Real Time Pore Pressure Calculation from Drilling Mechanics Data via Machine Learning Techniques

Authors: Matthew Reilly[1], John Thurmond[1], Koda Chovanetz[1], Orlando de Jesus[2], Muhlis Unaldi[2], Gareth Taylor[2]. Hess[1] Quantico Energy Solutions[2]

Abstract

A method is proposed to calculate pore pressure at the bit while drilling using all data typically available in a modern drilling rig. This method utilizes a machine learning approach that can estimate pore pressures at the same or a narrower range of uncertainty as traditional methods and can do so at the bit in real-time. Traditional pore pressure estimation while drilling utilizes a combination of data sources most of which are detected from sensors placed 100's of feet behind the drill bit (where resistivity, sonic, density etc. tools are commonly placed). Furthermore, smoothing algorithms are often used to average the detection data thus increasing the offset from the drill bit to the estimated pore pressure calculation. The result of this is that the pore pressure calculation while drilling is only relevant to the formation that has already been penetrated and not being actively drilled. In hole sections where minor pore pressure changes can have significant impact on operational decisions this has obvious disadvantages. However, while drilling a well, multiple sources of sensory data are being received from the drill bit itself and yet are typically left unused in pore pressure calculation. Whereas traditional methods give an estimate of pore pressure after the well has already experienced a change in pressure, this method can calculate pore pressure at the bit, as the change is experienced. Another benefit of applying a machine learning model to pore pressure calculation while drilling is that the computational time is almost instantaneous.

Introduction

'Traditional' approaches to pore pressure estimation from subsurface logging tools are primarily based on the concept of porosity preservation with depth which compares measured porosity with a 'normal compaction trend' which assumes hydrostatic conditions. There are a variety of methods to conduct 'traditional' log data to pore pressure translations but, it could be argued that most of these originate from Eaton's 1975 Paper "The equation for geopressure prediction from well logs" (Eaton 1975). In this paper Eaton proposed an equation for calculating pore pressure from resistivity, conductivity, sonic velocity and the D-exponent. These empirically-derived equations are based on a similar theoretical basis; that overburden stress is equal to the matrix stress in addition to the pore pressure and uses the concept of a normal compaction trend to calculate pore pressure. Compaction of sediments due to overburden stress was described by Terzaghi and Peck in 1948 and demonstrated that increased fluid pressure reduces the 'effective stress' on the matrix thus reducing compaction. Therefore, increased porosity with depth can be taken as a proxy for undercompaction related to fluid retention and thus overpressure (Athy 1930, Terzaghi 1948, Hubert and Rubbey 1959, Eaton 1975). In order to calculate pore pressure from subsurface logging tools a 'normal compaction trend' must be defined in order to reference the calculation to a fixed fluid gradient. This trend line is usually defined by a power law, or exponential equation based either off log trends from known normal pressure intervals and measurements, or by back calculating from a plot of known fluid pressure gradients vs. observed log values. The definition of a normal compaction line is subjective as measured data rarely forms a tight trend with a very high R^2 value. In effect the normal compaction trend is a geological interpretation subject to data availability, quality and experience of the analyst.

There are several other mechanisms of pore pressure generation other than compaction disequilibrium, clay diagenesis, unloading, etc., however, the detection of pore pressure while drilling is still primarily practiced via the use of subsurface logging tools and gas data. Subsurface logging sensors (Sonic, Resistivity, Density etc.) can be used as a proxy for porosity measurements and are collected while drilling using Logging While Drilling (LWD) tools. A bottomhole assembly (BHA) can be configured in a multitude of ways but the LWD tools are always some distance behind the bit itself and typically placed 100-200ft behind the bit.

The use of drilling data to detect pore pressure is rooted in the fundamental assessment that drilling a well disturbs the stress state of the subsurface. As pore fluid pressure is an innate component of the stress state in the rock it is therefore a determining factor in the energy required to break the rock. The obvious advantage of drilling-based pore pressure is that the calculation is based on data at the bit; as opposed to being offset by LWD placement on the BHA. In effect, the drill bit becomes the sensor.

The effect of drilling responses to pore pressure is not a new technique; Jorden and Shirley first proposed this in 1966 and has been commonly known as the 'D-exponent' (Jorden & Shirley 1966). The 'D-exponent' method relies on the concept that an increased drilling rate is related to drilling higher porosity medium to softer shales which are in turn assumed to be related to overpressure (Equation 1). The 'D-exponent' has since been modified by numerous authors for use in a variety of environments. Fundamentally the D-exponent is not relevant to modern drilling practices as Rate of Penetration (ROP) is often controlled (and thus not a variable result) and modern drill bits, notably PDC bits, cut the rock through applied torque as opposed to crushing the rock in compression through the Weight on Bit (WOB). Majidi et al (2016) significantly advanced the concept of using drilling parameters to detect pore pressure using Mechanical Specific Energy (MSE). MSE is based on a physical model which incorporates torque and WOB to calculate the energy required to drill the rock (Equation 2, Majidi et al 2016). Calculating pore pressure from MSE is done via the use of a 'drilling efficiency' component which relates Confined Compressive Strength (CCS)to MSE. Pore pressure is then calculated in a similar way to traditional methods (e.g., Eaton) via the change in drilling efficiency compared to a 'Normal' drilling efficiency trend line (Equation 3, DE trend) which is assumed via a power law relationship, similar to that of Athy's porosity decline with depth (Majidi at al 2016, Athy 1930). What defines the 'normal' drilling efficiency trend is, of course, subject to analog data availability and potentially fallible human decision making.

$$D\ exponent = \frac{\log\left(ROP/60RPM\right)}{\log\left((12 * WOB)/(10\text{^}6\ D)\right)}$$

Equation 1

Where ROP is rate of penetration, RPM is revolutions per minute, WOB is weight on bit and D is the Diameter of the drillbit.

$$MSE = \frac{480 * Tq * RPM}{D_{bit}^2 * ROP} + \frac{4 * WOB}{\pi D_{bit}^2}$$

Equation 2

Where MSE is mechanical specific energy, Tq is torque, ROP is rate of penetration, RPM is revolutions per minute WOB is weight on bit and D is the diameter of the drillbit

$$Pp = P_n + \Delta DE * MSE * \frac{1 - Sin\ \theta}{1 + Sin\theta}$$

Equation 3a

Where Pp is pore pressure, P_n is normal hydrostatic pressure, DE is drilling efficiency, and θ is the angle of internal friction

Where:

$$\Delta DE = DE_p - DE_{trend}$$

Equation 3b

Where DE_p is a pseudo drilling efficiency and DE_{trend} is a manually defined trend in drilling efficiency.

The described method herein uses machine learning techniques which allow incorporation of all data available while drilling and does not require a 'Normal' trend to be pre-defined.

Pore pressure data is, however, somewhat unique in terms of machine learning datasets. Several previous authors have used machine learning engines to create pseudo logs from wells where a particular logging tool was not run (Gan, et. al., 2019) this has been termed 'Log to Log transformation'. However, applying the same methodology to pore pressure curves is not as clear-cut because the pore pressure curve used to train the model is an interpretation of an amalgam of datasets and not an exact quantitative measurement. A standard procedure to test the validity of any machine learning method is to compare the output from the model to the measured parameter. However, in the case of pore pressure, the output is being compared to an interpretation which may not be 100% correct as it is not a definitive physical measurement like the aforementioned 'Log to Log transformation'.

One of the great challenges in pore pressure interpretation and prediction is knowing precisely what are the mudrock formation pressures. While the majority of the formations drilled are mudrocks, quantitative pore pressure measurements are only taken in permeable formations (sands, etc.). Attempts have been made to use downhole pressure sampling tools in shales (Flemings et al 2008) but this is not common practice. Even if every sand in a wellbore is tested for formation pressure the pressure in the shale is still an interpretation.

The traditional approach to pore pressure interpretations is to calculate pore pressure in mudrocks using empirically or lab-derived normal compaction trends (Eaton 1975) and then use the quantitative measurements taken in sands to calibrate the resultant pore pressure. However, it should also be noted that sands and mudrocks are often not in equilibrium due to lateral transfer or other effects and this can be a source of error in the pore pressure interpretation (Reilly & Flemings 2010, Flemings et al 2002). Wellbore stability events can also be used to help calibrate the interpretation in mudrocks (wellbore break out, cuttings analysis, etc.) but this is often imprecise and, at best, provides a window of possible pore pressures over a depth interval.

By using machine learning techniques this method allows incorporation of all data available while drilling the well including those that are not part of the above *equations 1 & 2* (inclination, standpipe pressure, vibration etc.) and does not require a 'Normal' trend to be pre-defined. Furthermore, additional 'soft data' types can be incorporated such as seismic horizons and biostratigraphy. Machine learning techniques can also be trained to detect the differences between sands and shales and provide estimates of pore pressure in sands which is not possible in traditional techniques from either log or drilling based data. When appropriately trained and QC'd this method can predict pore pressure with the same or greater accuracy as previous techniques while doing so at the drill bit and almost instantly.

Methodology & QLog Workflow

This method utilizes an Artificial Neural Network (ANN) to generate pore pressure from a collection of drilling mechanics measurements including, but not limited to, Rate of Penetration (ROP), Torque (Tq), Revolutions per Minute (RPM), Bit Size (BIT) Weight-on-Bit (WOB), wellbore survey (Azim & Inc), At-Bit Logging tools (GR & Res) and can also include seismic horizons or real-time biostratigraphy. For this study, multiple models with different inputs were tested but the emphasis was on data transmitted while drilling and sensed as close to the bit face as possible.

This method has four main components: data conditioning, feature engineering, training optimization and operation.

During data conditioning, drilling data, rock properties, well surveys, logs, pre-drill tops (from seismic or logs) and calculated pore pressure values, from offset wells are processed to create a central database with data for each well. Time-based data (generally from drilling) is processed to be converted to depth-based measurements. Calculated variables like Mechanical Specific Energy (MSE) and VShale are calculated in this stage.

The next stage is feature engineering which defines a sliding window of input data for each target value and associated scaling of inputs and targets. Hyperparameter optimization is used to find an optimized model configuration. Generally, the principal neural network parameters that are typically optimized include the number of layers, number of neurons, activation functions, use of dropout layers and learning rate.

Model training optimization is based on selected inputs and targets using the results from hyperparameter optimization to create optimum models. An ensemble of different models is created to allow their combined result to better generalize to unseen data. It does this by reducing variance, hence eliminating the overfitting of model. Genetic optimization is then applied where a testing well is used to find the best ensemble of models that will eventually be able to better predict the blind well. Based on the ensemble generated in the previous stage, results are applied to interpret the veracity of the model's predictions. In the training phase, a model is built using selected data from a group of offset wells in the target area. The training model embodies the relationship between input and output data to the model; the input data to the model are the drilling data, together with the Gamma ray and Resistivity logs, Tops and the survey data; the output data are the pressure curves for the training wells. The training model output should match the pore pressure curves to be predicted, if this is not the case, further iteration via optimization and quality control must be done. An example of a training well and model creation is shown in Figure 1. This model (green line) has an excellent match to the training data (dark blue line) and thus is ready to progress to the next stage.

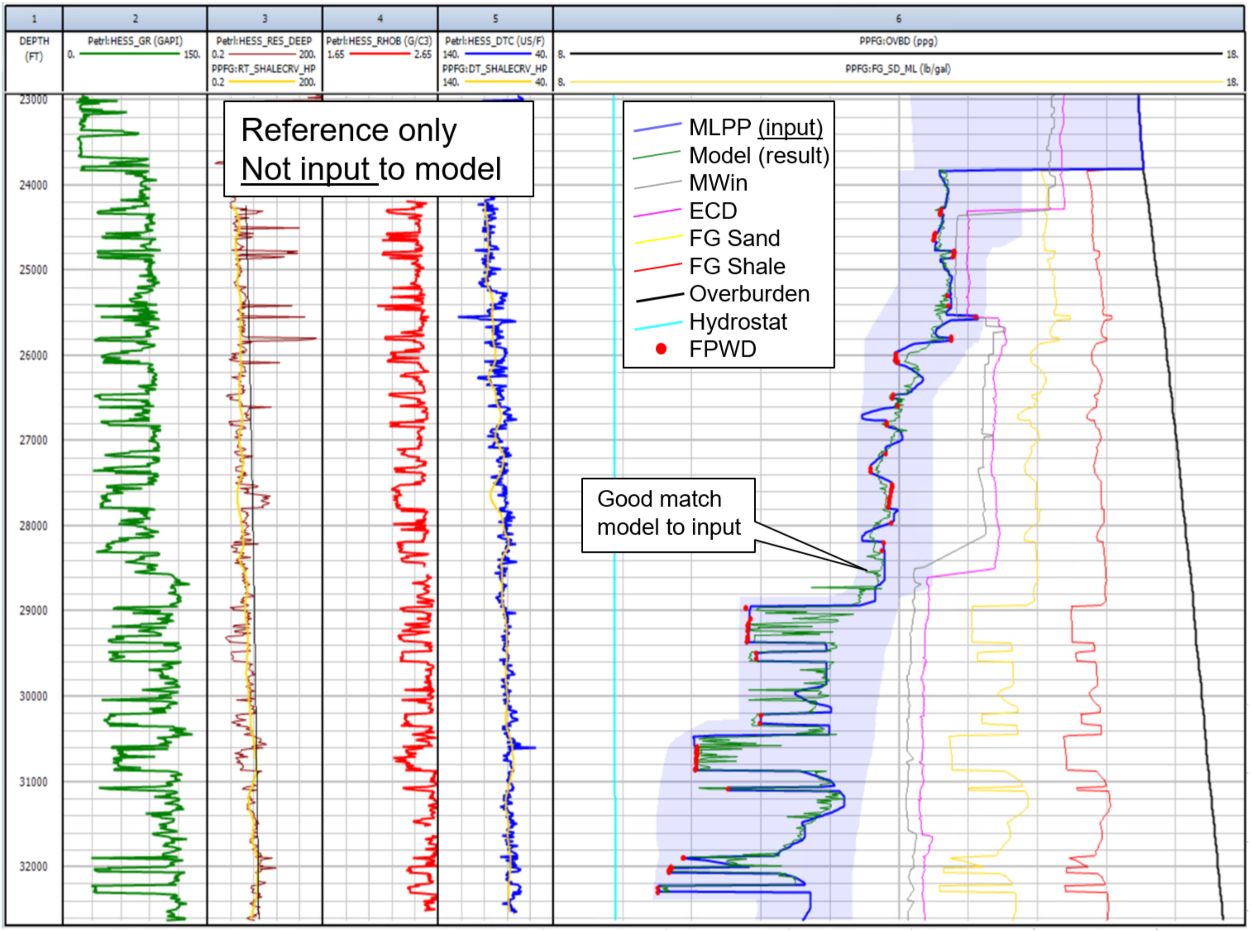

Figure 1: Example of training model results compared to input data. This well has high confidence in the pore pressure interpretation as multiple Formation Pressure While Drilling (FPWD) measurements were taken throughout the well. Resistivity, Density and Sonic curves are shown for reference but were not included in the trained model result (green curve in track 6). The Model result is an excellent match to the MLPP (Most Likely Pore Pressure) interpretation and so this can be considered appropriate to be used in blind well simulation. Also shown are MWin – Mudweight in, ECD – Equivalent Circulating Density of the mud, FG Sand – Fracture Gradient for Sand dominated lithologies, FG Shale -Fracture Gradient for mudrock lithologies, Overburden and Hydrostatic gradients.

The final step in the process is the operation or simulation phase. Here the pre-existing or 'trained' model is used to generate the output logs (pore pressure) for the test well from the input logs (gamma ray, resistivity and drilling data). The test well is withheld from the training data and commonly referred to as a blind well. The model is given only the inputs and calculates the output (pore pressure). The Training and Simulation phases can be somewhat iterative; if the blind result is not satisfactory, modifications can be made to the training optimization process and, if necessary, the input data. Once a satisfactory 'blind' result is received a further test on a second, or more, well(s) is suggested. If the model is performing satisfactorily, the training and blind wells can be rotated to test the influence of different wells on the result. Feature performance evaluation can be applied to determine relative importance of input datasets and to ensure the model is primarily focused on the input data deemed to have the most physical relevance (e.g., Torque instead of Bio Strat) and not taking shortcuts by learning to 'draw' a curve between tops or at given depth values.

In some cases, the output from the blind wells can provide a result that forces the user to challenge the initial assumptions or interpretation of the input data. An example of this can be seen at 30,000ft MD in Figure 3 where the post-drill assessment of the formation pressure was lower than the output from the machine learning model but also, after re-checking, from the traditional acoustic velocity to pore pressure translation. Adjustments from the model to the input data can, as in this example, be updated in the training model. Therefore, it is critical that the output of these models is verified by a subject matter expert before accepting or rejecting the final results.

Field Use

This method has been applied as a case study to the Stampede field in the Gulf of Mexico (GoM). Stampede is a subsalt Miocene oil field located in the Green Canyon protraction area (Figure 2) in the deep water GoM with reservoir depths of 28-30,000ft TVDss. The wells exit salt at roughly 21,000ft and encounter overpressured sands and shales. From ~26,000 ft, sands have significantly regressed pressures when compared to the surrounding shale sequences. 14 Wells have been drilled in the Stampede structure including several sidetracks (STs). A total of 4 exploration wells were drilled by Hess and Unocal in 2005. The field was unitized and sanctioned in 2014 and 7 production and 3 water injection wells were drilled. These development wells were drilled between 2015 and 2019 and operated by Hess. The wells were drilled using the Black Lion and Black Rhino drill ships and the tool provider was largely consistent. For this reason, the creation of the neural network utilizes the 10 production and injection wells to train the models and 2 blind wells were withheld to test the method. As previously described the blind wells were rotated to test model validity and influence of certain wells in the model.

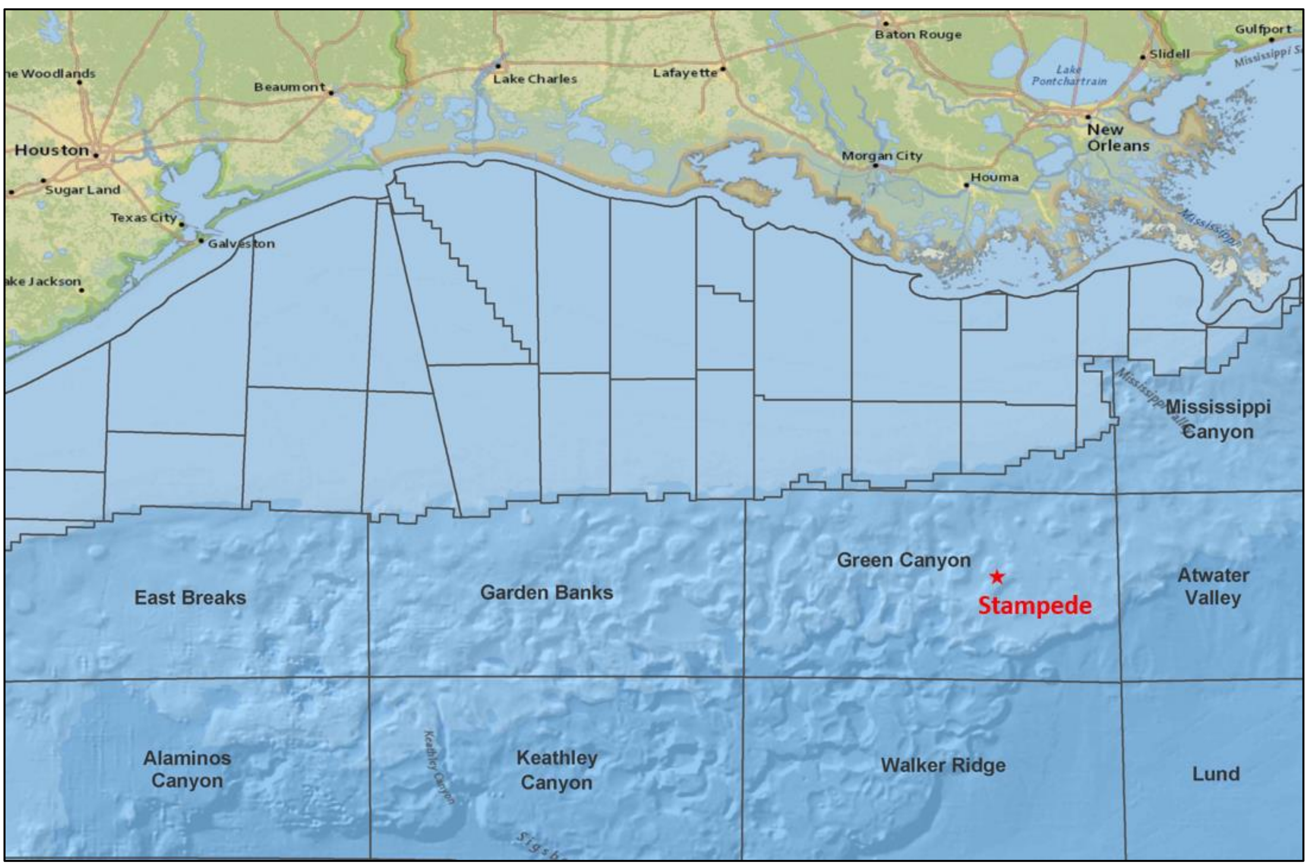

Figure 2 – Location Map of Stampede, Gulf of Mexico, USA

The result from the machine learning derived pore pressure shows good agreement with the post drill assessment (PDA) of pore pressure (Figure 3). In the upper section of the wells (21,000 – 27,000ft) where sands are interpreted to exhibit similar pressures to the shales, the output from the machine learning model has consistently, through multiple iterations of the model with varying input datasets, matched the PDA. The regressed section, from 27,000ft and deeper, required a more detailed investigation of the input data and feature performance. To match the regressed sand pressures a lithology classification tool was required, in this example the At Bit Gamma Ray (ABGR) tool was selected due to its proximity to the drill bit. The output of the machine learning models that are shown throughout this paper are not smoothed or filtered; they are the direct results of the calculation at each sample point. Smoothing is common practice in traditional log-based pore pressure estimation to be able to see trends more clearly and to eliminate small lithological variations in the formations. Examples of unsmoothed or 'raw' resistivity and acoustic models can be seen in figure 4 tracks 3&4.

Figure 3 shows the output (Model 3) from the machine learning model for the SA002 well and compared to the PDA assessment of pore pressure.

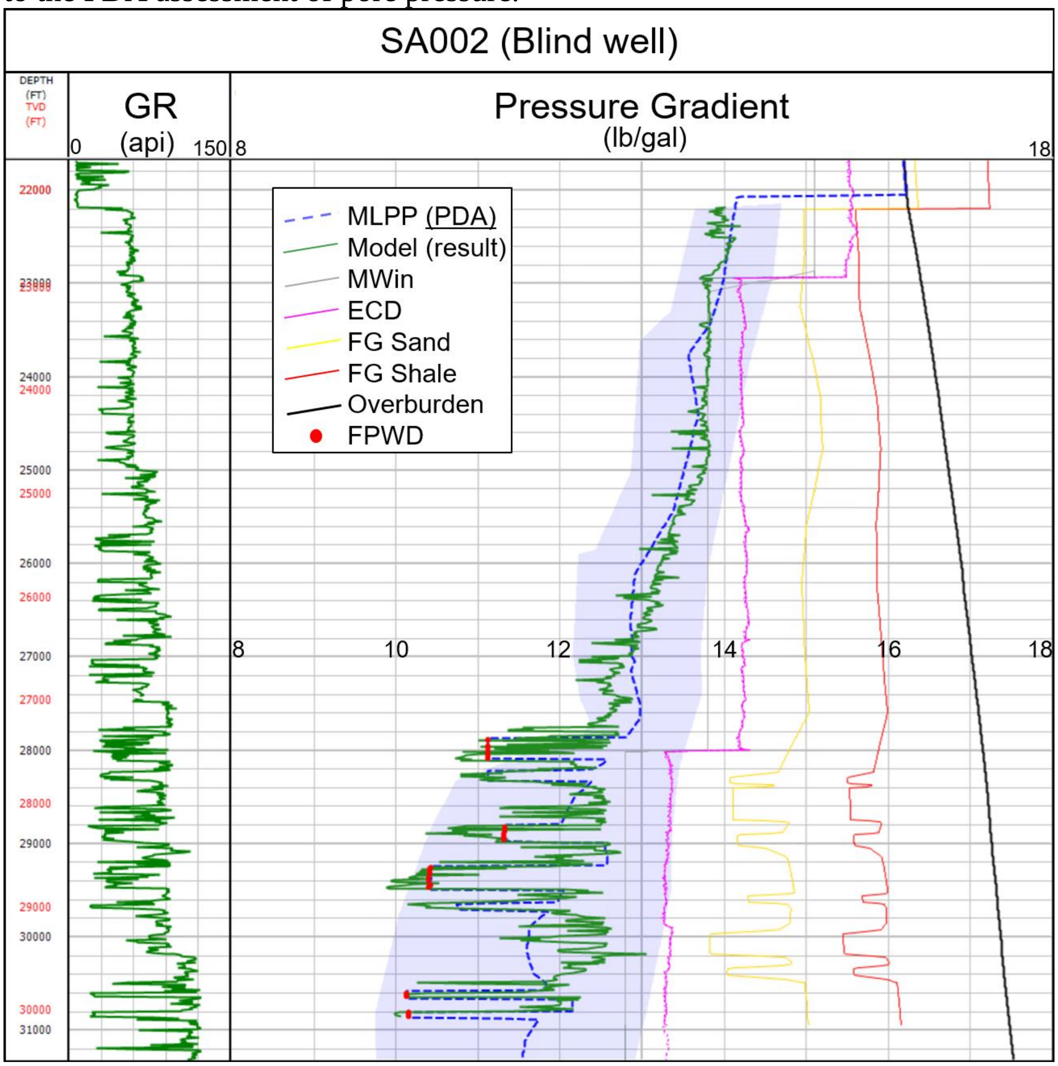

Figure 3 – Result of the Machine learning model on a well that was held blind (not included in the training dataset). The model output (solid green) which also used the ABGR as a lithology discriminator input is overlain on the post drill assessment (PDA) of pore pressure. The processed Gamma Ray is included in the plot for reference only but was not used in the calculation. Model output is in good agreement with the PDA curves and closely matches the measured MDT pressures in the sands. Around 30,000ftMD the model over-predicted the pressure vs. the PDA however on closer inspection of the pressure interpretation, both sonic and resistivity-derived pore pressures also indicate a higher pressure than the PDA interpretation. These higher than previously interpreted values initiated a re-examination of the previous assessment, and most data would suggest changing the PDA to be closer to the pressures calculated via the Machine Learning Method.

To fully evaluate the model's function and performance three models are shown each with different input datasets:

Model 1: real-time drilling mechanics data.
Model 2: real-time drilling mechanics data and at-bit Gamma (ABGR)
Model 3: real-time drilling mechanics data, at-bit Gamma (ABGR), pre-drill Tops (from seismic) and TVD.

The 'shale only' Model 1 was created using only drilling or rig-based data sets for the input (No Tops, no TVD, no ABGR) but the training data was edited so that only shale values were used as the input (Figure 4 Track 6). A shale only model is important because it allows this method to be more faithfully compared to traditional pore pressure and fracture gradient (PPFG) calculation methodologies where 'muting' of sand data is common practice. Examples of traditional methods can be seen in Figure 4 Tracks 2-5. Here the editing was done to the training data set using a simple cut-off value for the Gamma Ray log to determine Pore Pressure values in Shale only regions.

Model 2 allows all 'hard data' sources to be used as well as ABGR on input (Figure 4, Track 7). As the ABGR log is used in the calculation as a lithology indicator, all calculations are performed at the depth of the ABGR sensor (5ft behind the bit). This model allows us to use all the input data in training with no muting for lithology type. It also allows the model to calculate, on output, a sand-based pore pressure which is an improvement over traditional pore pressure techniques. Results of Model 2 show excellent correlation with the PDA of pore pressure throughout the well except at 30,500ftMD where, interestingly, it appears closer to the result from Resistivity and Sonic-based methods.

Model 3 (Figure 4 Track 8) includes all data including soft data sources (Tops, TVD, etc.) and shows excellent correlation with the PDA of pore pressure throughout the well and a minor improvement over Model 2. The primary reason for separating Models 2 and 3 is to demonstrate that by establishing a good result with Model 2 the ANN is not reliant on soft data included in Model 3 i.e., the model is not taking shortcuts or learning to draw curves with depth tops or non-physical data.

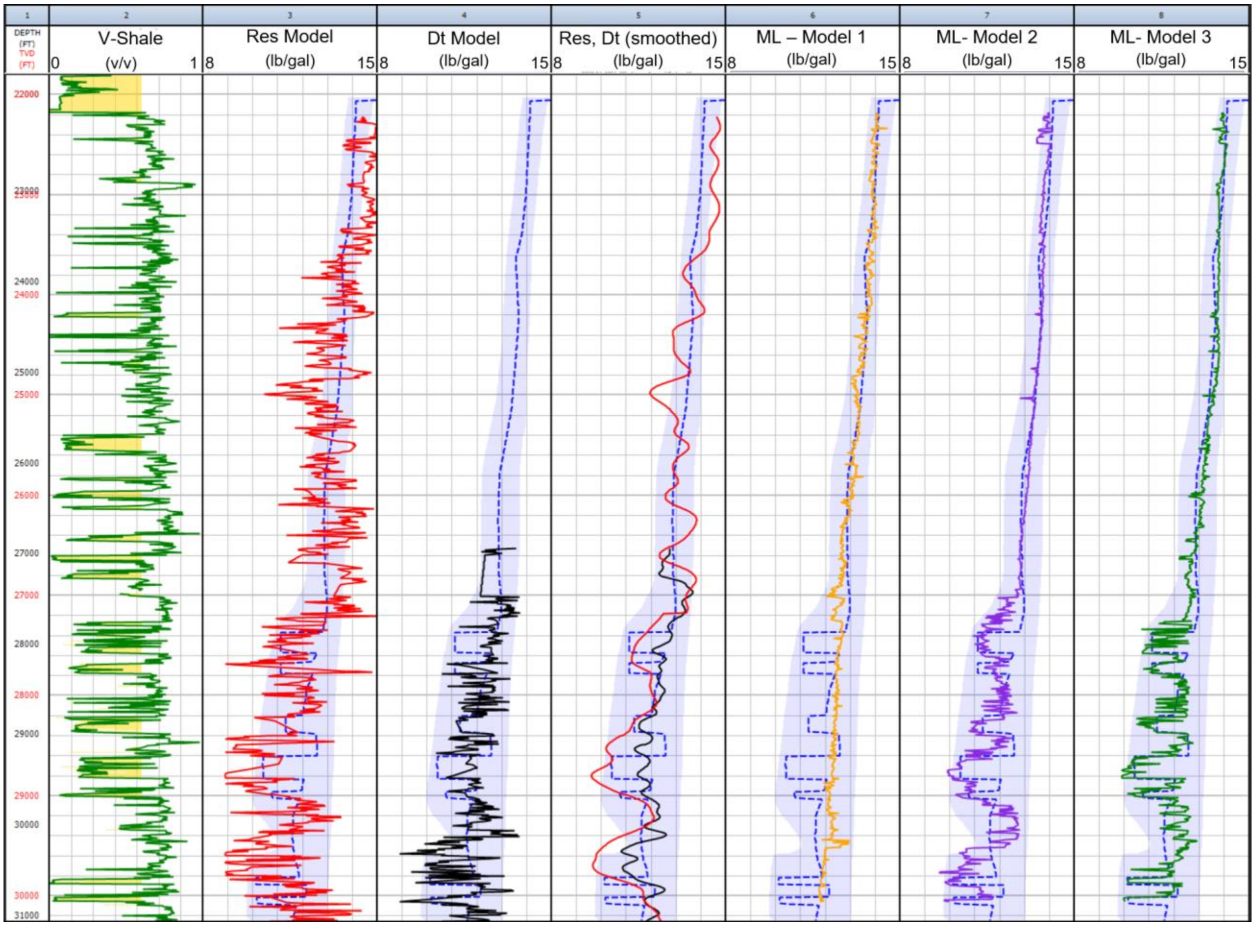

Figure 4 – Comparison of the 3 discussed model outputs with outputs from ‘traditional methods’ Resistivity and Sonic-derived pore pressures calculated via a modified Eaton methodology. ML-1 is a shale only model (no ABGR input) and is the most applicable to be compared to the resistivity and sonic-based methods. Both resistivity and sonic, even when smoothed, have a large degree of variability throughout the well whereas the output from the machine learning model (Track 6) has a much closer match to the PDA of pore pressure in the shale section. Tracks 7 and 8 show the changes to the model output when more ‘soft’ data (tops, lithology, etc.) are included in the calculation.

Discussion

This method, when correctly utilized, has the potential to provide accurate real-time pore pressure estimates at the bit and has been shown to be as or more accurate than traditional log or drilling methods at Stampede. This method can be used in conjunction to traditional methods to enhance PPFG accuracy while drilling.

Machine learning models work well in areas that do not differ greatly from well to well. Unless the model has experienced a particular parameter it does not necessarily ‘know’ how to interpret new, unseen information. In the case of the Stampede field, the pore pressure regime is relatively consistent across the field and as such the success of this method is more likely. This is partly intentional as during concept selection for this experiment a consistent and relatively ‘simple’ dataset was selected in order to trial this method in this conceptual stage. A trial on a well drilled in another subsalt field more than 60 miles away was not able to reproduce the PDA pore pressure. This may be due to a wide variety of reasons, but it reinforces the concept that machine learning models are most accurate near their training data set. In its current state this method would be most applicable to appraisal and development drilling (often the majority spend of field development drilling budget). While there are obvious limitations in this study, it has been demonstrated on a small scale to be accurate, useful, and worthy of additional investigation.

Future work is ongoing to assess the broader capabilities of these methods so the use case can be expanded beyond a single field and can predict pressure over entire plays and for exploration wells. In practice the largest challenge is not the creation or implementation of machine learning models, it is in data discovery, standardization, and aggregation of multiple datasets with varying formats.

The machine learning output more closely matches the PDA of pore pressure when compared to single curves created via 'traditional' methods (Figure 4). As discussed in the introduction of this paper fundamentally we do not actually 'know' in the shales which of these methods is the most accurate. A machine learning model needs to be trained on data nearby and thus when the pore pressure regime does not vary significantly it does an excellent job at reproducing the pore pressure. However, it should be noted that this is not too dissimilar to the sort of calibration required to calculate pore pressure via 'traditional' methods. When designing a pore pressure profile for an exploration well it is common to look at a variety of offset wells and calibrate the input parameters or equation constants (Eaton exponent, Bowers A & B etc.) to obtain a match between the output log-based pore pressure curve and the known pressure data (MDTs, kicks, gas events, etc.). When preparing for the subsequent exploration well a pre-drill model is created using the log to pressure translations calibrated to the offset wells. While drilling the well it is common to then re-calibrate these logs to pressure parameters or equation constants to measurements or observations seen in the well. Essentially the same process is occurring here with the exception that the user is not manually calibrating or conditioning the model.

This method has deliberately not included the use of standard subsurface logging tools to focus on the hypothesis that it would be possible to detect pore pressure using the drill bit and drilling mechanics data alone. Improved functionality could be possible with utilization of subsurface logs; one approach to this that was considered was using the drilling mechanics data to create an initial prediction that could then be modified as the subsurface logs move over the formation. Integration with seismic data would also be interesting to explore as this could potentially give one the ability to predict pore pressure ahead of the bit with the model 'learning' as the well is drilled.

While the concept for this method is new and requires further testing this method has the power to potentially transform the manner that pore pressure is estimated and communicated to the drilling rig leading to cost savings and improved rig safety.

References

1. Eaton B, A. "The Equation for Geopressure Prediction from Well Logs" SPE AIME 5544, 1975
2. Athy L. F, "Density, Porosity and Compaction of Sedimentary Rocks" AAPG Bulletin, 1930, Vol. 14, Number 1, 1-24
3. Terzaghi K and Peck R.B, 'Soil mechanics in Engineering Practice, John Wiley and Sons, Inc., New York (1948) 566.
4. Hubbert, M.K. and Rubey, W.W. "Role of Fluid Pressures in Mechanics of Overthrust Faulting: I. Mechanics of Fluid-Filled Porous Solids and Its Application to Overthrust Faulting", 1959, GSA Bulletin, 70, 115-166.
5. Jorden J. R, & Shirley, O.J., "Application of drilling performance data to overpressure detection." JPT 1966, 1387-1394
6. Majidi R, Albertin M, Last N, "Method for Pore Pressure Estimation using Mechanical Specific Energy and Drilling Efficiency", IADC/SPE 178842-MS, 2016,
7. Gan, T., Kumar, A., Ehiwario, M., Zhang, B., Sembrowski, C., de Jesus, O., Hoffmann, O., Metwally, Y., 2019. AI-Derived Logs for Formation Evaluation Using Case Studies in the Gulf of Mexico and Trinidad & Tobago

8. Flemings P.B., Long H., Dugan B., Germaine J., John C. M., Behrmann J.H., Sawyer D., "Pore pressure Penetrometers document high overpressure near the seafloor where multiple submarine landslides have occurred on the continental slope, Offshore Louisiana, Gulf of Mexico", 2008, Earth and Planetary science Letters, 269/3-4, 309-32
9. Flemings P.B., Stump B. B., Finkbeiner T., Zoback M., "Flow focusing in overpressured sandstones: theory observations, and applications", 2002, American Journal of Science, 302(10), 827-855
10. Reilly, M. J., & Flemings P.B., "Deep pore pressures and Seafloor venting in the Auger Basin, Gulf of Mexico", (2010, Basin Research, 22(4), 380-397