

## Research Article

# Geomechanical Properties of Unconventional Shale Reservoirs

Mohammad O. Eshkalak,<sup>1</sup> Shahab D. Mohaghegh,<sup>2</sup> and Soodabeh Esmaili<sup>3</sup>

<sup>1</sup>Department of Petroleum and Geosystems Engineering, The University of Texas at Austin, Austin, TX 78712, USA

<sup>2</sup>Department of Petroleum and Natural Gas Engineering, West Virginia University, Morgantown, WV 26506, USA

<sup>3</sup>Occidental Petroleum Corporation, Bakersfield, CA 93311, USA

Correspondence should be addressed to Mohammad O. Eshkalak; [eshkalak@utexas.edu](mailto:eshkalak@utexas.edu)

Received 12 July 2014; Accepted 12 October 2014; Published 3 December 2014

Academic Editor: Andrea Franzetti

Copyright © 2014 Mohammad O. Eshkalak et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Production from unconventional reservoirs has gained an increased attention among operators in North America during past years and is believed to secure the energy demand for next decades. Economic production from unconventional reservoirs is mainly attributed to realizing the complexities and key fundamentals of reservoir formation properties. Geomechanical well logs (including well logs such as total minimum horizontal stress, Poisson's ratio, and Young, shear, and bulk modulus) are secured source to obtain these substantial shale rock properties. However, running these geomechanical well logs for the entire asset is not a common practice that is associated with the cost of obtaining these well logs. In this study, synthetic geomechanical well logs for a Marcellus shale asset located in southern Pennsylvania are generated using data-driven modeling. Full-field geomechanical distributions (map and volumes) of this asset for five geomechanical properties are also created using general geostatistical methods coupled with data-driven modeling. The results showed that synthetic geomechanical well logs and real field logs fall into each other when the input dataset has not seen the real field well logs. Geomechanical distributions of the Marcellus shale improved significantly when full-field data is incorporated in the geostatistical calculations.

## 1. Introduction

Shale gas reservoirs, which are also called source rock reservoirs (SRR), have some unique attributes that make hydraulic fracturing an essential option in order to commence an economic level of the natural gas production. Unlike conventional gas reservoirs, insufficient permeability, ultra-low porosity of shale rock, and limited reservoir contact area, but vastly organic-rich formation, cannot offer production in a commercial value without stimulation processes. Many studies are conducted from shale pore-scale level to field scale reservoir simulations to improve the understanding of complex flow behavior that are developed and discussed through numerical, analytical, and semianalytical reservoir models for unconventional reservoirs [1–10]. However, in order to predict the performance of a shale gas reservoir, implementing accurate shale rock properties is essential for developing a geologic model for the entire asset. Hence, it is very critical to access more data while working on an unconventional reservoir. In this study, synthetic data

are generated using artificial intelligence and data mining techniques (AI&DM).

Principal stress profile of an oil and gas reservoir depends highly on the rock geomechanical properties. Geomechanical properties of reservoir rock include Poisson's ratio, total minimum horizontal stress, and bulk, Young, and shear modulus. These properties play significant role in current development plans of shale assets compared to conventional reservoirs that have established sufficient information available. Moreover, having access to geomechanical data can assist engineers and geoscientists during geomechanical modeling, hydraulic fracture treatment design, and reservoir simulation in shale gas fields across the U.S. and worldwide. Geomechanical well logs are one of the sources that secure such data. Running geomechanical well logs (in all wells in a shale asset) is not common practice among operators and the reason is attributed to the cost associated with running such logs. In this paper, data-driven models are developed to accurately determine the Marcellus shale rock properties.

Artificial intelligence and data mining have been used within last 20 years in reservoir modeling and characterization to perform analysis on formation of interest [11, 12]. Also, some studies indicated that artificial neural network (ANN) is a powerful tool for pattern recognition and system identification such as methodology developed by Mohaghegh et al. in 1998 [13] to generate synthetic magnetic resonance imaging (MRI) logs using conventional logs such as SP, GR, and resistivity. Their methodology incorporated an artificial neural network as its main tool to generate the target variable. The synthetic magnetic resonance imaging logs were generated with a high degree of accuracy even when the model developed used data not employed during model development. Moreover, Basheer and Najjar demonstrated that ANN is suitable to predict and classify soil compaction and rocks characteristics as well as determining some mechanical parameters such as Young's modulus, Poisson's ratio [14]. They mainly investigated the neural network capability in solving geotechnical engineering problems and they provided a general view of some neural network application in their field of research.

In this study, it is demonstrated that AI&DM technology is able to develop data-driven models for generating rock geomechanical properties. The overall work-flow includes development of synthetic geomechanical well logs from conventional logs such as gamma ray and bulk density that are commonly available. These data-driven models used around 30 percent of data (coming from geomechanical logs) of the entire asset, which were available to expand them for the rest of the field with conventional logs but no geomechanical logs. Data-driven models have been validated with blind wells. Blind wells are wells with actual data which are selected due to different locations in the asset of 100 horizontal gas wells. Moreover, the logs generated from data-driven models are used to build an integrated field-wide geomechanical distribution (maps and volumes) for rock geomechanical properties. In this work, the ultimate purpose is meant to propose a technique in order to omit the necessity of running geomechanical logs for the entire asset once such logs are obtained for some portion of the asset to be used in the data-driven models. The number of well logs required to run the data-driven models that leads to an accurate result is determined to be 30 percent of the wells in an unconventional asset. In this study, just 30 out of 100 horizontal wells in the Marcellus shale asset in southern Pennsylvania own actual geomechanical well logs that are provided by a major service company.

## 2. Methodology

In order to accomplish the objectives of this study, a methodology and thorough procedure are required to be defined. The methodology used to accomplish the objectives of this study includes four steps as indicated below. A detailed description of each step is explained afterwards:

- (a) data preparation,
- (b) data-driven model development using AI&DM,

- (c) validation of data-driven models,
- (d) geomechanical property distribution.

*2.1. Step (a): Data Preparation.* Data preparation is the most important step in developing data-driven models due to the fact that all the other steps are using the data prepared in this step. Data preparation involves checking the data for accuracy, entering data in a right format in a computer file, and developing and documenting a database structure that integrated the various properties used in the next steps. In this study, a dataset form the available well logs are required to be prepared that portrays specific property of the rock versus depth. First, production pay zone must be identified from the conventional well logs along with the horizontal wells trajectory. After identifying the depth of the producing zones for Marcellus shale from conventional well logs such as gamma ray, available information of each individual well is extracted from well logs in every one or half a foot according to its log characteristics and tools used. Also, in order for the models to understand the differences between production pay zones and while using these datasets, it is essential to specify the contrast between pay zones (upper Marcellus, Purcell, lower Marcellus, and Onondaga) and the adjacent rocks, nonshale. To account for this contrast, 50 feet depth of log data located above and below the pay zones of interest is also added to the main dataset.

The prepared dataset contains rows and columns consisting of following data that are recorded versus depth: the wells name, the well coordinates, the values for gamma ray (GR), bulk density (BD), sonic porosity, bulk modulus (BM), shear modulus (SM), Young's modulus (YM), Poisson's ratio (PR), and total minimum horizontal stress (TMHS) for each horizontal shale well. It must be emphasized that not all the wells include geomechanical well logs; thus geomechanical values are only recorded for the wells that have such real data, 30 wells.

*2.2. Step (b): Data-Driven Models Development.* In this step, the prepared dataset was processed using backpropagation algorithm of neural network into two main parts.

*Part 1 (Conventional Models).* In order to have consistent conventional logs data for all wells for the shale production pay zone, part 1 is defined and the conventional logs are generated for those wells that missed some conventional logs. As it is shown in Figure 1, the bulk density and sonic porosity for 30 wells were produced by using two different data-driven models. First model (neural network model 1) used gamma ray, depth, location, and well coordinates as an input to develop training, calibration, and verification segments for generating the bulk density of around 30 wells in the asset where bulk density and sonic data were missing. Second model (neural network model 2) used also bulk density as input beside inputs of first step to generate sonic porosity for the part of asset without this property (around 30 wells also used in this part).

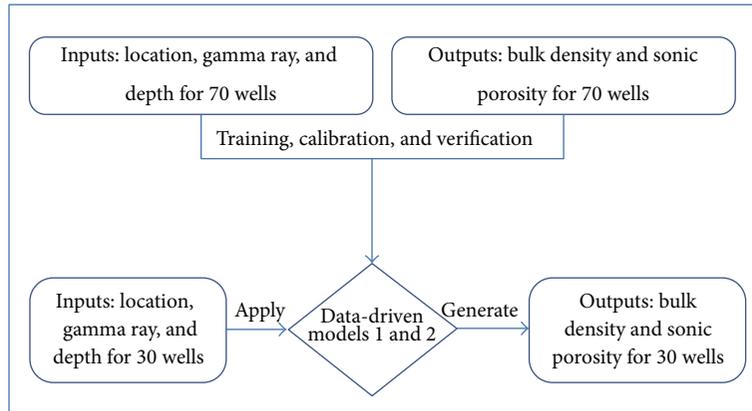


FIGURE 1: Part 1—data-driven model for conventional logs.

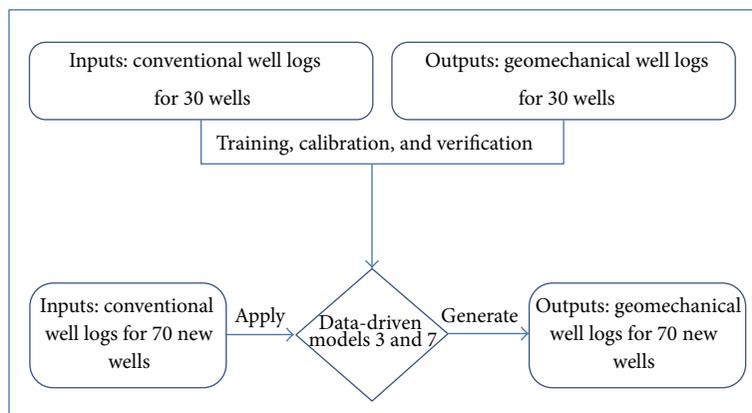


FIGURE 2: Part 2—data-driven models for geomechanical logs.

At the end of this step, all existing wells in the asset have the required conventional well log properties to be used in part 2.

*Part 2 (Geomechanical Models).* After completing the missing data for conventional logs from part 1 for all horizontal wells, another neural network structure is used to develop five different data-driven models (models 3 to 7) as shown in Figure 2 to generate the geomechanical well logs for all wells. As it is shown in Figure 2, this step consists of five neural network models in which the inputs were completed in each step by using the generated geomechanical property in the previous step. In detail, for each of these neural networks, the same conventional logs of 30 wells have been provided as the input and one geomechanical property was generated at a time. Then each generated geomechanical property was used as an input for the next neural network model and the process continued until all five geomechanical properties of interest were generated for the entire Marcellus shale asset.

All models are multilayer networks that are trained using a backpropagation method of neural network technology. In order to achieve the least error in backpropagation method, a different percentage of datasets are incorporated in the models and finally it is concluded that considering 80% of

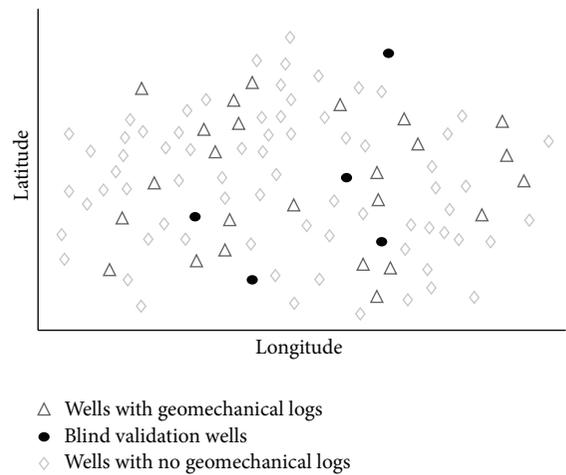


FIGURE 3: Marcellus shale gas field, southern Pennsylvania.

data used in the training process, and by considering 20% for the calibration process and the remaining for the last step, verification (10% for each), the universal backpropagation error is minimized. A general backpropagation scheme and formulation is explained in Appendix. In the next section, two different methods for error calculation s are explained.

*Error Analysis.* The mean absolute percentage error (MAPE), also known as mean absolute percentage deviation, is a measure of accuracy of a method for constructing fitted time series values in statistics, specifically in trend estimation. It usually expresses accuracy as a percentage and is defined by

$$\text{MAPE} = \frac{100\%}{n} \sum_{t=1}^n \left| \frac{At - Ft}{At} \right|, \quad (1)$$

where  $At$  is the actual value and  $Ft$  is the forecast value.

$$R\text{-squared} = 1 - \frac{\text{Sum of squared distances between the actual and predicted values}}{\text{Sum of squared distances between the actual and their mean values}}. \quad (2)$$

In our study the highest achieved  $R$ -squared was around 98 percent and the lower one in some cases around 89 percent, and in both situations, the results presented are highly acceptable. A higher level of  $R$ -squared reflects, in all three stages of training, calibration, and verification, a reliable correlation between actual and generated data. It is also important to mention that during the initial training of datasets; the results obtained were with low  $R$ -squared. Unsuccessful behavior of models was understood because of having some wells with log data for each 0.5 ft., which is in contrast with the rest of the wells with every available 1 ft. log data. Once piece of data of 0.5 ft. turned to 1 ft., which is considered as discrepancy that misleads the data-driven models; the results came out properly and the data-driven models showed rapid improvements. Further, second issue that resulted in smoother behavior of the data-driven models was related to the removal of nonshaly thin intervals log data from the pay zones that exists within the upper Marcellus, Purcell, lower Marcellus, and Onondaga. Once these layers were removed, the models converged much faster with very low backpropagation error.

*2.3. Step (c): Data-Driven Model Validation.* This step explains a robust method to analyze the accuracy and validity of the data-driven model's results, although the universal error of all the models was very low according to previous section. To examine the models validity, the well log data of some wells (which have both real conventional and geomechanical logs) was removed from the training dataset and it was attempted to regenerate the geomechanical logs. These removed wells are so called blind wells. Blind wells have been chosen from different location in the Marcellus shale asset under study. Then, data-driven models used this new dataset to be trained to generate geomechanical properties for blind wells. Data-driven models number 3 to 7 was separately validated to generate geomechanical properties. In each step, the generated property compared and plotted against the actual values which had been removed from main dataset. The results of this step are explained in Results and Discussions section.

Figures 4 through 8 demonstrate the actual well logs and generated logs for 5 blind wells that are chosen from different

The difference between  $At$  and  $Ft$  is divided by the actual value  $At$  again. The absolute value in this calculation is summed for every fitted or forecasted point in time and divided again by the number of fitted points  $n$ . Multiplying by 100 makes it a percentage error. Also,  $R$ -squared is defined and determined in (2) for all three steps, training, calibration, and verification. The higher the  $R$ -squared, the closest the results to the actual values:

location in the asset as illustrated in Figure 3. To compare the results, both actual and generated properties are plotted in the same figure like an actual well log. Properties such as bulk modulus, Young modulus, Poisson's ratio, shear modulus, and total minimum horizontal stress are presented, respectively. Blue line shows the actual value and the red line is for generated values by data-driven models. For well #1 to well #4, there is a perfect match between blue and red lines that shows the models have generated the exact actual data. These wells are in proximity of wells with actual geomechanical properties according to their locations and depths. As it was expected, results shown for these wells are accurate which demonstrate data-driven model's capability and accuracy in generating geomechanical properties of Marcellus shale.

*2.4. Step (d): Geomechanical Property Distribution.* The first objective of this paper is accomplished in the previous section and the geomechanical properties are generated for all existing wells in the Marcellus shale asset. To accomplish this step, different geostatistical methods, from Petrel commercial software, are considered to create geomechanical property distribution for the Marcellus shale field. Further, geomechanical well logs generated from the data-driven models are coupled with a commercial reservoir simulator in order to create geomechanical distributions for properties such as total minimum horizontal stress, Poisson's ratio, and Young, shear, and bulk modulus.

Sequential Gaussian simulation (SGS) is finally used to create distribution according to well locations for the entire field due to its very smooth and consistent surfaces and distributions (maps) obtained compared to other methods. Two types of maps were created. First map is only incorporated with 30 wells which already had actual geomechanical logs. The second map is related to entire field (70 wells with generated property and 30 wells with actual data). With comparing these two maps, significant difference between geomechanical property distribution with and without having full-field data is observed as shown in Figures 9 through 13. Ten maps that show distribution of five rock geomechanical properties in the Marcellus shale asset were created.

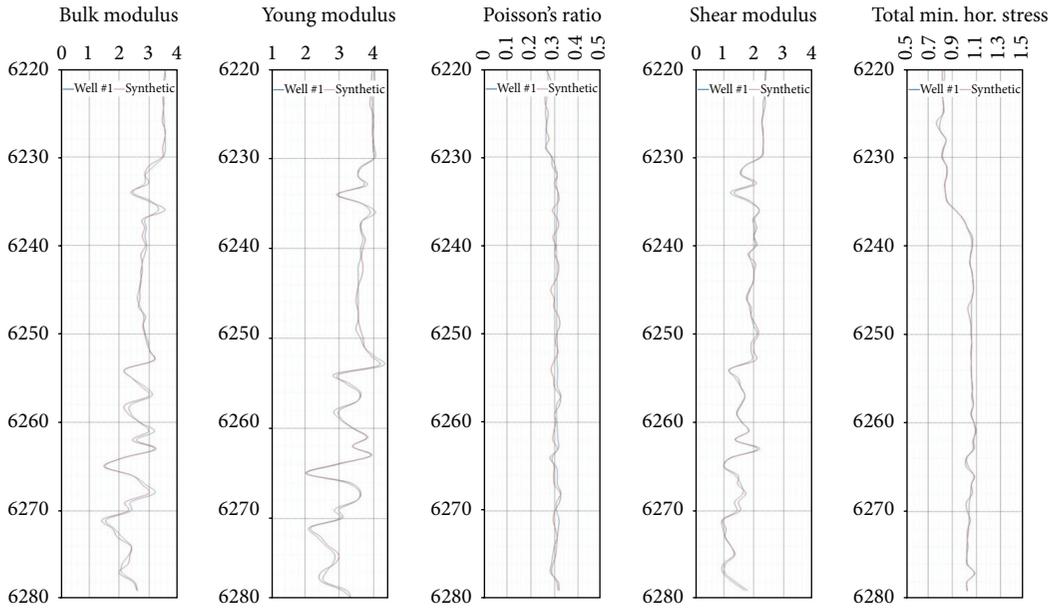


FIGURE 4: Well #1, actual versus generated well logs.

TABLE 1: Information used for databases for developing data-driven models.

Well identifier	Description	Conventional well logs	Geomechanical well logs	Number of wells
Circle	Blind validation wells	Yes	Yes	5
Triangle	Wells used for validation	Yes	Yes	25
Diamond	Wells with no geomechanical logs	Yes	No	70

### 3. Results and Discussions

The Marcellus shale under study consists of 100 multifractured horizontal wells. Figure 3 depicts the distribution of existing wells in the asset that is used in this study. Table 1 shows the information and number of wells that were used to develop data-driven models in different steps as well as the validation purpose in step (c).

In this study, a multilayer neural networks or multilayer perceptions are considered to develop the data-driven models. These networks are most suitable for pattern recognition specially in nonlinear problems neural network that have one hidden layer with different number of hidden neurons that are selected based on the number of data records available and the number of input parameters selected in each training process.

The training process of the neural networks is conducted using a backpropagation technique. In the training process, the dataset is partitioned into three separate segments. This is done in order to make sure that the neural network will not be trapped in the memorization phase. Moreover, the intelligent partitioning process allows the network to adapt to new data once it is being trained. The first segment, which includes the majority of the data, is used to train the model. In order to prevent the memorizing and overtraining effect in the neural network training process, a second segment of the data is taken for calibration that is blind to the neural

network and at each step of training process, the network is tested for this set. If the updated network gives better predictions for the calibration set, it will replace the previous neural network; otherwise, the previous network is selected. Training will be continued once the error of predictions for both the calibration and training dataset is satisfactory. This will be achieved only if the calibration and training partitions are showing similar statistical characteristics. Verification partition is the third and last segment used for the process that is kept out of training and calibration process and is used only to test the precision of the neural networks. Once the network is trained and calibrated, then the final model is applied to the verification set. If the results are satisfactory then the neural network is accepted as part of the entire prediction system [15, 16].

Figures 4 to 8 show the actual well logs and generated logs for 5 blind wells shown as black circles in Figure 3. To compare the results, both actual and generated properties are plotted in the same figure similar to an actual well log. In these plots, properties such as bulk modulus, Young modulus, Poisson's ratio, shear modulus, and total minimum horizontal stress are presented, respectively. Blue line shows the actual value and the red line is for generated values by data-driven models. For well #1 to well #4 (Figures 5, 6, and 7), there is perfect match between blue and red lines. These wells are in proximity of wells with actual geomechanical properties according to their locations and depths. As it was expected,

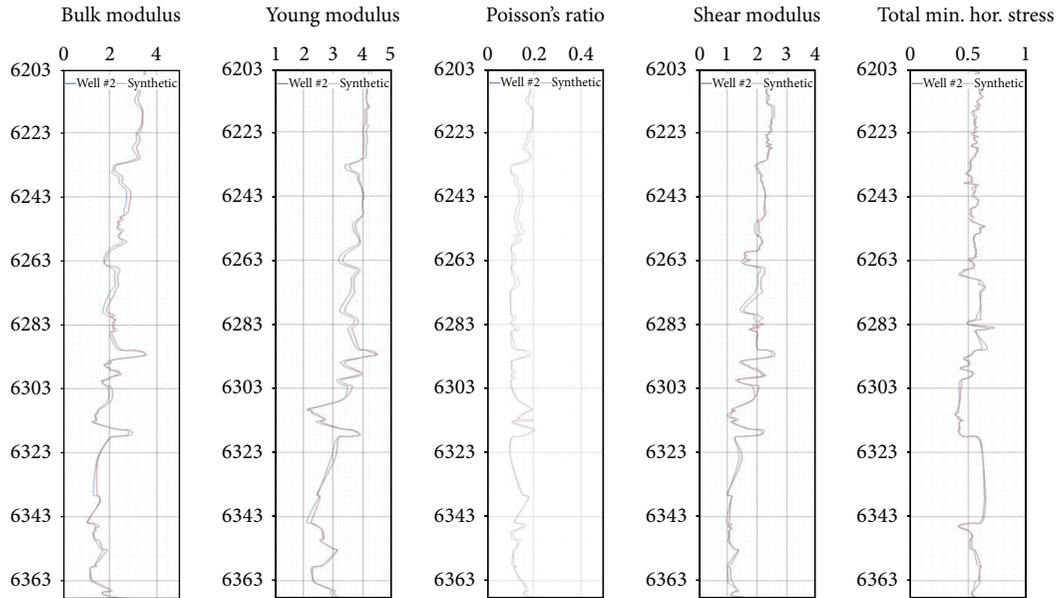


FIGURE 5: Well #2, actual versus generated well logs.

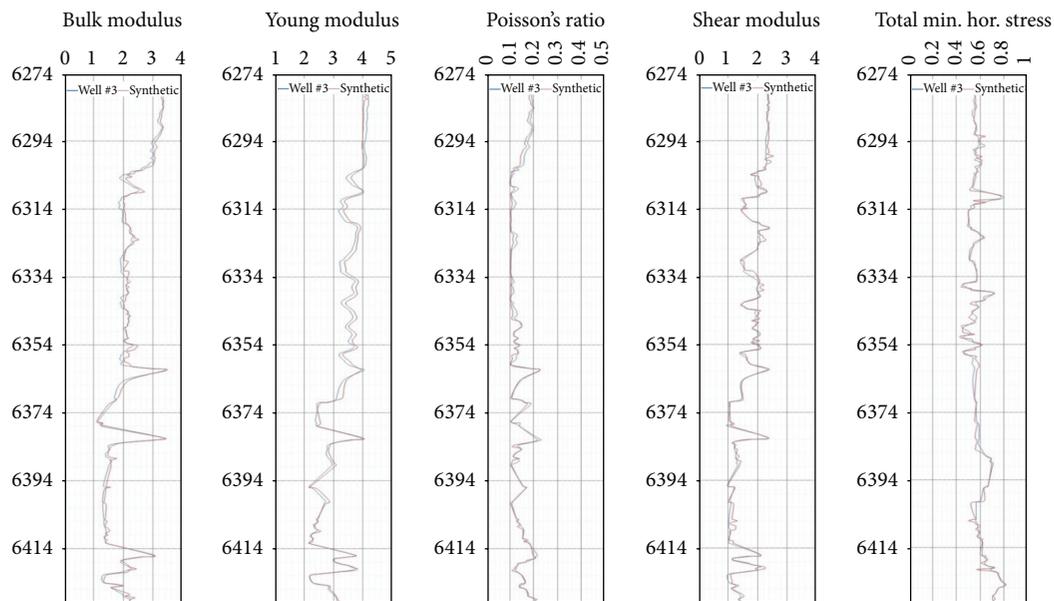


FIGURE 6: Well #3, actual versus generated well logs.

results shown for these wells are accurate which demonstrate data-driven models capability in predicting geomechanical properties.

For well #5 in Figure 8, the generated data is not in agreement with the actual logs and it is because of the location of the well that is far from (upper side of the asset in the field—Figure 2) the rest of wells in the asset that we have used for the training purposes. This fact indicates that the models could not predict the behavior of outlier wells and most importantly emphasizes on the fact that data-driven modeling is perfect for interpolations and not accurate for the extrapolation as it is in agreement with the neural network literature. Moreover,

it is found that the depth of producing pay zone of this well #5, compared to other four blind wells, is different (out of range) and it might be another reason related to the fact that models could not capture the behaviors very well.

Figures 9, 10, 11, 12, and 13 are showing distributions and maps for the five geomechanical rock properties of interest in this study. For each property, there are two distributions: one that is generated by using the actual data and the second that considered the information of both generated and actual data (full-field data). A comparison between maps for each property demonstrates that more reasonable and accurate distribution is achieved using more data for the asset.

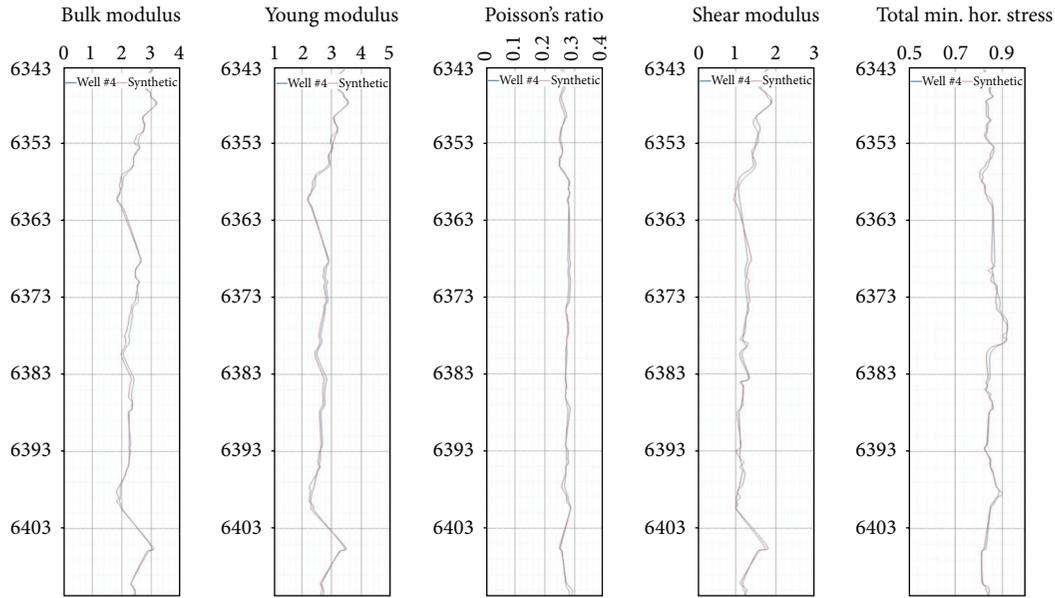


FIGURE 7: Well #4, actual versus generated well logs.

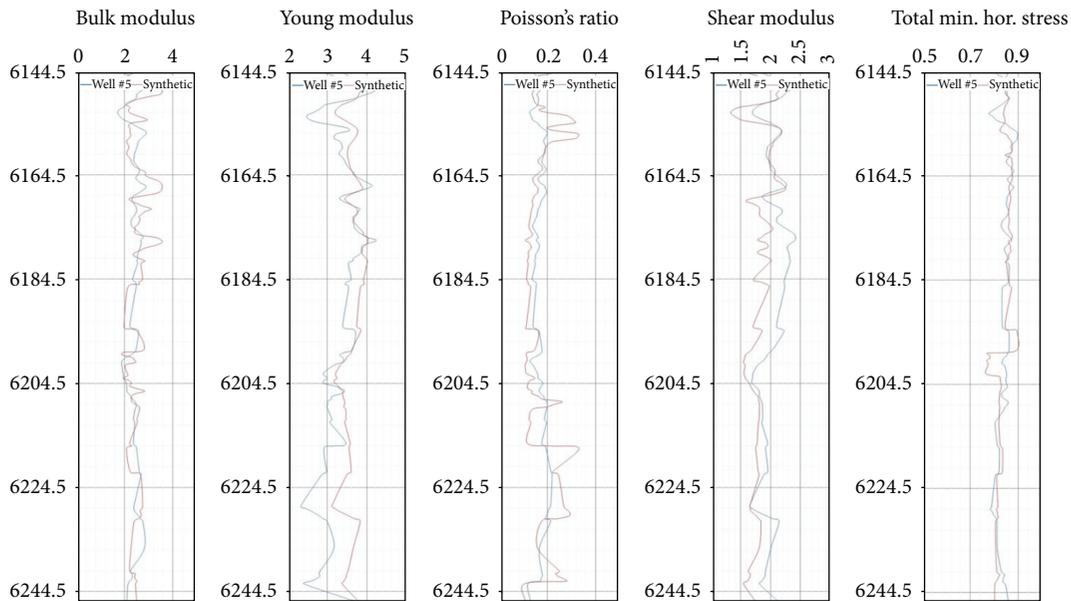


FIGURE 8: Well #5, actual versus generated well logs.

The sequential Gaussian simulation (SGS) algorithm was used in order to generate these maps. In the top distribution map, plus signs represent the wells with actual data which have been used in dataset for training, calibration, and verification during data-driven model development.

**4. Conclusions**

In this study, it is demonstrated that the data-driven modeling using AI&DM technology is a reliable and robust tool to obtain accurate results for generating synthetic geomechanical logs for unconventional shale resources. In simple terms,

we used conventional well logs to generate field-wide geomechanical properties and distribution maps of geomechanical properties for the entire asset, Marcellus shale in southern Pennsylvania.

Five data-driven models were designed, trained, and validated to predict five geomechanical properties of interest for Marcellus shale unconventional reservoir. First, data mining issue in this study, removing nonshaly intervals and adding 50 feet contrast zone, was successfully managed to lead a reliable prediction with least error calculated in backpropagation method. Also, second validation process, the use of 5 blind wells, was performed to show the robustness and accuracy of

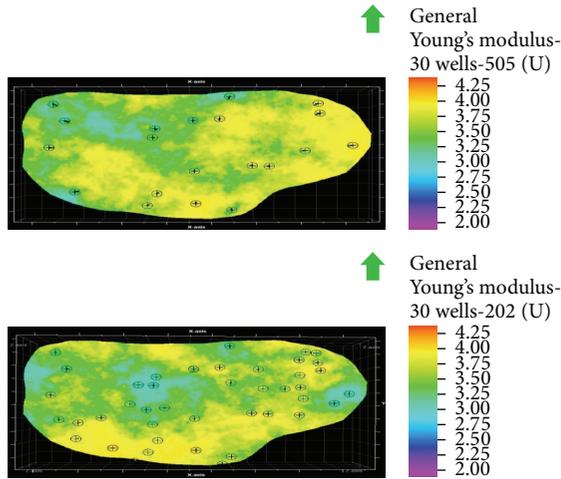


FIGURE 9: Young modulus.

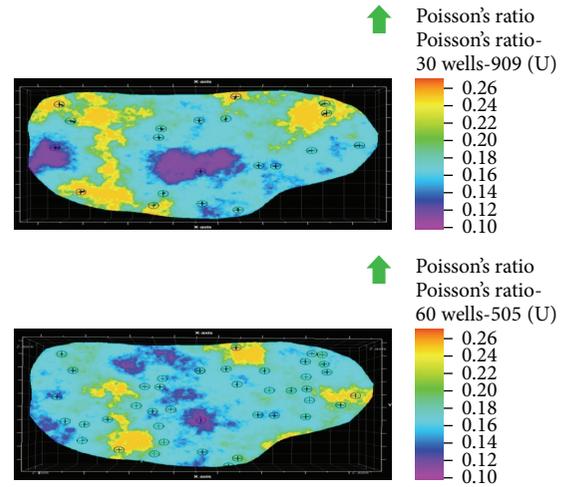


FIGURE 11: Poisson's ratio.

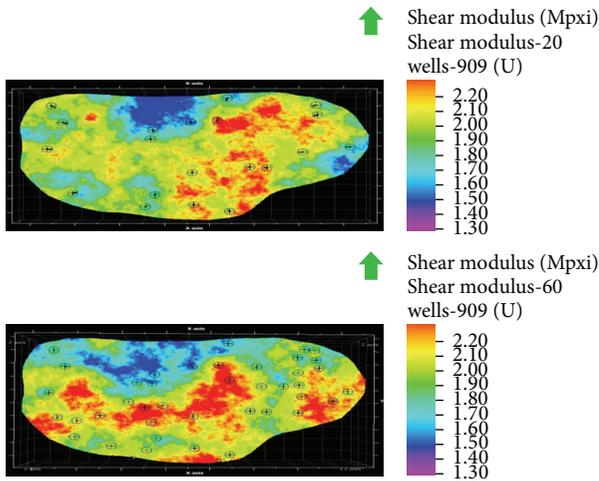


FIGURE 10: Shear modulus.

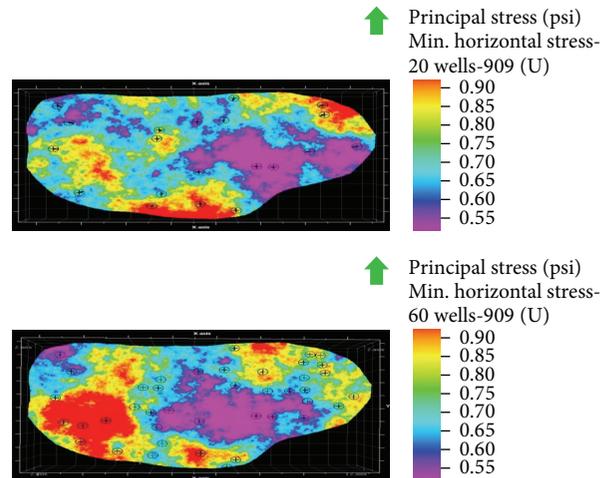


FIGURE 12: Total minimum horizontal stress.

data-driven models for predicting Young modulus, Poisson ratio, bulk modulus, shear modulus, and total minimum horizontal stress.

Geomechanical property distribution maps of the entire asset illustrated a significant difference between distributions when there are just a few available pieces of actual data rather than having access to the full-field data. These synthetic geomechanical logs and property distributions for Marcellus shale exhibit a great deal of assistance to better performing reservoir modeling, characterization and the optimization of hydraulic fracturing issues related to the current Marcellus shale development process. Authors expect these models will conclude also accurate results in other unconventional shale resources.

## Appendix

### Backpropagation Method Formulation

We now derive the backpropagation technique for a general case. The equations (A.1) through (A.8) show the

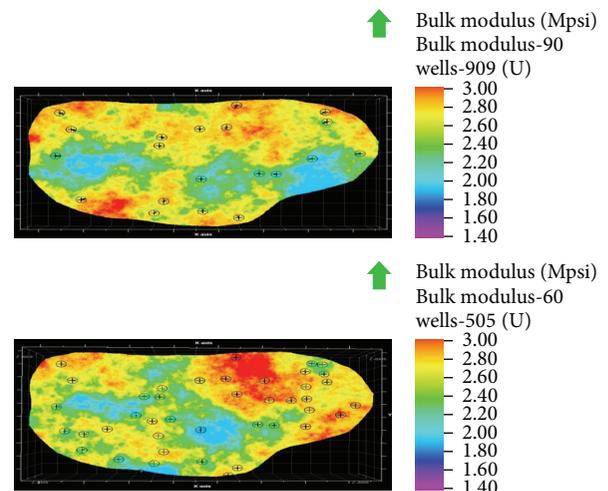


FIGURE 13: Bulk modulus.

mathematical representations of backpropagation method using the input dataset:

$$\begin{aligned}
 \bar{x}_j &= \text{input vector for unit } j, \\
 \bar{w}_j &= \text{Weight vector for unit } j, \\
 z_j &= \bar{w}_j \cdot \bar{x}_j, \quad \text{the weighted sum of inputs for unit } j, \\
 o_j &= \text{Output of unit } j \quad (o_j = \sigma(z_j)), \\
 t_j &= \text{target for unit } j,
 \end{aligned} \tag{A.1}$$

where  $t_j$  is the actual value or the target value that we wish to achieve using the backpropagation method. Since we update after each training example, we can simplify the notation somewhat by imagining that the training set consists of exactly one example and so the error can simply be denoted by  $E$ . Downstream ( $j$ ) = set of units whose immediate inputs include the output of  $j$ . Outputs = set of output units in the final layer.

We want to calculate  $\partial E / \partial w_{ji}$  for each input weight  $w_{ji}$  for each output unit  $j$ . Note first that since  $z_j$  is a function of  $w_{ji}$  regardless of where in the network unit  $j$  is located,

$$\frac{\partial E}{\partial w_{ji}} = \frac{\partial E}{\partial z_j} x_j \frac{\partial z_j}{\partial w_{ji}} = \frac{\partial E}{\partial z_j}. \tag{A.2}$$

Furthermore,  $\partial E / \partial z_j$  is the same regardless of which input weight of unit  $j$  we are trying to update. So we denote this quantity by  $\delta_j$ . Consider the case when  $j \in \text{Outputs}$ . We know

$$E = \frac{1}{2} \sum_{k \in \text{Outputs}} (t_k - \sigma(z_k))^2. \tag{A.3}$$

Since the outputs of all units  $k \neq j$  are independent of  $w_{ji}$ , we can drop the summation and consider just the contribution to  $E$  by  $j$ :

$$\begin{aligned}
 \delta_j &= \frac{\partial E}{\partial z_j} = \frac{\partial}{\partial z_j} \frac{1}{2} (t_j - o_j)^2 = -(t_j - o_j) \frac{\partial o_j}{\partial z_j} \\
 &= -(t_j - o_j) \frac{\partial}{\partial z_j} \sigma(z_j) \\
 &= -(t_j - o_j) (1 - \sigma(z_j)) \sigma(z_j) \\
 &= -(t_j - o_j) (1 - o_j) o_j,
 \end{aligned} \tag{A.4}$$

$$\Delta w_{ji} = -\eta \frac{\partial E}{\partial w_{ij}} = \eta \delta_j x_{ji}. \tag{A.5}$$

Now consider the case when  $j$  is a hidden unit. Like before, we make the following two important observations.

For each unit  $k$  Downstream from  $j$ ,  $z_k$  is a function of  $z_j$ .

The contribution of error by all units  $l \neq j$  in the same layer as  $j$  is independent of  $w_{ji}$ .

We want to calculate  $\partial E / \partial w_{ji}$  for each input weight  $w_{ji}$  for each hidden unit  $j$ . Note that  $w_{ji}$  influences just  $z_j$  which

influences  $o_j$  which influences  $z_k$  for all  $k \in \text{Downstream}$  each of which influence  $E$ . So we can write

$$\begin{aligned}
 \frac{\partial E}{\partial w_{ji}} &= \sum_{k \in \text{Downstream}(j)} \frac{\partial E}{\partial z_k} \cdot \frac{\partial z_k}{\partial o_j} \cdot \frac{\partial o_j}{\partial z_j} \cdot \frac{\partial z_j}{\partial w_{ji}} \\
 &= \sum_{k \in \text{Downstream}(j)} \frac{\partial E}{\partial z_k} \cdot \frac{\partial z_k}{\partial o_j} \cdot \frac{\partial o_j}{\partial z_j} \cdot x_{ji}.
 \end{aligned} \tag{A.6}$$

Again, note that all the terms except  $x_{ji}$  in the above product are the same regardless of which input weight of unit  $j$  we are trying to update. Like before, we denote this common quantity by  $\delta_j$ . Also note that  $\partial E / \partial z_k = \delta_k$ ,  $\partial z_k / \partial o_j = w_{kj}$ , and  $\partial o_j / \partial z_j = o_j(1 - o_j)$ . Substituting

$$\begin{aligned}
 \delta_j &= \sum_{k \in \text{Downstream}(j)} \frac{\partial E}{\partial z_k} \cdot \frac{\partial z_k}{\partial o_j} \cdot \frac{\partial o_j}{\partial z_j} \\
 &= \sum_{k \in \text{Downstream}(j)} \delta_k w_{kj} o_j (1 - o_j),
 \end{aligned} \tag{A.7}$$

thus

$$\delta_j = o_j (1 - o_j) \sum_{k \in \text{Downstream}(j)} \delta_k w_{kj}. \tag{A.8}$$

## Nomenclature

$X_j$ :	Input vector for unit $j$
$W_j$ :	Weight vector for unit $j$
$Z_j$ :	Weighted sum of inputs for unit $j$
$O_j$ :	Output of unit $j$
$T_j$ :	Laplace transform parameter
$E$ :	Calculated error for each unit
MAPE:	Mean absolute percentage error
$A_t$ :	Actual value
$F_t$ :	Predicted value by data-driven model.

## Highlights

Advanced artificial intelligence and data mining technique is used to develop data-driven models in order to generate synthetic geomechanical well logs.

Highly accurate results from data-driven models are achieved that are validated against blind wells that have actual file data in the Marcellus shale asset.

The geomechanical distributions created with field-wide data demonstrate much better consistency and improvement compared to using just partial field data.

## Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

## References

- [1] C. L. Cipolla, E. P. Lolon, J. C. Erdle, and B. Rubin, "Reservoir modeling in shale-gas reservoirs," *SPE Reservoir Evaluation & Engineering*, vol. 13, no. 4, pp. 638–653, 2010.

- [2] W. Yu and K. Sepehrnoori, "Simulation of gas desorption and geomechanics effects for unconventional gas reservoirs," in *Proceedings of the SPE Western Regional/Pacific Section AAPG Joint Technical Conference: Energy and the Environment Working Together for the Future*, pp. 718–732, Monterey, Calif, USA, April 2013.
- [3] S. Esmaili, A. Kalantari-Dahaghi, and S. D. Mohaghegh, "Forecasting, sensitivity and economic analysis of hydrocarbon production from shale plays using artificial intelligence & data mining," in *Proceedings of the Canadian Unconventional Resources Conference*, Calgary, Canada, October–November 2012, paper SPE 162700.
- [4] T. W. Patzek, F. Male, and M. Marder, "Gas production in the Barnett Shale obeys a simple scaling theory," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 110, no. 49, pp. 19731–19736, 2013.
- [5] U. Aybar, M. O. Eshkalak, K. Sepehrnoori, and T. W. Patzek, "The effect of natural fracture's closure on long-term gas production from unconventional resources," *Journal of Natural Gas Science and Engineering*, 2014.
- [6] M. O. Eshkalak, "Simulation study on the CO<sub>2</sub>-driven enhanced gas recovery with sequestration versus the re-fracturing treatment of horizontal wells in the U.S. unconventional shale reservoirs," *Journal of Natural Gas Science and Engineering*, 2014.
- [7] M. O. Eshkalak, U. Aybar, and K. Sepehrnoori, "An integrated reservoir model for unconventional resources, coupling pressure dependent phenomena," in *Proceedings of the the SPE Eastern Regional Meeting*, pp. 21–23, Charleston, WV, USA, October 2014, paper SPE 171008.
- [8] M. O. Eshkalak, U. Aybar, and K. Sepehrnoori, "An economic evaluation on the re-fracturing treatment of the U.S. shale gas resources," in *Proceedings of the SPE Eastern Regional Meeting*, Charleston, WV, USA, October 2014, paper SPE 171009.
- [9] M. O. Eshkalak, S. D. Mohaghegh, and S. Esmaili, "Synthetic, geomechanical logs for marcellus shale," in *Proceedings of the SPE Digital Energy Conference and Exhibition*, The Woodlands, Texas, USA, March 2013, paper SPE 163690.
- [10] M. Omidvar Eshkalak, *Synthetic geomechanical logs and distributions for marcellus shale [M.S. thesis]*, West Virginia University Libraries, West Virginia University, Morgantown, WV, USA, 2013.
- [11] L. Rolon, S. D. Mohaghegh, S. Ameri, R. Gaskari, and B. McDaniel, "Using artificial neural networks to generate synthetic well logs," *Journal of Natural Gas Science and Engineering*, vol. 1, no. 4-5, pp. 118–133, 2009.
- [12] S. Mohaghegh, R. Arefi, S. Ameri, K. Aminian, and R. Nutter, "Reservoir characterization with the aid of artificial neural network," *Journal of Petroleum Science and Engineering*, vol. 16, pp. 263–274, 1996.
- [13] S. Mohaghegh, M. Richardson, and S. Ameri, "Virtual magnetic imaging logs: generation of synthetic MRI logs from conventional well logs," in *Proceedings of the SPE Eastern Regional Meeting*, pp. 223–232, Pittsburgh, Pa, USA, November 1998.
- [14] I. A. Basheer and Y. M. Najjar, "A neural network for soil compaction," in *Proceedings of the 5th International Symposium on Numerical Models in Geomechanics*, G. N. Pande and S. Pietruszczak, Eds., pp. 435–440, Balkema, Rotterdam, The Netherlands, 1995.
- [15] A. J. Maren, C. T. Harston, and R. M. Pap, *Handbook of Neural Computation Applications*, Academic Press, San Diego, Calif, USA, 1990.
- [16] Y. Khazani and S. D. Mohaghegh, "Intelligent production modeling using full-field pattern recognition," *SPE Reservoir Evaluation & Engineering*, vol. 14, no. 6, pp. 735–749, 2011.



**Hindawi**

Submit your manuscripts at  
<http://www.hindawi.com>

