Geophysical Corner

Spectral Balancing of Seismic Data and Unsupervised Facies Analysis

A common problem with seismic data is their relatively low bandwidth. Significant efforts are made during processing to enhance the frequency content of the data as much as possible to provide a spectral response consistent with the acquisition parameters. Traditional seismic data usually preserve information with lower frequencies between 5 and 10 hertz and upper frequencies between 60 or 70 hertz. While such bandwidths are acceptable for thicker conventional reservoirs, they lack the needed resolution to map thinner reservoirs or thinner architectural elements within a thicker reservoir. Fortunately, the advancements in seismic data acquisition and processing, coupled with increased computer memory and speeds provide cost effective solutions for such objectives.

There are several methods that can be utilized to enhance the frequency content of the input seismic data. Here we discuss the application of a spectral balancing method discussed in detail in the May 2014 installment of Geophysical Corner. The workflow consists of suppressing crosscutting noise using a structure-oriented filtering algorithm, leaving mostly signal in the data. Next, the data are decomposed into time-frequency spectral components, followed by the computation of a smoothed average spectrum. If the survey has sufficient geologic variability within the smoothing window (that is, no perfect "railroad tracks"), this spectrum will represent the time-varying source wavelet. This single average spectrum is used to design a single time-varying spectral scaling factor that is applied to each and every trace. Geologic tuning features and amplitudes are thus preserved.

Demonstration

We demonstrate the application of this method on a 3-D seismic data volume from Smeaheia area in offshore Norway. Smeaheia has been considered as one of the potential areas to evaluate CO_2 storage.

The Smeahiea area lies about 30 kilometers east of the Troll gas field (figure 1), within the Norwegian continental shelf. The Smeahiea target is located in a fault block bounded by the Vette Fault to the west and the Øygarden Fault to the east and is raised about 300 meters relative to the Troll field. The Late Jurassic Sognefjord, Fensfjord and Krossfjord formations form the producing reservoir zones in the Troll gas field.

In the Smeaheia block, there are two four-way closure structures, the Alpha structure to the west and the Beta structure to the east. Two exploration wells, namely 32/4-1 and 32/2-1 have been drilled into these structures, and although the reservoir is good, both wells turned out to be wet, indicating that the Smeaheia area is not charged with hydrocarbons.

In the Smeaheia area, the Sognefjord Formation is the primary reservoir consisting of medium to coarse-grain, well-sorted, micaceous and minor argillaceous sandstone. Below this formation lies the Fensfjord Formation consisting of medium-grained, well-sorted sandstone with shale intercalations. Underlying the Fensfjord Formation is the Krossfjord Formation comprised of medium to coarse-grained, well-sorted sandstone.

As the Sognefjord, Fensfjord and Krossfjord are sandstone reservoir formations, there is concern that detailed mapping of their

Continued on next page ▶

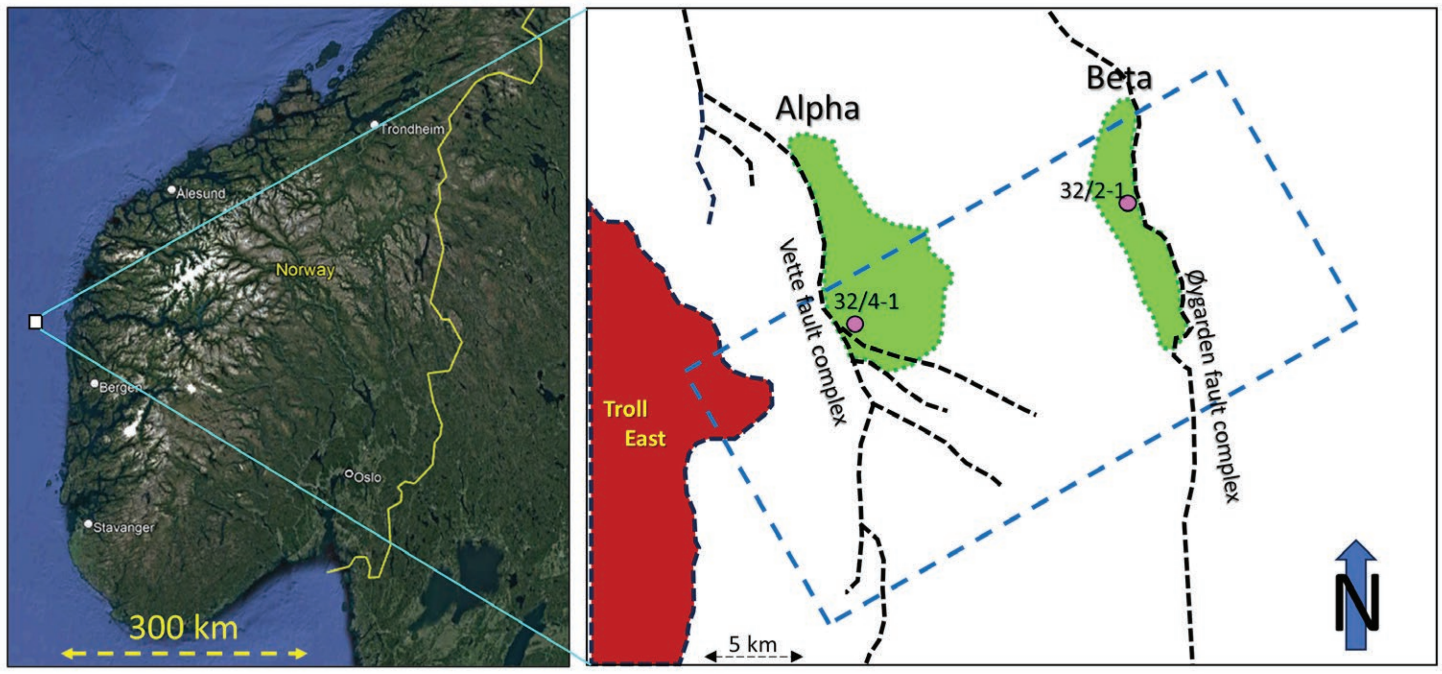


Figure 1: The Smeaheia area on the Norwegian continental shelf showing the Alpha and Beta structures. The image to the left was prepared with the use of Google Earth Pro. Modified after Furre et al., 2017.

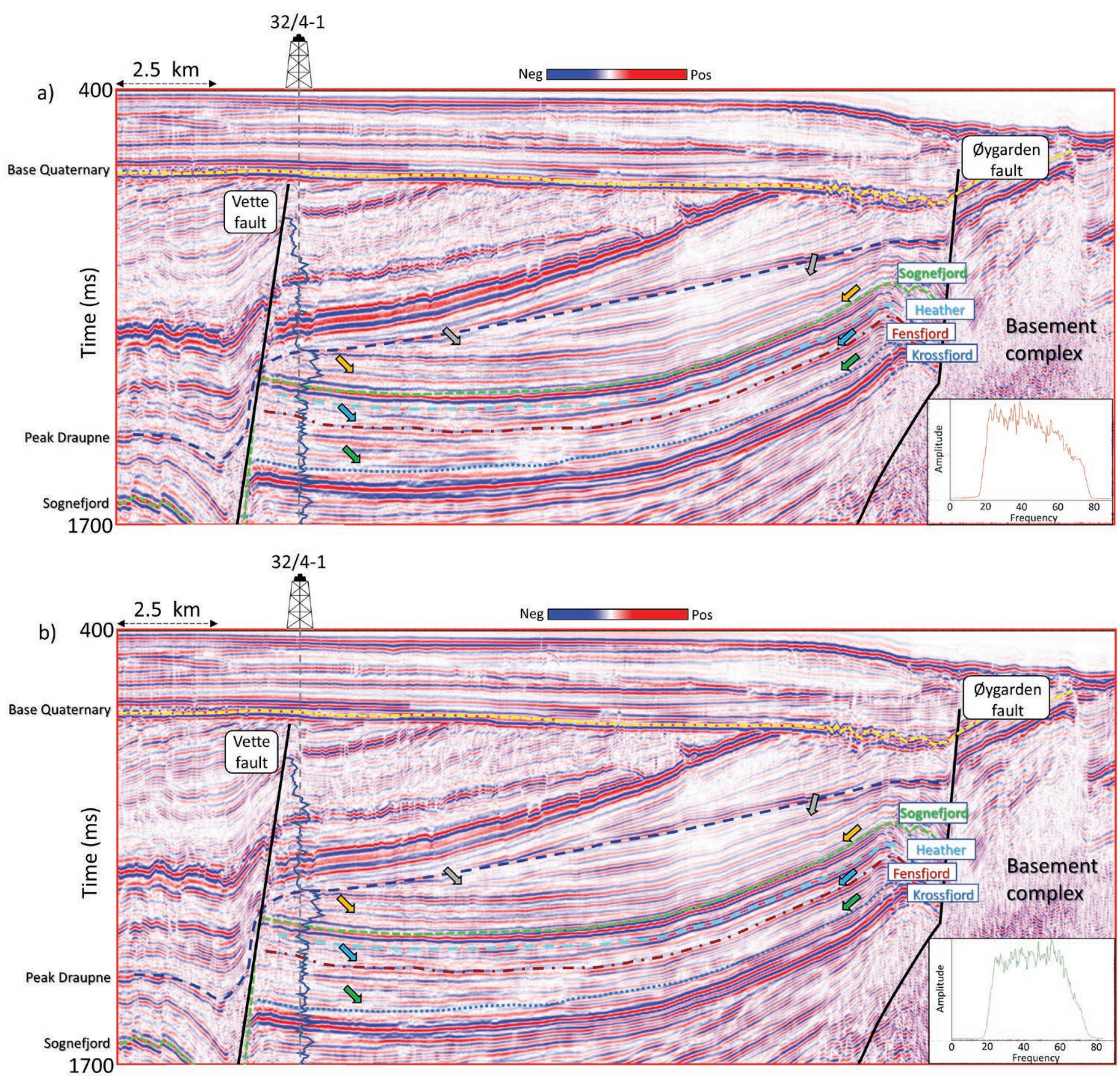


Figure 2: Segment of an inline extracted from (a) input seismic data volume, and (b) the same line after spectral balancing. Some relevant markers as well as the gamma ray curve for well 32/4-1 are overlaid on the two sections. The frequency spectral for the two data volumes in the indicated zones are also shown to the right. Notice the enhancement in the resolution of the reflections after spectral balancing, especially as indicated by colored block arrows in the broad zone of interest.

◀ Continued from previous page

properties is hampered by the limited seismic resolution, where embedded shale and carbonate stringers might introduce flow barriers. In addition, the existence of faults and fractures that fall below seismic resolution could provide pathways for CO_2 leakage. All these risks need to be evaluated and mitigated in the context of long-term CO_2 storage.

The seismic survey covering the Smeaheia (blue dashed rectangle shown in figure 1) was acquired by Gassnova in 2011 and made publicly available by Gassnova and Equinor. The bin size for the data is 12.5 x 25 meters, with a sample interval of 4 milliseconds. The interpreted horizons, well log data for the two 32/4-1 and 32/2-1 wells and well completion reports were also provided. These well logs included complete gamma ray curves, but with sonic and density logs that were not recorded for the shallower depths. The seismic data volume is of good quality.

Figure 2a and b shows a comparison of segments of seismic sections traversing the 3-D seismic volume before and after spectral balancing, along with their frequency spectra. Notice the well-defined appearance of the reflections in figure 2b as well as the flattened appearance of the frequency spectra as compared with the input seismic volume.

This data volume was then put through structure-oriented filtering and attribute computation. To bring out the advantage of spectral balancing and structure-oriented filtering, we compute relevant attributes on the data before and after the two data conditioning processes and compare the results in the next section.

Generation of a Suite of Attributes

With the improved vertical resolution seen in figure 2b, our next task is to determine if the spectrally balanced seismic data also enhances the lateral resolution as measured by seismic attributes. The following attributes were computed on spectrally balanced seismic data.

- Relative acoustic impedance is computed by continuous integration of the original seismic trace with the subsequent application of a low-cut filter. Because it assumes a zero-phase wavelet that is as close to a spike as possible, the improved resolution of spectral balancing will provide improved results over the original data. The impedance transformation of seismic amplitudes enables the transition from reflection interface to interval properties of the data, without the requirement of a low-frequency model. Figure 3 shows a comparison of stratal slices 32 milliseconds above the Sognefjord marker from the relative acoustic impedance attributes computed from input seismic data and input seismic data after spectral balancing. Notice the crisp definition of the faults as indicated by the highlighted areas in dashed purple outlines.

Likewise, the other attributes computed on the two seismic volumes are listed below along with their brief descriptions.

- Instantaneous amplitude is a measure of the reflection strength of the analytic seismic trace, independent of phase, and provides information on intensity of reflections. Similarly, instantaneous frequency provides information on attenuation and layer thickness. We use a smoother, more stable version of the instantaneous frequency usually obtained by weighting it by the envelope.
- Sweetness is a "meta-attribute" or one computed from others, which in this case is the ratio of the envelope to the square root of the instantaneous frequency. A clean sand embedded in a shale will exhibit high envelope and lower instantaneous frequency, and thus higher sweetness, than the surrounding shale-on-shale reflections.
- GLCM or grey-level co-occurrence matrix energy is a measure of textural uniformity in the data. If the reflectivity along a horizon is nearly constant, it will exhibit high GLCM energy.

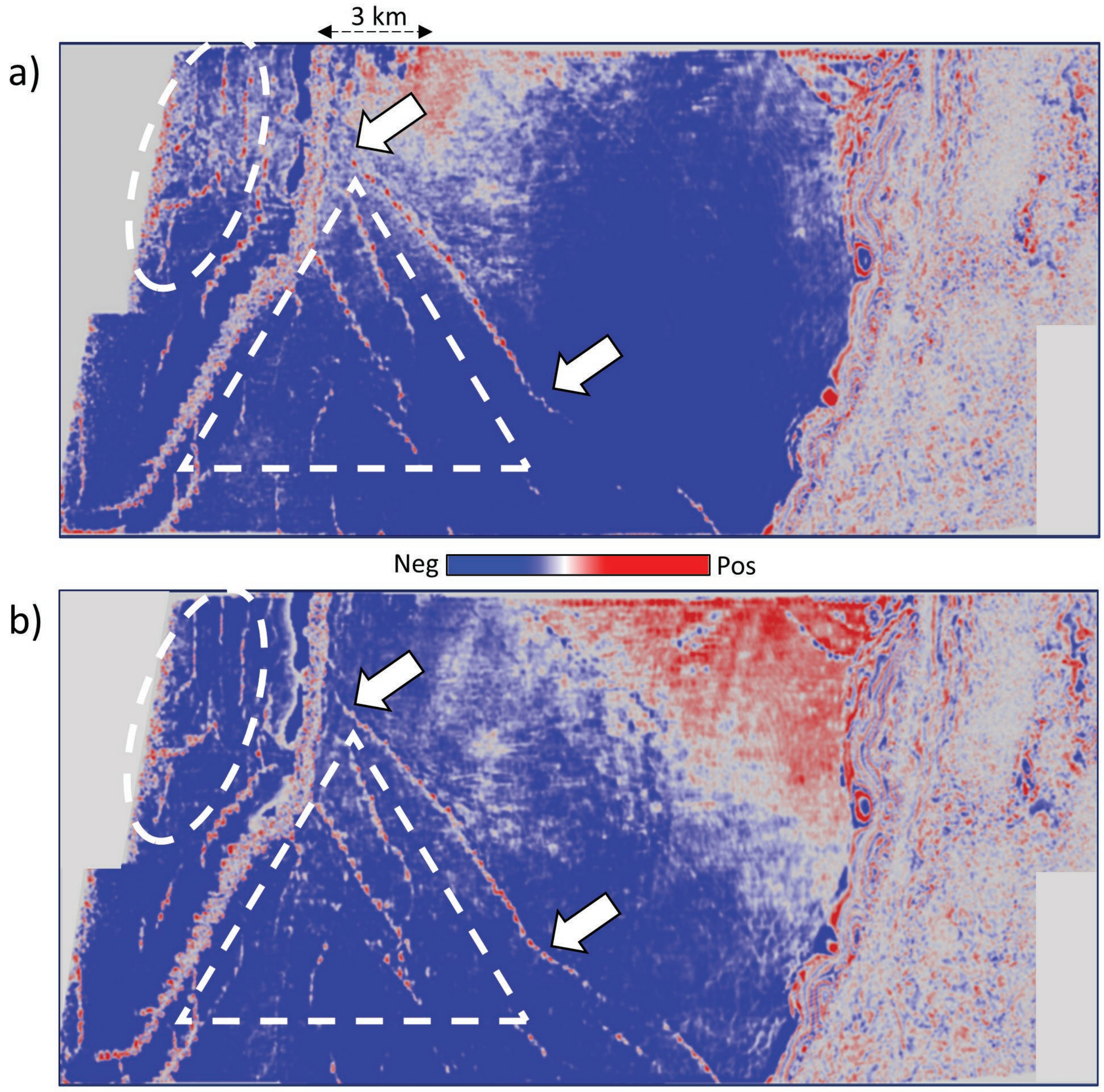


Figure 3: Stratal slice 32 milliseconds above the Sognefjord marker through the relative acoustic impedance attributes computed from (a) the original seismic data, and (b) the seismic data after spectral balancing. Notice the crisp definition of the faults as indicated by the white block arrows as well as the two highlighted areas in dashed white outlines and block arrows.

- Spectral magnitude: The magnitude of spectral components ranging from 20 to 70 hertz, which is the effective bandwidth of the input seismic data.

Specifically, the attributes used for the computation of seismic facies classification using some of the unsupervised machine learning methods were the relative acoustic impedance, envelope, sweetness, GLCM energy and spectral magnitudes at 25, 40 and 55 hertz.

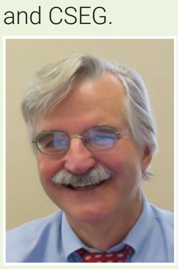
Two sets of these different attributes were generated – one from the original data and one from the data preconditioned using spectral balancing and structure-oriented filtering. Each dataset then served as input to unsupervised seismic facies classification using machine learning techniques described in the next section.

Unsupervised Seismic Facies Classification Using ML Techniques

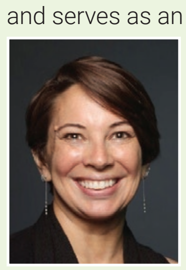
Unsupervised learning provides a means to determine if the seismic response can be related to flow units or rock types that can be calibrated with additional well control, but for which we do not understand the underlying petrophysical or geological theoretical support. Still, seismic interpreters face a perpetual challenge of extracting heterogeneous seismic facies on different generated attributes. Common analysis tools include corendering, crossplotting and visualization, which can help to an extent in terms of simultaneous display of the input attributes. Alternatively, projection techniques like principal component analysis represent the most common features seen in multiple attribute volumes by a more manageable smaller subset of linearly combined attributes. This smaller subset is then displayed by plotting three linear combinations against red, green and blue. In all these workflows, the human interpreter examines an image in color or in a crossplot, identifies a feature of interest and then defines its geologic meaning. Subconsciously, the human interpreter is defining clusters that have a given color or location in the crossplot. Unsupervised machine learning does the same thing but can use more than three independent attributes common to visualization and crosspplotting workflows. Whether interactive or through machine learning, clustering is only part of seismic interpretation. Equally important is the spatial relation between clusters, e.g., the change in cluster from channel fill to flood plain, or changes in clusters across a fault. In this article, we emphasize such lateral changes by corendering the images with coherence. In this manner, lateral changes in many of

See **GeoCorner** page 20 ▶

Ritesh Kumar Sharma is with Calgary-based SamiGeo. He has vast experience of working with 2-D/3-D, land, and marine seismic data with different applications such as AVO analysis, rock-physics analysis, frequency enhancement of seismic data, simultaneous inversion, extended elastic impedance inversion as well as geostatistical inversion. He is an active member of SEG and CSEG.



Kurt Marfurt has divided his career nearly equally between industry and academia, and is currently an emeritus professor of geophysics at the University of Oklahoma. For the past 20 years he has focused on seismic attributes and machine learning to aid the seismic interpreter. He has served as editor-in-chief for the SEG/AAPG journal Interpretation, has delivered two SEG distinguished instructor short courses, is the 2021-22 AAPG/SEG distinguished lecturer, and serves as an SEG director-at-large for 2020-22.



Heather Bedle received a bachelor's in physics from Wake Forest University and then worked as a system engineer in the defense industry for several years. She later received a master's and a doctorate in geophysics from Northwestern University. After nine years of working with Chevron, she moved to academia to pursue her dream of teaching and mentoring young researchers. Starting at the University of Houston in 2016, she then moved to the University of Oklahoma in 2018. Currently, she is PI of the AASPI Consortium, where her research group of 20 total students has published almost 20 peer-reviewed manuscripts in the last two years. Her primary research interests include seismic interpretation, machine learning and attribute analysis.

Alexandro Vera-Arroyo is currently a research assistant for the Subsurface Detective Agency and Attribute-Assisted Seismic Processing and Interpretation consortium at the University of Oklahoma. His research interests include wave propagation phenomena, optimization problems, quantitative reservoir characterization, machine learning and exploration risk management. He studied exploration geophysics at the National Autonomous University in Mexico City and then earned his master's in geophysics at OU and is now pursuing his doctorate there to better understand deep learning applications in different subsurface exploration challenges. He has worked as a geomechanical engineer at Schlumberger and a quantitative interpreter at PEMEX.

GeoCorner

from page 19

the clusters can be easily associated with an established geologic model. In contrast, lateral changes in clusters that are not delineated by coherence indicated more gradational changes in lithology or thickness that are easier to overlook.

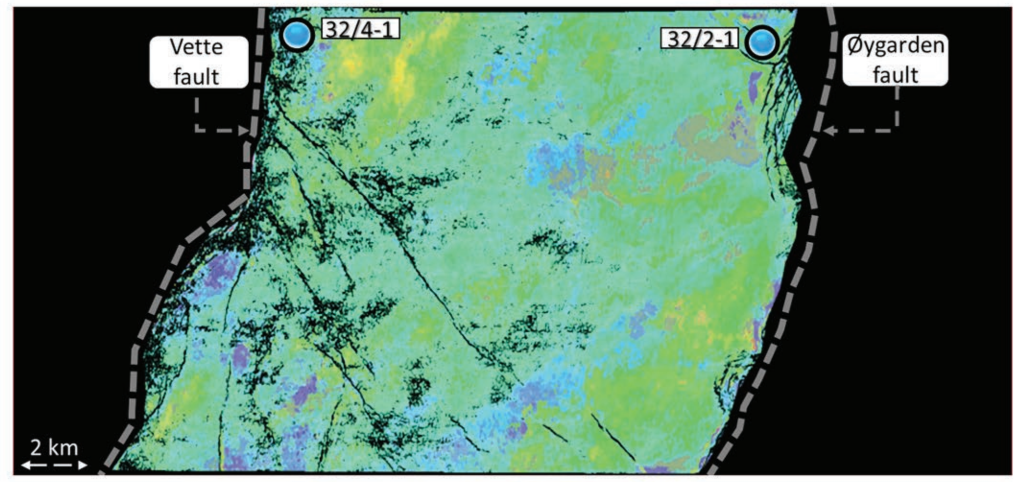
Here we compare the application of two established ML techniques, namely self-organizing mapping and generative topographic mapping. We find such an application promising as the facies results exhibit higher vertical and lateral resolution. More details about the methods and their applications can be picked up from the articles published in the November 2020 and January 2022 installments of Geophysical Corner.

Self-Organizing Maps

SOM is an unsupervised machine learning technique based on the clustering approach that generates a seismic facies map from multiple seismic attributes. In this technique the initial cluster centroids are defined in an N-dimensional attribute data space by fitting a plane defined by the first two eigen vectors of the covariance matrix to the data in a least-squares sense. With centroid still locked to this plane, it is iteratively deformed into a 2-D surface that fits the data still better. Once convergence is reached, the N-dimensional data are projected onto this 2-D surface. Thus, SOM may be considered as projection from a multidimensional attribute space to a 2-D space or "latent" (hidden) space. Usually, the output from SOM computation is obtained in the form of two projections on the two SOM axes, which can then be directly crossplotted and displayed using a 2-D RGB color bar.

Figure 4 shows the equivalent stratal displays (within the Fensfjord formation) extracted from the SOM crossplot volume computed for the input and spectrally-balanced versions of the seismic data, using a 2-D color bar. Some of the clusters seen on the display in figure 4b are better defined than the ones shown in figure 4a.

a)



b)

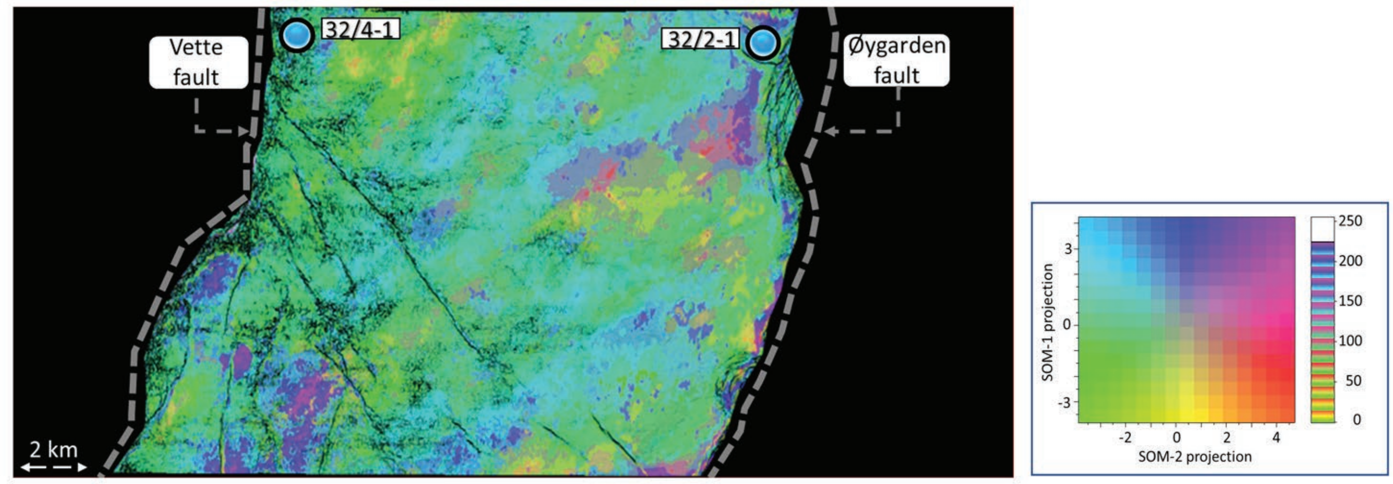


Figure 4: Stratal slice within Fensfjord Fm extracted from the SOM crossplot volume computed on attributes generated on (a) input seismic data volume, and (b) spectrally balanced input seismic data volume. The two volumes have been corendered with the respective multispectral energy ratio coherence attribute volumes. Better spatial resolution of the seismic facies is seen in (b) than in (a). Only the target area between the Vette and Øygarden was classified.

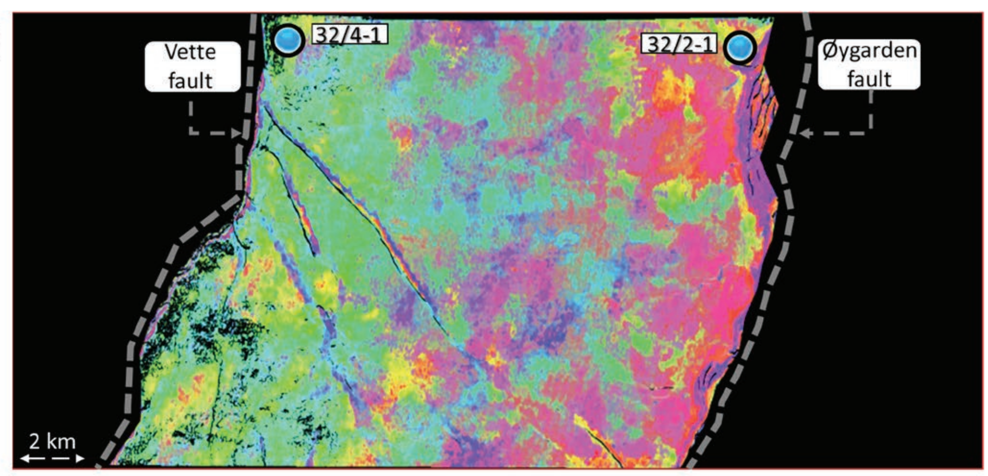
Generative Topographic Mapping

Though the Kohonen SOM method described above is easy to implement, is computationally inexpensive, and thus is a popular unsupervised clustering approach, it does have limitations. First, there is no theoretical basis for the selection of parameters such as training radius, neighborhood function and learning radius, as all of these are data dependent. Secondly, no cost function is defined in the method that could be iteratively minimized indicating convergence during the training process. Finally, as a measure of confidence in the final clustering results, no probability density is defined. An alternative approach to the Kohonen SOM method, called "generative topographic mapping," overcomes the above-stated limitations. It is a nonlinear dimension reduction technique that provides a probabilistic representation of the data vectors in latent space.

Thus, as the above descriptions suggest, the SOM and GTM methods project data from a higher dimensional space (7-D when seven attributes are used as input) to a lower dimensional space, which may be a 2-D plane or a 2-D deformed surface. Once projected on to these planes, the data can be clustered in that space, corendered with RGB or crossplotted using a 2-D color bar.

In figure 5 we show the equivalent displays (within the Krossfjord formation), where some of the clusters can be interpreted with ease with less background clutter and confusion The individual-colored patches or facies are crisper and could lead to more accurate interpretations.

a)



b)

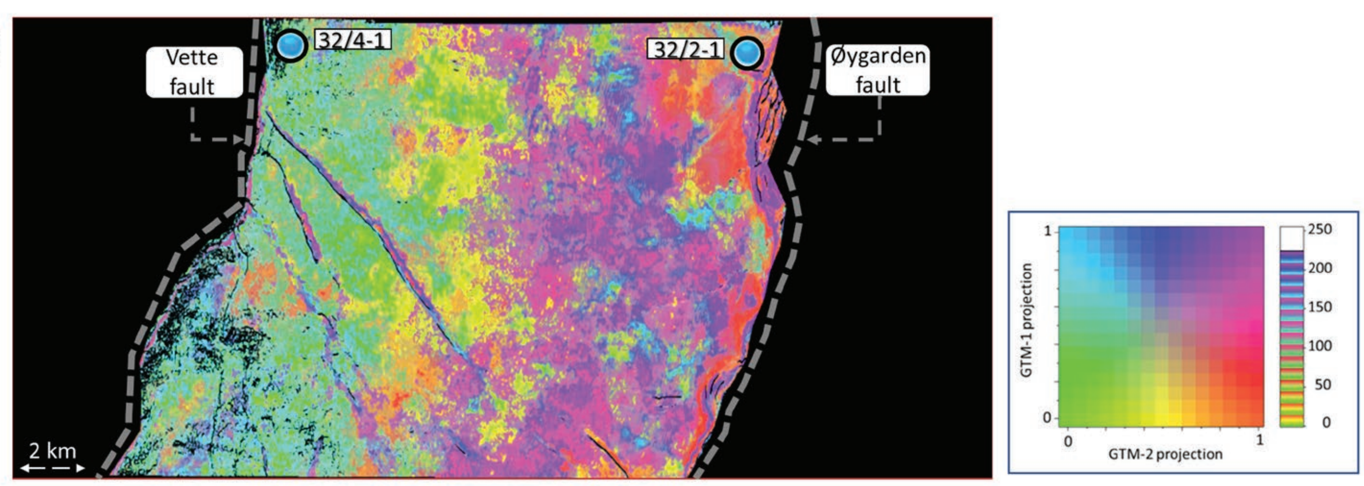


Figure 5: Stratal slice within Krossfjord Fm extracted from the GTM crossplot volume computed on attributes generated on (a) input seismic data volume, and (b) spectrally balanced input seismic data volume. The two volumes have been corendered with the respective multispectral energy ratio coherence attribute volumes. Better spatial resolution of the seismic facies is seen in (b) than in (a). As with SOM, only the target area between the Vette and Øygarden was classified.

Conclusions

We have found that spectral balancing of the input seismic data when used for attribute generation and further used in some of the multiattribute processes discussed here can significantly aid in accurate interpretation. Results obtained for the unsupervised machine learning applications employing both the input seismic as well as its spectrally-balanced version depict superior performance of the latter in terms of clarity of clusters as well as color variations within them, probably in response to the expected geologic variations as mentioned in the introduction.

Applications of SOM and GTM techniques to the same data allowed us to assess their relative strengths as well as their suitability. We found that both GTM and SOM show more promising results, with GTM having an edge over SOM in terms of the detailed distribution of seismic facies in terms of better resolution and distinct definition of the geologic features seen on the displays.

Usually, the seismic facies maps in the zones of interest are calibrated with the lithofacies information obtained from well cores and cuttings. As there is appreciable difference in resolution between the two types of data, it is advisable to enhance the resolution of seismic data by adopting a spectral balancing workflow. Such a workflow can narrow down the resolution gap between the facies data types (seismic and geologic) as well as help perform a better correlation/calibration between the two. This is thus the motivation for the work described in this paper. Though the analysis is qualitative at present, it paves the way for more detailed work as more well and other data become available.

Acknowledgements

The first author would like to thank the Attribute-Assisted Seismic Processing and Interpretation Consortium, University of Oklahoma, for access to their software, which has been used for all attribute computation as well as Subsurface AI (formerly Geomodeling Technology Corp.) Calgary, for making the Attribute-Studio software available. We wish to thank Gassnova and Equinor for access to the Smeaheia 3-D seismic and other associated data used in this exercise.

(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.)