Akshat Patni

Mud Motor Failure Analysis using Data Science & Machine Learning Methods

Department of Statistics & Data Science

August 2021

Abstract

Land based directional drilling operations centered around oil and gas wells use sophisticated downhole drilling assemblies that are prone to stalls and failures. One such element of these drilling assemblies are mud motors, which are vulnerable to failures that compromise drilling operations. Previous research has used data analytics to quantify sources of damage to mud motors during drilling. However, there is still much that is yet to be understood in order to make a comprehensive algorithm that can act as an early warning system for mud motor failures in real time. The purpose of this project was to analyze factors that contribute to mud motor failures and advance the effort for creating a program to predict these failures. Data from surface sensors provided by an oil and gas operator was used to formulate a quality rating to evaluate active drilling periods. I applied this Q rating to datasets from 4 oil and gas wells provided by the operator and assessed the correlation between low quality drilling periods and instances of mud motor failure. Furthermore, using analysis from previous research regarding the influence of differential pressure spikes on mud motor failures, we examined the instances of these spikes in the available datasets and analyzed them in greater depth. The results of the Q rating study demonstrated inconsistent success in predicting mud motor failures, though still reveals key information regarding drilling halts that damage mud motors. The study of differential pressure spikes was significant, showing consistent patterns in almost all occurrences. These results allow operators to look into the events corresponding to differential pressure spikes and possibly prevent them. The work done in this project has advanced our understanding of what would go into a predictive program for mud motor failures, though further work is required to effectively develop this program.

Introduction

Background & Problem

Mud motors are critical components of downhole drilling assemblies, and their failure during operation often leads to significant consequences such as downtime without productive drilling or the need for replacement. Mud motors connect directly to the drill bit and generate power during drilling to increase rotation speeds and the rate of penetration. There are multiple parts within mud motors such as the power section, connection subs to the drill string and drill bit, and a bearing assembly. However, the component that most commonly provokes mud motor failure due to wear and tear is the elastomer. An elastomer is a rubber seal that lines the stator within the power section of the motor as shown in **Figure 1**. Its purpose is to provide resistance to fatigue and abrasion forces to the power section during drilling, as well as protection against chemical degeneration. (Hendrik, 1997). Continuous operation of a mud motor degrades the elastomer due to corrosion from extended contact with mud and oil, as well as stress from excessive differential pressures. Extreme operating temperatures can also lead to thermal fatigue that negatively affects the elastomer. All of these factors cause the elastomer to deteriorate (chunking) and compromise the integrity of its seal with the metal stator (debonding). (Gandikota, 2016).

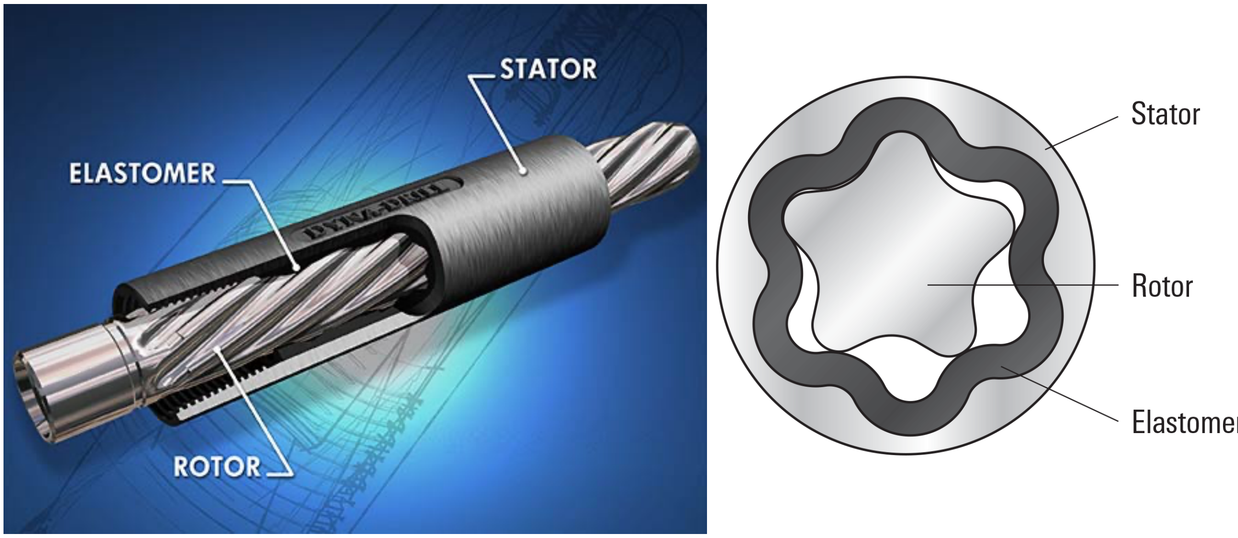


Fig 1. Components of the power section of a mud motor, demonstrating the role of the elastomer in lining the stator.

These outcomes are costly and disruptive to the drilling operation, requiring solutions that allow operators to predict and prevent these failures from occurring. Novel approaches have attempted to incorporate data science methods applied to sensor data to better understand mud motor failures and create solutions. It is possible to analyze high frequency data from downhole sensors, however this usage can be more complicated, costly, and time consuming. For the purposes of this project, more accessible information was preferred in the form of data from sensors at the surface of oil and gas wells.

Prediction & Conclusion

The objective of this research project was to deepen our understanding of the factors that contribute to mud motor failures and advance the effort to create a comprehensive warning program that can be used to signal possible mud motor failures in real time. We hope to add to the factors that are included in a 'cumulative damage index' and be able to effectively apply it to any oil and gas well with consistent results.

The work we have completed thus far has achieved reasonable results, augmenting prior research that has been done on this topic. It would be possible to achieve more significant progress with additional data to test our hypotheses, including temperature data that has been missing from all datasets. Greater consideration can also be given to the disparity in conditions that exist in different wells, including the types of rock formations that are encountered during drilling.

Materials

I worked on this project as part of a university research group in conjunction with a drilling company that will remain unnamed for purposes of confidentiality. In order to perform analysis, this operator provided us with historical drilling datasets of 4 oil and gas wells. The names of these wells, as well as locations and other specific details will also remain confidential. The source of the data in each well came from surface sensors that measured various physical parameters relevant to the operation of the downhole drilling assemblies. The datasets contain labelled/structured time-series data with a precision of 1 Hz. Each dataset contained over a million rows corresponding to several months of drilling operations. The relevant parameters in each dataset are listed below in **Figure 2**. The operator also provided us with daily drilling reports for each of the wells to give more comprehensive, qualitative details regarding their drilling operations.

Block Position	Top Drive RPM
Bit Depth	Top Drive Torque
Date	Total Depth
Time	Weight on Bit
Flow In	Differential Pressure

Fig 2. Relevant data channels provided in datasets for each of the four wells used in this study.

I used multiple softwares and programming libraries in order to perform analysis on the drilling datasets. All of the algorithms and data science applications were performed using Python and R Studio. The relevant data science libraries used include Pandas, NumPy, Matplotlib in Python and Dplyr, GGPlot2, and GLM in RStudio. Spotfire is an industry specific data visualization software that was used for initial visualization and overlay plotting. I also used Power BI for making more sophisticated visuals for presentation after analysis. Some of the algorithms that were utilized are sourced and adjusted from previous research projects for another operator in order to serve the purposes for this project.

Methodology

Data Transformation

The first steps of the project were to transform the datasets and prepare them for subsequent analysis. The time-series data provided to us was continuous over a period of months. However, during operations at each well there are points where the bottom hole assembly (BHA) is switched out and the components are inspected, repaired, or replaced to better serve the next stage of drilling operations. In some of these instances, the mud motor is replaced before the operation of the next BHA commences. Therefore, it was necessary to separate the datasets into time sections that correspond to the operations of each distinct BHA. This was imperative to properly evaluate distinct uses of the mud motors and accurately isolate instances of mud motor failures.

The procedure to separate these datasets consisted of opening the data in Spotfire for visualization, and locating the points where the bit depth went to ~0 ft indicating that they had pulled the assembly back to the surface and initiated a new BHA as demonstrated in **Figure 3** below. These timestamps were cross referenced with the daily drilling reports for the respective wells to confirm the day on which the BHA changed. I then sliced the datasets at the row indexes that matched each timestamp into separate dataframes using dplyr in RStudio. The date and time columns were then joined and converted into pd.datetime format in order to perform proper time series analysis.

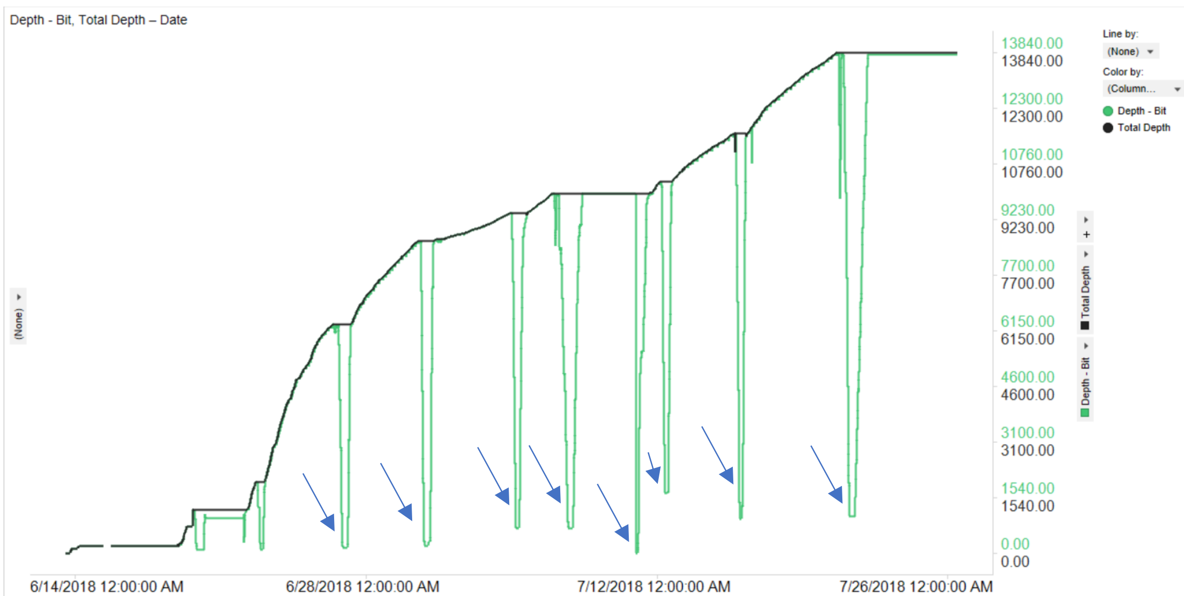


Fig 3. Visualization of where BHA changes occur in drilling datasets. CSV files are split at timestamps corresponding to drop in bit depth magnitude.

Analysis of Q Rating

Once these data transformation steps were completed, I was able to begin the analytical component of the project. In order to predict mud motor failures, we first had to identify adverse events during drilling that are hypothesized to cause damage to mud motors. During meetings with our team, the operator contemplated that periods of ineffective drilling create exceptional load on the drilling assembly and lead to wear and tear. These periods are characterized by active drilling operation of the assembly without a proportionate rate of penetration into the ground. This was expressed in the sensor data by periods where the magnitude of 'Weight on Bit' and 'RPM' were high, yet the rate of penetration stagnated near 0. This hypothesis was used to develop a quality or 'Q' rating formula that would spike when active drilling coincided with poor penetration.

$$Q = \frac{(Weight\ on\ Bit)^2 (RPM)^{0.7}}{ROP}$$

In the Q rating formula above, both the 'Weight on Bit' and 'RPM' are empirical quantities yet ROP is a derived quantity. I applied an algorithm from work on a prior mud motor project to derive the ROP at any given point and append it as a new column in the dataframe. This function calculated ROP by taking the differences in block position during active drilling and converting the units into ft/hr. Subsequently, I used the formula above to loop through all of the BHA's and calculate the Q rating as a new column derived from the columns of the other three variables.

Once the raw Q rating column was derived, there were multiple steps to filtering and smoothing the dataframes in order to make it suitable for analysis. Data from the sensors was fairly noisy and inconsistent. There were several instances where physical quantities such as 'Weight on Bit' and 'RPM' were recorded as negative values which is a physical impossibility. During the data cleaning process, I asserted negative values as '0' and filtered out rows where the values were either 'NA' or 'infinity' to focus the analysis on periods of active drilling.

The post-filtered data still required smoothing to accurately express the Q rating, due to the noisiness of the sensor data. I tested and applied several different smoothing methods to the data and evaluated which method was the best to follow through with. The objective was to smooth the data to emphasize significant observations without over-smoothing the data to flatten out important spikes since the entire purpose of the Q rating is to reveal values that are significantly above the norm. Smoothing methods such as LOESS, kernel regression, and simple/exponential moving averages over-smoothed the data and flattened out significant spikes. LOESS and kernel regression also proved to be extremely computationally intensive on datasets of this size (>1 million rows) and were inefficient. Other methods such as FFT were not appropriately applicable because the data was not periodic.

The most effective smoothing method proved to be the application of a Savitzky-Golay filter. This filter applies a least-squares fitted polynomial to the data within a fixed window size around each point. Since each of the spikes in the Q rating are defined by only a few points at a time, the Savitzky-Golay filter is advantageous to avoid over-flattening peaks. This method works well on data that is non-linear and non-periodic. Ultimately, the Savitzky-Golay filter had the best performance out of all of the methods to smooth the data while still preserving the significant spikes revealed by the Q rating. (Luo, 2005).

After developing the algorithm to calculate the raw Q rating and incorporate the filtering and smoothing conditions, I applied it to all of the BHA's in each well for analysis. I created an additional Boolean column in each well that returned as 'True' for all of the data points in the BHA's that contained instances of mud motor failure (sourced from the daily drilling reports) and returned 'False' for the BHA's that did not record any mud motor failures. I performed logistic regression analysis to determine whether high Q rating values correlated with instances of mud motor failure in each BHA. Logistic regression was used to evaluate the performance of the Q rating as a classifier between True/False cases of mud motor failure. I also plotted the Q Rating over time for each well in Power BI to visualize the metric and correlate whether or not BHA's with mud motor failure contained high Q rating spikes.

Analysis of Differential Pressure Spikes

Previous research into mud motor failures explored the relationship between spikes in differential pressure during drilling and cumulative damage to mud motors leading to failure. The study demonstrated that a high quantity of differential pressure spikes are the most significant contributor to mud motor failures compared to other empirical parameters. (Lawal, 2021). However, this research only analyzed the quantity of differential pressure spikes in each BHA and did not examine individual differential pressure spikes in detail. I proceeded to expand on this line of research by investigating the location of all recorded spikes as well as analyzing patterns and trends that occur amongst other key parameters around each differential pressure spike.

For this analysis, we decided to narrow our scope of analysis from BHA's in each well to individual drilling stands. Stands are sections of multiple pipes connected at joints in the well. Individual stands are expressed in the dataset as periodic variations in the 'Block Position' parameter, as demonstrated in **Figure 4** below. We utilized a Python function from the previous mud motor failure study to label each data point with the stand number that it coincides with.

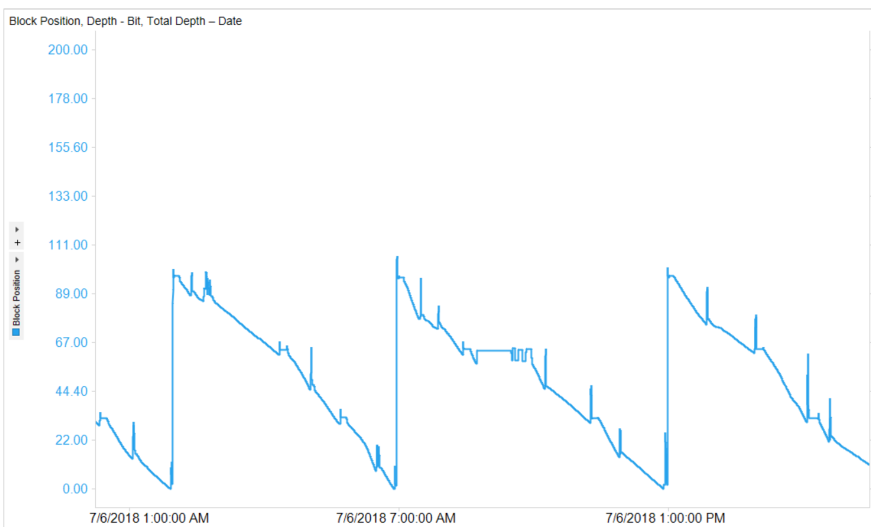


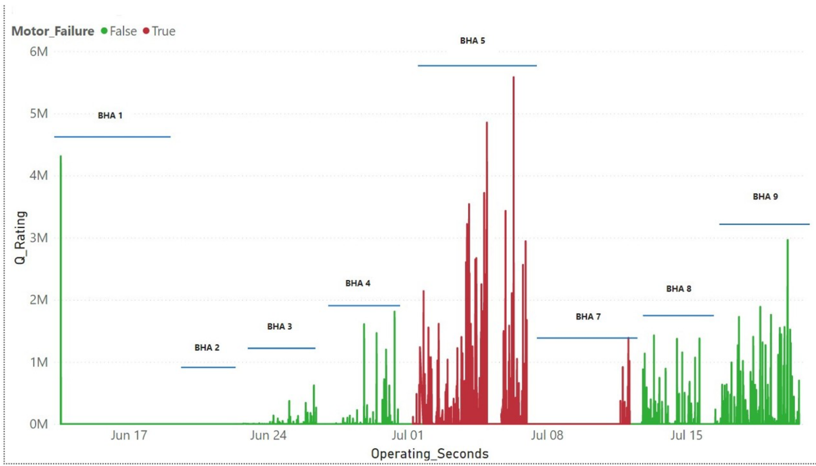
Fig 4. Example visualization of 3 separate drilling stands in a BHA. Stands are characterized by periodic movement in 'Block Position'. Stands normally have a length of approximately ~90 ft demonstrated on the Y axis.

I also adapted the two functions from the previous mud motor study used to calculate and record the magnitude of every differential pressure spike in the dataset. Since these functions only focused on cataloging the total number of spikes and their magnitude, I was required to alter the functions and add multiple parameters to extract the timestamps of each differential pressure spike at its peak and append it to the dataframe. (Lawal, 2021). This allowed us to locate the spikes within the data and analyze them in depth, compared to other empirical parameters in each stand. We used Spotfire to visualize the stands that contained differential pressure spikes in greater detail.

Results & Discussion

Analysis of Q Rating

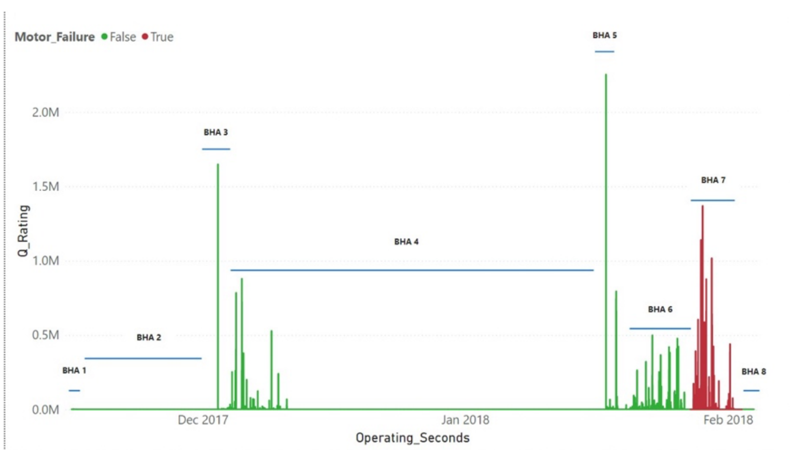
<u>Well #1</u>



Number of BHA	Mud Motor Failure	Total Depth Drilled by Bit (ft)	Average Q Rating	Cumulative Q Rating
1		1,097.00	580.87	2.52E7
2		756	62.26	1.57E6
3		4,357.00	236.53	5.27E7
4		2,309.00	1109.1	2.64E8
5	Yes	778	7733.83	2.37E9
6		538.00	Bad ROP Data	Bad ROP Data
7	Yes	313	3162.23	8.21E7
8		1,335.00	1764.11	3.66E8
9		2,237.00	2511.14	7.75E8

Fig 5. Significant Q Rating spike in BHA #5, corresponding to mud motor failure. BHA #7 also had instance of mud motor failure yet was not signaled by high Q rating. Average Q rating values were highest in the BHA's that contained instances of mud motor failure.

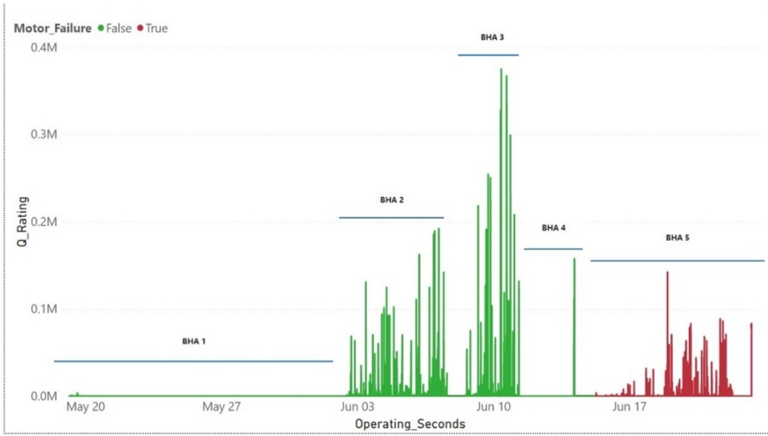
<u>Well #2</u>



Number of BHA	Mud Motor Failure	Total Depth Drilled by Bit (ft)	Average Q Rating	Cumulative Q Rating
1		355	33.04	5.28E5
2		768	0.00	0.00
3		4,499.00	522.64	5.02E7
4		4,475.00	387.21	1.08E8
5		9	1236.13	5.88E7
6		4,747.00	322.49	8.94E7
7	Yes	2,498.00	1150.51	2.05E8
8		862	0.00	0.00

Fig 6. BHA #7 had instance of mud motor failure and contained highest density of Q rating spikes. This BHA also had 2^{nd} highest average Q rating and highest cumulative Q rating.

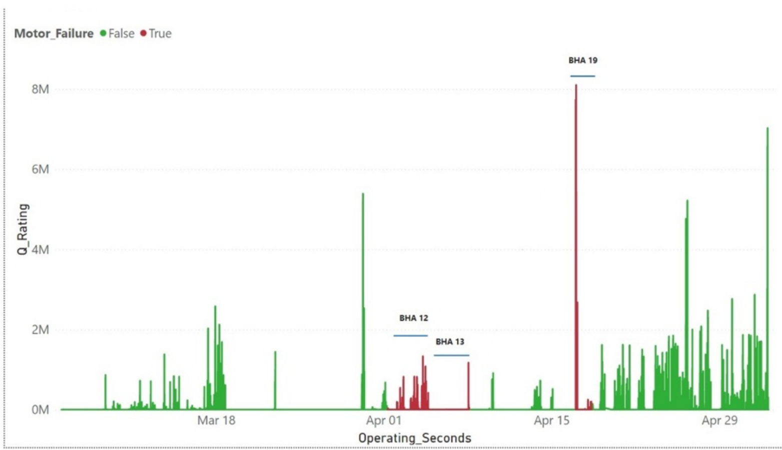
Well #3



Number of BHA	Mud Motor Failure	Total Depth Drilled by Bit (ft)	Average Q Rating	Cumulative Q Rating
1		1,130.00	9.20	5.70E9
2		6,865.00	272.53	9.58E7
3		1,970.00	655.64	1.32E8
4		10	4787.93	1.16E7
5	Yes	9,551.00	77.79	3.28E7

Fig 7. High Q values in BHA #2 and BHA #3 did not correspond to mud motor failure. BHA #5 had instance of mud motor failure yet did not show high Q rating values.

Well #4



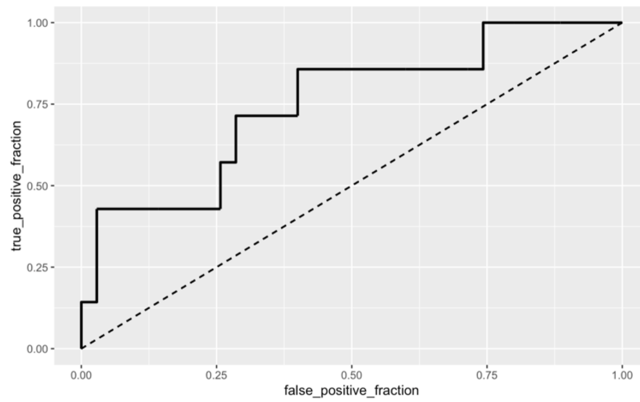
Number of BHA	Mud Motor Failure	Total Depth Drilled by Bit (ft)	Average Q Rating	Cumulative Q Rating
12	Yes	957	1862.21	3.27E8
13	Yes	1	11231.48	14.3E8
14		79	379.79	3.27E7
15		180	442.65	5.96E7
16		391	3676.31	3.37E7
17			14025.15	7.42E7
18		16	0.00	0.00
19	Yes	12.5	16614.32	1.83E9

Fig 8. Insufficient drilling depth in BHA #13 and BHA #19 to classify it as elastomer failure due to motor stalls. This well had faulty data and led to inconclusive application of Q rating.

The Q rating visualization analysis demonstrated that the metric is inconsistent in its correlation with instances of mud motor failures across all wells. Its application performed adequately in wells #1 and #2, while it showed poor correlation in wells #3 and #4. Well #4 cannot be deemed as a conclusive application due to the fact that two of the BHA's with mud motor failures had insufficient drilling depth to justify elastomer failure. Upon examination of the daily drilling reports, it was found that the mud motor failures were recorded before any active drilling operations took place. BHA's with the highest average Q ratings tended to correlate with instances of mud motor failure, leading me to investigate whether average Q rating could function as an effective classifier to predict whether or not a mud motor would fail in a particular BHA across all wells. There were 42 rows of data corresponding to available BHA's.

Logistic Regression Analysis

```
Call:
glm(formula = y ~ Q_Means, family = binomial(link = "logit"),
    data = logset)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.5462  -0.5069  -0.4395  -0.4200   2.2162

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept) -2.3838488  0.6024825  -3.957  7.6e-05 ***
Q_Means      0.0002295  0.0001046   2.193   0.0283 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 37.847  on 41  degrees of freedom
Residual deviance: 32.215  on 40  degrees of freedom
AIC: 36.215

Number of Fisher Scoring iterations: 4
```



	acc <dbl>	sens <dbl>	spec <dbl>	ppv <dbl>	auc <dbl>
1	0.8571429	0.2857143	0.9714286	0.6666667	0.7510204

Fig 9. Logistic regression application demonstrates that average Q rating correlates significantly with binary mud motor failure outcomes (p value < .05). ROC evaluation of the model shows an AUC value of .75 which is passable but mediocre. Overall accuracy of prediction was about 85.7%

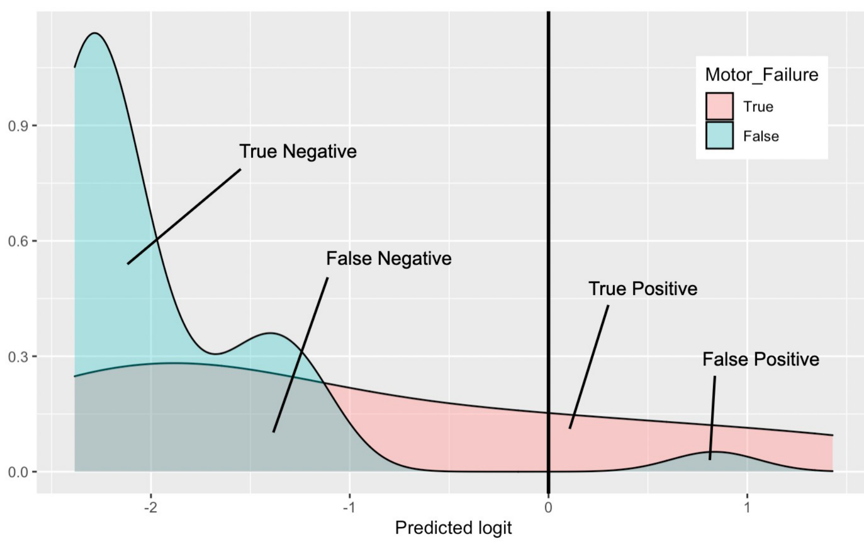


Fig 10. Colored areas under curves represent probability of average Q rating accurately classifying BHA’s as true/false for instances of mud motor failure. Distinct colored areas represent probability of true negative/positive classifications, whereas overlapping areas represent probability false positives/negative classifications.

The results of the logistic regression analysis between average Q rating and outcomes of mud motor failure were fairly positive. The model showed a significant p value of .0283 and adequately passed ROC evaluation with good accuracy and acceptable AUC values. **Figure 10** demonstrates that the average Q rating model is more likely to classify true positive/negative cases of mud motor failure than incorrectly classify false positive/negative cases.

Some of the limitations of the Q rating include the small sample size of BHA's with mud motor failure that were available to test the performance of the Q rating. The surface sensor data was also very noisy and possibly faulty at some points, skewing the analysis of the correlation. Some of the instances of mud motor failure in the data were also not believed to be caused by elastomer failure, due to insufficient drilling. It would also be valuable to differentiate between the different rock formations that exist in different wells, as that has a significant impact on overall drilling operations and the patterns that appear in the data.

While it cannot function as a standalone metric to predict mud motor failure, there is some valuable information that can be obtained from studying the components of the Q rating. For example, there are many periods of drilling in which the 'Weight on Bit' exceeds manufacturer recommendations, putting a great deal of stress on the drilling assembly that likely causes damage to the mud motor. Additionally, there are many periods where the drill is rotating and high weight on bit is applied yet the rate of penetration stalls. This is likely due to the fact that the drill bit encounters hard rock formations which prevent it from. Further analysis could be done to evaluate whether Q rating spikes coincide with differential pressure spikes.

Analysis of Differential Pressure Spikes

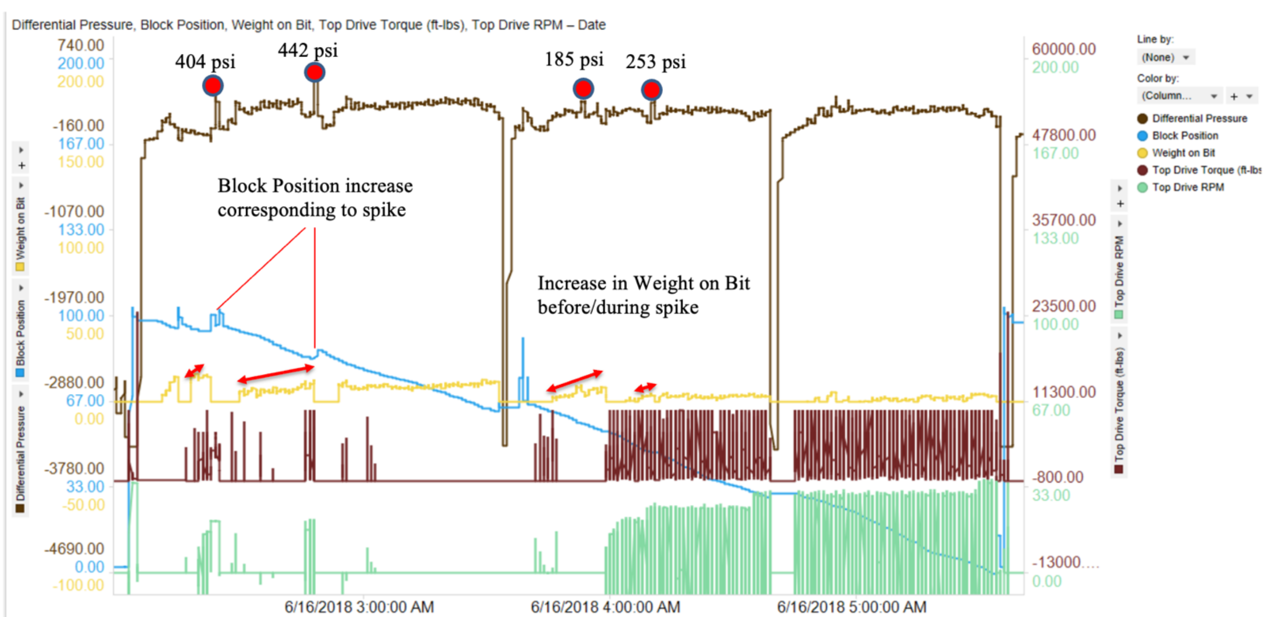


Fig 11. Locating high impact pressure spikes and overlaying with other parameters shows some trends. 'Weight on Bit' parameter tends to spike prior to differential pressure spikes.

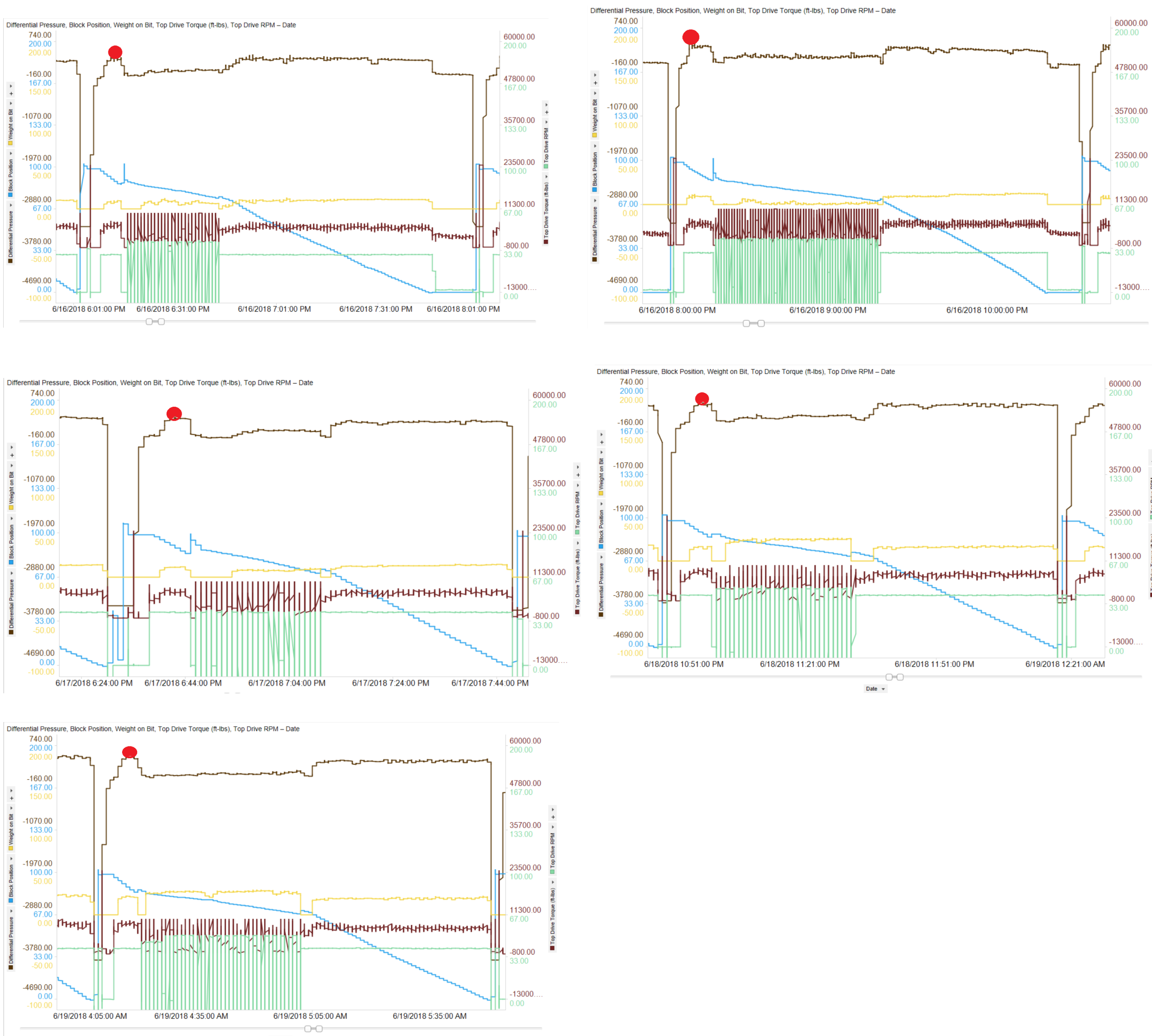


Fig 12. Extracted timestamps of all high impact pressure spikes shows consistent trend that almost all spikes occur within first 10-15 minutes of each stand.

The differential pressure spike analysis proved to be significant. Certain trends were found amongst confounding variables at the time of occurrence of high impact pressure spikes, such as a corresponding spike in the weight applied to the bit. **Figure 12** also demonstrates across all plots that the ‘Top Drive RPM’ and ‘Top Drive Torque’ parameters begin to oscillate very strongly after differential pressure spikes as the drillers transition from rotary drilling to slide drilling.

Analysis across 42 differential pressure spikes also demonstrated that 90% of the high impact pressure spikes occur within the first 10-15 minutes of drilling in each stand, with a mean of approximately 12 minutes. This is a significant observation that will certainly be considered in future use of differential pressure spikes to formulate a predictive damage index that signals mud motor failures. During our team meetings, it was hypothesized that these spikes could be caused by activation of the auto-driller unit. If that is the case, it is something that operators can directly target and aim to prevent during drilling operations.

Acknowledgements

I would like to thank the members of my research group for their help and cooperation throughout this project. Dr. Pradeepkumar Ashok provided valuable guidance to us regularly and gave us direction when we needed help with our ideas. I would also like to thank my fellow research assistant, Mohammed Alameer for help with understanding the technical aspects of petroleum engineering and drilling since it was out of my realm of my expertise when I began this project. I would also like to thank the companies that we worked on this project with for providing us with the datasets and dedicating time to work with us during team meetings.

Works Cited

1. Drilling Manual. 2021. "Drill Pipe Length Ranges, Design & Run." *Drilling Manual*. February 18. https://www.drillingmanual.com/2017/11/determining-drill-pipes-lengths-on.html.
2. Gandikota, Raju. 2016. "Improving Reliability of Drilling Mud Motors." *Improving Reliability of Drilling Mud Motors - 10-14-2016 : IADD News for International Association of Directional Drilling in Houston, Texas USA - International Association of Directional Drilling*. October. https://www.iadd-intl.org/news/improving-reliability-of-drilling-mud-motors/.
3. Hendrik, John. 1997. "Elastomers in Mud Motors for Oil Field Applications." *Corrosion97*.
4. Lawal, Tesleem. 2021. "Real-Time Prediction of Mud Motor Failures Using Surface Sensor Data Features and Trends." Thesis. The University of Texas at Austin.
5. Luo, Jianwen. 2005. "Savitzky–Golay Smoothing Anddifferentiation Filter for Even Number Data." *Elsevier*.
6. "PowerPak Steerable MOTORS." 2021. *Steerable Motors | Schlumberger*. Accessed August 9. https://www.slb.com/drilling/bottomhole-assemblies/directional-drilling/powerpak-steerable-motors.
7. DrillingFormulas.Com. 2018. "What Are Positive DISPLACEMENT Mud Motors in Drilling for Oil and Gas?" *Drilling Formulas and Drilling Calculations*. April 29. https://www.drillingformulas.com/what-are-positive-displacement-mud-motors-in-drilling-for-oil-and-gas/.