DATA-DRIVEN BRITTLENESS INDEX PREDICTION FROM ELEMENTAL SPECTROSCOPY AND PETROPHYSICAL PROPERTIES USING SUPPORT-VECTOR REGRESSION

Ardiansyah Negara, Syed Shujath Ali, Ali Al Dhamen, Hasan Kesserwan, Guodong Jin, Baker Hughes Inc.

This paper was prepared for presentation at the SPWLA 58th Annual Logging Symposium held in Oklahoma City, Oklahoma, USA, June 17-21, 2017.

ABSTRACT

Brittleness index is one of the critical geomechanical properties to understand the rock's drillability in drilling operations and screen effective hydraulic fracturing candidates in unconventional reservoirs. Brittleness index can generally be obtained from stress or strain based relationships. It can also be estimated from conventional well logs or rock mineralogical composition. Brittleness index measurements from stress/strain based relationships require laboratory tests, which are time-consuming and core samples are available only at discrete depths. While well logs can estimate a continuous profile of brittleness index along the borehole, it is derived from empirical correlation specific to a rock type. More recent advancements in logging tools have enabled the determination of elemental spectroscopy downhole. This information combined with petrophysical properties such as density and porosity can capture brittleness characteristics of rocks. This paper presents the use of support-vector regression (SVR) to construct a data-driven brittleness index prediction from the elemental spectroscopy and petrophysical properties.

The relationship of brittleness index with elemental spectroscopy, density, and porosity is often complex and nonlinear. The SVR described in this paper is used to correlate the elemental spectroscopy, density, and porosity to the brittleness index, thereafter building a data-driven brittleness index prediction model. The dataset of brittleness index, elemental spectroscopy, density, and porosity used in this study are based on various geological formations. Laboratory tests such as unconfined compressive strength, confined compressive strength, and Brazilian test were conducted. Brittleness indices were calculated based on data generated from these tests. Elemental spectroscopy data were obtained from X-ray fluorescence (XRF) analysis. The data are then separated into two categories: training and testing data. Training data are used to train the SVR and establish the brittleness index prediction model, while the testing data are used for validation.

In total, 28 cases were run with different combinations of petrophysical and mineralogical properties, number of training dataset, and SVR kernel functions. The results reveal that the SVR-based brittleness indices match very well with the laboratory-measured brittleness indices. Cross-correlation plots of regression models between the predicted and the measured brittleness indices show high values of coefficient of determination. The small error and high values of coefficient of determination denote the SVR models' good performances. The prediction accuracy improves as more data are included to train the algorithm. From the comparison of SVR-kernel-function-based models, we observe that the RBF-based model performs better than the polynomial-based model. The RBF-based model yields better accuracy than the polynomial-based model using the same number of training dataset. Referring to the RBF-based model with 80% training dataset, it was observed that elemental spectroscopy has more influence than the other rock properties on the prediction. The promising results stemming from this study confirm that SVR can be further applied to build a brittleness index prediction model based on mineralogy logs and petrophysical logs.

INTRODUCTION

Brittleness is defined as a property of material that ruptures or fractures under small deformation or strain (Hucka and Das, 1974; Neuendorf *et al.*, 2011). Brittleness of rock is a complex function of lithology, mineral composition, total organic carbon, effective stress, reservoir temperature, diagenesis, thermal maturity, porosity, and type of fluid (Wang and Gale, 2009). Brittleness has been extensively studied by many researchers due to its importance for wellbore stability analysis (Moos *et al.*, 2003; Zoback *et al.*, 2003), hydraulic fracture evaluation (Rickman *et al.*, 2008; Yang *et al.*, 2013; Jin *et al.*, 2014), sand production control (Weingarten and Perkins, 1995; Bradford *et al.*, 1998; Nicholson *et al.*, 1998) and ROP prediction (Kahraman, 2002; Liu *et al.*, 2014; He *et al.*, 2016). However, there is currently no universally standardized

and accepted formula to quantify brittleness index. In fact, there are more than twenty brittleness index formulations have been developed by various authors (Jin *et al.*, 2014). Current approaches of calculating brittleness include stress/strain based relationships. In addition, brittleness can be estimated from conventional well logs or rock mineralogical composition. Brittleness measurements from stress/strain based relationships require laboratory tests, which are time-consuming and core samples are available only at discrete depths. While well logs can estimate a continuous profile of brittleness index along the depth, it is derived from the empirical correlation specific to a rock type.

Recently, artificial intelligence techniques (*e.g.*, artificial neural networks, extreme learning machines, support-vector machines) have been proposed and used intensively to predict rock brittleness (Ghobadi and Naseri, 2016; Shi *et al.*, 2016a, 2016b). Ghobadi and Naseri (2016) discussed the brittleness index prediction of Hamekasi limestone during freeze-thaw cycles. Simple and multiple regressions were utilized to identify the correlation between geomechanical properties and brittleness index. Based on simple regression, it was observed that all geomechanical properties (unconfined compressive strength, tensile strength, P-wave velocity, porosity, quick absorption index, and dry density) have good correlations with brittleness index. Ghobadi and Naseri (2016) employed multiple regression (MR) and artificial neural network (ANN) to build the brittleness index prediction model. The results of the study reveal that the brittleness index model that has porosity, P-wave velocity and dry density as input variables produce the best prediction according to the statistical parameters. Increasing the input variables does not necessarily improve models performance. Shi *et al.* (2016a) presented data-driven brittleness index prediction approaches based on back-propagation artificial neural network (BP-ANN), extreme machine learning (ELM), and linear regression using conventional logging data and laboratory mineralogical-based brittleness. In total, 71 core dataset collected from Silurian Longmaxi marine shale, Jiaoshiba Shale Gas Field, Sichuan Basin, China. The brittleness index used in the study was derived based on mineralogy. The results of the study demonstrate that the brittleness index prediction model based on BP-ANN and ELM yield high accuracy and efficiency on the prediction process compared with the simple regression correlations. The BP-ANN-prediction model is better on the prediction accuracy, while ELM excels on the computational cost. Moreover, the BP-ANN and ELM-based models were further applied to estimate brittleness index from well logging data. Comparing to empirical correlations-based brittleness approaches, data-driven based approach using artificial intelligence techniques shows wider applications with better accuracy. Similar study had also been conducted again by Shi *et al.* (2016b) to estimate brittleness index from well logs using artificial intelligence techniques. In the later study, Shi *et al.* (2016b) build the brittleness index prediction model based on BP-ANN and least square-support vector regression (LS-SVR). Core samples and well logging data were acquired from a well in Silurian Longmaxi marine shale gas reservoir in the Sichuan Basin, China. A comparison of the BP-ANN and LS-SVR models demonstrate that the latter is more accurate than the first at the same conditions.

As mentioned earlier, brittleness of rock is a complex function of mineral compositions, lithology, total organic carbon, effective stress, reservoir temperature, diagenesis, thermal maturity, porosity, and type of fluid (Wang and Gale, 2009). Having most or even all of these parameters to build the brittleness model will enhance the prediction accuracy. However, those parameters are not always available. From the literature mentioned above, it can be seen that petrophysical properties such as density and porosity are mostly used as input parameters to predict brittleness index. Nevertheless, none of them considered mineral compositions as input in the model prediction. In fact, rock is composed of minerals that contribute to the brittleness. In this paper, we propose to build a data-driven approach of brittleness index prediction from elemental spectroscopy and petrophysical properties such as density and porosity using support-vector regression. There are 76 data were collected from different wells, formations, and geographical regions. Brittleness index was obtained from stress-strain based relationship. In this paper, we calculate the brittleness index defined by Hucka and Das (1974). Density was measured using pycnometer and then used to calculate porosity. Elemental spectroscopy was obtained from X-ray fluorescence (XRF) analysis. In total, 28 cases are run with different combination of input parameters, number of training data as well as SVR kernel types. We evaluate the prediction accuracy using coefficient of determination and mean absolute percentage error. Considering the limited number of data available in this study, the SVR-based brittleness model produces very good accuracy. The results reveal that SVR-based brittleness model has a great potential to be used for brittleness prediction. It can be further applied to build a brittleness index prediction model based on mineralogy logs and petrophysical logs.

SUPPORT-VECTOR REGRESSION

Support-vector regression is a type of support-vector machines (SVMs) that was further developed from SVM classification (Vapnik and Lerner, 1963; Vapnik and Chervonenkis, 1964) to solve nonlinear regression problems in high dimensional feature space by introducing kernel function and Vapnik's ε-insensitive loss function (Vapnik 1995, Vapnik *et al.*, 1997). In the following, we describe the concept of SVR and its application to predict brittleness index from elemental spectroscopy, density, and porosity. Let us consider the following dataset, $\mathcal{D}$, expressed by

$$\mathcal{D} = [(x_1, y_1), \ldots, (x_k, y_k)], \;\; x \in R^n, y \in R \tag{1}$$

with a linear function given by

$$f(\mathbf{x}) = \langle \mathbf{w}, \mathbf{x} \rangle + b, \quad \mathbf{w} \in X, b \in R \tag{2}$$

where $\langle \mathbf{w}, \mathbf{x} \rangle$ indicates the inner product between the weights, **w**, and the input vector, **x**. In this study, the input, **x**, consists of elemental spectroscopy, density, and porosity. The output, **y**, is the core measured brittleness index. There are some constraints to find a function $f(\mathbf{x})$ in Eq. (2): first, the deviation of estimating output values should not be greater than ε from the actual training data; second, the weight should be as flat as possible, implying that **w** is as small as possible by minimizing the Eucledian norm of **w** (i.e., $\|\mathbf{w}\|^2$). The optimal regression function is determined from the estimation of **w** and b by solving the following optimization problem:

$$\text{Minimize} \;\; \frac{1}{2}\|\mathbf{w}\|^2 + C\sum_{i=1}^{n}(\zeta_i + \zeta_i^*) \tag{3}$$

subject to

$$\begin{cases} y_i - \langle \mathbf{w}, \boldsymbol{x_i} \rangle - b \le \varepsilon + \zeta_i \\ \langle \mathbf{w}, \boldsymbol{x_i} \rangle + b - y_i \le \varepsilon + \zeta_i^* \\ \zeta_i, \zeta_i^* \ge 0 \end{cases} \tag{4}$$

where ε is the error accuracy used to measure the error between the actual and predicted data, ζ_i and ζ_i^* are the slack variables, which are used to penalize complex fitting functions. The constant C in Eq. (3) is the regularization parameter that determines how large the deviation from the desired accuracy is tolerated. The penalty is acceptable as long as the fitting error is not greater than ε. The SVR technique sets the ε-insensitivity loss function, $|\zeta|_\varepsilon$, by

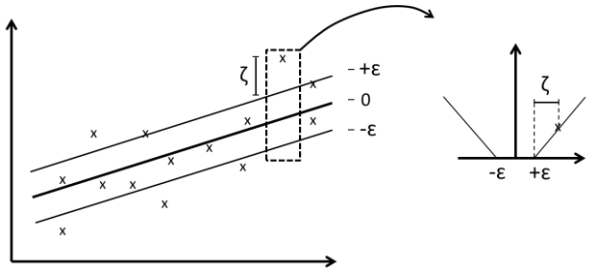

Fig. 1. The insensitive band for a linear SVR (Schölkopf and Smola, 2002).

$$|\zeta|_\varepsilon = \begin{cases} 0 & |\zeta| \le \varepsilon \\ |\zeta| - \varepsilon & otherwise \end{cases} \tag{5}$$

The ε-insensitivity loss function stabilizes the estimation and can be visualized as a tube-sized equivalent to the approximation accuracy in training data (**Fig. 1**). The optimization problem given by Eqs. (3) and (4) can be solved by a dual formulation by constructing a Lagrangian function from the objective function and constraints by introducing a dual set of variables. The regression-estimation model can then be expressed by

$$\mathbf{w} = \sum_{i=1}^{k}(\alpha_i - \alpha_i^*)\,\mathbf{x}_i \tag{6}$$

$$f(\mathbf{x}) = \sum_{i=1}^{k}(\alpha_i - \alpha_i^*)\langle \mathbf{x}_i, \mathbf{x} \rangle + b \tag{7}$$

where b is calculated using the Karush-Kuhn-Tucker (KKT) conditions (Karush 1939, Kuhn and Tucker 1951):

$$b = \begin{cases} y_i - \langle \mathbf{w}, \mathbf{x}_i \rangle - \varepsilon & \text{for } 0 < \alpha_i < C \\ y_i - \langle \mathbf{w}, \mathbf{x}_i \rangle + \varepsilon & \text{for } 0 < \alpha_i^* < C \end{cases} \tag{8}$$

In nonlinear problems, the SVR can preprocess the training patterns into a high-dimensional feature space by mapping $\boldsymbol{\Phi}$ to linearly estimate the regression. The linear regression model constructed in the feature space reflects a nonlinear regression model in the original input space. The flattest function in the feature space is derived as follows:

$$\mathbf{w} = \sum_{i=1}^{k}(\alpha_i - \alpha_i^*)\,\boldsymbol{\Phi}(\mathbf{x}_i) \tag{9}$$

$$f(\mathbf{x}) = \sum_{i=1}^{k}(\alpha_i - \alpha_i^*)k(\mathbf{x}_i, \mathbf{x}) + b \tag{10}$$

where $\mathbf{w}$ is implicitly given in the estimation function and $k(\mathbf{x}_i, \mathbf{x})$ is the kernel function. There are several types of kernel functions as presented in **Table 1**.

Table 1. Kernel functions and their corresponding mathematical formulations.

Kernel Function	Mathematical Formulation
Linear	$k(\mathbf{x}_i, \mathbf{x}) = (\mathbf{x}_i, \mathbf{x})$
Polynomial	$k(\mathbf{x}_i, \mathbf{x}) = (\mathbf{x}_i, \mathbf{x})^d$
Gaussian RBF	$k(\mathbf{x}_i, \mathbf{x}) = e^{\frac{\|x_i - x\|^2}{2\sigma^2}}$
Sigmoid	$k(\mathbf{x}_i, \mathbf{x}) = tanh[\kappa(\mathbf{x}_i, \mathbf{x}) + \upsilon]$

MODEL PERFORMANCE INDICATORS

Two statistical indicators are used to evaluate the performance of the brittleness index prediction model. They are coefficient of determination and mean absolute percentage error (MAPE). The coefficient of determination (R^2) is used to explain how much variability of one factor can be caused by its relationship to another factor. In the context of statistical model to evaluate the prediction of future outcomes, R^2 measures the degree of correlation between the actual and predicted data. The coefficient of determination ranges from 0 to 1 where the value of 1 implies a perfect correlation between the actual and predicted data and value of 0 means no statistical correlations between the actual and predicted data. Coefficient of determination is calculated as follows:

$$R^2 = \left[\sum_{i=1}^{N}\left(D_p - \bar{D}_p\right)(D_m - \bar{D}_m)\right]^2 \times \left[\sum_{i=1}^{N}\left(D_p - \bar{D}_p\right)^2 (D_m - \bar{D}_m)^2\right]^{-1} \quad (11)$$

where N is the number of observations, D_m is the measured data, D_p is the predicted data, $\bar{D}_m$ and $\bar{D}_p$ are the mean of measured and predicted data, respectively. MAPE is a measure of the size of the error in percentage terms. MAPE is used in quantitative forecasting methods because it yields a measyre of relative overall fit. The smaller the MAPE value, the better the prediction. MAPE is calculated by

$$MAPE = \frac{100}{N}\sum_{i=1}^{N}\left|\frac{D_m - D_p}{D_m}\right| \quad (12)$$

METHODOLOGY

Fig. 2 shows the flowchart of the computational experiment, where in the first step we acquired the core-measured brittleness index data from the laboratory test. Next, we check the data to remove outliers, if any, and build cross-plot correlation between each input parameter with the output. In this case, we plot the cross-correlation between elemental spectroscopy, density, and porosity, with brittleness index separately. Then, we separate the data into two categories: training and testing data. In general, training dataset consists of more than half of the total dataset and the remaining dataset is used to validate the predicted results.

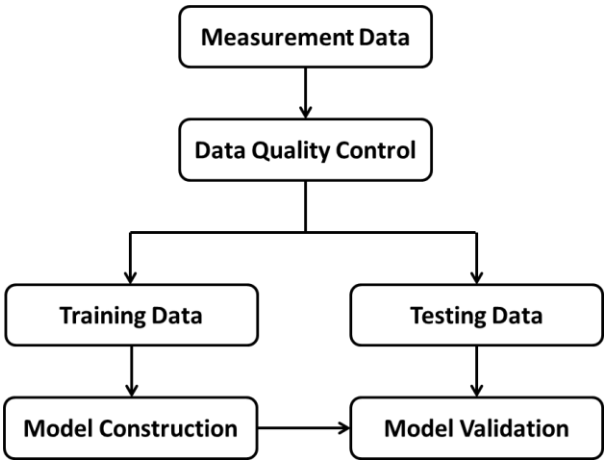

Fig. 2. Flowchart of the brittleness index model construction and validation.

DATA PROCESSING AND DATA ANALYSIS

A set of 76 brittleness index, density, porosity, and elemental spectroscopy data from different wells, formations, and geographical regions were collected. Laboratory measurements were conducted to obtain the brittleness index, density, porosity, and elemental spectroscopy data. In this paper, brittleness index is calculated as follows (Hucka and Das, 1974):

$$BI = \frac{\varepsilon_r}{\varepsilon_t} \quad (13)$$

where ε_r and ε_t are the recoverable and the total strain. Density was measured using pycnometer and then used to calculate porosity. Elemental spectroscopy was obtained from X-ray fluorescence (XRF) analysis. **Table 2** presents the statistical feature of the 76 dataset including minimum, maximum, range, and mean for each parameter. The dataset used in this study includes

rocks with low and high brittleness indices. The lowest brittleness index is 0.248 and the highest one is 1. Porosity and grain density range from 0.761% to 10.323% and 2.447 kg/m^3 to 2.866 kg/m^3, respectively. Laboratory XRF analysis shows that calcium oxide (CaO) content is the highest and ranges from 57.066% to 96.626%.

Table 1. Statistical features of the 76 dataset.

Parameters	Range	Min	Max	Mean
BI	0.752	0.248	1	0.788
Porosity (%)	9.562	0.761	10.323	4.995
GD (kg/m^3)	0.419	2.447	2.866	2.612
BD (kg/ m^3)	0.616	2.225	2.840	2.483
Elemental:				
Na_2O	0.029	0	0.029	0.001
MgO	23.808	0.801	24.609	1.939
Al_2O3	7.576	0	7.576	1.823
SiO_2	18.115	0	18.115	5.512
P_2O_5	0.375	0.035	0.410	0.099
SO_3	21.830	0.256	22.086	3.966
K_2O	0.853	0.069	0.922	0.399
CaO	39.560	57.066	96.626	83.898
TiO_2	0.274	0.107	0.381	0.213
MnO	0.038	0	0.038	0.011
Fe_2O_3	1.737	1.068	2.805	1.840
ZnO	0.174	0.006	0.180	0.059
ZrO_2	0.397	0	0.397	0.020
SrO	1.407	0	1.407	0.211

*) BI: Brittleness Index
*) GD: Grain Density
*) BD: Bulk Density

RESULTS AND DISCUSSION

In this section, we present 28 cases with different combinations of input parameters, number of training dataset, and SVR kernel functions. We begin to consider a case with single input parameter (only grain density or porosity or elemental spectroscopy) and then add the number of input parameters gradually. Training data are set to be 60% and 80%. In general, the performance of the SVR model strongly depends on the selection of the kernel function and its parameters. In this work, we use RBF and polynomial kernel functions with kernel parameters are C = 30.299 and σ^2 = 1.173. **Tables 3-6** show the 28 cases with the following details: **Table 3** presents the cases with 60% training dataset and RBF as kernel function; **Table 4** displays the cases with 60% training dataset and polynomial as kernel function; **Table 5** shows the cases with 80% training dataset and RBF as kernel function; and **Table 6** contains the cases with 80% training dataset and polynomial as kernel functions. First, we want to investigate the behavior of the prediction accuracy with different input parameters. The results presented in **Table 3** demonstrate that the RBF-based model with 60% training data yields relatively good prediction for all cases (Cases 1-7). However, it is not easily observed which formation property has more influence on the prediction. As an example, when the input parameter is only grain density (Case 1) or only porosity (Case 2), both cases have smaller errors compared to the others, however, the coefficients of determination are lower. In addition, the case with elemental spectroscopy being the only input parameter produces the highest coefficient of determination but also the highest error. If we look at the polynomial-based model with the same number of training dataset (**Table 4**), similar behaviors could be identified, for example the cases with only grain density (Case 8) or only porosity (Case 9) have smaller error compared with others.

When we feed the model with more training data for the SVR algorithm to learn, then we begin to identify the behaviors of the prediction especially for RBF-based model. **Table 5** shows the cases with 80% training dataset using RBF kernel function. From the results presented in this table, we clearly see that the cases with elemental spectroscopy being one of the input parameters have the highest coefficients of determination and the smallest errors compared with the other cases without elemental spectroscopy as an input parameter. This indicates that mineral elements have more influence on brittleness index prediction than the density and porosity. The errors for the cases without considering elemental spectroscopy (Cases 15, 16, and 18) are about 9-10%, which is almost double than the errors for the cases that consider elemental spectroscopy (Cases 17, 19, 20, and 21). Furthermore, the coefficients of determination for the cases that include elemental spectroscopy are higher than the ones without elemental spectroscopy. As mentioned earlier, the performance of the SVR model strongly depends on the selection of kernel function. Sometimes certain kernel function does not perform well especially when the function to be solved is highly complex. As an example, although the number of training dataset has increased, the behavior of the prediction results using polynomial-based model (**Table 6**) are not easily identified. This implies that the polynomial kernel function is not capable of representing the nonlinear correlation between density, porosity, elemental spectroscopy, and brittleness index. From the prediction results presented in **Tables 3-6**, the RBF-based model with 80% training dataset is the best to conduct sensitivity analysis of the rock parameters on the

brittleness index prediction because this model produces consistent behavior on the prediction results.

Next, we compare the prediction results generated by the RBF-based model and the polynomial-based model. Both models have linear trends between the number of training dataset and the statistical parameters. The RBF-based model with 60% training dataset has about 15-17% errors on the prediction, while the errors generated by the polynomial-based model are about 21-24%. In average, the error difference between the RBF-based model and the polynomial-based model is 7%, which implies that the RBF-based model produces 30% more accurate result than the polynomial-based model. When the number of training dataset increases to 80%, the error on the prediction decreases to be 5-9% (**Table 5**). It indicates that the error has reduced by 43-66% compared to the cases with 60% training dataset. Similarly, the error on the prediction generated by polynomial-based model has reduced from average of 23% to be 14.5% (37% improvement on the accuracy). In overall, brittleness index prediction using RBF-based model has better estimation than the polynomial-based model. In average, brittleness index prediction using RBF-based model is 30% and 50% more accurate than polynomial-based model for 60% and 80% training dataset, respectively. **Figures 3-6** display the measured against predicted brittleness index for selected cases.

CONCLUSIONS

SVR-based brittleness index prediction model has been built based on the 76 dataset, which were collected from different wells, formations, and geographical regions. In total, 28 cases were run with different combinations of petrophysical and mineralogical properties including density, porosity, and elemental spectroscopy. Furthermore, the combinations include the number of training dataset and the SVR kernel functions. The results of this study can be summarized as follows:

- The SVR-based model is capable of estimating brittleness index from density, porosity, and elemental spectroscopy especially when using RBF kernel function.
- The RBF-based model performs better than the polynomial-based model. The RBF-based model yields higher coefficient of determination and smaller error than the polynomial-based model using the same number of training dataset.
- Referring to the RBF-based model with 80% training dataset, it is observed that elemental spectroscopy has more influence than the other rock properties on the brittleness index prediction.
- The number of training has a linear relationship with the prediction accuracy.

Table 3. Cases with 60% training dataset and RBF kernel function.

Case	GD	Porosity	Elements	R^2	MAPE
1	✓	-	-	0.798	15.384
2	-	✓	-	0.740	15.733
3	-	-	✓	0.910	17.217
4	✓	✓	-	0.800	16.107
5	✓	-	✓	0.880	16.112
6	-	✓	✓	0.889	16.887
7	✓	✓	✓	0.867	15.973

Table 4. Cases with 60% of training dataset and polynomial kernel function.

Case	GD	Porosity	Elements	R^2	MAPE
8	✓	-	-	0.634	21.560
9	-	✓	-	0.663	21.525
10	-	-	✓	0.534	24.505
11	✓	✓	-	0.611	22.434
12	✓	-	✓	0.591	23.382
13	-	✓	✓	0.541	24.397
14	✓	✓	✓	0.594	23.354

Table 5. Cases with 80% of training dataset and RBF kernel function.

Case	GD	Porosity	Elements	R^2	MAPE
15	✓	-	-	0.805	9.603
16	-	✓	-	0.828	9.596
17	-	-	✓	0.915	5.902
18	✓	✓	-	0.828	9.877
19	✓	-	✓	0.921	5.979
20	-	✓	✓	0.929	5.310
21	✓	✓	✓	0.918	5.851

Table 6. Cases with 80% of training dataset and polynomial kernel function.

Case	GD	Porosity	Elements	R^2	MAPE
22	✓	-	-	0.761	14.485
23	-	✓	-	0.765	14.705
24	-	-	✓	0.524	16.301
25	✓	✓	-	0.766	14.688
26	✓	-	✓	0.720	12.402
27	-	✓	✓	0.507	16.590
28	✓	✓	✓	0.715	12.550

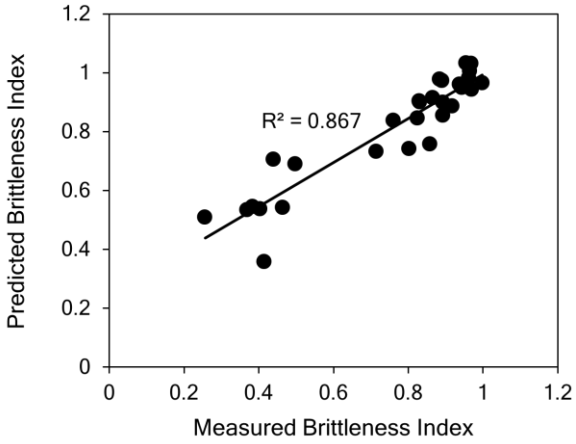

Fig. 3. Measured against predicted brittleness index for Case 7.

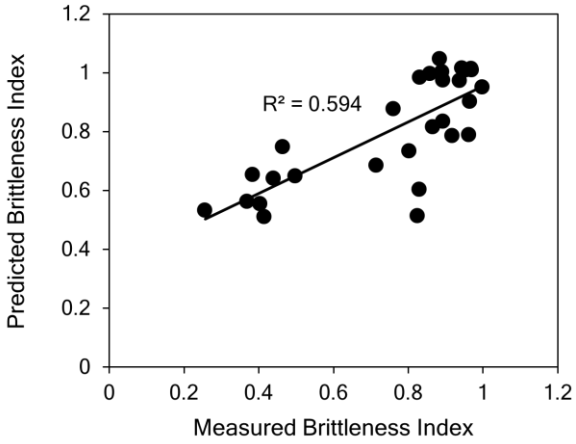

Fig. 4. Measured against predicted brittleness index for Case 14.

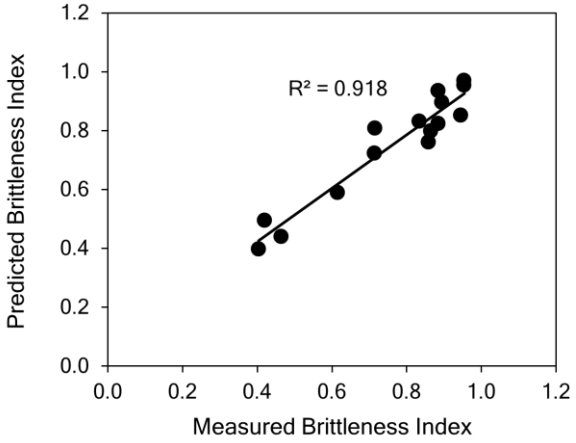

Fig. 5. Measured against predicted brittleness index for Case 21.

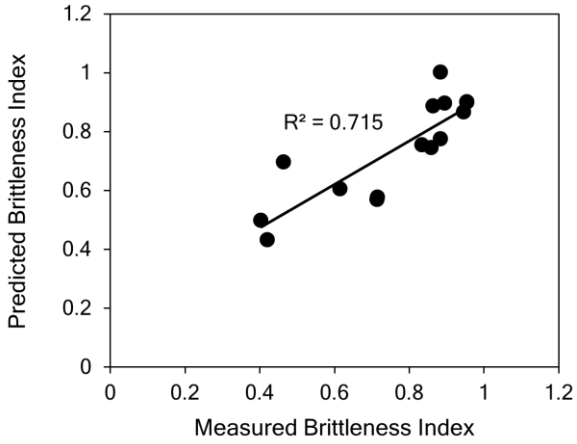

Fig. 6. Measured against predicted brittleness index for Case 28.

ACKNOWLEDGEMENTS

The authors thank Baker Hughes for the permission to publish this paper.

REFERENCES

Bradford, I.D.R., Fuller, J., Thompson, P.J., and Walsgrove, T.R., 1998, Benefits of Assessing the Solids Production Risk in a North Sea Reservoir using Elastoplastic Modelling, SPE-47360-MS, presented at the SPE/ISRM Rock Mechanics in Petroleum Engineering Conference, Trondheim, Norway, 8-10 July.

Ghobadi, M.H. and Naseri, F., 2016, Rock Brittleness Prediction Using Geomechanical Properties of Hamekasi Limestone: Regression and Artificial Neural Networks Analysis, *Geopersia*, 6(1), 19-33.

He, J., Chen, Y., Zhengchun, L., and Samuel, R., 2016, Global Correlation of Rock Brittleness Indices with Petrophysical and Geomechanical Properties and Its Application to the Prediction of Rate of Penetration (ROP), SPE-180518-MS, presented at the IADC/SPE Asia Pacific Drilling Technology Conference, Singapore, 22-24 August.

Hucka, V. and Das, B, 1974, Brittleness Determination of Rocks by Different Methods, *International Journal of Rock Mechanics and Mining Sciences & Geomechanics Abstracts*, 11(10), 389-392.

Jin, X., Shah, S., Truax, J., and Roegiers, J-C., 2014, A Practical Petrophysical Approach for Brittleness Prediction from Porosity and Sonic Logging in Shale Reservoirs, SPE-170972-MS, presented at the SPE Annual Technical Conference and Exhibition, Amsterdam, The Netherlands, 27-29 October.

Kahraman, S., 2002, Correlation of TBM and Drilling Machine Performances with Rock Brittleness, *Engineering Geology*, 65(4), 269-283.

Karush, W., 1939, Minima of Functions of Several Variables with Inequalities as Side Constraints, MS Thesis, Department of Mathematics, University of Chicago, Chicago, Illinois.

Kuhn, H.W. and Tucker, A.W., 1951, Nonlinear Programming, In Proceedings of 2nd Berkeley Symposium on Mathematical Statistics and Probabilistics, 481-492, Berkeley, California: University of California Press.

Liu, Z., Marland, C., Li, D., and Samuel, R., 2014, An Analytical Model Coupled with Data Analytics to Estimate PDC Bit Wear, SPE-169451-MS, presented at the SPE Latin America and Caribbean Petroleum Engineering Conference, Maracaibo, Venezuela, 21-23 May.

Moos, D., Peska, P., Finkbeiner, T., Zoback, M., 2003, Comprehensive Wellbore Stability Analysis Utilizing Quantitative Risk Assessment, *Journal of Petroleum Science and Engineering*, 38(3-4), 97-109.

Neuendorf, K.K.E., Mehl, J.P., and Jackson, J.A., 2011, Glossary of Geology, Fifth Edition revised. Alexandria, Virginia, USA: American Geological Institute (reprint).

Nicholson, E.D., Goldsmith, G., and Cook, J.M., 1998, Direct Observation and Modeling of Sand Production Processes in Weak Sandstone, SPE-47328-MS, presented at the SPE/ISRM Rock Mechanics in Petroleum Engineering Conference, Trondheim, Norway, 8-10 July.

Rickman, R., Mullen, M.J., Petre, J.E., Grieser, M.V., and Kundert, D., 2008, A Practical Use of Shale Petrophysics for Stimulation Design Optimization: All Shale Plays Are Not Clones of the Barnett Shale, SPE-115258-MS, presented at the SPE Annual Technical Conference and Exhibition, Denver, Colorado, USA, 21-24 September.

Schölkopf, B. and Smola, A., 2002, Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond, Cambridge, Massachusets: Adaptive Computation and Machine learning, MIT Press.

Shi, X., Liu, G., Cheng, Y., Jiang, H., Chen, L., Jiang, S., and Wang, J., 2016a, Brittleness Index Prediction in Shale Gas Reservoirs Based on Efficient Network Models, *Journal of Natural Gas Science and Engineering*, 35, 673-685.

Shi, X., Liu, G., Jiang, H., Chen, L., and Yang, L., 2016b, Brittleness Index Prediction from Conventional Well Logs in Unconventional Reservoirs Using Artificial Intelligence, IPTC-18776-MS, presented at the International Petroleum Technology Conference, Bangkok, Thailand, 14-16 November.

Vapnik, V. and Lerner, A., 1963, Pattern Recognition using Generalized Portrait Method, *Automation and Remote Control*, 24, 774-780.

Vapnik, V. and Chervonenkis, A., 1964, A Note on One Class of Perceptrons. Automation and Remote Control 25.

Vapnik, V., 1995, The Nature of Statistical Learning Theory, New York: Springer.

Vapnik, V., Golowich, S., and Smola, A., 1997, Support Vector Method for Function Approximation, Regression Estimation, and Signal Processing, In Advances in Neural Information Processing Systems 9, edition M. C. Mozer, M. I. Jordan, and T Petsche, 281-287, Cambridge, Massachusetts: MIT Press.

Wang, D.B. and Gale, J.F., 2009, Screening Criteria for Shale-Gas Systems, *Gulf Coast Association of Geological Societies Transactions*, 59, 779-793.

Weingarten, J.S. and Perkins, T.K., 1995, Prediction of Sand Production in Gas Wells: Methods and Gulf of Mexico Case Studies, *Journal of Petroleum Technology*, 47(7), 596-600.

Yang, Y., Sone, H., Hows, A., and Zoback, M.D., 2013, Comparison of Brittleness Indices in Organic-rich Shale Formations, ARMA 13-403,

presented at the 47th U.S. Rock Mechanics/Geomechanics Symposium, San Francisco, California, 23-26 June.

Zoback, M.D., Barton, C.A., Brudy, M., Castillo, D.A., Finkbeiner, T., Grollimund, B.R., Moos, D.B., Peska, P., Ward, C.D., and Wiprut, D.J., 2003, Determination of Stress Orientation and Magnitude in Deep Wells, *International Journal of Rock Mechanics and Mining Sciences*, 40(7-8), 1049-1076.

ABOUT THE AUTHORS

Ardiansyah Negara is a Petroleum Engineer at Baker Hughes Dhahran Technology Center. His primary research interests are in reservoir modeling and simulation and machine learning applications for formation evaluation and drilling optimization. He received his MS and PhD in Earth Science and Engineering from King Abdullah University of Science and Technology. He holds a BEng in Civil Engineering from Institut Teknologi Sepuluh Nopember. He is an active member of SPE and SPWLA.

Syed Shujath Ali is a Geomechanics Research Scientist at Baker Hughes based in Dhahran Technology Center. His area of interest includes rock mechanical properties characterization with special interest in unconventional shales. He is also involves in building 1-D mechanical model with application to wellbore instability. He holds a Master of Science degree in Mechanical Engineering from King Fahd University of Petroleum Minerals, Saudi Arabia.

Ali Al Dhamen is a GeoMechanics specialist at Baker Hughes with experience in wellbore instability. He also is experienced in laboratory Petrophysics and Rock mechanics. He received his Master's degree in Geology and Geophysics from Louisiana State University, Baton Rouge in 2014.

Hasan Kesserwan is currently working as a research scientist in the Baker Hughes Dhahran Global Technology center located in Saudi Arabia. He received his MS and PhD in condensed matter physics and nanotechnology from University of Strasbourg, France. His focus was on experimental and theoretical investigation of laser induced magnetization dynamics in magnetic nanoparticles based nano-devices for data storage application. He joined the Spintronics Theory Group at King Abdullah University of Science and Technology (KAUST) in 2011 as a postdoctoral and was involved in quantum mechanical modelling of the ultrafast dynamics phenomena and thermal control of magnetization dynamics, Spin-Caloritronics.

Guodong Jin is Petrophysics Advisor and Lead of Petrophysics & Geomechanics Group at Baker Hughes Dhahran Global Technology Center, Saudi Arabia. He has been actively involved in reservoir characterization and evaluation, logging interpretation, digital rock physics, geomechanics and core analysis. Prior to his current position, he was a research scientist and Project Lead at Baker Hughes Houston Technology Center. He received a PhD in Reservoir Engineering and Petrophysics, and a MS in Geomechanics, both from the University of California at Berkeley. He holds a BEng in Petroleum Engineering from China University of Petroleum.