue 3 is used in this classification as it highlights the zones in the mud slide that are very variable. The structurally oriented semblance process computes semblance along reflector dip and azimuth and is sensitive to faults that are expressed as sharp offsets and clear discontinuities. The structurally oriented component of the calculation ensures the attribute is not strongly influenced by steeply dipping reflectors. This attribute is useful to understand the structural elements of the dataset to determine if there was any associated toe thrusts with the mud slide.

The texture attribute provides a statistical measure of regions of different texture within the seismic data. It is a combination (in this case multiple) of the standard deviation of the reflectivity data and the envelope of the data, enabling one volume to represents both amplitude and variability. Texture is designed to highlight large-scale deep-water turbidite channels and has been found to have application in many situations when an object is visible due to differences in both the amplitude values and the stability of the seismic response. The texture delineated the low amplitude chaotic regions of the mud slide.

Using these three attributes places the gross mud slide within the blue class (Fig. 10, Table 4). Within the mud slide there is a sub-class that is higher in amplitude (the green class) in the vertical center of the gross mud slide that moves to the west with increase in time. The faults zones have been classified as the yellow class and the surrounding strata as the pink class.

Conclusions

Multiattribute classification techniques are becoming more widespread as a result of the increased integration of attribute analysis into the interpretation work flow. Many of these techniques produce classified volumes that are hard to interpret due to the lack of input by the user. The technique presented here allows the interpreter to be central to a data-driven classification process enabling geologic knowledge to guide the process. This technique has proven to be successful in classifying many different geological settings, and the results presented here illustrate its applicability and versatility. The regional geological classification of the Teapot Dome dataset allows rapid understanding of key geological elements. The detailed classification enables the Tensleep reservoir to be isolated from the nonproducing Tensleep sandstone and to investigate the variations in well production. A previously difficult-to-extract channel in the Stratton study has been classified successfully and isolated, permitting thickness variations to be determined. The North Sea case study not only has classified the mud slide but shows variations within it, as well as identifying the faulting.

Acknowledgments

All examples and images were created using ffA's SVI Pro software. RMTOC. The US Department of Energy and Statoil are acknowledged as the data providers.

References

Hardage, B.A., R.A. Levey, V. Pendleton, J. Simmons, and R. Edsen, 1994, A 3-D seismic case history evaluating fluvially deposited thin-bed reservoirs in a gas-producing property: Geophysics, v.59, p. 1650-1665.

Henderson, J., S. J. Purves, and C. Leppard, 2007, Automated delineation of geological elements from 3D seismic data through analysis of multichannel, volumetric spectral decomposition data: First Break, v. 25, p. 87-93.

Leppard, C., S. Purves, J. Henderson, and A. Eckersley, 2009, Quantifying the temporal and spatial extent of depositional and structural elements in 3D seismic data using spectral decomposition and multiattribute RGB blending: http://www.searchanddiscovery.com/abstracts/html/2009/annual/abstracts/leppard.htm.

Milliken, M. and B. Black, 2006, Surface geologic mapping, seismic, and well data resolve complex faulting at Teapot Dome field, Natrona County, Wyoming: Presentation 2006 WGA Field Conference.

Table 1. Teapot regional geology classes.

Class	Color	Feature
1	Red	Basement, thrust deformation
2	Blue	Predominately shale or dolomite
3	Yellow	Sandstone

Table 2. Teapot Tensleep reservoir classes.

Class	Color	Feature
1	Pink	Reservoir sandstone
2	Turquoise	Nonproducing sandstones
3	Blue	Dolomites
4	Red	Surrounding strata

Table 3. Stratton channel classes.

Class	Color	Feature
1	Yellow	Channel
2	Green	Over banks
3	Red	Over burden
4	Blue	Surrounding strata

Table 4. Mud slide classes.

Class	Color	Feature
1	Blue	Mud slide
2	Green	Mud slide sub class
3	Yellow	Fault zones
4	Pink	Surrounding strata

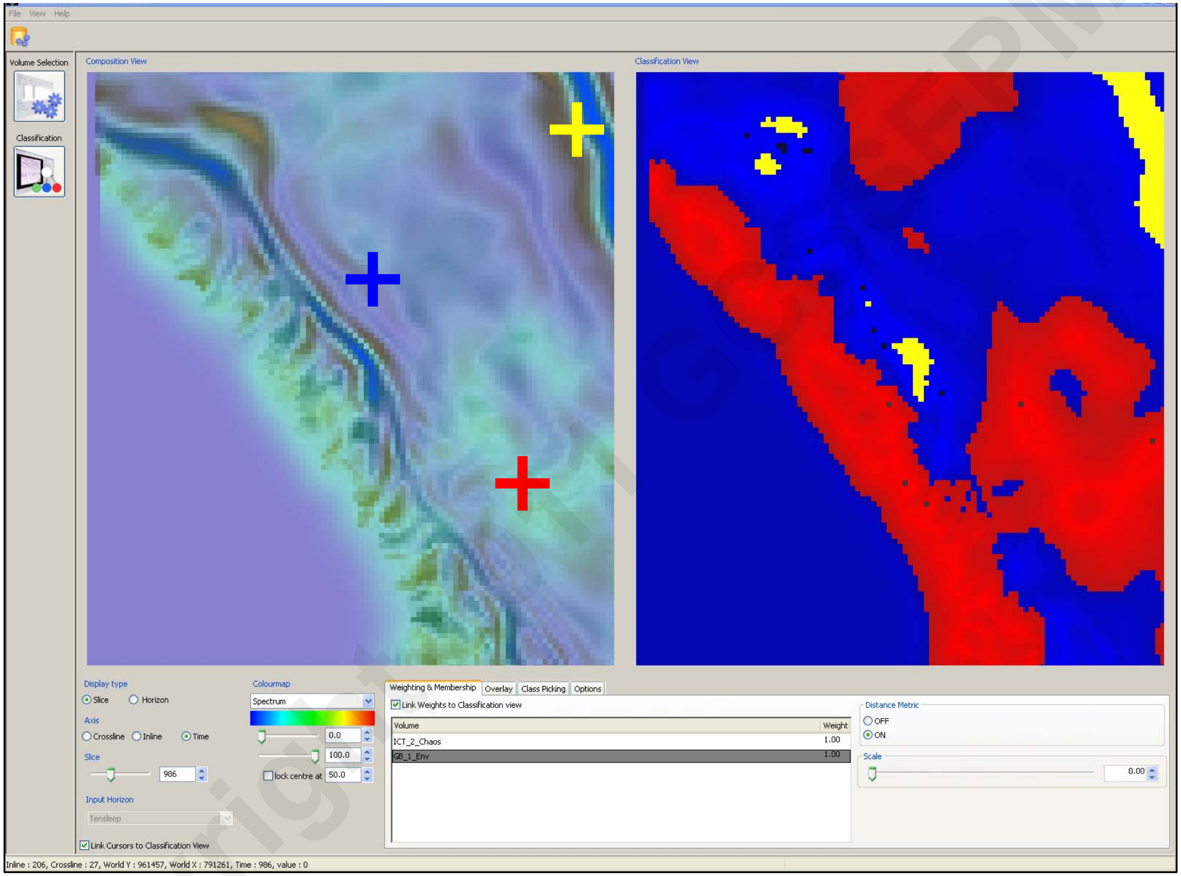

Figure 1. The Interactive Facies Classification Tool user interface. The composition view is in the left panel and the resulting classification view is in the right panel. Points are picked in the composition view and are the centers of the classification view in the right panel.

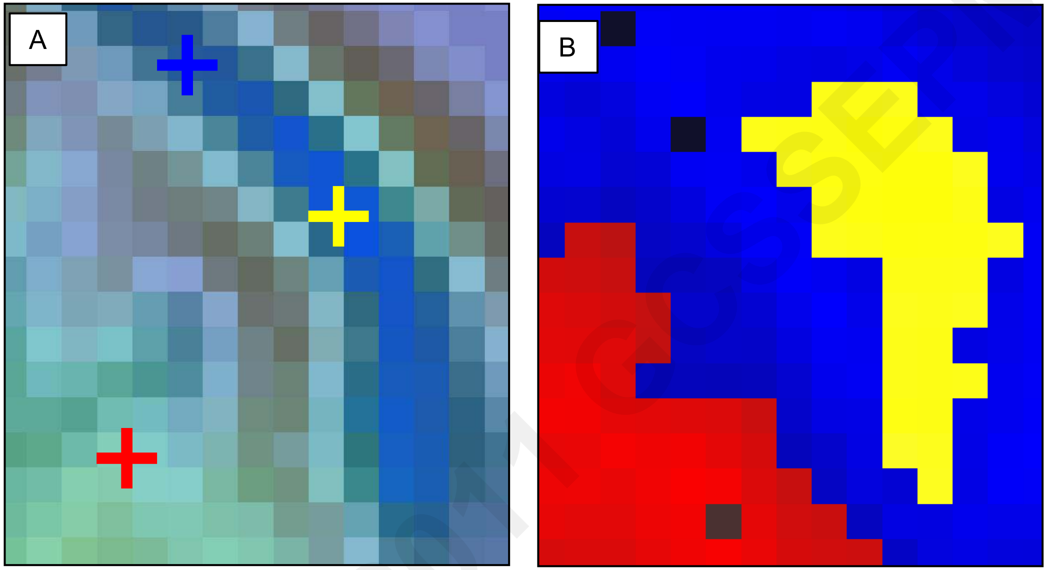

Figure 2. Distance metric where classes picked in (A) mark the class center and all other voxels are assigned to one of three classes. The saturation of color in (B) shows how well the voxel fits in the class; the lighter the color, the better the fit.

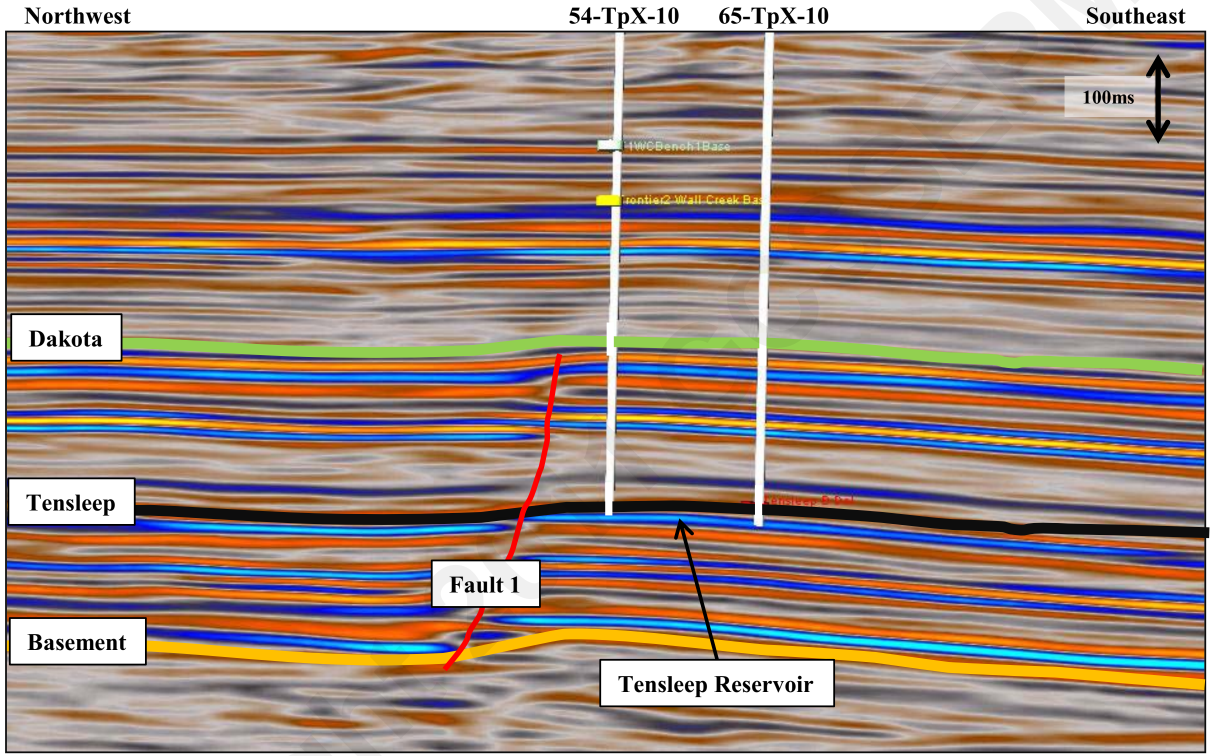

Figure 3. Northwest-Southeast arbitrary line through Tensleep reservoir; 54-TpX-10 and 65-TpX-10 are wells. The Tensleep horizon is picked as black, Dakota as green, and basement as yellow. Fault 1, located by the reservoir, is marked.

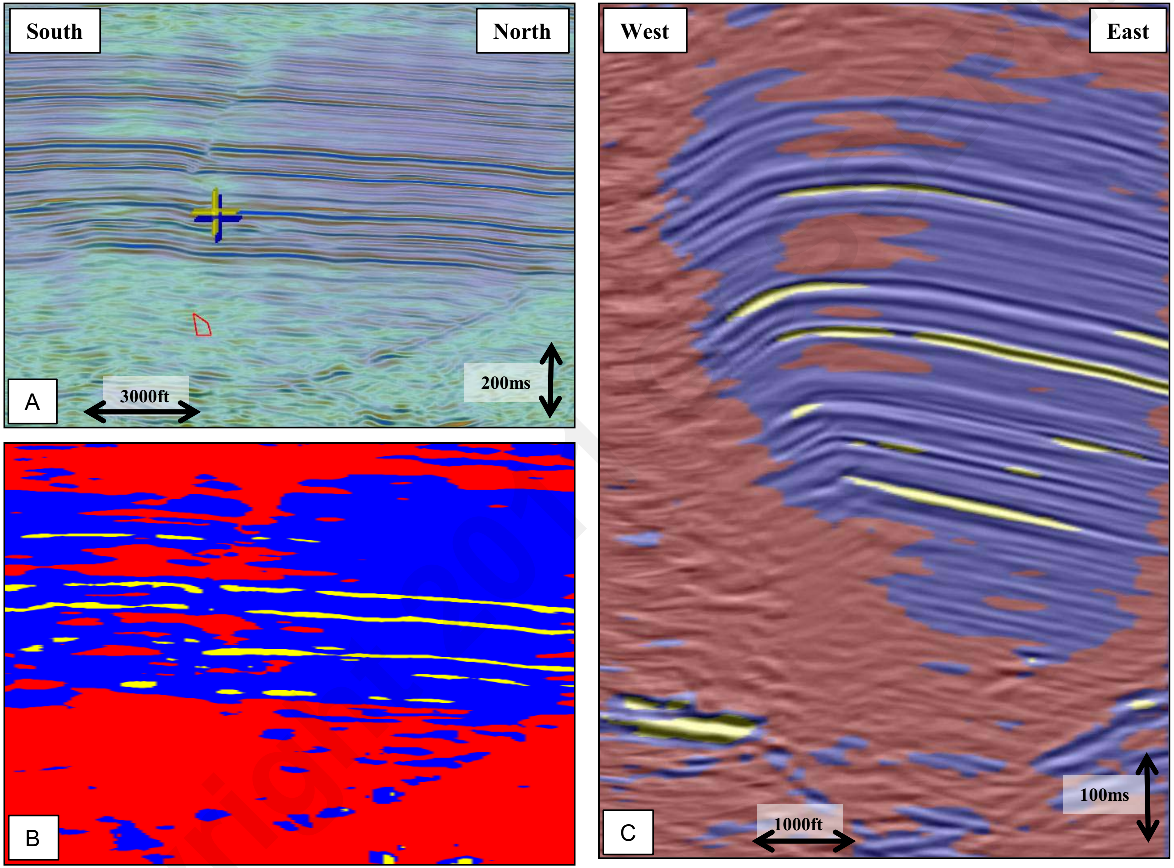

Figure 4. Classification of regional geology for Teapot Dome. (A) Composite view with interpreter picks. (B) The resulting classification. (C) Classification of regional geology; reflectivity is opaquely overlain. In the classification view, red is thrust deformation and basement, blue, is predominantly shale or dolomite facies, and yellow is sandstone.

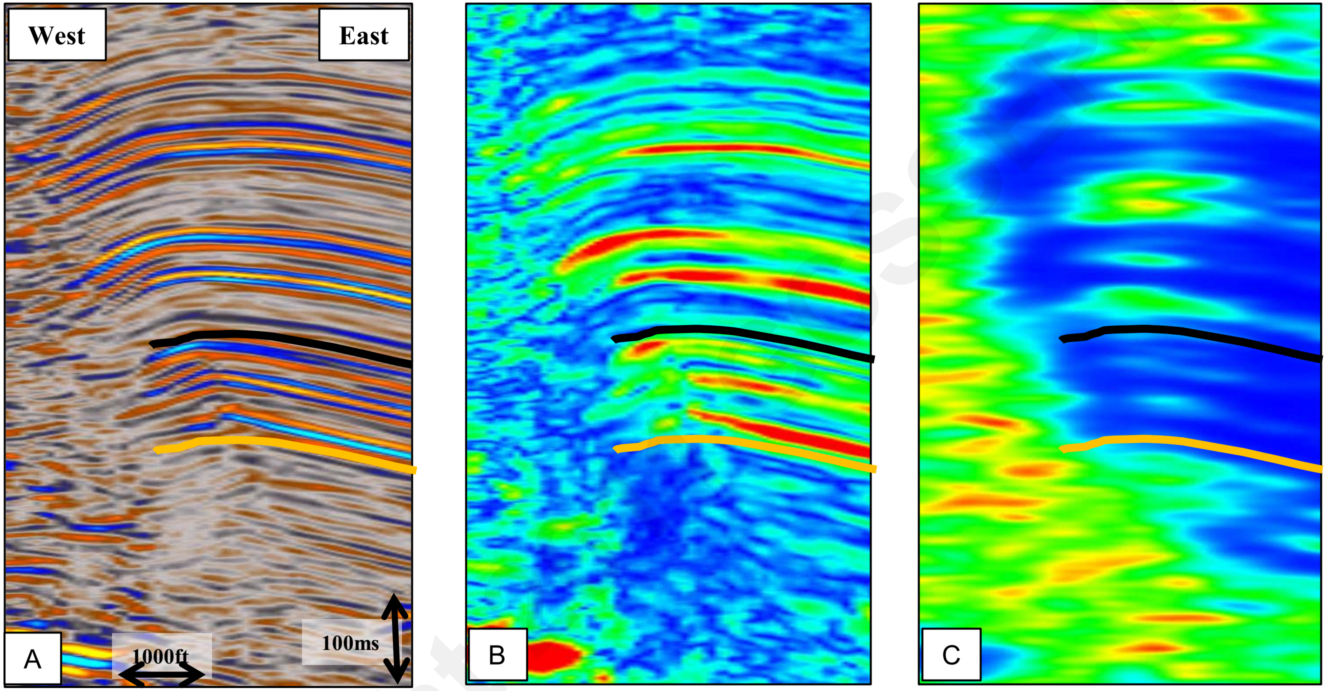

Figure 5. Attributes used in the classification of Inline 109. Tensleep horizon has been marked in black and basement as yellow. (A) Reflectivity. (B) Amplitude; red is high amplitude and blue is low amplitude. (C) Chaos; red is high chaos, blue is low chaos.

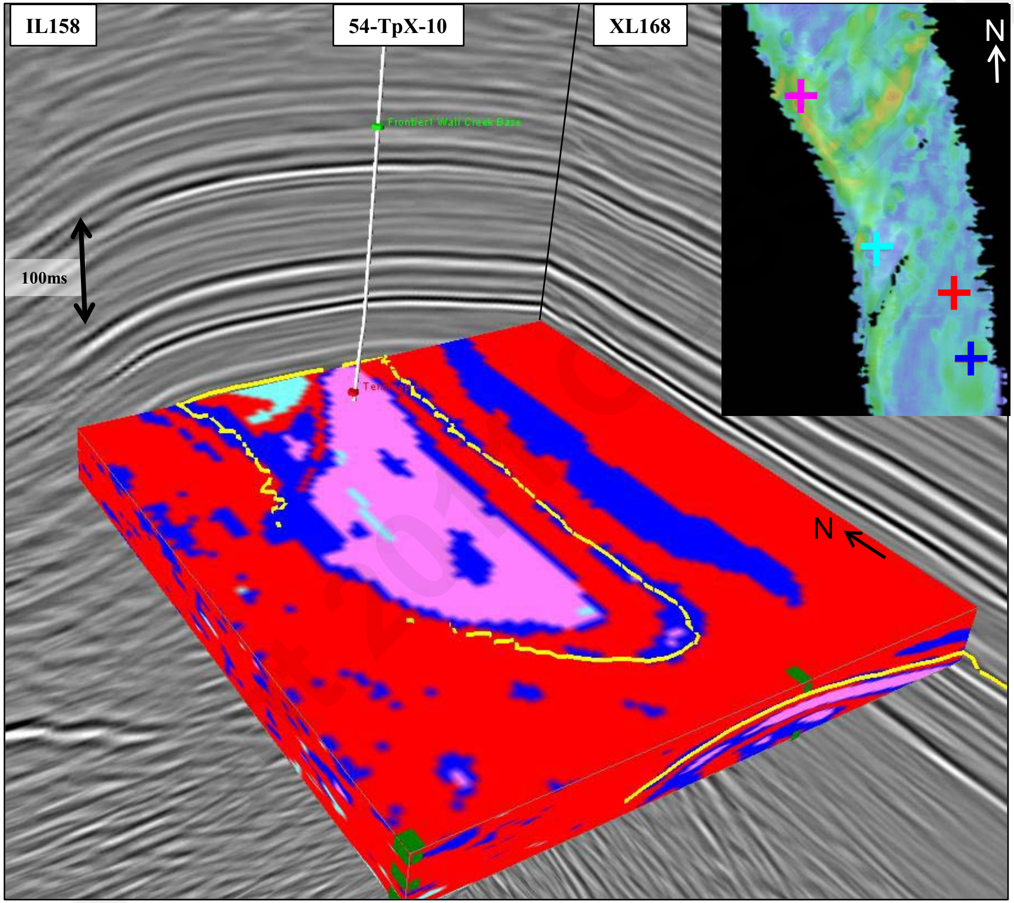

Figure 6. Classification of Tensleep horizon (insert) to identify reservoir. Probe shows Time Slice 1058. Pink class is reservoir; red is surrounding strata; blue is dolomite; and turquoise is nonreservoir sandstone. Yellow is the Tensleep horizon.

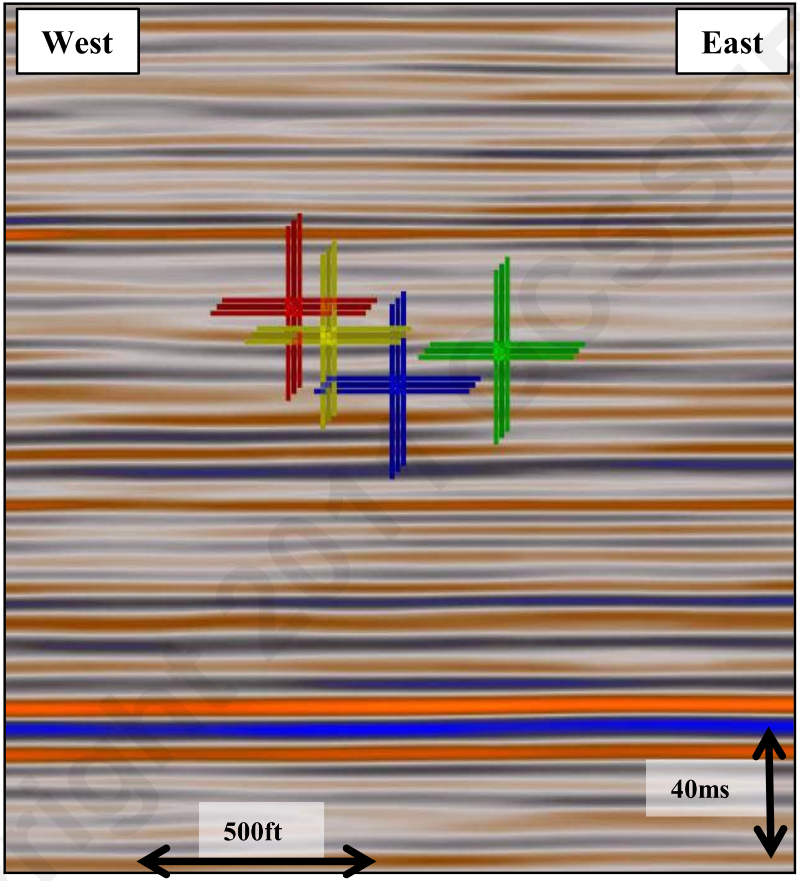

Figure 7. Classification picks of upper Frio channel in Stratton dataset, Cross Line96. Yellow pick is the channel, green pick is channel overbank, red pick is channel overburden, and blue pick marks surrounding strata.

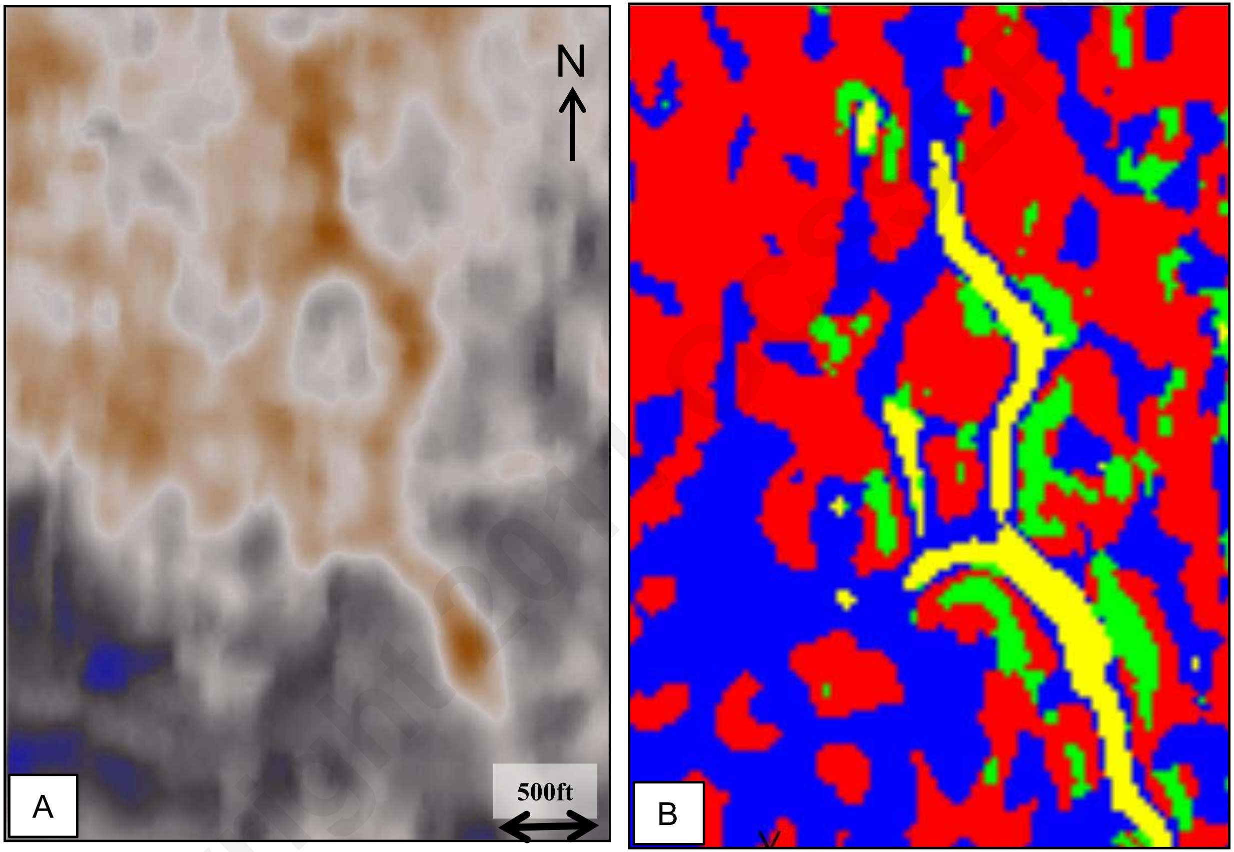

Figure 8. Classification of upper Frio channel in Stratton dataset on Time Slice 834. (A) is reflectivity and (B) is classification. Yellow class is on the channel; green class is on channel overbank; red class is on channel overburden; and blue class is on surrounding strata.

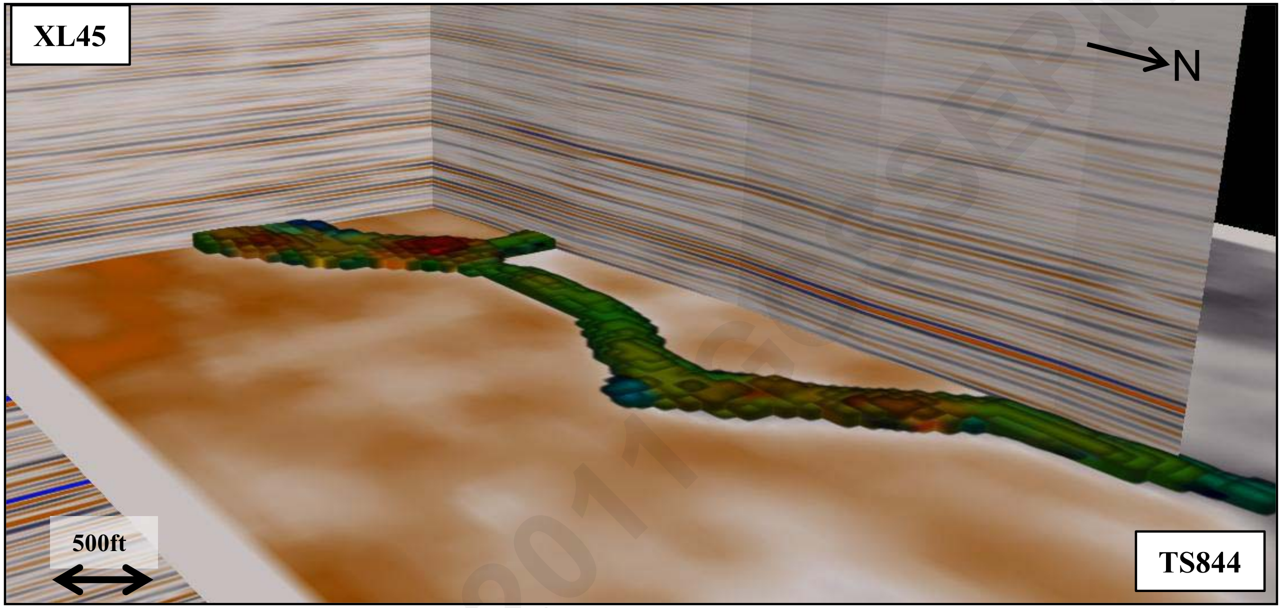

Figure 9. Thickness analysis on the classified upper Frio channel, Time Slice 844. Thickness varies from blue to red.

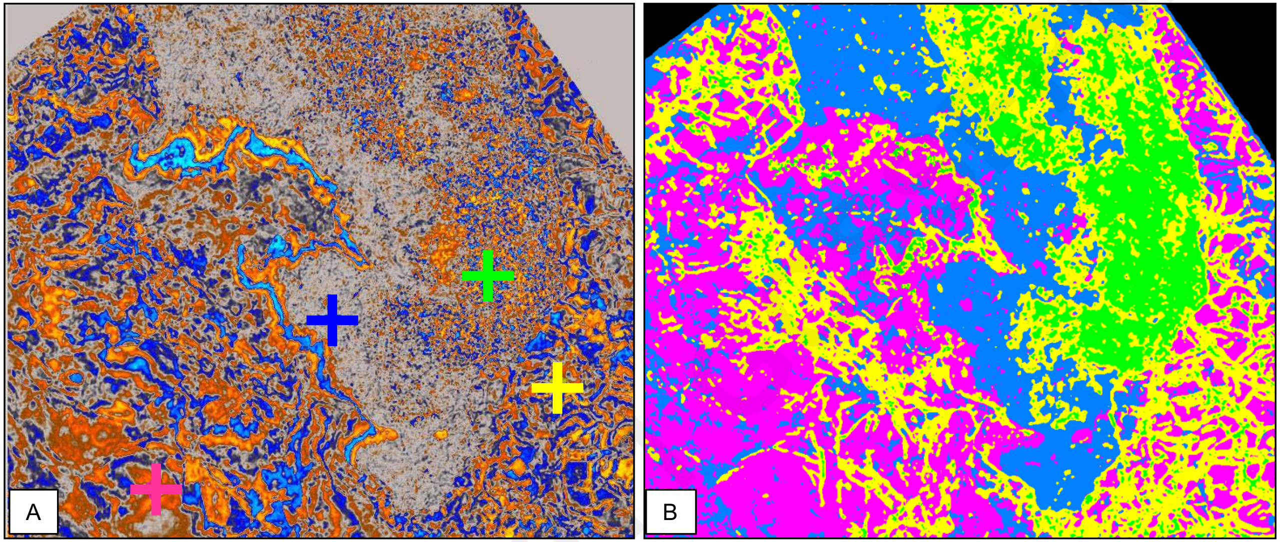

Figure 10. Classification of Norwegian North Sea mud slide. (A) is reflectivity and (B) is classification. Blue class is mud slide; green class is a subclass of mud slide; yellow class represents fault zones; and pink class is surrounding strata.