A Simple Statistical approach to guide fault interpretation

Kamal Hami-Eddine,Bruno de Ribet , Paradigm*

SUMMARY

In recent years, the use of statistics, machine learning and artificial intelligence has attracted a lot of interest. However most of the recently developed techniques are not readily applicable to geophysics, while some of the basic statistical methods, which can add significant value to the process of seismic interpretation, remain underutilized. The advent of new seismic structural attributes and technology to improve their visualization has brought about great improvements in the quality of supporting information available to the seismic interpreter. Calculating many attributes introduces a concomitant problem of how to use them efficiently. The use of multivariate analysis methods can provide a good compromise to get optimal information out of several attributes without overwhelming the interpreter with a surfeit of subtly different sources of information. We present an example case study showing the successful application of these techniques to fault-enhancing attributes in a 3D seismic dataset from New Zealand.

INTRODUCTION

Seismic attributes for fault delineation have seen major improvements in the last 20 years. The increase in quality of seismic imaging now makes it possible to get discontinuity attributes which contain clear indications of faults. However, different seismic attributes naturally exhibit different characteristics, resulting in different estimations of lateral and vertical fault extent and connectivity. Attributes are also impacted by noise contained in data, and by discontinuities other than faults. Automation of fault picking is therefore not trivial. The careful use of multiple attributes should provide better estimation of fault locations, and increase the accuracy of their interpretation. As with all multi-attribute analysis, it becomes tedious to analyze more than 3 in a manual way, and automated processes do not provide satisfying solutions to consider several attributes at once.. This paper shows how a simple, fast track approach utilizing Principle Component Analysis and seismic facies clustering can provide significant improvements to the use of conventional attributes such as semblance, coherency and curvatures.

STATISTICS AND CLUSTERING FOR FAULT DELINEATION

Seismic attributes such as dip, azimuth, coherency, curvatures and others provide highly valuable information on fault identification, extension and connectivity although they all generate different images, and contain different value ranges. Multivariate analysis is a way to consider all these attributes simultaneously to isolate fault character from areas of noise and other discontinuities. Principal Component Analysis (PCA) (Figure Saporta (2011)) is one of the simplest multivariate methodologies. A PCA projection represents a dataset in terms of the orthonormal eigen-vectors of the datasets co-variance matrix. PCA finds the orthonormal eigenvectors of the covariance matrix as the basis for the transformed feature space. Higher eigenvalues in the covariance matrix indicate lower correlation between the features in the data set. PCA projections seek uncorrelated variables. For a set of variables $(\mathbf{X}_1,...,\mathbf{X}_n)$, such as the features of a data set, we can create a matrix which represents the co-variance between each pair of variables X_i and $\mathbf{X}_j$ where i and j are indices of the feature vector.

For PCA, we subtract the means $\bar{\mathbf{X}}_i$‘ from each $\mathbf{X}_i$ before constructing the co-variance matrix so that each $\bar{\mathbf{X}}_i$ has a mean of zero. Subtracting the means allows us to rewrite the covariance matrix as the following matrix multiplication:

$$\mathbf{\Sigma} = \frac{1}{n}\mathbf{X}\mathbf{X}^T$$

Then, by the spectral decomposition theory, we can factor the matrix above into:

$$\mathbf{\Sigma} = U \Lambda U^T$$

where $\mathbf{\Lambda} = diag(\lambda_1, \cdots, \lambda_n)$ is the diagonal matrix of the eigenvalues of the co-variance matrix ordered from highest to lowest. The principal components are the row vectors of $\mathbf{U}^T$. We call $\mathbf{U}^T$ the projection weight matrix $\mathbf{W}$ and the transformed data matrix $\mathbf{S}$ can be obtained from the original data matrix $\mathbf{X}$ by:

$$\mathbf{S} = \mathbf{W}\mathbf{X}$$

Note that if we choose not to use eigen-vectors that correspond to lower eigen-values so that $\mathbf{W}$ has fewer rows, then each S will have lower dimensionality than its corresponding X. Discarding these eigen-vectors can be thought of as discarding noise from the data, since these eigen-vectors represent highly correlated, and thus uninformative variables.

The total variation in a PCA transformation of a data set is the sum of the eigen-values of the co-variance matrix. Since these eigen-values are contained in $\mathbf{\Lambda}$,

$$\sum_{i=1}^{n} Var(PC_i) = \sum_{i=1}^{n} \lambda_i = trace(\mathbf{\Lambda})$$

So the fraction,

$$\sum_{i=1}^{k} \frac{\lambda_i}{trace(\mathbf{\Lambda})}$$

gives the cumulative proportion of the variance explained by the first k principle components. PCA therefore transforms the initial set of attributes into a collection of uncorrelated one, consisting in a linear combination of the original ones. PCA is then ordering the amount of information contained in each of its component. Figure 1, shows the comparison of seismic attributes and their first PCA combination. PCA provides a better image of the fault as combining somehow the best of the input attributes.

Using simple clustering techniques, on top of PCA result (Figure 1), to qualify fault zones helps to simply the image. The distance between the different clusters, can be used to analyze the certainty of fault character. It may also provide a way to quantify fault connectivity likelihood in noisy areas, if combined with probabilistic models.

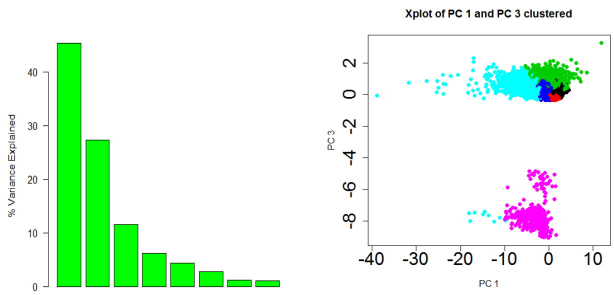

Figure 1: The 3 first PCA component explain 75% of explained variance. The clustering of these 3 components provides a good data segmentation as shown on the cross-plot between principal component 1 and 3.

EXAMPLE ON A NEW ZEALAND OFFSHORE FIELD

The Opunake prospect is located in Taranaki basin, New Zealand. The evolution of the Taranaki basin started with Late Cretaceous extensional faulting associated with the break-up of Gondwana and the formation of the Tasman Sea (Kamp, 2004). The Neogene saw a period of compressive tectonic events at a regional level, resulting in complexly faulted structures. The seismic imaging of these faults is of good quality, but fault positioning and contacts are difficult to assess based on conventional seismic attributes such as coherency, as shown in figure 2.

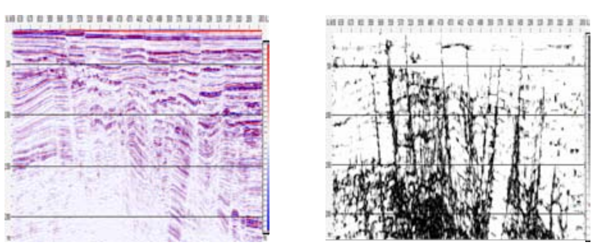

Figure 2: Left image shows a seismic amplitude line in a faulted zone, the same inline is shown on the right with coherency attribute displayed. The complexity of faulting, and the obvious impacts of data noise and complex stratal geometries makes the precise location of faults difficult.

PCA utilization has previously been demonstrated to highlight geological features, including by highlighting the seismic manifestation of potential faults on horizon-based extracted data (Priezzhev and Scollard, 2012). The same type of approach was used to delineate channel edges using curvature attributes and spectral decomposition attributes (Chopra et al., 2014). However the 3D volumetric delineation and interpretation of faults using PCA attributes has not yet become a common practice despite its ease of implementation,and simplistic concept.

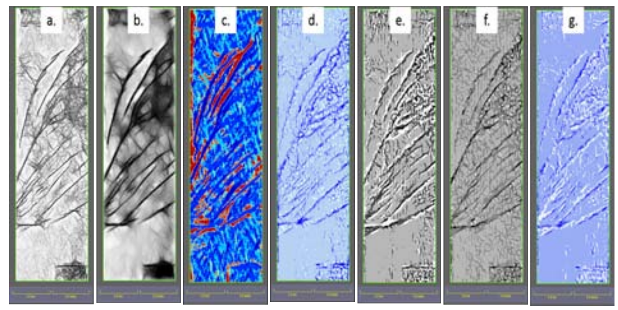

Figure 3: 7 attributes have been used to calculate PCA components. Coherency, Fault enhanced, azimuth, and a set of curvature attributes have been used in this example.

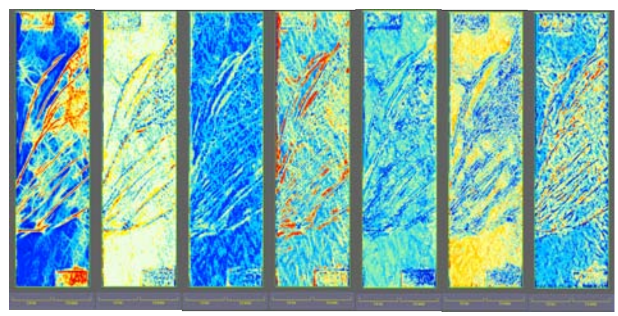

Figure 4: The 7 PCA components highlight different type of information. PCA 1 (left) seem to propose the best compromise to isolate faults, from noise and lineaments

On this specific data set, the objective is to combine the information from different attributes and to try to isolate fault characteristics from other discontinuities and noise. Considering attributes one by one does not provide sufficient information, as some tend to accentuate noise while improving fault connectivity images, and others do the opposite. Merging attributes through color blending such as RGB, HSV or CMR does not add significant value in such case, as even if it may help for visualization, it does not provide a support for picking. Running a principal component analysis on structural attributes (Figure 3) is providing an interesting compromise to help seismic interpreters in their structural interpretation process. The new attributes, result of a PCA, show several advantages compared to each single initial attribute. Firstly they appear, as expected, to remove regular noise related to acquisition trend. Then the first PCA component, corresponding to the highest eigenvalue, seems to be the best projection of input attributes to isolate fault features. First they appear as expected to remove regular noise related to acquisition trend. Then the first PCA component, corresponding to the highest eigenvalue, seems to be the best compromise between input attributes to isolate fault features from the rest (Figure 4). It shows good continuity of fault segments, as well as a higher-definition image that more clearly illustrates fault connectivity (Figure 5). Finally, PCA is exhibits a broader, and more discriminatory dynamic spectrum enabling an easier customization of color mapping to better visualize faults.

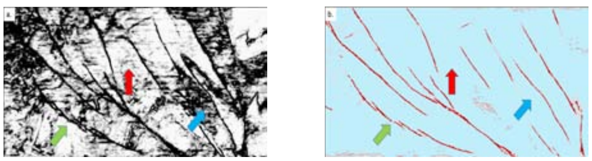

Figure 5: a. Fault enhanced attribute on a time slice, b. Fault cluster dissimilarity extracted on the same time slice. Continuity and details are better highlighted in b. Step faults indicated by the green arrow are more evident to interpret as such on b.

A simple clustering on top of carefully selected PCA attributes helps to isolate even more the fault features as shown in figure 6 and figure 7. This views shows how the contrast between faults and noise is not only impacted by color palette description, but by data transformation through linear combination of attributes.

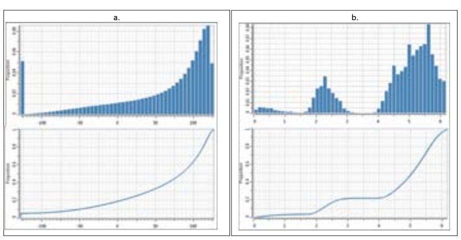

Figure 6: The comparison between coherency histograms and fault cluster dissimilarity one shows a multi-modal distribution, isolating fault features.

Considering the principal components we can observe that the spread of points on a cross-plot matrix proposes as well an obvious segmentation on some components.

The clustering of the PCA components provides a good picking support for semi-automatic fault tracker as shown in figure 8.

CONCLUSION

Principal Component Analysis provides an alternative methodology for fault identification. Using the new attributes derived from the PCA process provides higher definition images for fault visualization. This approach significantly reduces the impact of noise on the seismic image, and limits the time spent on fine-tuning attribute parameters. Because the input attributes are all related to structure related information, the application

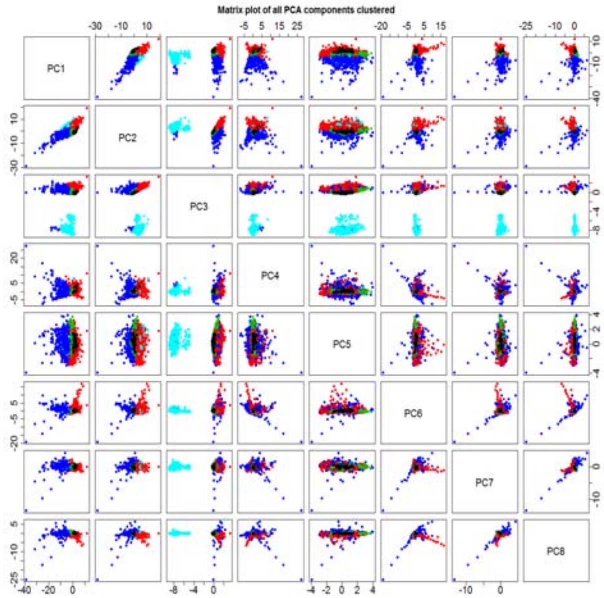

Figure 7: principal Components scatter plot showing clusters isolationg fault features in Cyan. All components are not needed to isolate faults.

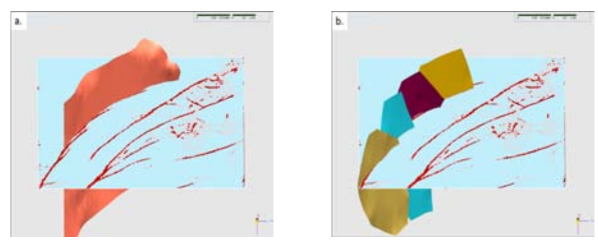

Figure 8: a. The propagation of a fault surface from coherency attribute extracts one fault in that zone. b. The propagation of fault in the same zone using PCA distance to clusters, is providing a quick way to extract step faults accurately.

of the PCA technique emphasizes the characteristics of the fault pattern. The generated PCA attributes could be used as input data for creating fault clusters to highlight the 3D complexity of fault network in a very efficient way. Moving forward, we can easily foresee how improved machine learning techniques such as deep learning, will benefit from such data segmentation, to enable accelerated fault interpretation using convolutional networks.

ACKNOWLEDGEMENT

We would like to thank Paradigm for authorizing us to publish this abstract. Data courtesy of Ministry of Economic Development Unpublished Petroleum Report, New Zealand.

EDITED REFERENCES

Note: This reference list is a copyedited version of the reference list submitted by the author. Reference lists for the 2017 SEG Technical Program Expanded Abstracts have been copyedited so that references provided with the online metadata for each paper will achieve a high degree of linking to cited sources that appear on the Web.

REFERENCES

Chopra, S., and K. J. Marfurt, 2014, Churning seismic attributes with principal component analysis: 84th Annual International Meeting, SEG, Expanded Abstracts, http://dx.doi.org/10.1190/segam2014-0235.1.

Kamp, P., 2004, Neogene stratigraphy architecture and tectonic evolution of Wanganui, King Country, and eastern Taranaki Basins, New Zealand: New Zealand Journal of Geology & Geophysics, **47**, 625–644, http://dx.doi.org/10.1080/00288306.2004.9515080.

Priezzhev, I., and A. Scollard, 2012, Faults and fracture detection based on seismic surface orthogonal decomposition: 74th Annual International Conference and Exhibition, EAGE, Extended Abstracts, http://dx.doi.org/10.3997/2214-4609.20148740.

Saporta, G., 2011, Probabilites, analyse des donnees et statistique: Technip.