ASEG-PESA 2015
Geophysics and Geology together for Discovery
24th International Geophysical Conference and Exhibition
15 – 18 February 2015 Perth, Western Australia

# Using Multivariate Data Classification on Frontier Exploration Basins to Enhance the Information Value of Suboptimal 2D Seismic Surveys for Unconventional Reservoir Characterization

**David Handwerger**
*Schlumberger*
*1945 S. Fremont Dr.*
*Salt Lake City, UT 84104 USA*
*dhandwerger@slb.com*

**Roberto Castañeda-Aguilar**
*Schlumberger*
*1945 S. Fremont Dr.*
*Salt Lake City, UT 84104 USA*
*RAguilar2@slb.com*

**Don Stachiw**
*Northern Cross (Yukon) Ltd.*
*#850, 700-4th Avenue SW*
*Calgary, AB T2P 3J4 Canada*
*dstachiw@northerncrossyukon.ca*

**Roberto Suarez-Rivera**
*Schlumberger[1]*
*1945 S. Fremont Dr.*
*Salt Lake City, UT 84105 USA*
*roberto@wdvglab.com*

[1]*Now with Van Gonten and Associates, College Station, TX USA*

## SUMMARY

We developed a workflow that allows integrating legacy 2D seismic surveys with modern log and core data, validating their consistency, classifying them into rock classes with consistent properties, propagating material properties across each of these rock classes, and using this information to improve reservoir characterization and the assessment of their hydrocarbon resource potential. As proof of concept, we analyzed two intersecting 2D seismic lines shot in 2001 in a frontier basin in Canada to determine the distribution of reservoir quality. Each of these had been separately prestack inverted, but have modern core and log data (as well as legacy log data) which were integrated with the inverted attributes.

Results identify a most prospective class for reservoir quality within the zone of interest, and show that it increases in thickness to the south in the seismic section.

**Key words:** classification, unconventional, seismic, characterization

## INTRODUCTION

Many 2D seismic surveys that were shot decades in the past remain an invaluable tool for early hydrocarbon reservoir exploration. Those surveys were typically shot to gather basic geologic and structural information for understanding basin architecture prior to drilling initial wells. As unconventional reservoirs develop into resources of interest, these old surveys are being reinterpreted for additional information.

One useful mechanism for integrating data, especially among sources with different inherent scales (such as core, log and seismic) is to associate data to consistent patterns of inputs that represent consistent properties within the reservoir. For example, log data are influenced by rock material properties, which vary as a function of the rock texture and composition. Therefore, if one’s chosen inputs for a classification model are sensitive to variability in material texture and composition (as logs typically are), then the resulting ‘classes’ should isolate zones of consistent material properties. Such an approach was used previously for modern 3D seismic volumes (Suarez-Rivera et al. (2013), Handwerger et al. (2014)) in regions with extensive log and core data. In this study, we employ similar techniques on older 2D seismic data with fewer logs and core.

## METHOD AND RESULTS

Our goal was to interpret two intersecting legacy 2D surveys from an unconventional reservoir in a frontier Canadian basin (a situation not dissimilar to what is faced in Australia today) for rock material properties, following a process similar to Suarez-Rivera et al. (2013) and Handwerger et al. (2014) but modified to account for the differences in data type and quality. The core and log data that provide material properties for integration with the seismic data came from new wells, but the seismic data were collected in 2001, and were inverted separately prior to the recent drilling using legacy logs.

One well was drilled and logged recently on each of the lines (two total), and each of these was cored. The cores were primarily analyzed for petrophysical properties needed to make preliminary assessment of the reservoir quality. Data presented here include total porosity, hydrocarbon-filled porosity, water saturation, organic content and permeability.

The inversions produced only acoustic impedance and shear impedance attributes due to insufficient signal in the data for density extraction. Since the 2D lines were shot and inverted separately, the attribute values are not consistent between them; exemplified by a lack of agreement at their intersection. However, using multivariate classification, based initially on the log data, and integrated with the seismic we were able to associate consistent material properties to the seismic data in spite of this.

We first created a unified rock class model using unsupervised classification of the log data from both wells (figure 1). The input log data consisted of gamma ray, bulk density, neutron porosity, deep resistivity and P-wave travel time. The logs from each well were extracted over the same zone of interest and these data appended into a composite well for classification.

One advantage of an unsupervised class model is that it partitions the input data into classes with maximum uniqueness and minimal distribution of input data per class, if the number of classes chosen is optimized. This facilitates using the class model as input for supervised classification of subsequent logs, where one tries to recognize in those new logs the same previously defined classes in the model set. Once the classes are recognized in subsequent wells, their material properties, as determined through core analyses, are attributed to these same zones.

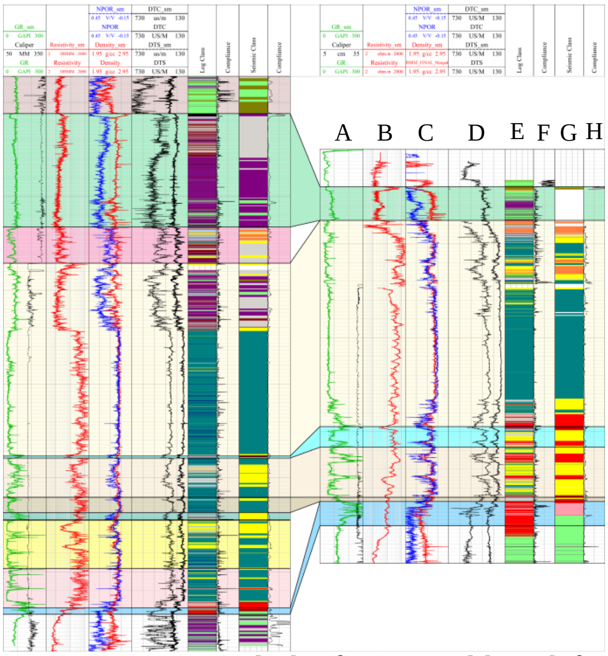


**Figure 1. Unsupervised classification model used for integrating core and log data with the seismic inversion. Two wells are shown. In each, the data shown are: A. Gamma ray and calliper log, B. Deep resistivity log, C. Neutron porosity (blue) and bulk density (red) logs, D. DTC and DTS logs, E. the unsupervised classification at log scale, F. the compliance to the log-scale model, G. the seismic-scale model, and H. the compliance to the seismic-scale model. For each log, both native resolution and smoothed are shown. The logs are shown at 1:3000 scale.**

We integrated the core data to the unsupervised class model by upscaling them to log resolution, then collecting the statistical distributions of each property for each class. The upscaling consisted of building relational models between core and log data for each class, where the samples were strategically chosen to maximize representation of the constrained distribution of logs within each class. These models were then used to predict each core property, class by class, from the multivariate logs at the sampling scale of log data (6" or 0.1524 cm). By mapping the core data to the log data in this manner, the core data are *de facto* upscaled, as the logs become the template for core property propagation. This also serves to distribute the core properties at a much greater sample density, thus improving the robustness of each class' distribution statistics.

The classification model from the logs, described above, was then upscaled to seismic resolution and used to provide training data for supervised classification of the inverted seismic attributes (AI and SI). The log data were first upscaled to seismic resolution using a smoothing filter, and then the classification was rerun using the smoothed logs (figure 1).

We trained the attributes to recognize each upscaled class at the well trace and its adjacent trace in each direction along the 2D line (to boost the training attribute population). We performed extensive QC and data filtering as part of the training scenario to constrain the outcome to unique responses for each class, given that we change data types from logs, with which we define the classes, to the acoustic attributes (AI and SI) used to recognize the classes in 2D away from the wellbores.

To propagate the classes through the full 2D sections, we applied the training rules to each 2D line's inversion results (figure 2) separately, since the inversions are independent, and inconsistent. Despite this, at the intersection trace of each line the classes are consistently aligned. However, one needs to recognize that not all classes were available to train both lines. This is likely due to the fact that the wells were drilled in different depositional environments (one on the self, the other entirely in the off-shelf basin), and so show different rock properties. For example, orange in one line was not available to train the other, as is also the case for grey and purple. However, the teal colour exists in the training data of both lines and meets conformably at the intersection of the two lines. This will be discussed further, below.

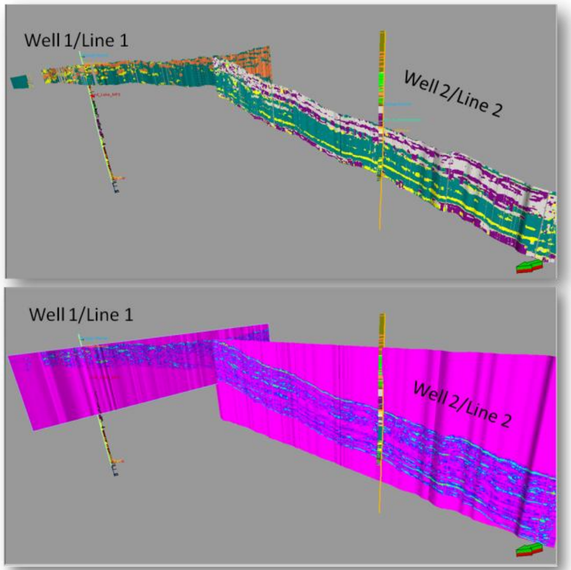


**Figure 2. Propagation of the upscaled log-derived classification model through the two 2D lines used for this study. The upper image shows the classification, the lower image shows the compliance to the model (see text, below). Only the thin wisps of light-bluish color exceed the threshold suggestive of a new class.**

When supervised classification of additional wells (or seismic data) is employed for recognizing the defined classes in additional areas, there are three possible outcomes. First, the unsupervised model is perfectly descriptive of the classes that are found in the prediction well. All the log responses in the prediction well are accounted for by classes defined in the model set. Second, there are more classes in the model set than are found in the prediction well. In this instance, the extra classes are simply ignored by the supervised classification and do not show up in the prediction well. Third, and most problematic, is when classes that should exist in the prediction well are not defined by the current model. To account for this possibility, a separate metric is used to evaluate the goodness-of-fit of each class assignment in the prediction well to the chosen class from the model set. If the fit is good, the class assignment is compliant (low error). If the fit is poor (high error), then there should be a new class assigned to those responses in the prediction well.

Any supervised classification is only capable of recognizing classes it is trained to recognize. Thus a 'basinal only' well, with its inherent basinal rock textural and compositional defined classifications, wouldn't necessarily provide training data to identify shelfal rock textures and compositional classes. The output anywhere in a prediction well can only be the closest class match from the training model set, but this does not mean that the class assigned is statistically "close" to the reality; it's just the closest fit among a bunch of bad fits. The metric we employ, a stratified Mahalanobis distance, measures a variance normalized Euclidean distance to the closest possible class (the one ultimately chosen) and if that distance exceeds a given threshold, it signifies a "none of the above" or "N+1" solution, regardless of the class presented. When this occurs, the new log responses can be iteratively added to the previous model, and then become recognized going forward. The same metric is used to assess the compliance of supervised classification with the seismic data to the input model as well. The result for this study is shown below the classification result in figure 2.

In the context of the above discussion on the vagaries of supervised classification, as applies to this study, if there are classes in Well 1 that do not exist in Well 2, then Well 2's seismic line cannot be trained to recognize that unique class in Well 1. Consequently, such a class cannot show up in Well 2's 2D line. However, where each well contains the same class, both 2D lines will also contain that same class, even if each line is independently classified to its own well. If both wells contain the same class, both lines have means to recognize it.

Therefore, one way to then "normalize" the lines, given that they have inconsistent attribute values from the original inversions, is to see if the similar classes between the two training wells converge at the intersection trace. Other classes, unique to each training well, do not converge because they do not have representation in the opposite line's training well. This can be seen in figure 2. At the intersection trace, the teal class exists in each line, because it exists in both training wells, but more interestingly it exists over roughly the same interval. This suggests that despite the differences in the inversions, the training regimes for each line allow for a compatible result at the intersection of the two lines – each line is at least internally consistent, if not offset in their relative magnitudes. Other classes, such as the grey class from Well 1/Line 1 and orange from Well 2/Line 2 propagate to, and meet at the intersection trace, but the alternate line for each has no training data to allow reconciliation of this discrepancy.

Since the teal class generally meets conformably at the trace intersection, despite the differences in the actual inverted attribute values, the implication in terms of material properties is the same, as the material properties come from the integration of the core data to the class. In this manner, the two lines are quasi-unified in terms of interpreting their attribute responses for material properties.

With recognition of the classes away from the wellbores along the 2D lines comes identification of the expected material properties from the core calibration of the log model. Log data are a more complete recorder of the changes in material properties of the formation. They contain more measurements and more degrees of freedom than seismic data, which tend to be highly correlated, and only consist of Vp, Vs, density, and/or combinations of them (e.g. Poisson's Ratio or Young's Modulus calculations). In the case of these two lines, the acoustic data are restricted to AI and SI.

As a result, the log data are integrated with the core data via the rock classification, and serve as the model set for propagation through the seismic. The five log tracks chosen are each uniquely sensitive to some variability in rock properties, regardless of possible correlations. We are not concerned exactly how because we are classifying with all five logs simultaneously and attributing core-derived properties to each class.

Figure 3 shows the relationship between the log-derived classes and the material properties measured from the core. On this star plot, a series of axes radiate from a central point. Each axis represents a single parameter. The nodes of each coloured polygon lie at the mean value of each parameter plotted on its respective axis. The polygon colour itself reflects the given class whose properties are being plotted. The uniqueness of each class, therefore, is reflected by the difference in each polygon's shape. The classification scheme shown is the upscaled version used for integration with the seismic.

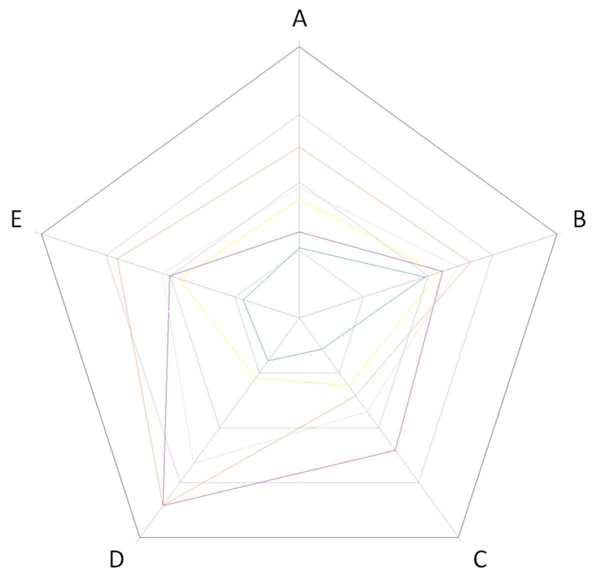


**Figure 3. Relationship between seismic upscaled classes and certain petrophysical properties from core. For legibility, the axes are given index labels. The labels translate to: A=hydrocarbon-filled porosity (0-12 p.u.), B=permeability (10 nD-1000 nD), C=$H_2O$ saturation (0-100%), D=TOC (0-8 wt.%), E=total porosity (0-12 p.u.). The lowest value in the plot range on each axis is the centre of the star plot. The colour of each polygon represents the class colour shown in the seismic propagation (figure 2).**

What becomes apparent in this display is that there are properties where several classes share similar values (e.g. TOC for the purple and orange classes). This is an unavoidable mathematical outcome. It is not possible to develop a classification unique to multivariate patterns of one input (logs, in this case), and assure that it is simultaneously unique in all possible associated properties (i.e. core data). One could isolate an individual property, and tailor a classification scheme strategically to maximize the uniqueness of that property per class (this is, in a sense, what forward modelling does), but in the context of bulk material properties, one cannot insure uniqueness in both the model and every possible material property. It is the collection that is unique, not any one property.

We see from the data, then, that the orange class contains the highest hydrocarbon-filled porosity (it is the outermost node on the HFP axis), at ~5 ± 1.8 p.u. This class also has the highest TOC, at ~3.4 ± 1.4 wt.%, and intermediate total water saturation (which includes loosely bound clay water), at ~35 ± 12 %. The purple class has similar TOC content to the orange class, but lower hydrocarbon-filled porosity and total porosity and higher water saturation, making it potentially less prospective, and illustrating the point that some parameters can be similar between classes, but the groupings different. The next best reservoir quality class, from these data, could be the gray class, which from the 2D seismic lines is the one that meets the orange class at the intersection trace, given the discrepancies in the inversions discussed earlier. The teal class has the lowest hydrocarbon-filled porosity, TOC and total porosity, but is similar in permeability to the other classes, possibly due to it also having the lowest water saturation.

If the orange class becomes the target class for production (which may or may not be ideal, given that we have no rock mechanics data, for example), then one can see from the seismic data that it thickens towards the south, behind the other 2D line from the visual perspective of figure 2, perhaps making this a more attractive region for future drilling. It could also extend at relatively uniform thickness parallel to the direction of the intersecting seismic line that shows the gray class. From figure 3, these two classes are more similar in properties to each other than the rest of the classes, and they share what seems to be a similar horizon. However, the line with the orange class has no training data for gray, and vice versa so we can only speculate that, at seismic resolution, and considering the discrepancies in the inversion, that they are roughly equivalent rock. Additional well control could rectify this mathematical paradox by providing training data to each line with both classes present. This would also clarify where these two classes may coexist, rather than forcing the interpretation to choose one or the other due to limited training data.

## CONCLUSIONS

We have applied a workflow similar to Suarez-Rivera et al. (2013) and Handwerger et al. (2014) to a pair of intersecting legacy 2D seismic lines from a frontier basin in Canada. The previous studies had been applied to modern 3D surveys with multiple modern logs available for the pre-stack inversions. This study differed in that we used legacy 2D lines inverted with a limited population of legacy logs. The inversions of the two lines were conducted separately, and the results showed discrepancies between the lines.

Each 2D line, however, did have one modern well with modern logs set and associated core. These data were collected after the inversions were done. A multivariate unsupervised classification model was built on a collection of modern logs collected in the field, both from the two wells on the two 2D lines in this study as well as elsewhere in the basin. This model was then upscaled to seismic resolution and used to train a supervised class model so that the same classes could be propagated across the 2D seismic lines.

Since each 2D line was trained to its respective log separately, but the logs themselves are unified in their class model, we were able to reconcile discrepancies in the inverted parameters by identifying the common classes on each line. At the common trace tied at the intersection of the 2D lines, a common class isolated in the log data meet over a common range. This suggests that despite the quantitative difference in attribute value at the intersection trace, the training to the classification was qualitatively accurate, and propagation of the class was consistent for the set of attributes used to carry this out.

Results suggest that the pattern of logs that define the orange class, speculatively identified as having the highest reservoir quality (based on available ground-truth data) in the region thickens south of the well control.

Each class from the log-based model was associated to a collection of core measured properties. Since we have established that the class assignments are appropriate for each line, despite the quantitative differences between them, we were able to populate the lines with consistent properties through the core-log-seismic integration. Recognition of the same classes calibrated with core properties away from the wellbore (albeit upscaled) along each 2D line allowed for the estimation of those same properties and improved characterization of the resource via the seismic data.

## ACKNOWLEDGMENTS

We would like to thank Northern Cross Yukon Ltd.'s Exploration Team for their insight, willingness to undertake this pilot project with us, and assistance, and for permission to publish these results. We would also like to thank Schlumberger Reservoir Laboratories in both Salt Lake City, UT USA and Calgary, AB Canada for their diligent work in making the core measurements and creating the log-scale HRA model. Lastly, we would like to thank Schlumberger Stavanger Research for their experimental codes to carry out the seismic classification, and their technical expertise and insight.

## REFERENCES

Suarez-Rivera, R., Dahl, G.V., Borgos, H.G., Paddock, D.R., and Handwerger, D., 2013, Seismic-Based Heterogeneous Earth Model Improves Mapping Reservoir Quality and Completion Quality in Tight Shales: SPE Unconventional Resources Conference – USA, The Woodlands, TX USA, SPE-164544.

Handwerger, D., Martinez, C., Castaneda-Aguilar, R., Dahl, G.V., Borgos, H., Ekart, D., Raggio, F., Lanusse, I., Benedetto, M., and Suarez-Rivera, R., 2014, Improved Characterization of Vaca Muerta Formation Reservoir Quality in the Neuquén Basin, Argentina by Integrating, Core, Log and Seismic Data Via a Combination of Unsupervised and Supervised Multivariate Classification: 9$^{th}$ Congress on Exploration and Development of Hydrocarbons, Mendoza, Argentina, *in press*.