Garethtaylor12
ARMA 20–1917

Seeing-ahead-of-the-bit:
A game changer enabled by Machine Learning

Gareth Taylor, Barry Zhang, Muhlis Unaldi, Edward Tian, Eduardo Alvarez, Altay Sansal
Quantico Energy Solutions, Inc., Houston, Texas, USA

This paper was prepared for presentation at the 54th US Rock Mechanics/Geomechanics Symposium held in Golden, Colorado, USA, 28 June-1 July 2020. This paper was selected for presentation at the symposium by an ARMA Technical Program Committee based on a technical and critical review of the paper by a minimum of two technical reviewers. The material, as presented, does not necessarily reflect any position of ARMA, its officers, or members. Electronic reproduction, distribution, or storage of any part of this paper for commercial purposes without the written consent of ARMA is prohibited. Permission to reproduce in print is restricted to an abstract of not more than 200 words; illustrations may not be copied. The abstract must contain conspicuous acknowledgement of where and by whom the paper was presented.

ABSTRACT:

Numerous attempts have been made to assist drilling engineers with the assessment and mitigation of pre-drill uncertainties by providing "seeing-ahead-of-the-bit" capabilities while drilling. For several reasons, such attempts have generally failed to achieve the much sought-after panacea for reducing operational risk and non-productive time (NPT) while improving drilling efficiency at reduced operational costs.

We present an artificial intelligence (AI) based workflow driven by a proprietary machine learning algorithm to deliver a near real-time 3D earth model while the well is being drilled. The model can be populated with reservoir properties of interest. In addition to improving the accuracy of wellbore positioning, the process has clear implications for risk mitigation while drilling such as shallow water flow, shallow gas, anomalous high-pressured pockets, lost circulation, wellbore instability, stuck pipe, top/base target formation and salt/sediment proximity to delineate salt boundaries, etc.

1. INTRODUCTION:

Numerous attempts have been made to assist drilling engineers with the assessment and mitigation of pre-drill uncertainties by providing "seeing-ahead-of-the-bit" capabilities while drilling. These attempts range from calculating resistivity several meters ahead of the drill bit (Constable, et al., 2016) to "ahead-of-the-bit" pore pressure prediction using VSP while drilling (Xi et al., 2010), to the deployment of "seismic-while-drilling" technology for salt flank and fault plane imaging (Dethlof and Petersen, 2007). These solutions are either too complex, too costly, too slow, too inaccurate or of too low resolution to achieve the much sought-after drilling panacea for reducing operational risk and non-productive time (NPT), while improving drilling efficiency at reduced operational costs.

Numerous Artificial Intelligence systems have been deployed by the E&P industry over the past several decades. These include well log property prediction from seismic data (Hampson, et al., 2001); lithofacies prediction from well log data (Aliouane, et al., 2013); deep learning for identifying geologic features from seismic attributes (Huang,L., et al., 2017), seismic fault detection using CNNs (Pochet, A., 2018) and AI for seismic interpretation (Lowell, J., et al., 2019).

We present a workflow (QEarth) driven by a proprietary digitalization engine that utilizes a machine learning algorithm to deliver a near real-time 3D earth model while the well is being drilled.

The model may be populated with any property that is logged or measured in the wellbore; more specifically, this could include: Elastic Properties: e.g. DT, DTS, Vp/Vs, etc., Petrophysical Properties: e.g. Density, Gamma Ray, Neutron, etc., Geomechanical Properties: e.g. Pore Pressure, Principal Stresses, Young's Modulus, Poisson's Ratio, etc. and Drilling Properties: e.g. Wellbore Stability, UCS, etc.

The workflow includes the use of Logging While Drilling (LWD) data that is passed via WITSML to a Cloud-based computation environment where a locally calibrated machine learning model is updated and used to rebuild the entire 3D seismic volume, explicitly in the depth domain, for any of the properties listed above. In the example cited, the time to rebuild the earth model is approximately 30-minutes per 50ft. of new LWD data, although the precise time may vary with computational resources.

The first step in the process involves the generation of a pre-drill earth model populated with reservoir properties of interest. Inputs into the machine learning software include migrated seismic data (full stack and angle stacks) in the time or depth domain and well logs. In the second

step, LWD data is ingested in real-time via WITSML. In the third step, the machine learning algorithm autonomously updates the elastic, petrophysical geomechanical or drilling properties for the entire seismic volume, thus delivering ahead-of-the-bit predictions. This can translate into hundreds of feet or several days of advance time so drilling engineers and operational geoscientists can better anticipate potential drilling issues and optimize well delivery with the latest earth model information.

This entire process is demonstrated using a 98-square mile 3D seismic dataset recorded in the Midland Basin of West Texas in the US.

2. THE MIDLAND BASIN - GEOLOGY:

The Permian Basin of West Texas and Southeast New Mexico has generated hydrocarbons for about 100 years and supplied more than 33.4 billion barrels of oil and about 118 trillion cubic feet of natural gas as of September 2018. Implementing hydraulic fracturing, horizontal drilling and completion technology advancements during the past decade has reversed the production drop in the Permian, and the basin has exceeded its previous peak in the early 1970s. In 2017, it accounted for 20% of the total U.S. crude oil production and about 9% of the total U.S. dry natural gas production. For 2016, EIA estimates remaining proven reserves in the Permian Basin to exceed 5 billion barrels of oil and 19.1 trillion cubic feet (Tcf) of natural gas, making it one of the largest hydrocarbon-producing basins in the United States and the world (EIA, 2017).

2.1. Regional Tectonic Setting and Geologic Framework.

The Permian Basin is a complex sedimentary system located in the foreland of the Marathon–Ouachita orogenic belt. It covers an area of more than 75,000 square miles and extends across 52 counties in West Texas and Southeast New Mexico. The Permian Basin was developed in the open marine area known as the Tobosa Basin in the middle Carboniferous period, approximately 325 million–320 million years ago (Galley, 1958). The ancestral Tobosa Basin was formed by an asymmetric structural flexure in the Precambrian basement at the southern margin of the North American plate in late Proterozoic time (Beamont, 1981; Jordan 1981). During consequent phases of basin development, sediments eroded from the surrounding highlands and were deposited in the basin (Brown et al., 1973; Yang et al., 1995).

The Permian Basin is now an asymmetrical, northwest to southeast-trending sedimentary system bounded by the Marathon-Ouachita orogenic belt to the south, the Northwest shelf and Matador Arch to the north, the Diablo platform to the west and the Eastern shelf to the east (Gardiner, 1990; Ewing, 1991; Hills, 1985). The basin comprises several sub-basins and platforms: three main sub-divisions include the Delaware Basin, Central Basin Platform and the Midland Basin (EIA, 2018). Figure 1.

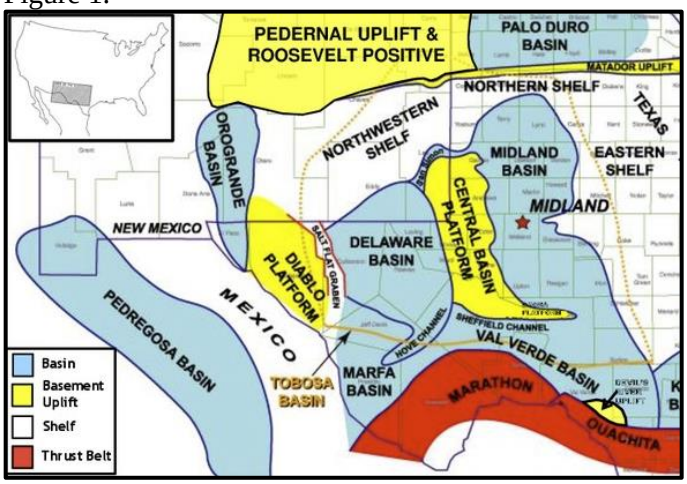

Fig. 1. Major structural and tectonic features in the Permian Basin (Pioneer Natural Resources).

The Midland Basin is bounded to the east by the Eastern shelf through a series of north-south trending fault segments and to the north by the Northwest shelf. Southward, Midland Basin formations thin out into the Ozona Arch, an extension of Central Basin Platform, which separates the Delaware and Midland Basins (Figure 1), (EIA, 2018).

2.2. Regional Stratigraphy

The age of sedimentary rocks underlying the Permian system in West Texas to Southeast New Mexico ranges from Precambrian to Pennsylvanian. Typically, the oldest rocks immediately underlie Permian rocks in uplift areas such as the Central Basin Platform and the Ozona Arch. Pennsylvanian rocks are common across the Delaware and Midland Basins and on the Northwestern and Eastern shelves.

Regional stratigraphic relationships for upper Carboniferous to upper Permian strata in the Permian Basin are shown on a generalized stratigraphic schema (Figure 2), (EIA, 2018).

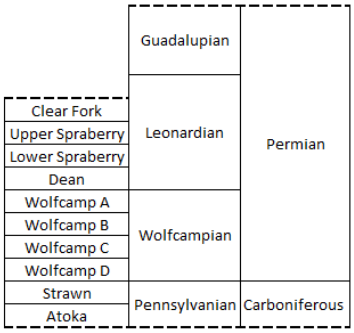

Fig. 2. Generalized stratigraphic schema of upper Carboniferous through upper Permian intervals for the Midland Basin (adapted from EIA).

Upper Pennsylvanian and Wolfcampian strata spread across the entire Permian Basin; however, the thickest accumulations are in the central and southern parts of the Delaware Basin. This stratigraphic interval quickly thins out to the Central Basin Platform in contrast with the more gradual thickness decrease toward the western part of the Delaware Basin and eastern part of the Midland Basin (Figure 3), (EIA, 2018).

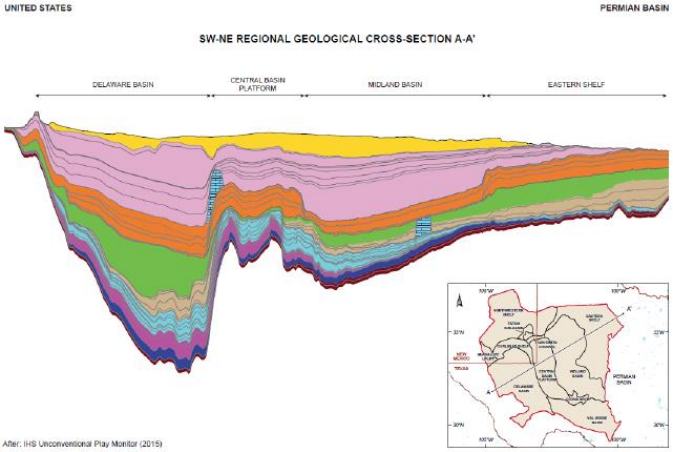

Fig.3. Greater Permian Basin cross section.

Upper Carboniferous Pennsylvanian rocks that range in thickness from 0 feet to 3,000 feet generally occur in the depth between 5,000 feet and 15,000 feet. Pennsylvanian formations including Atoka, Strawn and Cisco, predominantly consist of limestone, shale and minor quantities of sandstone and siltstone. An extensive development of reef facies accounts for a large percentage of the limestone deposits in shallow peripheral areas of the Delaware and Midland Basins (Dolton et al., 1979; Hills, 1984), (EIA, 2018).

3. DATA SCIENCE:

3.1. Methodology

In this study a shallow feed-forward neural network (FFNN) is used. The FFNN predicts rock properties at the reservoir scale from a network trained on seismic attributes as inputs and logged attributes as targets. The FFNN can update a 3D earth model and provide high-resolution predictions hundreds of feet from the drill-bit. This section will discuss the basic anatomy of the FFNN and the data science techniques used to train a set of models that can best generalize over unseen data.

3.1.1. Training Data Preparation

The training data are prepared for use by filtering noisy seismic and log data. Seismic angle stacks are checked for alignment and interpreted seismic horizons are inspected for interpretation "busts" and smoothed, if warranted.

The conditioned inputs to the FFNN are seismic amplitudes and, optionally, derived attributes (e.g. Hilbert Transform, AVO, inversion and seismic velocities, etc.). For each point in space, several seismic amplitude values are presented to the neural network, both above and below a given point of analysis for any given seismic sample. Well log data that have been previously tied to the seismic data are input as training targets. The final input provided to the FFNN is the relative position of the seismic trace sample with respect to interpreted horizons, as well as the horizon position itself. Optionally, the XY position of the trace and/or seismic inline/crossline position is supplied particularly if many wells are included in the training phase. Following selection of the input data, normalization scaling is performed using min-max or standardization scaling.

3.1.2. Model Architecture and Hyperparameters

The complexity requirements of a neural network are largely dependent on the size (number of samples) and shape (dimensions of the input vector) of the training data. In general, there is no analytical way to calculate the number of layers or nodes to use in network construction. These are just a few of the available options that are referred to as hyperparameters. Table 1 depicts the hyperparameters tested in this study.

Table 1. The first column shows the different hyperparameters tested in this study. The second column shows the different options provided for the automated grid search algorithm.

Hyperparameter	Grid search Options
Number of Hidden Layers	1, 2, or 3
Size of Hidden Layers	10, 20, or 30
Architecture Type	Sequential
Training Epochs	25, 50, 75, or 100
Optimizer	SCG, LM, BFG, and BR

Another key hyperparameter is the number of training epochs. An epoch refers to a single pass of the entire training dataset through the FFNN during training. If the model is trained for too few or too many epochs, the resulting network will not learn the system or overfit, respectively. An overfit model is one that is unable to make general predictions on unseen data because it has only memorized the training dataset.

Although not often considered a hyperparameter, the type of model architecture is another critical element of the FFNN. The decision between graphical or sequential model, for the purposes of this study, depends on whether the model is designed to target single or multiple outputs. Figure 3 is a sequential model with 10 input neurons and a single output neuron. The data flows through the model sequentially, from layer to layer, without any need to split in different directions. Figure 4 is a graphical model with 10 input neurons and two output neurons. What defines a graphical model is the split between the last hidden layer and the output layer. This allows the network to predict multiple outputs. In the near real-time portion of this study, a single-output sequential FFNN was used to generate a real-time acoustic impedance volume.

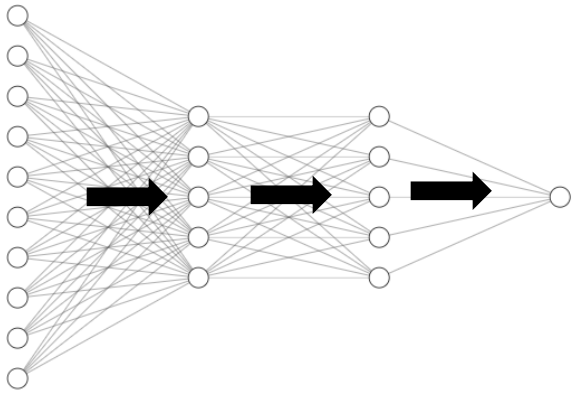

Fig. 3. Sequential Architecture

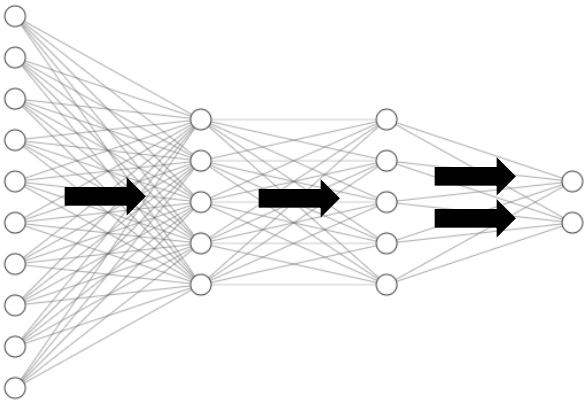

Fig. 4. Graphical Architecture

Backpropagation algorithms or "optimizers" are names used to describe a wide range of algorithms that update the weights and biases of hidden layers in the network. This process of optimizing weights and biases is often described as the learning phase during training. The optimizer works by computing the gradient of the loss function with respect to the weights found at each layer by the chain rule. In simpler terms, this algorithm helps the model navigate the multi-dimensional decision space in an attempt to locate the global-minimum.

The learning algorithms tested in this study are Scaled Conjugate Gradient (SCG), Broyden Fletcher Goldfarb–Shanno (BFG), Levenberg-Marquardt (LM) and Bayes Regularized (BR).

3.1.3. Hyperparameter Optimization

The search for an optimal combination of hyperparameters is a process called hyperparameter optimization. In this study, a combination of systematic experimentation and automated grid search was used to construct a generalized FFNN.

The automated grid search begins with the construction of a hyperparameter grid followed by an iterative process of model testing. The process is illustrated in Figure 5 using a four part process diagram. Part 1 automatically selects a permutation from the grid and uses those parameters to build a FFNN in part 2. In part 3, the model is trained. The resulting model is then tested on unseen data and ranked, with the other resulting models, by MSE. This process continues until all permutations in the hyperparameter grid have been exhausted.

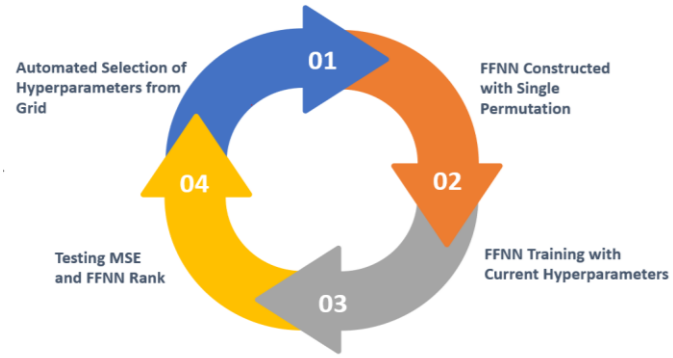

Fig. 5. Process diagram illustrating the iterative process of hyperparameter of optimization through automated grid search.

3.1.4. Training

The hyperparameters from the highest ranking model (lowest MSE) were used to systematically train 20 FFNNs. For each training batch, a random number generator (RNG) selected a new set of initial weights and withheld a different 25% of the dataset for validation. At a high level, the training process for a FFNN, involves the following steps:

1. Forward pass of training and validation data through the network
2. Calculation of error/loss function after the inference/simulation is complete.
3. Backpropagation to update weights and biases.

Steps 1-3 are executed once per batch. Batches are subsets of the training data that are passed through the model. The batch size determines the number of iterations of steps 1-3 required to complete an entire epoch. Once the hyperparameter epoch setting is reached or the process is early stopped (to prevent network overfitting or underfitting), the training for a single FFNN is complete. The resulting models are then fed into a process called genetic algorithm-based ensemble selection.

3.1.5. Automated Ensemble Selection

Genetic Algorithm (GA) Optimization is implemented to define the optimal combination of FFNNs from the suite of models that are generated by the training process. As the name implies, the GA optimization search process is inspired by Darwinian evolution; looking at how chromosomal combinations change over genetic generations. The GA optimization search is not perfect since it may not necessarily find the objectively "best" ensemble. However, it is very efficient in traversing the decision space and returns a good result in a reasonably short time.

3.1.6. Simulation

The optimal set of trained models, as determined by the GA, are then used to infer rock properties over an entire volume, from the input seismic attributes. This process involves feeding in the seismic attributes as inputs to the FFNNs, ensembling the results of each FFNN in the GA ensemble, and outputting rock properties at every inline/x-line associated with an input seismic attribute.

4. STUDY EARTH MODEL GENERATION:

Conventional earth models that employ seismic data may be created using a seismic inversion process which requires appropriately conditioned seismic data (e.g. angle stacks or range-limited offset stacks) and processed well logs. The process is often time-consuming and requires expertise to generate and interpret. Consequently, conventional seismic inversion-based earth model usage is somewhat uncommon in the subsurface analysis of unconventional resource plays where elastic property variability is often subtle. By contrast, machine learning technologies offer a completely different approach to seismic and well log-derived earth models that are significantly more efficient and often of higher resolution. In addition, results are achieved in weeks rather than months.

Since machine learning techniques can "learn" from various types of data, the input data set is significantly more flexible than a conventional seismic inversion. It can include well logs, seismic reflectivity volumes (both prestack and poststack), derivative attributes, geologic maps, isochrons, isochores, production data and other geophysical data, e.g. EM, CSEM, gravity/magnetics, resonance, etc.

More specifically, the QEarth process used in this study commenced with data preparation and network architecture selection. Input variables were next selected; these included poststack seismic data, interpreted horizons and six acoustic impedance well log curves (a seventh well was retained as a "blind well" for cross validation purposes). Next, missing or bad data in the well logs was addressed. The candidate inputs were then scaled and/or normalized using the methods described in 3.1.1.

Training and validation data subsets were then built. Randomly withheld subsets (25%) were used for validation to reduce overfitting and for model generalization. A shallow feed-forward ANN with two hidden layers was chosen, followed by modification of the network architecture (e.g. number of neurons and activation functions). Back propagation was employed to optimize the ANN weights and biases using the Levenberg-Marquardt optimizer. Sensitivity and performance analysis were conducted next; involving error analysis of multiple realizations (candidate networks) whereby poorly performing networks and realizations were discarded. At that point, ensemble networks were automatically generated for a set of 20 most-likely outputs by model aggregation methods.

Final simulation for a time step was achieved using individual networks or ensembles and tested on target seismic inlines and crosslines for efficiency purposes. Confidence metrics were produced that were associated with the direct simulation of, in this study, Acoustic Impedance and Unconfined Compressive Strength (UCS). However additional elastic, petrophysical, geomechanical and drilling properties can be simulated using the QEarth methodology (Table 2).

Table 2. Properties that may be simulated in near real-time.

Elastic Properties	Petrophysical Properties	Geomechanical Properties	Drilling Properties
Velocities	Lithology	Pore Pressure	Mud weight
Impedances	Porosity	Stresses	WBS
Derivatives (LMR, etc.)	Density	Young's Modulus	Fracture Gradient
	S_w	Poisson's Ratio	MSE/UCS
	Gamma Ray		
	Resistivity		

Finally, post-processing tasks were performed (e.g. structure-oriented filtering, scaling, clipping, etc.) and the simulated property data were compared with the input data (Figure 6).

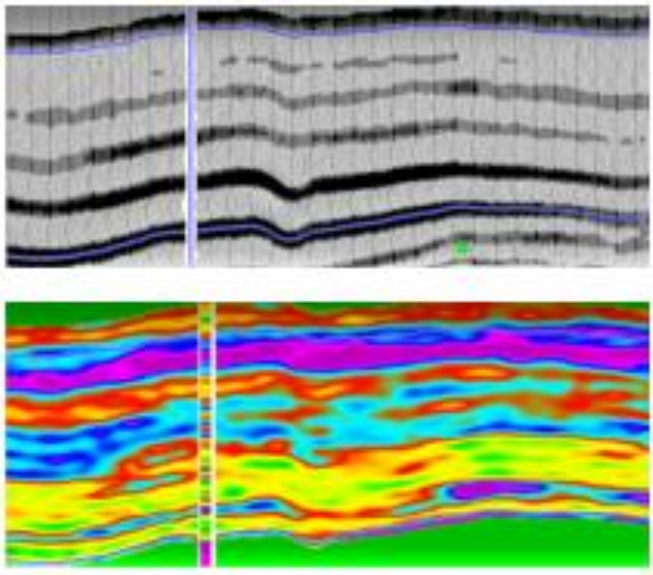

Fig. 6. Comparison between input seismic amplitude data and a QEarth high-resolution, unconfined compressive strength (UCS) property volume. Also shown is inserted log data from a "blind" test well not used in the training.

4. RESULTS:

The images below (Figure 7a, 7b, 7c) are movie frames showing the progressive simulated update of an acoustic impedance earth model and acoustic impedance well log.

Note this could be any logged or measured property as shown in Table 2.

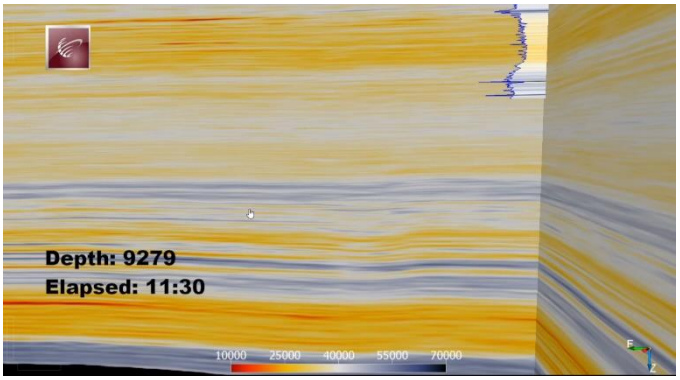

Fig. 7a. QEarth snapshot at 9,279ft. after 11 hours 30 minutes of drilling.

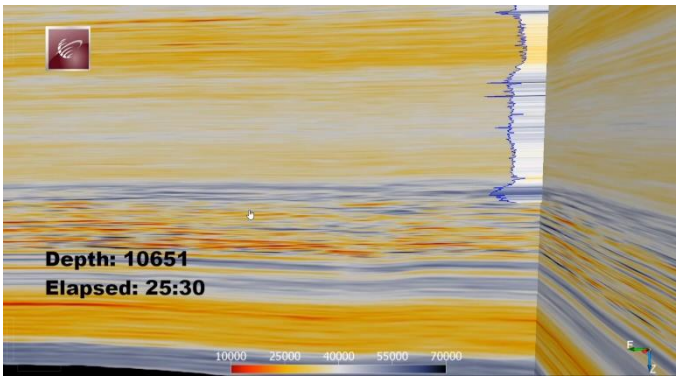

Fig. 7b. QEarth snapshot at 10,651ft. after 25 hours 30 minutes of drilling.

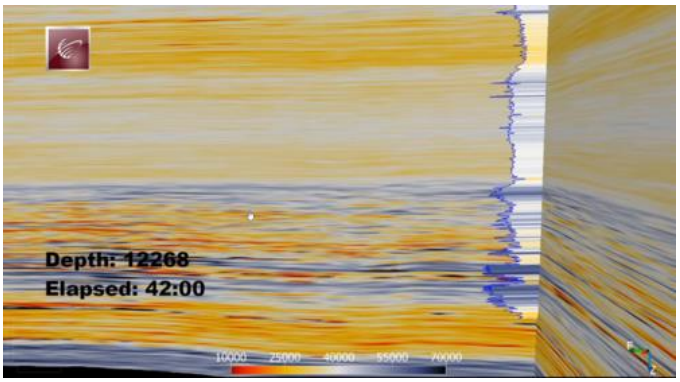

Fig. 7c . QEarth snapshot at 12,268ft. after 42 hours of drilling.

The 3D seismic volume shown in Figure 7a, 7b, 7c is 14 x 7 miles (98 sq. mi. or 259 km^2) in area and the vertical span is 3,000ft (914m). The distance from the wellbore location to the left-hand edge of the seismic inline is approximately 1 mile (1,609m). As discussed, six offset wells were used to train QEarth to map Acoustic Impedance across the entire volume with a seventh well withheld as a blind test (depicted in Figures 7a, 7b, 7c). The nearest well is 1.5 miles (2.4km) away from the well that is being "drilled" in the snapshots. Depth (in feet) and time (in hours/mins) are annotated.

The well log was segmented into 50ft (15.2m) intervals and was presented to the machine learning algorithm as a simulated LWD feed. For the avoidance of doubt, this is a computer simulation of a real time WITSML feed. As new well log information was received, the ANN was re-trained on-the-fly. It took approximately 25 minutes to retrain the entire 98 sq. mi. (259 km^2) volume and 5 minutes to produce an updated inline and crossline. The mouse (hand symbol) is located at the top of the Wolfcamp formation.

Within the shallower Clearfork and Spraberry formations, a stable prediction of acoustic impedance has been reached. However, as new acoustic impedance log data is received from the more complex Wolfcamp sequence, the acoustic impedance model is revised autonomously and in near real-time. Important updates can be discerned from the retrained and resimulated earth model after each 30-minute sequence. Properties ahead of the drill bit are also updated as a by-product of revising the entire seismic volume.

5. CONCLUSION:

QEarth represents a complete step change in the manner that real-time data is presented to the drilling engineer and operational geoscientist. After the initial pre-drill earth model is established, the automation of machine learning re-training and simulation delivers the latest information at a fraction of the cost and cycle time of traditional, manually updated workflows.

The near real-time Earth model provides the most accurate assessment of "ahead-of-the bit" geologic structure so that well delivery teams can more precisely position the well in the desired landing location. In addition, by providing drilling engineers and geoscientists with up to thousands of feet of early awareness, they are better able to mitigate potential drilling hazards. These include shallow water flow, shallow gas, anomalous high-pressured pockets, lost circulation, wellbore instability, stuck pipe, top/base target formation and salt/sediment proximity to delineate salt boundaries, etc.

In additional to the clear HSE benefits identified above, the potential financial impact of not deploying "ahead-of-the bit" technology could be dizzying when failing to recognize or adapt to challenges that may be result in increasing operational risk and non-productive time (NPT).

A machine learning-driven process that has a potentially profound impact on the way wells are drilled represents yet another compelling use case that advocates for the faster adoption of Artificial Intelligence by the oil and gas E&P industry. Manifestly, it is no longer just an interesting thing to do, it is an absolute imperative!

6. ACKNOWLEDGEMENT:

We would like to thank Geophysical Pursuit, Inc. for permission to show the seismic data.

7. REFERENCES:

1. Constable, M.V. et. al., 2016, SPWLA: Petrophysics, Vol.57, Issue:05
2. Xie, Y.H. et. al., 2010, Ahead of the Bit Pore Pressure Prediction using VSP – a Case Study in South China Sea. SPE Conference Paper.
3. Dethloff, M.H. and Petersen, S.A., 2007, Seismic-While-Drilling Operation and Applications. SPE Annual Technical Conference and Exhibition.
4. Hampson, D, Schuelke, J.S. Quirein, J.A., 2011, Use of multiattribute transforms to predict log properties from seismic data. GEOPHYSICS, VOL. 66, NO. 1.
5. Aliouane, L., Ouadfeul,S.A. , Noureddine Djarfour, N. and N.Boudella, A. Lithofacies Prediction from Well Logs Data using Different Neural Network Models. Proceedings of the 2nd International Conference on Pattern Recognition Applications and Method
6. Hung, L., Dong, X., Clee, E., 2017, A scalable deep learning platform for identifying geologic features from seismic attributes: The Leading Edge,**36**,249–256.
7. Pochet, A., Diniz, P., Lopes, H.B., Gattass, M., 2018, Seismic fault detection using convolutional neural networks trained on synthetic poststacked amplitude maps: IEEE Geoscience and Remote Sensing Letters,**16**,352–356
8. Lowell, J., Erdogan, V. Combining artificial intelligence with human reasoning for seismic interpretation. 2019, SEG International Exposition and Annual Meeting.
9. Brown, L. F., J. R. , Cleaves, A. W., Erxleben , A. W. , 1973, Pennsylvanian depositional systems in North-Central Texas, a guide for interpreting terrigenous elastic facies in a cratonic basin: Austin, The University of Texas at, Austin, Bureau of Economic Geology, Guidebook 14, 122 p.
10. Yang, Kenn-Ming, Dorobek, Steven, 1995, The Permian basin of West Texas and New Mexico: Tectonic history of a "composite" foreland basin and its effects of stratigraphic development, p. 149-174, 10.2110/pec.95.52.0149.
11. Dolton, G.L., Coury, A.B., Frezon, S.E., Robinson, K., Varnes, K.L., Vunder, J.M., and Allen, R.V., 1979, Estimates of undiscovered oil and gas, Permian basin, Vest Texas and Southeast New Mexico: U.S. Geological Survey Open-File Report 79-838, 72 p.
12. Hills, J. M., 1984, Sedimentation, tectonism, and hydrocarbon generation in Delaware basin, west Texas and southeastern New Mexico: American Association of Petroleum Geologists Bulletin, v. 68. p. 250- 267.
13. Bengio, Y., Deep Learning of Representations for Unsupervised and Transfer Learning, JMLR: Workshop and Conference Proceedings 27:17–37
14. Bishop, C.M. (2006), Pattern recognition and Machine Learning, Springer.
15. Permian Basin, Wolfcamp Shale Play, Geology Review, October 2018, US EIA.