Seismic fault proximity to production

Jesse Lomask, Toby Burrough, and Allison Gilmore, S&P Global; Michael Pyrcz, The University of Texas at Austin*

Summary

We present a semi-automated workflow to statistically evaluate relationships between faults and production. First, we estimate 3D faults and fault orientations using a 3D-CNN. Next, we apply fast marching for a shortest path calculation to faults. This creates a 3D volume of distances with each value being the distance to the closest fault, along with a variety of associated fault attributes including fault magnitude, dip, and strike attributes. We also integrate seismically derived fluid-flow attributes into the shortest path calculation. The attribute volumes are extracted along wellbores and statistically compared and ranked relative to production. Finally, we use this information to calculate volumes that indicate the most prospective well locations.

Introduction

Knowing the location of the closest fault and its orientation may be useful information for determining optimum well placement. Spatial data analytics and machine learning can be utilized to learn and model the relationships between well performance, proximity to faults and the various associated fault attributes in the reservoir in general.

Several authors have attempted to automate and model the relationship of seismic faults to well performance in both conventional and unconventional reservoirs. Kolisnyk et al. (2013) determine well orientation and fault orientation are highly correlated with production in a gas field in Brazil through qualitative inspection. Other authors study the number of intersections of faults to wells (Marek et al,2020). Reine & Dunphy (2013) calculate overall fault statistics to create fault attributes that are correlated to performance. Lomask et. al (2017) link fault statistics to fracture models that are then related to production via simulation. However, none of these approaches attempt to automatically identify the closest fault and its orientation to each well and relate these to production.

We propose a spatial data analytics and machine learning-based semi-automated workflow to integrate seismic fault information with well information in order to identify optimal well locations. The foundation of this method is to calculate the shortest path (and first arrival travel-times with fast marching) from the faults to all other locations within the seismic volume. In a similar manner, fault attributes are also inferred throughout the seismic data. This data is compared to well data to determine any statistically significant relationships through feature ranking and then applied to assist with the optimal placement of new wells. We demonstrate this method on a field dataset.

Method

We first estimate the faults and fault orientations from seismic. While we demonstrate the application of deterministic faults, probabilistic faults may also be applied. We also apply a minimum fault size threshold to the fault solution to eliminate noise and small faults.

This method calculates two types of automatically derived attributes. The first, which we refer to as "fault proximity" is a volume of minimum distances (or in some cases fastest travel-times) from the faults. A value at any point in one of these volumes give the shortest travel-time to the nearest fault. The second type is a volume of nearest fault attributes, such as fault strike, dip, or magnitude. We refer to these as "fault strike proximity", "fault dip proximity", and "fault magnitude proximity". For example, a value at any point in an extended strike volume is the strike of the closest fault.

We apply the Dijkstra algorithm (1959) to compute a volume of shortest travel path by assigning the fault locations (above a threshold) as initial zero distance nodes. We also used a fast marching Eikonal solver (Adalsteinsson & Sethian, 1999). The resulting volume contains the distances along the shortest path to the closest fault node.

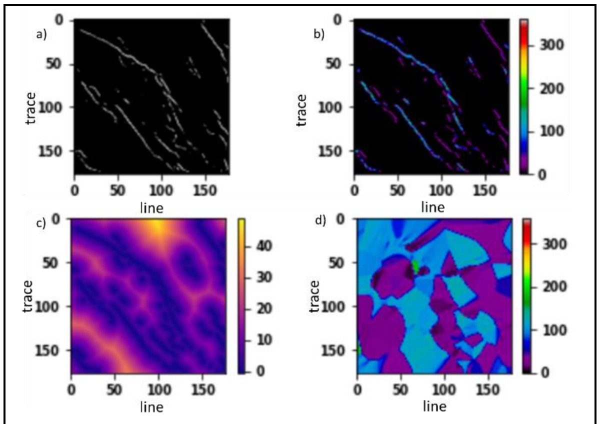

Figure 1: Time-slice comparison of (a) faults (b) strike (c)minimum distance and (d)extended strike.

10.1190/image2022-3746974.1

Figure 1a is a time-slice of an automatically created fault attribute. Figure 1b is the strike of the faults. We set the faults in Figure 1a to be the source location (zero distance) and run an Eikonal solver to get the minimum distance map displayed Figure 1c. In this case, the data is in depth and the speed for the Eikonal solver is set to one, so every location in the distance map is the shortest distance to the closest fault.

Figure 1d is the interpolated strike map. We use the fast marching method to extend the strike (Figure 1b) using the same method as extension velocities. In this simple case of constant speed, this is essentially a nearest neighbor interpolation. However, as will be discussed further, the speed away from the faults can be a variable function of seismic attributes, in which case the interpolated strike (and other attributes) is an earliest arrival.

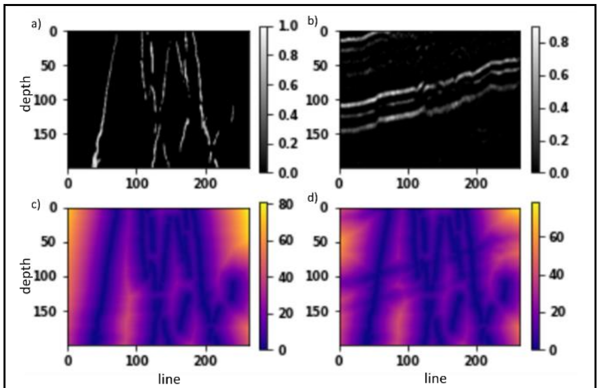

Figure 2: Section display comparing (a)faults (b)a seismically derived attribute related to zones of increase flow (c) the travel using only the faults as source locations. (d)the travel time using the combination of the faults as a source with the seismic attribute.

It is worth noting that in this case, the faults, the fault strike, and the fault dip are automatically generated using deep learning with convolutional neural networks. This is not a requirement of this method and any fault solution and associated fault attributes can be used.

This method can incorporate additional information into the travel-time calculation. Additional information is often available such as pre-stack seismic that can easily be included in the travel-time calculation by adjusting the speed proportional to the local attribute. This could include zone information and pre-stack seismic inversion results.

Figure 2a is a crossline section of the same fault solution in Figure 1a. Figure 2c is the minimum travel time using the faults in Figure 2a as the source. Figure 2b is an example of a seismically derived attribute that could be related to flow. In this case it is just the thresholded seismic amplitude. Figure 2d is the travel-time where the speed is a function of the seismic amplitude. Comparing Figures 2c and 2d, the impact of attribute on the travel-time is easily observed.

The next volumes we create are nearest fault strike, nearest fault dip, and nearest fault strength. To efficiently calculate these volumes, we propagate fault attributes away from the fault during the fault distance calculation. We adapt the extension velocities calculation from level set methods (Adalsteinsson & Sethian, 1999) by solving

$$\nabla T \cdot \nabla F_{attribute} = 0,$$

where T is the travel time and $F_{attribute}$ is strike, dip, and strength.

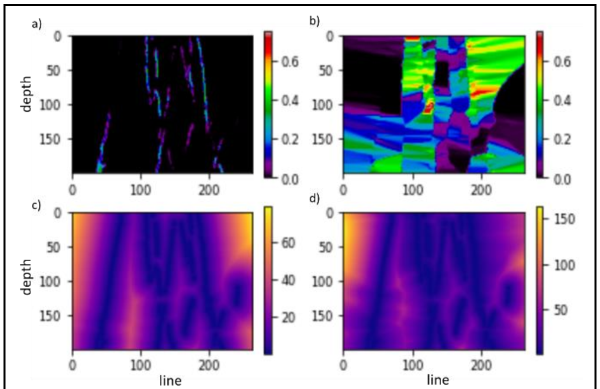

Figure 3: Section display comparing (a)fault strength (b)extended fault strength (c) the travel time using only the faults as source locations. (d)the travel time using the extended fault strength speed.

We also calculate travel time using the extended fault strength as speed. Then when we subsequently extend the strike and other attributes, the faults with greater strength will extend further. Figure 3a is fault strength. Figure 3b is the extended fault strength. Figure 3d is the travel time using the extended fault strength as speed. The faults with greater strength have reduced travel time therefore causing extended attributes to have greater influence when compared to travel time using constant speed, Figure 3c.

A coblended volume of seismic data, fault distance, fault strike proximate, and fault strike is displayed in Figure 4. Inspection of this volume with wells is helpful in determining which faults are likely most impactful on a given well's production as well as which fault orientations.

10.1190/image2022-3746974.1

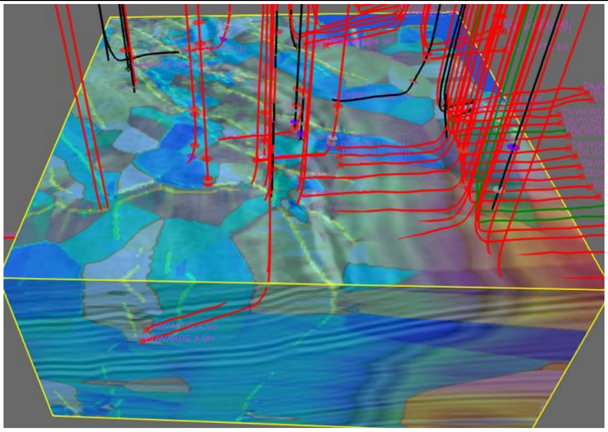

Figure 4: Coblended volume of the seismic data, the fault distance, nearest strike.

To identify any statistically significant relationships between these seismic attributes and well production we extract the attributes from the seismic data along the wellbores and compare them to the production data for each well using statistical feature ranking and determination. The relationships between the nearest fault attributes and production are evaluated with correlation analysis, rank and partial correlation coefficients, feature importance based on random forest from predictive machine learning and mutual information. Once these relationships are determined, these new volumes are applied to build spatial models to support determination of optimum well location.

Rank correlation coefficients is the Pearson correlation coefficient of the rank transform of the response ($\boldsymbol{R_y}$) and predictor ($\boldsymbol{R_{x_i}}$) features.

$$\boldsymbol{\rho_{R_x,R_y} = \frac{\sum_{i=1}^{n}(R_{x_i} - \overline{R_x})(R_y - \overline{R_y})}{(n-1)\sigma_{R_x}\sigma_{R_y}}},$$

Partial correlation coefficient calculates the linear correlation between each predictor feature and the response feature after linearly modeling and removing the influence of all other predictor features.

Random forest feature importance summarizes and standardizes the reduction in model error by including each predictor feature in each of the decision trees, providing a measure that is not limited to linear and accounts for interaction between the predictor features.

Mutual information is free of any assumptions such as linearity or Gaussianity. It summarizes the departure from the fundamental independence relationship between each predictor feature (X_α) and the response feature (Y).

$$I(X_\alpha;Y) = \int_Y \int_{X_\alpha} P_{X_\alpha,\,Y}(x_\alpha,y) \cdot log\left(\frac{P_{X_\alpha,\,Y}(x_\alpha,y)}{P_{X_\alpha}(x_\alpha)\cdot P_Y(y)}\right) dx_\alpha\, dy$$

These statistics are shown in Figure 5 and indicate significant relationships between the nearest fault attributes and well production for the field dataset.To further evaluate the relationship between faulting and well productivity we built several machine learning models and calculated the error of the model, Figure 6a. Random Forest was selected as the algorithm of choice for this study as it produced the highest correlation given the data used.

After a model is trained and error calculated, we endeavor to understand which input features have the most influence on the prediction. We apply a technique to the model known as Permutation Feature Importance to rank and sort the input features based on importance, Figure 6b. Based on this analysis, the most import variables are the median distance to the nearest fault, median strike of the nearest fault and median dip of the nearest fault.

After observing the importance of the input attributes on the

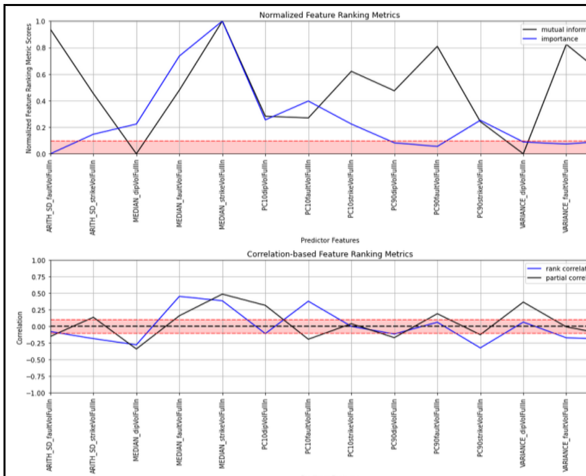

Figure 5 - Nearest fault attributes to production wells features ranking metrics to learn and model relationships between faults and production. All related to well oil production.

model we apply Accumulated Local Effects (ALE) to better understand the optimal influence of the fault schema on well production. Figures 6 c, d, and e show the impact of the top three attributes on the model and how each variable impacts production. Below the line graphs are the distributions of each variable in the data set. What can be observed from these graphs is the optimal distance to the nearest fault for production is 30 feet or more, with a fault striking 95 degrees and a dip less than 26.50.

From this analysis, we can easily generate sweet spot maps or volumes identifying drilling locations.

Using this approach, we found which attributes and data ranges were the most important to production prediction on this dataset. However, on another dataset the impact of these attributes may be significantly different. The key to this methodology is to let the data speak for itself.

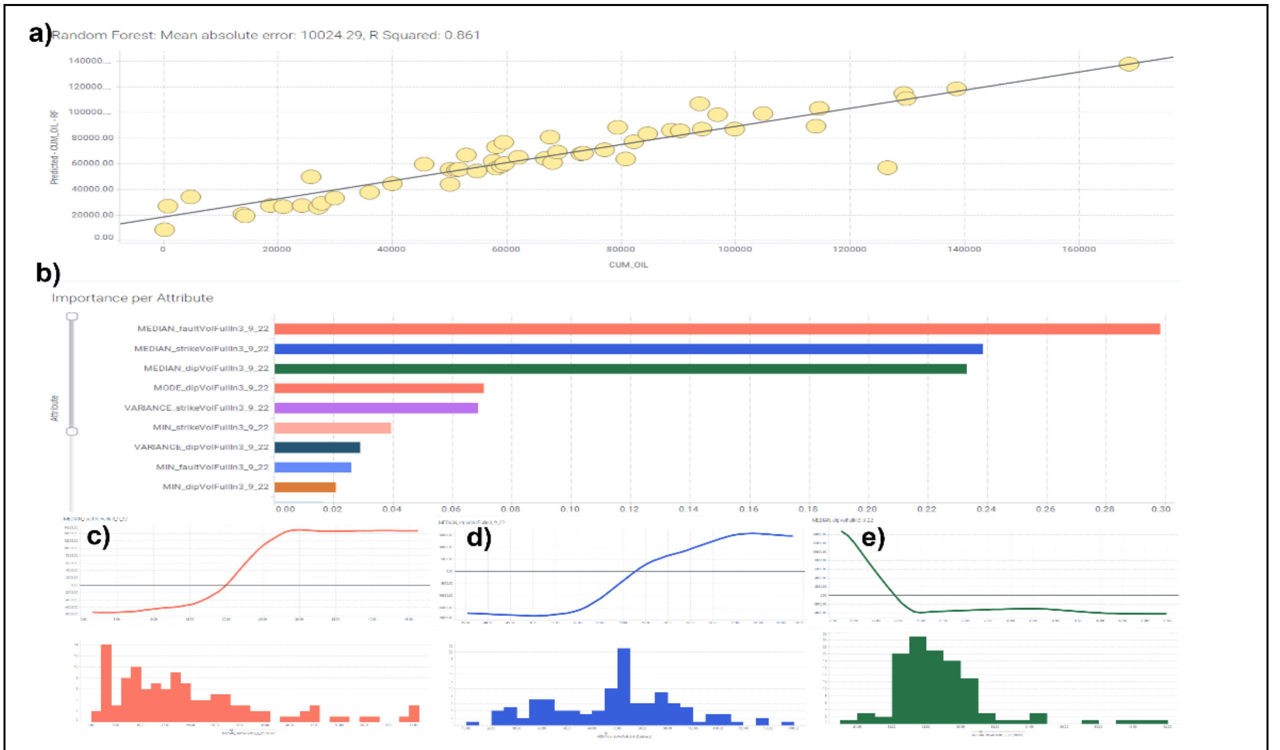

Figure 5: (a) Results from applying machine learning on well control to predict 12-month cumalative oil production. (b) Ranking of the input features in the model in order of which features were most important when predicting production. (c) (d) and (e) graphs are showing the impact of the top three features on 12-months oil production with the distribution of the input data below. It is observed from these series of graphs optimal fault orientations and distances on production.

Conclusion

Machine learning has reinvigorated interpretation by enabling the automated evaluation of new attributes. For any reservoir new attributes can be generated and then evaluated against production data in largely automated end-to-end workflows.

We presented a novel set of fault attributes integrated into a complete workflow to reduce well placement risk. This largely automated, data-driven method leverages several machine learning algorithms to collate disparate datatypes and glean new insight into their reservoir. To get the same information with traditional interpretation methods would be extremely time-consuming.

Acknowledgements

We thank Jim Bock and Vladimir Rybakov for useful discussions.

.

REFERENCES

Adalsteinsson, D., and J. A. Sethian, 1999, The fast construction of extension velocities in level set methods: Journal of Computational Physics, **148**, 2–22, doi: https://doi.org/10.1006/JCPH.1998.6090.

Dijkstra, E. W., 1959, A note on two problems in connexion with graphs: Numerische Mathematik, **1**, 269–271, doi: https://doi.org/10.1007/BF01386390.

Kolisnyk, A., R. Castellani, E. de Abreu, L. T. L. e Sá, A. Pimentel, C. da Cunha, V. da Silva, R. Kraft, C. M. A. Silva, and D. C. Moraes, 2013, Connecting seismic attributes with production performance at Camarupim Gas Field, Espirito Santo Basin, Brazil: 13th International Congress of the Brazilian Geophysical Society.

Lomask, J., L. Hernandez, V. Liceras, A. Kumar, and A. Khadeeva, 2017, A seismic to simulation unconventional workflow using automated fault-detection attributes: Interpretation, **5**, no. 3, SJ41–SJ48, doi: https://doi.org/10.1190/INT-2016-0148.1.

Marek, S., and G. Spence, 2019, Fault analysis using 3-D data reduces water production, H2S risks in Delaware Basin: The American Oil & Gas Reporter.

Reine, C., and S. Lovric, 2013, Geophysical insights into completions and production predictability for a shale gas reservoir using fault intensity and AVO inversion: Unconventional Resources Technology Conference, 1672–1678, doi: https://doi.org/10.1190/urtec2013-171.

10.1190/image2022-3746974.1