

Marius Berge-Skillingstad & Viktor Hamre
Anderssen

drillWiz: An application for data visualization and drilling optimization

Master's thesis in Petroleum Geosciences and Engineering

Supervisor: Bjørn Astor Brechan

June 2019

Marius Berge-Skillingstad & Viktor Hamre
Anderssen

drillWiz: An application for data visualization and drilling optimization

Master's thesis in Petroleum Geosciences and Engineering
Supervisor: Bjørn Astor Brechan
June 2019

Norwegian University of Science and Technology
Faculty of Engineering
Department of Geoscience and Petroleum



Preface

This thesis was carried out during the spring semester of 2019 and completes the Master of Science in Petroleum Engineering with a specialization in Drilling Technology for the authors.

This thesis is carried out at the Norwegian University of Science and Technology at the Department of Geoscience and Petroleum.

The thesis was conducted with background from a fall semester project "MSE: Next Generation Digital Drilling Optimization", and is a continuation of this work through the development of a web-based application for data visualization and drilling optimization.

Trondheim, 5/6/2019



Marius Berge-Skillingstad



& Viktor Hamre Anderssen

Acknowledgements

First and foremost, we would like to pay gratitude to our supervisor Bjørn Astor Brechan for his guidance, both throughout our master thesis and the semester project. With guidance throughout our semester project, we became ready to conduct this thesis on a more individual basis. In addition to this, we appreciate the effort to set us up with the necessary industry contacts to obtain the data for our thesis.

We would like to give a special thanks to Fred Dupriest, our former professor at Texas A&M, who taught us the concepts of MSE, and motivated and inspired us into seeking physics-based answers in our work. We are tremendously appreciative for your willingness to support us on the more theoretical part of our thesis, and for always being there to help us with our questions.

Another big thanks goes to associate professor Sigve Hovda for his insight and guidance on how to best develop our application. Sigve provided constructive feedback about the more programmatic part of our thesis, and helped us better understand change detection and pattern recognition.

We would also like to thank our previous employers that have, through internships, provided us with unique insight and experiences on drilling engineering in industry applications. That is why we want to emphasize the need for drilling engineering students to get out there and witness operations first-hand, to get an idea of how theory can be applied in practice.

Last but not least, a huge shout-out goes to our fellow students at NTNU for five good years of laughs, friendships, and good beer. It's been a blast! A special remark goes to our office colleagues Ivar Gravdal and Joakim Nornes for all the good class discussions and always motivating us to do our best.

We would like to thank the readers of this report for taking their time and interest. For this application to achieve its full potential and fulfill its purpose, it is crucial that we all get it out there and spread the word, enjoy!

Sincerely,

Marius Berge-Skillingstad & Viktor Hamre Anderssen

Abstract

An application has been developed to identify bit dysfunction and drilling inefficiency from expected responses in MSE and operating parameters. This work is a continuation of a semester project mapping the potential of automated physics-based practices and MSE-based surveillance techniques through smart algorithms, ref. (Berge-Skillingstad & Anderssen, 2018). The application is essentially split into two main functionalities; a parameter optimization module used to identify trends in parameter selection through lithology specific surface regression, and a diagnostic tool used to identify intervals with drilling dysfunction. Click this [link](#) to test the application, or [here](#) for a video walk-through of the app.

Through implementation of fundamental drilling theory, statistics and mathematical modelling, successful results have been generated to validate the concept of automated drilling diagnostics and lithology specific road-maps. To produce the results in this thesis, data have been acquired from contact with an anonymous operator and its associated service companies. The results have been validated using external information such as bit pictures, downhole memory subs, end-of-well reports, service company reports, and manual interpretations. As a result, extensive data preparation has been necessary to parse data from different data silos.

The purpose of the application is twofold; to prove the concept that MSE-based surveillance and diagnostics may be automated for post-drill evaluation, and secondly provide a scalable cross-platform web application to assist the drilling engineer in reference well studies and well planning phases. How the application can create value is a focus area throughout the thesis.

Since the application is still in the concept-phase with great potential, focus has been on establishing core functionalities, software architecture & GUI, and smart algorithms for automatic drilling diagnostics. As well as proving the concept, challenges in implementation and level of confidence have been assessed for further development. The application serves as a post-drill analysis tool, but together with the successful results in this thesis, a way-forward for next-level implementation is presented to move from offline to online change detection, i.e. towards real-time implementation.

Sammendrag

En applikasjon er utviklet for å identifisere boredysfunksjon-og-ineffektivitet fra forventede responser i MSE og driftsparametere. Dette arbeidet er en videreføring av et semesterprosjekt som kartla potensialet for automatisering av fysikkbaserte praksiser og MSE-baserte overvåkningsteknikker gjennom smarte algoritmer, (Berge-Skillingstad & Anderssen, 2018). Applikasjonen er i hovedsak delt inn i to hovedfunksjoner; en parameteroptimaliseringsmodul som brukes til å identifisere trender i parametervalg gjennom litologi-spesifikk overflate-regresjon, og et diagnostiseringsverktøy som brukes til å identifisere intervaller med boredysfunksjon. Klikk denne [koblingen](#) for å teste programmet, eller [her](#) for en videogjennomgang av appen.

Gjennom implementering av grunnleggende boreteori, statistikk og matematisk modellering, har vellykkede resultater blitt generert for å validere konseptet om automatisk bore diagnostikk og litologi-spesifikk parameteroptimalisering. For å produsere de ulike resultatene er data ervervet gjennom kontakt med en anonym operatør og tilhørende serviceselskaper. Resultatene er validert ved hjelp av eksternt informasjon, som for eksempel bilder av borekroner, nedihulls vibrasjonsmonitorering, brønn- og servicereporter, samt manuelt tolkningsarbeid. Omfattende dataforberedelser er blitt gjort for å samle inn og strukturere data fra forskjellige databaser.

Hensikten med applikasjonen er todelt; bevise konseptet at MSE-basert overvåking og diagnostisering kan automatiseres for post-analyse etter boring, samt etablere en skalerbar internett applikasjon til å bistå ingeniører med referansebrønnstudier i brønnplanleggingsprosessen. Hvordan applikasjonen kan skape verdier er et fokusområde gjennom hele avhandlingen.

Siden applikasjonen fortsatt er i konseptfasen med stort potensial, er det blitt satt fokus på å etablere kjernefunksjoner, programvarearkitektur & brukergrensesnitt, samt smarte algoritmer for automatisk bore diagnostikk. I tillegg til å bevise konseptet, har utfordringer ved implementering, samt konfidensnivå av diagnostikk blitt vurdert og diskutert for videreutvikling. Applikasjonen fungerer som et analyseverktøy etter boring, men sammen med de vellykkede resultatene i denne oppgaven, presenteres en visjon for å gå fra offline til online endringsanalyse, dvs. mot nåtidsimplementasjon underveis i boreprosessen.

Contents

Preface.....	i
Acknowledgements	iii
Abstract.....	v
Sammendrag.....	vii
List of figures.....	xiii
List of tables.....	xv
Nomenclature & abbreviations.....	xvii
1 Introduction	1
2 Theory.....	3
2.1 Mechanical specific energy.....	3
2.1.1 Mechanical specific energy & efficient drilling operations.....	5
2.1.2 Regions of drilling efficiency	6
2.1.3 Analyzing trends in MSE.....	7
2.2 Drilling a linear borehole	8
2.3 Detecting change	10
2.3.1 Heuristic vs. analytical solutions	10
2.3.2 Basics on change.....	12
2.3.3 Changepoint detection theory	13
2.3.4 Methods for detecting optimal changepoints	17
2.3.5 PELT	17
2.3.6 Convolution to smooth noisy data	19
2.4 Regression for optimization	20

2.4.1	Adjusted coefficient of determination: error analysis.....	22
3	Method.....	23
3.1	The choice of programming language: Python	23
3.1.1	External Python packages	24
3.2	Software architecture.....	28
3.2.1	Software user interface	32
3.2.2	Software functionality list.....	35
3.3	Data gathering: acquisition and availability.....	36
3.4	Data preparation	37
3.4.1	Preparation of depth-based data.....	38
3.4.2	Preparation of time-based data.....	40
3.5	Changepoint detection scheme.....	42
3.6	Picking out data points of interest	44
3.7	Establishing lithology specific road maps for optimal parameter selection.....	45
3.8	Diagnostics	48
3.8.1	Establishing a linear trend of “expected” MSE	48
3.8.2	Whirl	50
3.8.3	Stickslip.....	52
3.8.4	Torsional severity.....	53
3.8.5	Stringer detection & damage assessment.....	55
4	Results.....	57
4.1	Well information	59
4.2	Creating baselines	63
4.3	Concept: establishing an overview.....	64

4.4	Concept: lithology specific road-maps.....	67
4.5	Case study: lithology specific road-map for drilling the Apollo Limestone.....	69
4.5.1	Drilling with a RSS.....	70
4.5.2	Drilling with a bent motor.....	75
4.6	What results will you show your manager?.....	78
4.7	Proof of MSE with simple diagnostics.....	80
4.7.1	Bit whirl.....	80
4.7.2	Stickslip.....	82
4.8	drillWiz Diagnostics Scenarios.....	84
4.8.1	Case: Well C4 diagnostics study.....	84
4.8.2	Case: Well C3 diagnostics study.....	90
4.9	Stringer detection & bit forensics.....	93
5	Discussion.....	96
5.1	Uncertainties and the impact of this.....	96
5.1.1	Corruption of MSE data through parameter uncertainties.....	96
5.1.2	Uncertainties when determining order of surface regression.....	101
5.1.3	Confidence-level and uncertainty in dysfunction analysis.....	104
5.2	Further work.....	108
5.2.1	Strategy for validation of diagnostics.....	108
5.2.2	Automated data parsing to match all files and databases.....	109
5.2.3	Proposed bit balling diagnostics scheme for next level implementation.....	110
5.2.4	Potential markets for commercialization: NOVOS application platform – The Reflexive Drilling System.....	112
5.2.5	Real-time implementation.....	113

6 Conclusion 115
Bibliography 117

List of figures

Figure 2.1: Characteristic WOB vs. ROP drill-off curve.....	6
Figure 2.2: Example of different straight-line fits to an input signal.....	11
Figure 2.3: Change in mean (Left) & Change in frequency (Right).....	12
Figure 2.4: Negative log-likelihood response to changes in penalty.	16
Figure 2.5: Binary segmentation method visualized.....	17
Figure 2.6: PELT method outlined step-by-step.....	19
Figure 3.1: From Python (Foundation)	24
Figure 3.2: Software architecture on how functions are initiated after starting drillWiz.	30
Figure 3.3: Flow chart of function architecture visualized.	31
Figure 3.4: Software interface after starting drillWiz.....	32
Figure 3.5: The top section of the application comprises the user input interface.	33
Figure 3.6: User interface of drillWiz after engaging Surface Road-Map mode.....	34
Figure 3.7: Road to Data Preparation	38
Figure 3.8: Snippet from Excel template	39
Figure 3.9: First stage	40
Figure 3.10: Second stage.....	40
Figure 3.11: Third stage.....	41
Figure 3.12: Flowchart overview of creating baselines.....	42
Figure 3.13: PELT sub-functions overview.....	43
Figure 3.14: Baseline generator overview.	43
Figure 3.15: PickPoints function illustrated.....	44
Figure 3.16 Parameter optimization scheme summarized.	47
Figure 3.17: Generic diagnostics procedure strategy used in this thesis.	48

Figure 3.18: Whirl diagnostics scheme overview.....	50
Figure 3.19: Stickslip diagnostics scheme overview	52
Figure 3.20: Torsional severity diagnostics overview	53
Figure 3.21: Stringer detection and damage assessment scheme overview.....	56
Figure 4.1: Typical well illustration and associated lithology depths	62
Figure 4.2: Baselines consequently reduces the date over several magnitudes	63
Figure 4.3: Well C4 Overview.....	65
Figure 4.4: Well C4 8.5 Section, shale in green, sand in yellow and chalk in blue.....	66
Figure 4.5: Filtering of baselines results in less data points for analysis.....	68
Figure 4.6: Introducing additional data points.....	69
Figure 4.7: Baselines generated with sufficient accuracy.....	72
Figure 4.8: 3 wells are aggregated to fill the operating area with only representable data.	73
Figure 4.9: Surface generated from aggregated data points	73
Figure 4.10: A second and thrid angle of the generated surface regression for the RSS case.....	74
Figure 4.11: 3 wells are aggregated to fill the operating area with only representable data.	76
Figure 4.12: Generated surface from bent motor drilling	77
Figure 4.13: Whirl scenario depicted.....	80
Figure 4.14: ROP vs. WOB proportionality shift at higher WOB. Indicates whirl.....	81
Figure 4.15: Whirl scenario taken from an arbitrary well. Whirl intervals colored in red.	82
Figure 4.16: Stickslip scenario from well A3 with a bit trip on the black line.....	83
Figure 4.17: Well C4: 8 ½’’ section diagnostics	85
Figure 4.18: Well C4: 8 ½’’ section diagnostics: Zoomed-in.....	87
Figure 4.19: Well C4: Apollo Limestone diagnostics	89
Figure 4.20: Well C3: 8 ½’’ section diagnostics	91

Figure 4.21: Well C3: 8 ½’’ section diagnostics: Zoomed-in.....	92
Figure 4.22: Bit pictures with clear IFS damage.	94
Figure 4.23: Damaging stringer detected, indicated by red fill.	94
Figure 4.24: Differential pressure spiking along with MSE, i.e. typical stringer response.	95
Figure 4.25: High load on bit shoulder	95
Figure 5.1: Different orders of surface regression yield different solutions.....	103
Figure 5.2: Weakness in Diagnostics: Bit balling event in the Gemini Sands, well C4.....	106
Figure 5.3: Bit balling response seen.....	110
Figure 5.4: Proposed bit balling diagnostics procedure.....	111

List of tables

Table 1: Plot legends.....	58
Table 2: Diagnostic Legend.....	58
Table 3: Well Information	60
Table 4: Formation List	61
Table 5: Case data for RSS.....	70
Table 6: Baseline parameters to achieve results	70
Table 7: Case data for the bent motor drilling scenario.....	75
Table 8: Summarized results from the two road-map cases.	78
Table 9: Economic potential for well B6 and well C3.....	78
Table 10: Case data for stringer detection	93

Nomenclature & abbreviations

MD	Measured depth
TVD	True vertical depth
TD	Target depth
WOB	Weight on bit
RPM	Rotations per minute
ROP	Rate of penetration
TOR	Torque
MSE	Mechanical specific energy
Diff	Differential pressure
DOC	Depth of cut
DOCC	Depth of cut control
BHA	Bottom-hole assembly
MWD	Measurements while drilling
PDC	Polycrystalline diamond compact
PELT	Pruned exact linear time
RSS	Rotary steerable system
GUI	Graphical user interface

1 Introduction

In 2015, McKinsey & Company conducted a large independent market research and concluded that drilling and completion costs account for about 40-50% of total capital expenditure for an average offshore operator. This cost may be as high as 65% for many onshore operators. 70-80% of these costs are time related, meaning that any reduction in well delivery time may have a substantial impact on capital spending (Brun, Aerts, & Jerkø, 2015).

The oil industry has seen countless technological innovations throughout the years in an effort to increase efficiency and reach ever-more challenging prospects. The technological innovations from the first well drilled in 1859 until today have been compared to technological feats such as placing a man on the moon (Mau & Edmundson, 2015). Even though there are countless innovations and new tools being developed in the oil industry every year, perhaps the most valuable resource in driving performance is us, and the way we work.

Performance cultures in some E&P companies tend to be empirical, rather than coming from physics-based approaches. It is true that empiricism and making operational decisions based on prior experiences, statistics and optimization are very important. However, empirical cultures intrinsically resist change in practices. Without change in practices there can be no fundamental change in performance. Most operators have already optimized performance vs. risk with best-practices. Many of these best-practices are often empirical, meaning that limits are set from previous experiences. The problem with such best-practices are that few people challenge to understand the root problem, from a physics-based point-of-view. Without understanding what is limiting us, there can be no real gain in performance without risking unintended consequences. After all, root limiters are considered risks. Redesigning limiters often overlaps with redesigning risk.

In 2005, (Dupriest & Koederitz, 2005) introduced how MSE surveillance could serve as a physics-based diagnostic tool to continuously detect changes in drilling efficiency. After all, MSE-based practices and surveillance techniques provide the basis for physics-based practices. A physics-based practice is essentially a workflow regime designed to identify what is limiting us from increasing performance, from a physics-based point-of-view. Introducing such a

workflow facilitates a continuous innovation scheme, by identifying root causes limiting performance, and safely redesigning the current system to extend the performance limiter at a higher operating setting. In 2004, ExxonMobil launched a three-month pilot putting many of these MSE-based practices and surveillance techniques into action. The result exceeded expectation as several performance limiters in increasing WOB were identified, and average ROP was increased by 133% on six selected rigs, establishing field records on 10 of 11 wells (Dupriest & Koederitz, 2005). Right now, there lies a great opportunity in “operationalizing” what we already know in physics-based practices and MSE-based surveillance workflows into automated processes and algorithms to assist drilling optimization engineers and achieve differentiating performance gains.

This thesis will look into the concept of a web-based application to support the drilling engineers in post-drill analysis. The application will have a foundation in MSE-based surveillance techniques to identify optimal drilling parameters through lithology-specific surface regression, and attempt to automatically identify intervals with bit dysfunction from expected responses in MSE and other operating parameters. Important aspects such as a detailed understanding of MSE and dysfunction diagnostics (MSE theory), statistics and mathematical modelling of representable trends in data (Detecting change), and several pseudo-codes for automated dysfunction diagnostics and associated scoring systems (Methods chapter) will be covered in this thesis. In addition, a substantial focus is set on data acquisition & preparation, software architecture & user interface (Methods chapter). Conclusively, results and discussion from automated dysfunction diagnostics and optimal parameter selection from a set of selected wells provided by an anonymous operator are presented. Creating a tool which programmatically assists the engineer in identifying performance limiters may effectively lessen the workflow when scanning through reference wells. The tool may also quantify data and allow the engineer to initiate a cost justification process for a potential BHA redesign to extend the performance limiters in the next well. Digitalization does not belong to the future anymore and becoming digitally enabled might be the key factor for operating competitively in the oil & gas industry.

2 Theory

2.1 Mechanical specific energy

A comprehensive literature study of mechanical specific energy was conducted in 2018 by (Berge-Skillingstad & Anderssen, 2018). The following subchapters contain a summary of the most relevant findings in MSE theory for this thesis.

Teale defined in 1965 specific energy being “the energy required to excavate a unit volume of rock” (Teale, 1965). Mechanical specific energy or (MSE), is derived by calculating the required torsional and axial work over the drill bit to remove a volume of rock, and dividing that energy by the volume of rock removed (1). Through Teale’s experiments, the minimum baseline value of MSE was found to numerically equal compressive rock strength in atmospheric drilling conditions. This would develop into a strong diagnostic tool for measuring rock cutting efficiency in laboratory environments. The derived equation for MSE is depicted in equation (1) & (2) below, and the full derivation may be found in the Appendix of (Berge-Skillingstad & Anderssen, 2018).

$$MSE = \frac{\text{Input Energy}}{\text{Rock Volume Drilled}} \quad (1)$$

$$MSE = \frac{4 \times WOB}{\pi \times Dia^2} + \frac{480 \times TOR \times RPM}{ROP \times Dia^2} \quad (2)$$

Whereas the metric units are given below:

MSE	Mechanical Specific Energy	Pa
WOB	Weight on Bit	N
TOR	Torque	Nm
RPM	Rotations per Minute	1/min
ROP	Rate of Penetration	m/hr
Dia	Hole Diameter	m

Mechanical specific energy is further separated between downhole and surface MSE, depending on the origin of measurements. Surface torque measurements contain frictional drill string losses, which may be considerable in elongated wells. Downhole torque provides a superior valuation of MSE contrary to surface, by decoupling the energy losses in the string and only depicting the actual energy spent in the rock cutting process. Downhole MSE with bit torque is the preferred measure for inclined wells. If no downhole torque data is available and a motor is present, bit torque should be estimated using motor performance factors and motor differential pressure. Equations for bit MSE with and without motor factors are presented in (Berge-Skillingstad & Anderssen, 2018), and bit MSE with a motor present is also depicted in equation (3) below.

$$MSE_{bit} = \frac{480 \times (F_{TOR} \times \Delta P_{Motor}) \times (RPM_{TpDr} + (F_{MotorRPM} \times Q))}{Dia^2 \times ROP} + \frac{4 \times WOB}{\pi \times Dia^2} \quad (3)$$

Whereas the additional metric units are given below:

F_{TOR}	Motor torque factor	Nm/Pa
RPM_{TpDr}	Top drive RPM	1/min
$F_{MotorRPM}$	Motor RPM factor	1/m ³
ΔP_{Motor}	Pressure difference over motor	Pa
Q	Pump rate	m ³ /min

2.1.1 Mechanical specific energy & efficient drilling operations

MSE has evolved from a proven diagnostic tool in laboratory experiments, to also being capable of evaluating drilling efficiency and diagnosing bit dysfunction in field applications (Dupriest & Koederitz, 2005). For the purpose of this thesis, we define drilling efficiency as a measure of required energy to drill a given volume of rock. In other words, MSE is the underlying metric for quantifying efficient drilling operations. If the bit is drilling 100% efficient, MSE should roughly equal the confined compressive rock strength in laboratory drilling conditions.

Drilling efficiency is, however, highly dependent on the downhole pressure environment, and may reduce by a factor of two- to threefold compared to atmospheric laboratory drilling conditions (Rafatian, Miska, Ledgerwood, Yu, & Ahmed, 2009). The reduction in rock cutting efficiency under increasing downhole pressure conditions is recognized by several factors. Chip hold down forces, large confined compressive rock strengths and transition from brittle to ductile failure mode being among them. The transition from brittle to ductile failure mode is recognized as the most severe factor, responsible for reducing efficiencies by up to 70% (Rafatian et al., 2009). Ductile failure mode is induced by a combination of overbalanced mud, dilatant expansion from compressing and shearing the underlying rock of the bit, and poor mud filtration. The compressive strength of the rock matrix in the underlying rock becomes so large from differential effective stress, that the cutter must plow through the rock, completely destroying the grain structure and reconstituting the rock on a granular level, creating the characteristic PDC ribbons.

Conclusively, the minimum baseline value of MSE from efficient drilling operations in field applications will rarely approach the same compressive rock strength such as in laboratory drilling conditions. Instead, in efficient drilling operations, downhole MSE is expected to approach and stabilize to level more equal to a two- or threefold increase of the confined compressive rock strength.

2.1.2 Regions of drilling efficiency

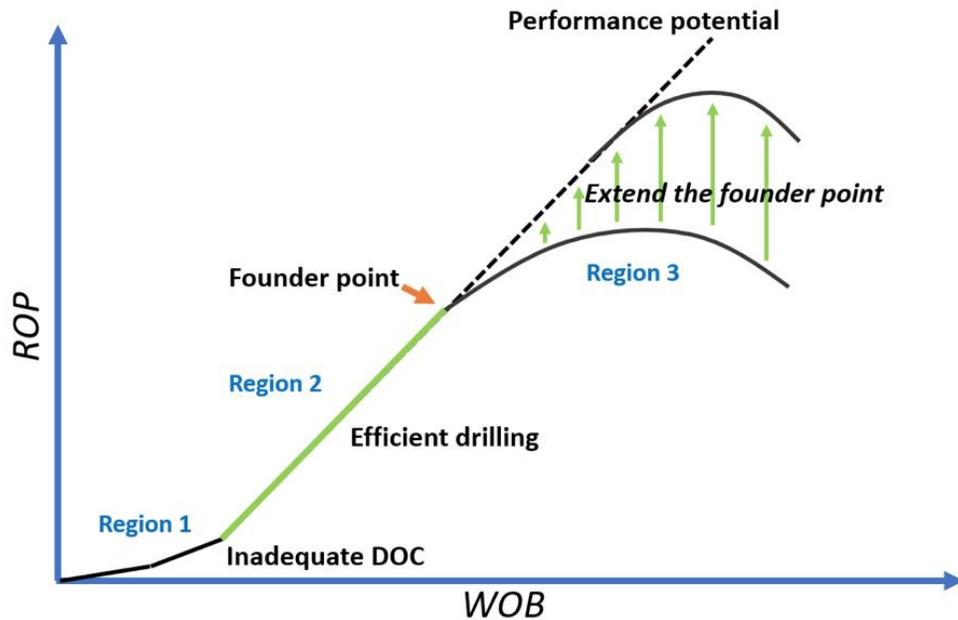


Figure 2.1: Characteristic WOB vs. ROP drill-off curve
(Berge-Skillingstad & Anderssen, 2018)

The traditional “drill-off” curve, depicted in Figure (1) above, neatly represents three regions of drilling efficiency, which is uniquely tied to bit dysfunction and characteristic MSE responses (Dupriest & Koederitz, 2005).

Region 1 is dominated by inefficient drilling from an inadequate depth of cut (DOC). An inadequate DOC causes low bit aggressiveness from “chamfer” drilling. Bit whirl is likely to dominate this region due to the bit cutters not being fully buried into the formation. MSE in this region is very high and erratic, indicating that large amounts of energy put into the system are being dissipated elsewhere in the form of excessive vibrations. Any additional energy input into the system will disproportionately increase the output of the system (ROP).

Region 2 represents efficient drilling. Drilling efficiency improves dramatically with sufficient WOB to bury the cutters, adequately constraining lateral bit movement and preventing bit whirl.

Maximum drilling efficiency is achieved in region 2 as any additional energy input into the system will always increase the output of the system (ROP) proportionately. In other words, ROP will always increase linearly to any increase in WOB or RPM in an 100% efficient system. MSE and drilling efficiency remains constant in region 2, up until the founder point.

Region 3 is beyond the system founder point. Onset of the founder point represents the maximum WOB setting that allows for efficient drilling in the current system. Parameter settings beyond the system founder point is dominated by bit dysfunction. Characteristic of this region is very large, disproportionate, and often erratic changes in MSE. A constraint in the energy transfer from the bit to the rock formation causes large amounts of energy to be dissipated elsewhere. Any additional energy input into the system will disproportionately change the output of the system. It is vital to avoid bit dysfunction due to excessive misplaced energy from inefficient drilling that may severely damage the bit and BHA. Drilling in Region 3 may consequently cause premature bit trips due to excessive wear.

2.1.3 Analyzing trends in MSE

Analyzing and baselining trends in MSE will conveniently identify which region of the drill-off curve the bit is being operated in. Real-time MSE surveillance facilitates a continuous detection of changes in drilling efficiency which allows for an optimum selection of drilling parameters by sufficient parameter exploration or so-called “step-tests”. In other words, MSE will indicate if a change in a drilling parameter is moving you closer to, or further away, from the maximum expected performance. Post-drill MSE analysis may provide quantitative data to identify drilling inefficiencies and bit dysfunction in historical reference wells, enabling a cost-justification process to propose changes in the current system to extend the founder point of the next well (Dupriest & Koederitz, 2005). Extending the founder point which onsets a bit dysfunction in any well, may improve drilling performance and BHA tool and bit longevity considerably.

Operational considerations and the ability to diagnose several bit dysfunctions from characteristic MSE responses and analysis are presented in detail in (Berge-Skillingstad & Anderssen, 2018).

MSE is expected to change adaptively as formation lithology and compressive rock strength changes. However, the incremental changes in MSE from different compressive rock strengths pales in comparison to the large (and often erratic) fluctuations in MSE from transitioning between efficient and inefficient drilling (Dupriest & Koederitz, 2005). Baselining trends in MSE combined with pattern recognition and analysis to determine drilling efficiency, onset and type of bit dysfunction, proves to be a strong operational diagnostic tool in field applications.

The basis of this thesis will be to computerize several aspects within the MSE surveillance workflow process through smart algorithms and data analysis to begin to automate the process of drilling.

2.2 Drilling a linear borehole

Another purpose of this thesis will be to discuss how a smart diagnostic application with a foundation in MSE can improve conventional well planning and execution phases, and ultimately assist in automation of future drilling processes. At the heart of every automation process is reproducibility. Any non-linear response to a change in an operating parameter can be considered non-reproducible. As previously mentioned, drilling efficiency is characterized as a measure of required energy to drill a given volume of rock, where ROP has a linear response to any change in WOB or RPM in the traditional drill-off curve Figure 2.1. In other words, any additional energy input into an efficient system should always increase the output of the system proportionately. This is a reproducible effect. A proportionate response to any change in an operating parameter dictates a predictable outcome, which may be programmed and automated. Inefficient drilling operations cause non-linear responses to changes in operating parameters, ref. region 1 & 3 in Figure 2.1. Non-linear responses cannot be automated in an effective manner due to non-uniqueness of solutions. In conclusion, efficient drilling conditions should be the backbone of every autonomous operation.

Inefficient drilling operations do not only create non-linear responses to changes in drilling conditions, but are also responsible for considerable borehole quality issues. E.g. bit whirl which may create significant spiral borehole patterns from excessive bit tilting and side cutting. Spiral borehole patterns may result in anything from significant weight transfer issues to the bit, inability to run casing or pipe to bottom from tight hole responses, or poor hole cleaning and cementing operations, just to mention a few. In effect, inefficient drilling operations may also create non-gauge boreholes, which may also add to the specter of problems associated with automating “non-linear” boreholes.

“*A linear borehole is defined as a borehole with linear responses to changes in operating conditions.*”

Due to MSE being the underlying metric for quantifying drilling efficiency, MSE should also be the underlying metric for facilitating drilling automation, being a unique reference point in indicating borehole “linearity”.

2.3 Detecting change

2.3.1 Heuristic vs. analytical solutions

Automatically detecting changes in continuous streams of data invokes the need for sophisticated change detection tools. Fundamentally, one may separate between analytic or heuristic methods when detecting changes in data sets. Analytical methods rely on statistics and mathematical ingenuity when assessing changes in data sets, while heuristic methods rely more on the human ability to eye-ball a decent fit using an empirical approach. One may say a heuristic method is a trade-off between a more “accurate” solution with a more “efficient” solution, i.e. a shortcut to a faster, but sufficient solution.

For instance, imagine a signal “S” is to be fitted with a single baseline which best represents the data set, ref. Figure 2.2. Choosing such a baseline may either be done by a heuristic method or an analytical method. In this case, heuristically “eye-balling” different values of “H” can after a certain amount of tries result in a line which may seem to be a good fit for this data set (e.g. “H” = 5). On the other hand, choosing such an empirical fit to this data set may not be as favorable to other data sets. Analytically choosing the baseline by calculating the average of the data set may provide a slightly inferior answer (e.g. “A” = 4.8) to this data set, however may prove more accurate in future data sets. In this simplistic example it is not clear which method is superior, the heuristic or analytical approach, as the nature of the future data sets are unclear.



Figure 2.2: Example of different straight-line fits to an input signal

In theory, analytical methods should be superior to heuristic methods as they produce standardized solutions and are statistically correct. However, with enough identifiable similarities between data sets and large databases of “training” data, heuristic methods may provide computational favorable algorithms as composed to the often more complex analytical solutions. For some data sets it will be quite easy to assess the best method for an analytical solution, especially if the solution is as simple as a straight-line, however, complex situations might occur when the solution requires more sophisticated aspects than a single-value straight-line. This is where the discussion of heuristic vs. analytical solutions occur (Pearl, 1984). Below is a set of key factors to consider when choosing between heuristic and analytical methods as described by (Kunche, 2016).

- **Optimality:** When several solutions exist for a given problem, does the heuristic method guarantee a best-fit solution? Is it actually necessary to find the best solution?
- **Completeness:** If more than one solution exists for a given problem, will the heuristic method find all of them? Are all solutions actually needed?
- **Accuracy and precision:** Will the heuristic method provide a sufficient confidence to the solution? What is the error to the real solution?
- **Execution time:** Some methods converge faster to the solution than others, is the current method significantly slower or faster than others?

2.3.2 Basics on change

Building data interpretation algorithms to manage and understand the nature of the different drilling dynamics measurements relies on requirements and functionalities. (Basseville, Espiau, & Gasnier, 1981) stated the need to define a set of primary functionalities and properties. Below is a list of the main attributes needed for the drilling optimization application discussed in this thesis:

- Detecting a change in the mean of a signal
- Detecting a change in signal frequency
- Detecting asymmetry in signal cycles
- Filtering edge points of a signal
- An effective algorithm, i.e. low computational time

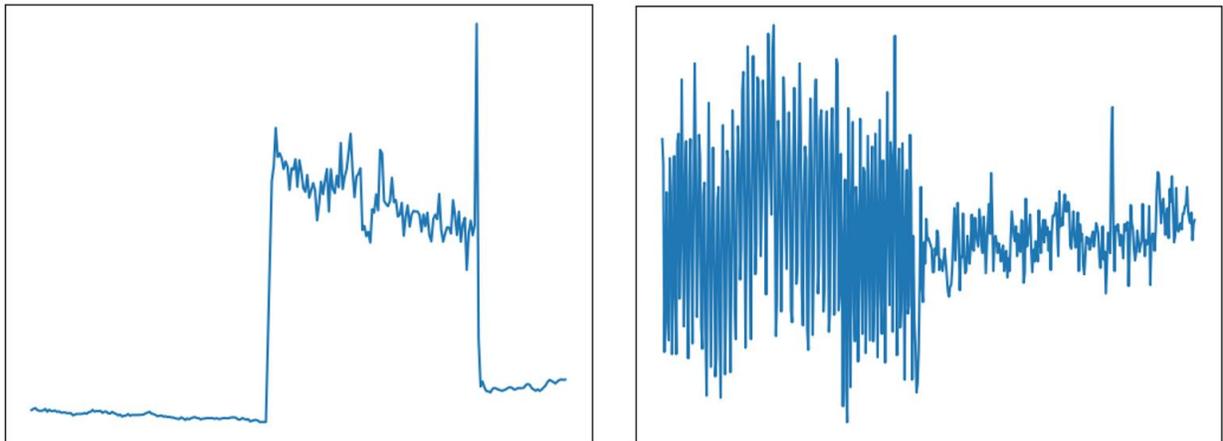


Figure 2.3: Change in mean (Left) & Change in frequency (Right)

When building algorithms for detecting change, it is important to consider the format and structure of the data evaluated. Change detection is split into offline or online detection, dependent on if the data is received as a whole or continuously as a time series. Online or offline change detection will affect the requirements of the algorithm significantly.

Online change analysis:

- Data is received as individuals or in bulk data points.
- Data is received continuously and depending on the sample rate, a requirement of low execution time is necessary, meaning one must aim for quick detections once new data is received to the existing stream.
- Since the focus is on how the data changes real-time, focus is often on recent data points.

Offline change analysis:

- All data is received and processed at the same time.
- As current stage is post-analysis, one can tolerate slower execution time for higher accuracy.
- Both big-picture and detailed changes may be of interest.

2.3.3 Changepoint detection theory

From a statistical point of view, one may separate between two different techniques for change detection, Bayesian analysis, explored by (Smith, 1975) and Maximum Likelihood techniques, (Hinkley, 1969). (Basseville, 1981) examined how the Techniques of Hinkley could be used for detecting a change in mean of a signal of independent variables, as illustrated in Figure 2.3, and which coincides with the primary attribute of the algorithm used in this thesis.

When discussing change detection, it is convenient to introduce the principle of changepoints and the theory behind the maximum likelihood technique. For a given ordered sequence, $Z_{1:N} = (z_1, z_2, \dots, z_N)$ on a time interval $\tau \in (1, 2, \dots, N)$ one may say that a changepoint exists on the sequence if the statistical properties of (z_1, \dots, z_τ) and $(z_{(\tau+1)}, \dots, z_N)$ are different, meaning a change has occurred. For M changepoints, each with a position on the time sequence (τ_1, \dots, τ_M) , by definition they must all be an integer $M \in (1, \dots, N - 1)$. The boundary conditions $\tau_0 = 0$ and $\tau_{M+1} = N$ and the time series is defined as strictly increasing, i.e. $\tau_i < \tau_j$ for $i < j$, allowing for a segmented data set split into $M+1$ segments. The i th segment of the sequence will consequently contain the data points $Z_{(\tau_{i-1}+1):\tau_i}$. Each segment is defined by its corresponding parameters, which is estimated as (θ_i, φ_i) , where θ_i represents the data points of interest and contain a statistical change and φ_i represents the nuisance parameters (Killick & Eckley, 2014).

“Definition: Changepoints are abrupt variations in time series data.”

With increasing lengths of data comes the need for identifying multiple changepoints in a single data stream. As presented by Killick et. al., one of the most common approaches to multiple-changepoint analysis comes from minimization of a predefined cost function and a penalty term, ref. (4)

$$\sum_{i=1}^{M+1} [C(Z_{(\tau_{i-1}+1):\tau_i})] + \beta f(M) \quad (4)$$

Where,

Z	Signal function of independent data points
C	Cost function
τ_{i-1}	Time step for previous changepoint
τ_i	Time step for current changepoint
M	Number of changepoints
$f(M)$	Penalty function
β	Penalty constant

For this thesis, a method of negative log-likelihood is chosen to define the cost function through an empirical mean. One could equivalently choose the opposite, i.e. the maximum log-likelihood as applying the logarithm of a function will still result in the same increasing or decreasing trend. Eq. (5)-(7) below illustrates how the maximum log-likelihood method differs from a negative log-likelihood method. If a maximum log-likelihood method was used, a maximizing technique must have been implemented. As most methods for optimizing function are based on minimization, using the negative log-likelihood is more convenient (Baum & Sylvester, 1990).

$$\text{Max}(\log LR_+) \equiv \text{Min}(\log LR_-) \quad (5)$$

$$LR_+ = \frac{\text{sensitivity}}{1 - \text{specificity}} \quad (6)$$

$$LR_- = \frac{1 - \text{sensitivity}}{\text{specificity}} \quad (7)$$

Where,

LR_+	Positive likelihood ratio
LR_-	Negative likelihood ratio
Sensitivity	Ratio of true positives to those of a sequence that are actual as such
Specificity	Ratio of true negatives to those of a sequence that are actual as such

The typical approach to evaluate changepoints on a sequence is through the logarithm of the likelihood ratio by each consecutive interval. By using the logarithm, the computation is more efficient as the logarithm of products of probability densities facilitates an additive equation rather than a set of multiples. By principle, the procedure for negative log-likelihood compromises two steps, where the probability density of the intervals are first computed and then the ratio is calculated (Aminikhanghahi & Cook, 2017). More specifically for this thesis, the logarithm of an empirical mean multiplied with the length of the interval is used to define the cost function which is minimized to obtain optimal changepoints, ref. Eq. (8).

$$C(Z_I) = |I| \log \bar{\mu}_I \quad (8)$$

Where,

I	Length of the interval defining the sub-signal
Z_I	Signal function
C	Cost function

$\bar{\mu}_I$ Empirical mean of the interval I

To understand more intuitively the role of the cost function and how this models the input signal, one may consider the cost function as the “cost” of using different magnitudes of penalty. Using a high penalty will result in a high “cost” which can be represented as a model which does not fit the input signal very well. Using a lower penalty will then result in a more correct model but will also decrease the computational efficiency. Figure 2.4 depicts how the negative log-likelihood function responds to changes in the penalty parameter, i.e. displaying how a penalty must be included to balance the need for an accurate model, but avoid overfitting.

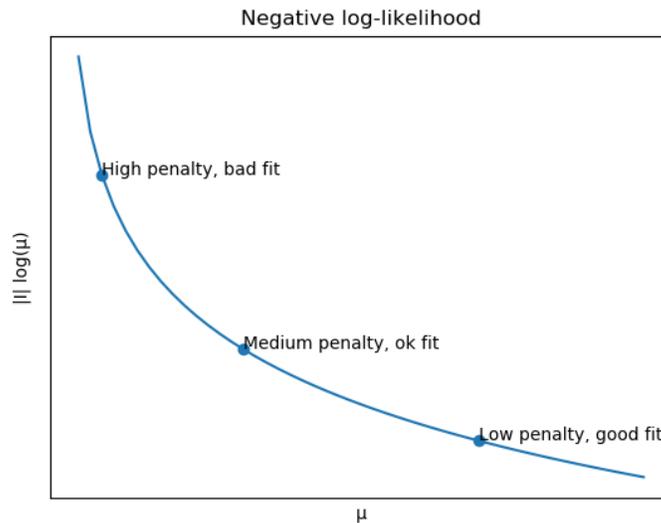


Figure 2.4: Negative log-likelihood response to changes in penalty.

2.3.4 Methods for detecting optimal changepoints

Considering the fact that there are several different options for multiple changepoint detection, it may be interesting to compare some of the different methods to understand the reasoning behind the choice of method for this thesis.

Due to its intuitive nature of comprehension, the binary segmentation method for changepoint detection, presented by (Scott & Knott, 1974) is one of the more established methods. As visualized in Figure 2.5 it applies one changepoint at a time for each sub-interval, by finding the best fit on each segment. By doing this to each sub-interval until a penalty condition is satisfied, a set of optimal changepoints are identified with an efficiency of $O(n \log(n))$, however according to (Killick, Fearnhead, & Eckley, 2012) it has a flaw in accuracy as it may not approach the global minima of optimal changepoints, ref. (4).

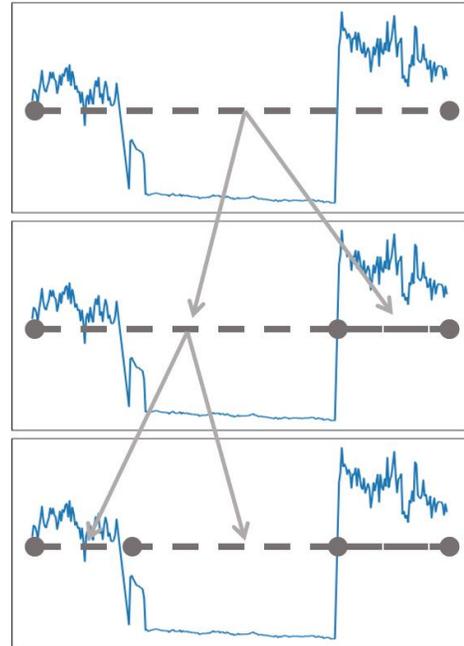


Figure 2.5: Binary segmentation method visualized. Segmentation performed according to magnitude of change.

Another method, presented by (Auger & Lawrence, 1989), known as the segment neighborhood method, comprises a less efficient, but exact method in accuracy.

By using dynamic programming, it is able to search the entire segmentation space by setting a maximum number of segments, denoted Q . Through calculation of the cost function of all possible segmentations, the global minimum can be established and the optimal changepoints picked out. This with an efficiency of $O(Qn^2)$ (Killick et al., 2012).

2.3.5 PELT

For this this thesis, a pruned exact linear time method (PELT) is chosen, which fundamentally builds on the work of (Jackson et al., 2005) through the optimal partitioning method. The

primary idea behind the PELT method is to prune the interval by removing those values of time points, τ , which can never be a minimum of the solution obtain through minimization of eq. (4) (Multi changepoint detection). By doing so and still using an exact method, one obtains a method that is both efficient and accurate. According to (Killick et al., 2012) PELT is asymptotically faster than the binary segmentation method and under certain conditions it can perform with $O(n)$. Although for most signals it will be slower, but due to its exact nature, PELT is the preferred method.

For the interest of the reader, (Killick et al., 2012) presents in further detail how the PELT method builds on the optimal partitioning method. Below is a summarized theorem by Killick et al. to simplify how the pruning is performed.

Theorem 1.1.:

- *“If we assume that a changepoint is introduced to a data sequence, the cost of total cost of the sequence is reduced.”*

Let $F(s)$ be denoted as the minimization of the data points $Z_{1:s}$.

$$F(s) = \min_{\tau \in \tau_s} \sum_{i=1}^{M+1} [C(Z_{(\tau_{i-1}+1):\tau_i}) + \beta] \quad (9)$$

Formally, a constant K exists, such that for all $t < s < T$,

$$C(Z_{(t+1):s}) + C(Z_{(s+1):T}) + K \leq C(Z_{(t+1):T}) \quad (10)$$

If the following expression holds at a future time $T > s$,

$$F(t) + C(Z_{(t+1):s}) + K \geq F(s) \quad (11)$$

, t can never be the optimal last changepoint prior to T .

Below is a flowchart describing how the PELT algorithm works in principle, ref. Figure 2.1. Using the predefined cost function to measure the data fit and penalty term to define the iterating parameters and boundary conditions, the optimal changepoints can be calculated.

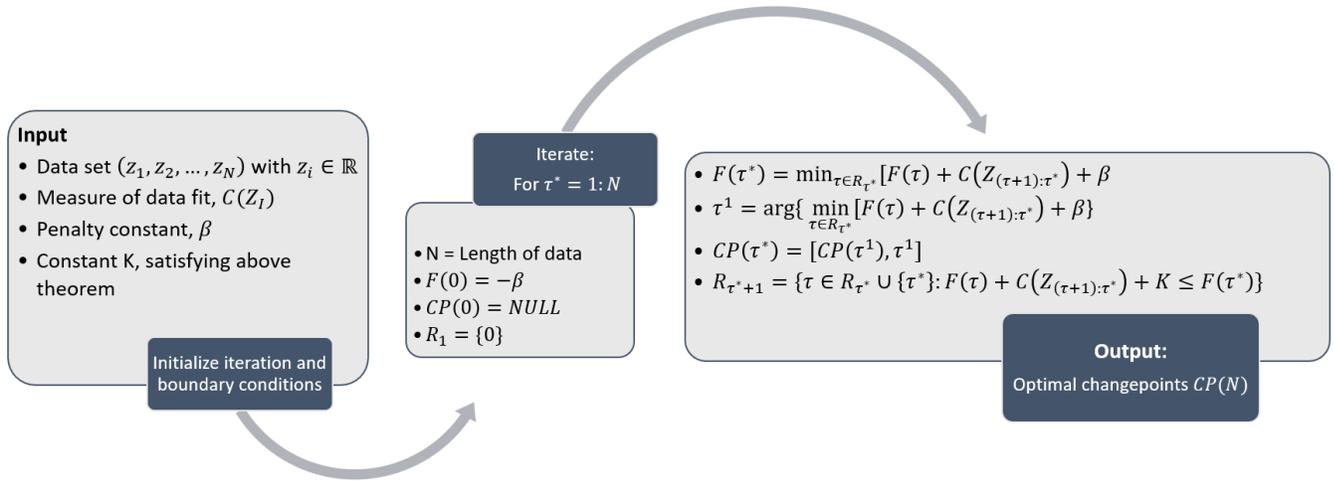


Figure 2.6: PELT method outlined step-by-step.

2.3.6 Convolution to smooth noisy data

An alternative or possibly supplement to the changepoint algorithm is to use a noise reduction filter. A smoothing function may be used to obtain filtered data points, removing potential outliers and noise from physical measurements. A noise filter function is implemented in the application, however is not used to obtain any results in the *Results* chapter.

2.4 Regression for optimization

In this thesis, a quadratic model of surface regression is applied, defined in the same way as for a straight line. The basis of the regression function is derived from the quadratic solution, or polynomial function with (k=2), depicted in eq. (12) below.

$$y = a_0 + a_1x + \dots + a_kx^k + \varepsilon^k \quad (12)$$

Where,

(x_i, y_i)	Input data points
a_0, \dots, a_k	Polynomial coefficients
n	Total number of data points
k	Degree of polynomial
ε^k	Error

The residuals in the models are defined in eq. (13) as,

$$R^2 = \sum_{i=1}^n [y_i - (a_0 + a_1x + \dots + a_kx^k + \varepsilon^k)]^2 \quad (13)$$

For further reading on the method, the reader is encouraged to look to (Weisstein, 2019). On matrix form, the above equation leads to the following set of equations for the quadratic version,

$$\begin{bmatrix} 1 & x_1 & \dots & x_1^2 \\ 1 & x_2 & \dots & x_2^2 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x^n & \dots & x_n^2 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \quad (14)$$

To simplify the operation, the equation can be written as,

$$\mathbf{y} = \mathbb{X}\mathbf{a} + \boldsymbol{\varepsilon} \quad (15)$$

Multiplying the transpose of the solution, \mathbb{X}^T allows to solve the solutions as,

$$\mathbb{X}^T\mathbf{y} = \mathbb{X}^T\mathbb{X}\mathbf{a} \quad (16)$$

The above matrix equation can be solved numerically or inverted directly to the solution vector,

$$\mathbf{a} = (\mathbb{X}^T\mathbb{X})^{-1}\mathbb{X}^T\mathbf{y} \quad (17)$$

Even though the method is considered non-linear in the independent variables $x^{k=2}$, it can still be considered a linear method due to linearity in the dependent polynomial coefficients,

a_0, \dots, a_k .

2.4.1 Adjusted coefficient of determination: error analysis

When performing regression, one must check the residuals of the model, i.e. how well the model fits the input data set. The adjusted coefficient of determination, commonly spoken as adjusted R-squared, can be a good statistical measure to do just this. The reason for using the adjusted R-squared and not the standard R-squared error is that the R-squared error increases with increasing explanatory variables, i.e. predictors. Using the adjusted R-squared accounts for this, as well as problems that may occur with noisy data (Theil, 1961).

Adjusted R-squared, denoted \bar{R}^2 , is outlined in equation (18) & (19):

$$\bar{R}^2 = 1 - \frac{\text{Variation of residuals}}{\text{Total variation}} \quad (18)$$

$$\bar{R}^2 = 1 - \frac{\frac{1}{n-p-1} \sum_{i=1}^n (y_i - \hat{y})^2}{\frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2} \quad (19)$$

Where:

\bar{R}^2	Adjusted R-squared error
n	Total data points on interval
p	Number of predictors, i.e. free variables
y_i	Data point, i
\hat{y}	Predicted value of y_i
\bar{y}	Mean of data points across interval

(Miles, 2014)

3 Method

A comprehensive literature study about the physics of each bit dysfunction and detailed procedures of dysfunction diagnostics, MSE surveillance techniques, and physics-based practices, are presented in (Berge-Skillingstad & Anderssen, 2018). The report maps the potential of operationalizing physics-based practices through automation and smart algorithms. From the theory outlined in this project comes application in practice presented below. In this section, the fundamental steps necessary to develop a smart application for improved reference well studies and automated dysfunction diagnostics are outlined. The section contains the choice of programming language, software architecture, data acquisition, data preparation, workflows, flow charts, pseudo-codes, and the application itself.

3.1 The choice of programming language: Python

From the authors point of view and proficiency, the high-level programming language Python provides intuitive tools for analysis and several easy-to-access open-source code packages with existing functions. A flexible environment is paramount when dealing with the complexity of different file types, requirements, analysis & visualization, and connectivity between existing work and ongoing work. The indentation requirements of Python, meaning the significant use of whitespace, provides the necessary clarity to comprehend and keep order on long strips of code. This is also achieved through Visual Studio Code, the chosen editor for building and debugging code, which facilitates the use of smart extensions to code cleaner and more efficient. In addition to the clarity provided, Python features a dynamic system with automatic memory management, which improves efficiency by allocating only the required memory per operation, for a given process. An example of this is the package system implemented in Python, requiring a code to import only the necessary packages and functions to run, preventing unnecessary processes to be initiated. Multiple paradigms are also supported, e.g. object-oriented programming, functional and imperative programming, consequently different ways of programming can easily be integrated together (Kuhlman, 2011).



Figure 3.1: From Python (Foundation)

The following fundamental concepts define how Python is structured:

- Module*** A file with a .py ending which contains variables and Python functions, i.e. a piece of executable code which may be run through a Python execution terminal.
- Package*** A package contains sub-modules used in similar settings. Packages contain an `_init_.py`-file, which allows the user to split larger packages into sub-packages or sub-modules. The package principle is what makes Python ideal for simple and quick operations.
- Library*** Often used in similar settings as packages. Conceptually, packages and libraries are referred to as the same, meaning for this thesis there is no difference between them.
- Framework*** A collection of libraries is referred to as the framework, i.e. what defines the structure of the code.
- Pip*** The package-management system utilized to install and upgrade packages associated with Python.

(Foundation, 2019)

3.1.1 External Python packages

In addition to the standard packages included in Python, some external packages must be imported for more complex operations. The main packages and the associated functions used in this thesis are presented below, together with their functionalities.

3.1.1.1 NumPy

NumPy's main function for this thesis is scientific computations and operations involving array objects, used for storing data streams. The package comes with several useful functions such as

variance, median, mean, and simple optimization functions. The package creates the environment needed for most functions implemented in this thesis. Converting data into N-dimensional NumPy arrays is both efficient and provides easy use of operations during calculations.

3.1.1.2 Matplotlib

For most of the results generated as stand-alone figures and graphs, the Matplotlib package and its associated functions in the PyPlot branch provides the graphing tools in a framework with several similarities to MATLAB. The object-oriented way of programming provides easy-to-use graphing tools where the plot parameters can be generated with only a few lines of code. PyPlot provides a pleasing basis for visualization in an offline plotting mode. PyPlot is however not well suited for HTML, and is for that reason not implemented in the web application itself.

3.1.1.3 Dash

Dash is a library and framework for creating analytical web applications. The Dash library has a declarative and reactive architecture which makes building complex applications with elements of cross filtering, data analysis, and data visualization with ease. Dash is capable of comprehensive customizable data visualization with CSS in the web browser which makes it highly suitable for creating interactive web applications with the full power of Python behind it, and is why we chose it for developing our application.

Dash runs on Flask and React libraries, which is the JavaScript user-interface maintained by Facebook and is deployable in multi-user applications. Most cloud server providers today are tailor made for Flask apps which ultimately makes deployment and scalability easy (Parmer, Johnson, Parmer, & Sundquist, 2018).

3.1.1.4 Plotly

Plotly is the graphing library used for HTML, and for interaction with Dash. Plotly provides superior online plotting functionalities, such as the web-service option for hosting graphs. Graphs in Plotly are organized with a ‘data’ structure, defining the type of graph and the data points, in addition to a ‘layout’ which defines the graph aesthetics and properties.

3.1.1.5 *Pandas & CSV*

Data files exported from different drilling data software or Wellsite Information Transfer Specification (WITS) data records may typically organized as Comma Separated Values (.csv). The Pandas and CSV python libraries provide basic functions for editing and restructuring these data files. CSV contains reading and writing functions for .csv data files. Pandas extracts and structures data from .csv-files into easy to access data frames which are stored in the memory. Pandas provides an easy interface for reading and editing data elements in data frames, which can easily be converted into lists or arrays for further operations.

3.1.1.6 *SciPy*

As a supplement to more complex and sophisticated scientific computing, SciPy provides a library with a whole range of different functions for optimization, regression, interpolation, linear algebra and more. For the application of this thesis, the SciPy package is used for linear algebraic calculations.

3.1.1.7 *os*

This module makes operating system dependent functionalities, independent of operating systems. In other words, if the application is to manipulate paths or read files from certain database folders, this module may navigate pathways regardless of operating system, i.e. Windows, Linux or Mac.

3.1.1.8 *Json*

This module enables encoding and decoding of json-files. In other words, if the application needs to temporary store data in the browser session, the json module allows for encoding and dumping data in the web-browser, which may be decoded and loaded into the application at a later stage.

3.1.1.9 *Ruptures*

From the theory on changepoint algorithms, defining the optimal sub-intervals used in changepoint detection is a complex process. The Ruptures library provides the necessary modules for offline changepoint analysis and is fundamentally organized around the two base classes, *BaseCost* and *BaseEstimator*. As outlined in the theory, a changepoint algorithm relies

on analysis of a cost function which defines the cost of a sub-segment and defines the optimal changepoints. When importing the Rutures package, the `_init_.py`-file initializes an estimator, i.e. the BaseEstimator.

The estimator has the following arguments:

- 'model'*** The module defines the cost function used to approximate the error, i.e. the “cost” on the segment.
- 'cost'*** Customized cost function used in the detection scheme with a BaseCost instance.
- 'jump'*** Reduces computational time spent on the algorithm. Defines the possible indexes that a changepoint may occur on a multiple of ‘jump’.
- 'min_size'*** To reduce the number of changepoints on a short interval, ‘min_size’ defines the amount of data points between each changepoint.

(Truong, 2018)

3.2 Software architecture

In this sub-section we will first introduce basic concepts in building a web application, then the analytical web application itself, its architecture, user interface, functionalities, and different available modules for data analysis.

The application is first initialized by defining the application layout or graphical user interface (GUI). The GUI is built up by Dash dependencies, which are primarily HTML components. The HTML components compose the layout of the application by defining a set of HTML objects in a structural hierarchy which determine the layout position. Each HTML or Dash component is essentially a Python class containing a list of distinct properties describing the attribute, style and class of the component. In other words, Dash dependencies convert Python classes into HTML code, enabling to compose a web application layout using only Python. Below is an example of a set of HTML components written with Dash dependencies. Note that these components define the “skeleton” or layout of the application.

```
#Well Dropdown Menu

html.Div([

    html.Div([

        dcc.Dropdown(

            id='choose_well',

            options=[{'label': i, 'value': i} for i in database_list],

            value=database_list[0] ),

        ],style={'width': '50%','float':'right'},className='six columns'),

#Selected Well Textbox

    html.Div([

        html.Div(id='chosen_well'),

        ], style={'margin-left':5,'color': '#FFFFFF'},className='six columns'),

    ],className='row'),
```

Note that the significant use of whitespace and tabbing defines the structural hierarchy of the HTML components. Each HTML object is defined as a '*html.Div*', which may contain the set of attributes and class names as discussed above. The class name defines the spacing of the objects in the layout. For instance, the top *html.Div* in the piece of code above contains a class name equal to *row*, defining that every component under this level of hierarchy will be defined within a row in the application. The next two objects (*html.Div*) are *six column* class names, meaning they will fill six columns spacing each within the overlying defined row in the application.

The aesthetic design of each HTML component is defined through the *styles* property which initializes custom CSS builds. The components are actually React *propTypes*, which are written in JavaScript but encoded and unpackaged into Python through Dash dependencies. The strength of unpackaging JavaScript components into Python through Dash dependencies is the ability to combine Front-End and Back-End web development into one language and holistic system, making analytical web development easier for non-software engineers (Dash.plot.ly, 2019).

Furthermore, Dash components represent interactive functionalities within the application, and can be anything from a dropdown menu, range sliders, graphs, buttons and more. The components have a unique identification key which are accessible by callback functions that execute a Back-End process or pre-defined function. The callback functions are continuously fired when any input element changes or when prompted by the user in the GUI. Below is an example of such a callback function.

```

#EXTRACT GLOBAL DATA IN DATAFRAME:

@app.callback(
    dash.dependencies.Output('intermediate-value', 'children'),
    [dash.dependencies.Input('choose_well', 'value')])
def update_global_data(well_chosen):
    # extract data

    df = pd.read_csv(current_dir+database+well_chosen,low_memory=False).to_json(orient='split')

    # read pandas to json & temporary store data in browser

    return df

```

The callback function above accesses the predefined dropdown menu from the layout by calling its identification key *'choose_well'*. Upon selecting a well from the dropdown database list, the callback fires automatically, reading a .csv file from the database and temporary storing the data as a .json file in the browser memory to be accessed at a later time through different callbacks in the application.

To put it short, the HTML components in the application layout define the user interface. The callback functions make the application interactive and allows for computationally intensive processes in the web browser to utilize the full power of Python behind it.

Below is a flowchart of the application architecture. The application is essentially split into two main functionalities. A parameter optimization module, to identify trends in parameter selection through lithology specific surface regression, and a diagnostics part to identify intervals with bit dysfunction. Depending on the desired functionality and output of the application, the user is prompted to follow certain steps depicted by the flowchart below in order to interact and fine-tune the results of the analysis.

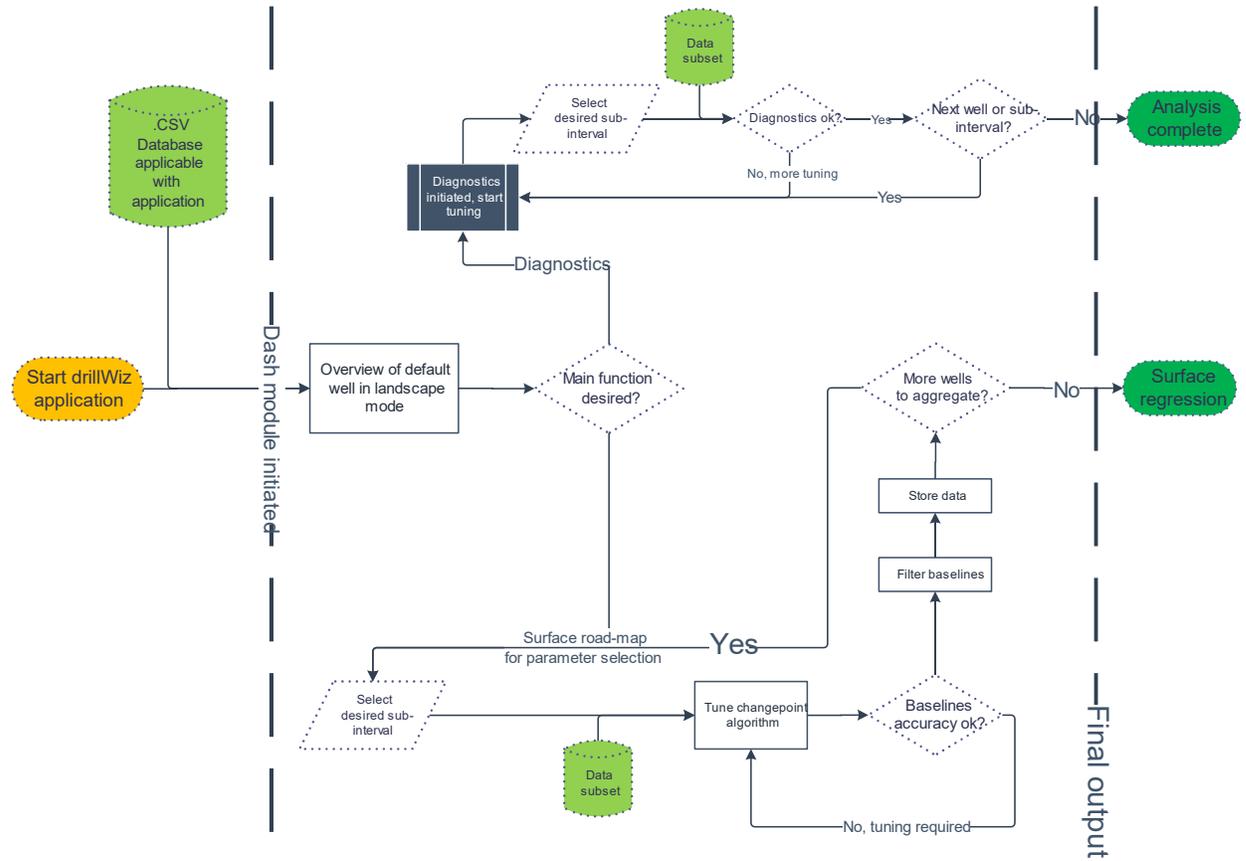


Figure 3.3: Flow chart of function architecture visualized.

Figure 3.3 visualizes the architecture of the application. The first step after initiating the desired data set for analysis is to decide the functionality, either diagnostics or lithology-specific road-amps for parameter selection. Each of the main functionalities will be described in detail below.

3.2.1 Software user interface

An introduction to the software and user interface is presented below. However, in order to fully familiarize oneself with the application, the reader is advised to play around with the software him- or herself. The application may be accessed by clicking the following [link](#). Furthermore, a video tutorial has been created to give an easy and effective introduction to the software. The video may be accessed by clicking [here](#).

The software user interface is basically split into two compartments or sections. The first section contains the input and interactive components, whereas the second section contains the results or plots. An overview of the software interface is depicted in Figure 3.4 below.

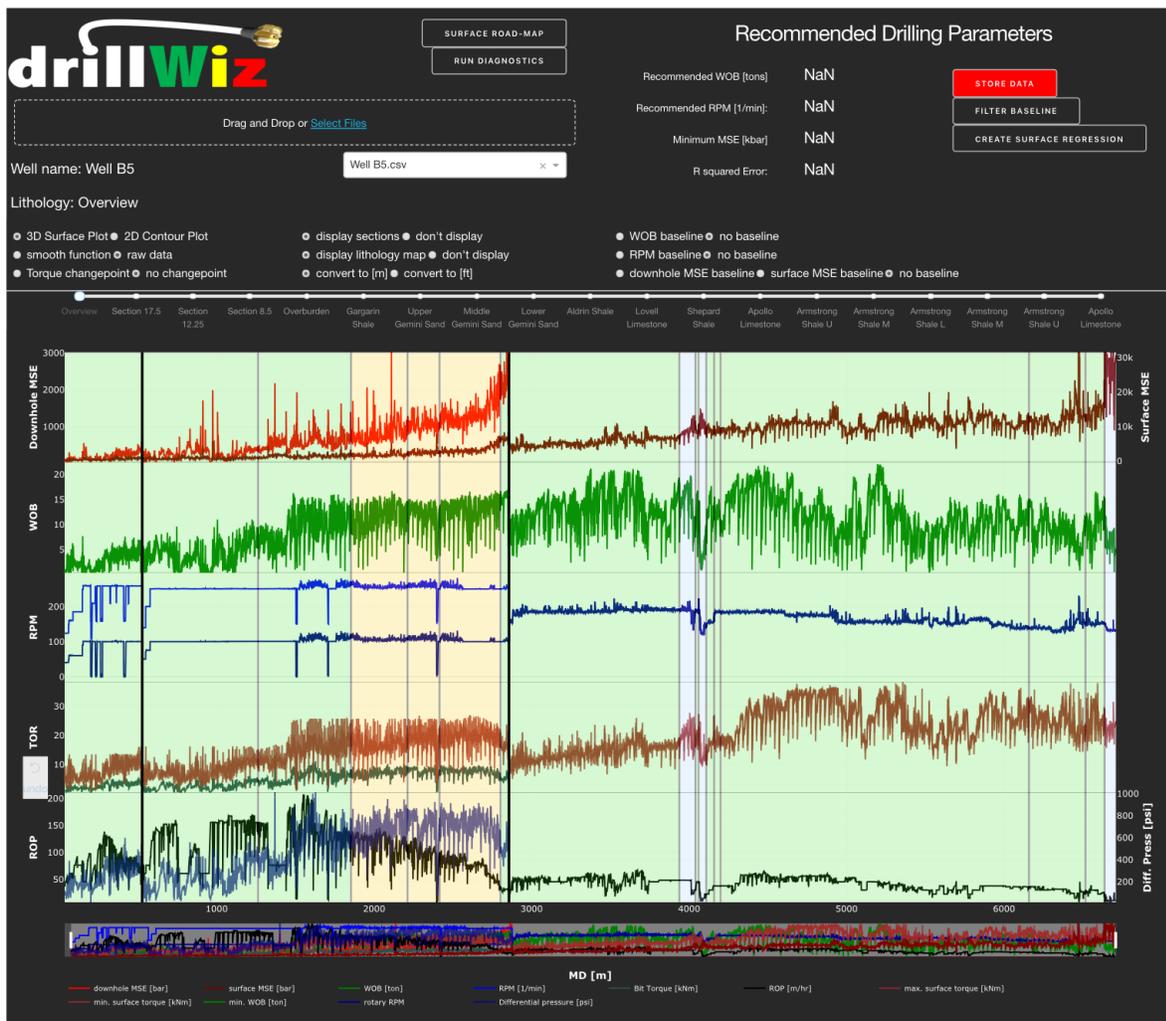


Figure 3.4: Software interface after starting drillWiz.

Upon loading the application, the default state will visualize the well data in a landscape mode such as depicted in the figure above. The default state establishes an initial overview of the input data, which will be discussed in more detail in the *Results* chapter. The “input section” of the user interface is depicted by the top compartment in the figure above, separated by a thin white line from the “output section” or plot area. The input section provides the user with many different options for data visualization and analysis, depicted in the zoomed-in snippet below, ref. Figure 3.5.

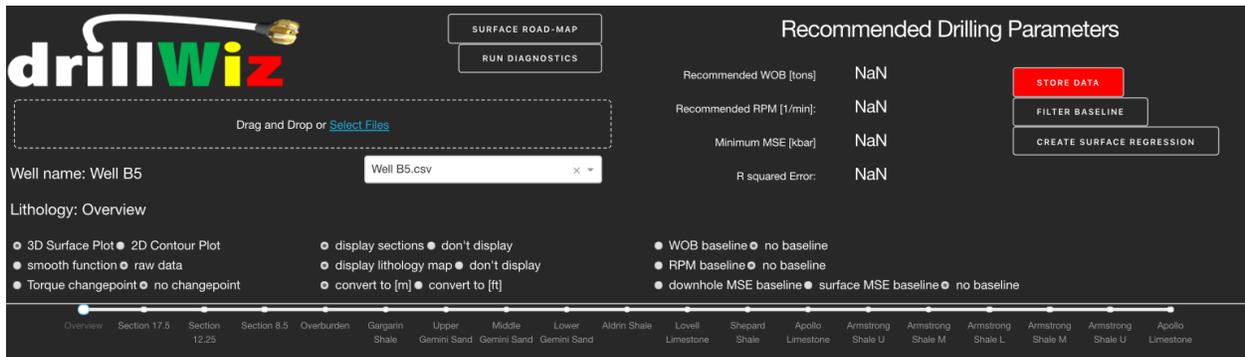


Figure 3.5: The top section of the application comprises the user input interface.

The application [link](#) is initially initiated with a demo well, for visualization and testing purposes only. New well data may be entered into the application by either dragging and dropping new csv files directly into the drop box depicted in the figure above, or by selecting files through the file explorer. Note that multiple csv files may be selected into the application simultaneously. The application contains a dropdown window, depicted in Figure 3.5, which is set to automatically access and select between the uploaded csv files in the application.

The range slider in Figure 3.5 indicates which data subset is currently being viewed. The range slider is divided into ‘Overview’, ‘Sections’, and ‘Lithologies’, in descending order (from RKB to TD). The default mode of the range slider is an ‘Overview’ of the entire well data. The range slider gives the user an easy interface tool to “slide” or select between data subsets to view or analyze separately.

Above the range slider are several different radio points representing different functionalities in the application, depicted in Figure 3.5. These include 3D surface plots, contour plots, smoothing

functions, changepoint baselines, and visualization options such as e.g. overlaying lithology maps or casing depths to the plot.

Once data are initialized, the user may choose between the two main functionalities of the application, either by selecting the “*Surface Road-Map*” or “*Run Diagnostics*” button depicted in the figure above, ref Figure 3.5.

Diagnostic results are added as additional traces in the landscape window. If the parameter optimization scheme by surface regression is initiated, the output section of the application is split into two graphs, depicted by Figure 3.6 below. The left-hand graph depicts the 3D surface plot or scatter plot, whereas the right-hand graph contains the overview plot to fine-tune changepoint baselines and pick representable points, which will be discussed in more detail below in the *Lithology-Specific Road-map* and *Pick Points* subchapter. Note that the selectable *Filter Baselines*, *Store data*, and *Create Surface Regression* buttons from the flowchart in Figure 3.3 are presented in the input section of the user interface.

After identifying optimal parameters from surface regression, the recommended parameters are presented in the *Recommended Drilling Parameters* window, replacing the initial *NaN* values, depicted in Figure 3.5.

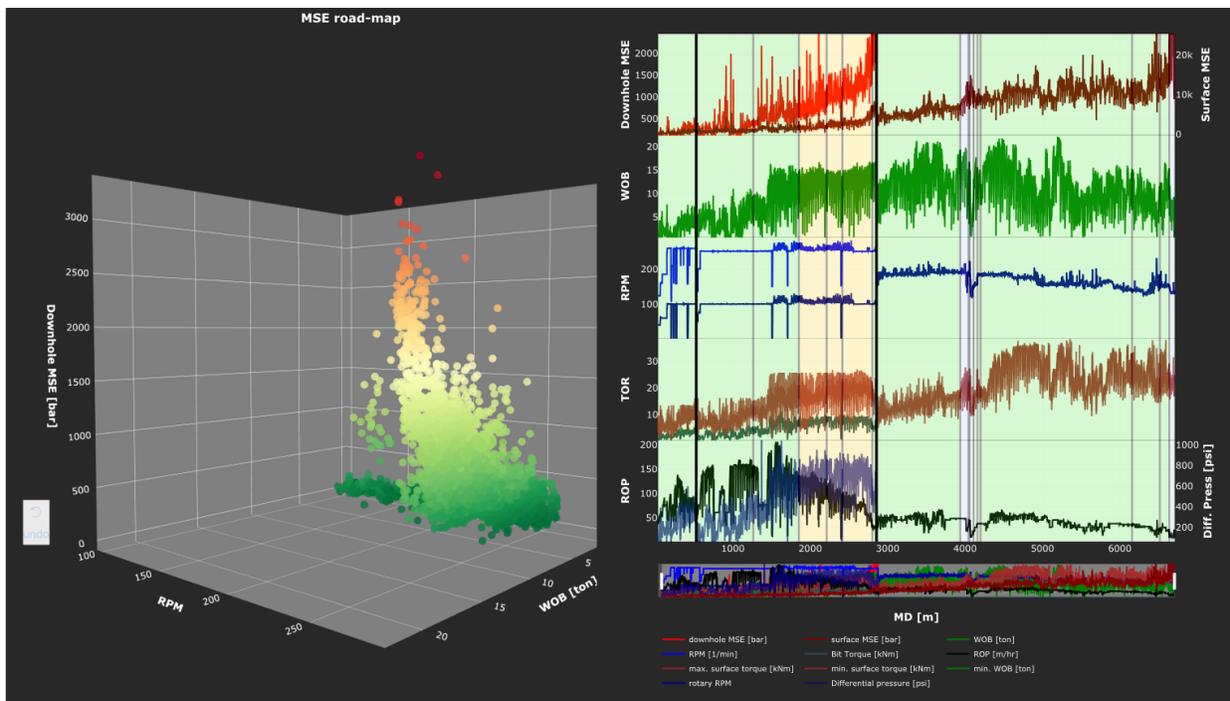


Figure 3.6: User interface of drillWiz after engaging Surface Road-Map mode.

Recall that the application is continuously fired through callbacks, such that any interactive change by the user in the application will automatically trigger callback functions to run and produce outputs. E.g. changing changepoint sensitivities will automatically generate new baselines in the plot. The main functionalities of the application are presented below in each respective sub-chapter with executable pseudo-codes.

3.2.2 Software functionality list

- Data parsing algorithm for input csv. files
- Drag 'n drop functionality for adding well data
- Changepoint algorithm for detection of baseline trends in operating parameters
- Visual lithology mapping of intervals:
 - Sand = yellow, shale = green, limestone = blue
- Display casing setting depths
- 3D scatter function of MSE for combinations of WOB and RPM
- Filtration functionality of baseline values
- Surface regression for optimal parameter selection
- Storing function to aggregate multiple wells with same lithology for parameter optimization
- Automated diagnostics mode with associated confidence level and severity
 - Whirl
 - Stickslip
 - Torsional severity
 - Stringer detection and damage assessment
- Unit conversion option (ft. or m)
- Smoothing function using convolution to smooth data for reducing noise

3.3 Data gathering: acquisition and availability

A significant part of this thesis has been to acquire the necessary data to produce the results for the thesis. Data liberation and accessibility is key in order to unlock external expertise and accelerate continuous innovation. This is especially true in the digital transformation era of today. While there is a consensus of mutual benefit from data liberation between internal players or company joint-ventures, there is still a way to go for making data liberation efficient between companies and third-party independent institutions, such as scholars or universities.

As a basis for data acquisition, all data was pre-defined in a project scope for an efficient dialogue with the operators. Below is a generalized list of the data requested from contact with industry companies.

- **MWD logs**
 - Depth-based data (high frequency, if possible)
 - Time-based data, preferably one-second surface data
- **BHA data**
 - Type of configuration and ratings for maximum load, etc.
- **Memory sub data**
 - Shock & vibration subs; axial, lateral, torsional
- **Final well reports:**
 - Summary of drilling experiences
 - Trip-risk log with formation tops
 - Bit pictures
 - Incident reports

Through the supervisor of the thesis, the proposed application and project scope was pitched to different industry engineers and sub-division leaders which triggered a fertile collaboration with a major international operator. To fully understand the operators drilling procedures and data reporting regimes, the authors of the thesis travelled abroad to the operator's office in the area of interest to spend some time with the operator, learning about their experiences with drilling performance. Having established a solid connection with the operator, associated service

companies were contacted on behalf of the operator to initiate data extraction and access to the operators internal reporting schemes on a well-specific basis for the desired areas.

3.4 Data preparation

Data preparation is perhaps one of the most important aspects with regards to developing an accurate model. To put it simply, data preparation is the act of processing, structuring, filtering and consolidating data for use in analytics and data visualization applications. The process of preparing data to fit a model is often called feature engineering and is an important step to ensure quality in data analysis. A good model can only produce a result as good as the data it is provided. In machine learning, training data must comply with the model's ability to read and analyze, just as the input drilling data must comply with the application model in order to be read and understood.

This all comes down to generating the highest possible confidence level in the interpretations. To be able to generate a good confidence level, one must integrate as much different drilling data as possible, in order to remove non-uniqueness. One may summarize the data preparation process in three key words:

- ***Cleanness***
 - Identifying actual and false measurements are critical in an optimization process, where erroneous measurements must be omitted to avoid any errors.
- ***Consistency***
 - To make sure the data is not tampered with by different models, all acquisition is done using the same template for extraction and processing.
- ***Accuracy***
 - For maximum accuracy, the highest possible sampling rate is utilized. In addition, both maximum and minimum values on each recorded depth is extracted where relevant.

More specifically, the idea behind this drilling application is to integrate lithology with drilling data to add an extra level of confidence to diagnostics and parameter selection. Due to the way the different data are organized with the operator and service company today, the need for certain

data preparation steps is crucial. Due to the origin of the data, i.e. if it comes from the operator or the service company, the data is organized differently and will require different types of data preparation.

3.4.1 Preparation of depth-based data

Data is either extracted as depth-based data or time-based, meaning each depth has an assigned data point, regardless of the time spent drilling on this depth or assigned to each time. For depth-based data, extracting information is easier as all data can be assumed to be gathered while drilling and not during other operations off-bottom. The process of data extraction and processing for the depth-based data is summarized in Figure 3.7.

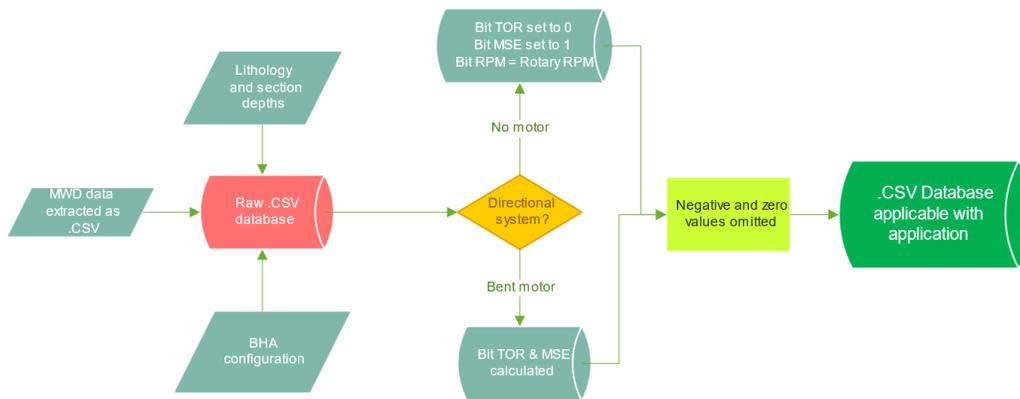


Figure 3.7: Road to Data Preparation

All data is extracted from each source and gathered in a tabular form in an Excel spreadsheet, which serves as a template for data processing. In this spreadsheet the data is processed based on the input configurations. As presented in Figure 3.8, MWD data is first extracted as a .csv file and inserted into an Excel spreadsheet. Excel proves an efficient tool for this job as it can both read and edit .csv files.

Input sheet										Calculation sheet							
MD	WOB	ROP	WOB min	TOR max	TOR min	Rotary RPM	Diff pressure	Pump rate	Well data		Surface MSE	Downhole MSE	Motor RPM	Bit RPM	Bit TOR	Motor = 0	RPM factor
[m]	[tonnes]	[m/hr]	[tonnes]	[kNm]	[kNm]		[psi]	[gal/min]	Section	Start depth	RPM factor	TOR factor	[bar]	[bar]	[kNm]		[1/gal]
0.5	0	52.51	0	0	0	0	0.65	0	0	0	0.166	7.9	0	0.0	0	0	0.16
1	-999.25	113.55	-999.25	0	0	0	0.91	0	17.5	0	0.166	7.9	-631	-631.5	0	0	0.16
1.5	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	12.25	527	0.166	7.9	-28937	-631.5	-165.876	-1165.13	0.00
2	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	8.5	2848	0.28	7.9	-28937	-631.5	-165.876	-1165.13	0.00
2.5	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	Interval without motor				-28937	-631.5	-165.876	-1165.13	0.00
3	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	Start	End			-28937	-631.5	-165.876	-1165.13	0.00
3.5	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	2848	4877			-28937	-631.5	-165.876	-1165.13	0.00
4	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	Calculations performed				-28937	-631.5	-165.876	-1165.13	0.00
4.5	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	0 (No type)	1 (Sand)	2-2.4 (Shale)	3 (Chalk)	-28937	-631.5	-165.876	-1165.13	0.00
5	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	Lithology information				-28937	-631.5	-165.876	-1165.13	0.00
5.5	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	Formation	Top [m]	Type		-28937	-631.5	-165.876	-1165.13	0.00
6	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	Overburden	0	2		-28937	-631.5	-165.876	-1165.13	0.00
6.5	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	Gemini Sand	1854	1		-28937	-631.5	-165.876	-1165.13	0.00
7	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	Aldring Shale	2802	2		-28937	-631.5	-165.876	-1165.13	0.00
7.5	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25		3938	3		-28937	-631.5	-165.876	-1165.13	0.00
8	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25		4065	2		-28937	-631.5	-165.876	-1165.13	0.00
8.5	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25					-28937	-631.5	-165.876	-1165.13	0.00
9	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25					-28937	-631.5	-165.876	-1165.13	0.00
9.5	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25					-28937	-631.5	-165.876	-1165.13	0.00
10	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25					-28937	-631.5	-165.876	-1165.13	0.00
10.5	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25					-28937	-631.5	-165.876	-1165.13	0.00

Figure 3.8: Snippet from Excel template, displaying the input section to the left and the output section to the right.

Formation tops and section depths are gathered from the end-of-well reports with confirmed depths from geologists. After extracting the drilling data, it is first important to assess which of the data is relevant and what is surplus. Before calculating the necessary parameters, the Excel template must be fed the configuration of the directional system, i.e. whether a downhole motor has been utilized. The wells analyzed in this thesis are either drilled with a RSS and no motor, or with a bent housing and motor. If a motor is used, one can estimate the downhole torque at the bit by using the differential pressure over the motor and the motor torque factor, meaning downhole MSE can also be calculated. In the case where a RSS with no motor is utilized, no differential pressure across a motor is available and only surface MSE may be calculated for analysis. For simplicity of later calculations, bit torque is therefore set to zero and bit MSE set to 1 bar. Using the acquired section depths, MSE is calculated for each depth, ref. equation (3). After performing the necessary calculations, the .csv file is saved into the well database, ready for the last stage of the data processing which is done using a Python script.

This stage omits all the data points of a depth if either WOB, RPM, or torque is zero or negative. The .csv files are organized as columns with associated headers describing the content of the data column below. When extracting information from a column, one can simply just pick out the header and acquire the column, independent of the relative column place, hence one may add as much information as desired in different columns as long as they are all assigned with the same header in each file. Based on the available data and logging subs on each well, different vibration data may also be extracted for later validation and assigning a confidence level to diagnostics.

3.4.2 Preparation of time-based data

For some of the wells analyzed in this thesis, the data was extracted as time-based data from the service company’s database. In practice this means the data is sampled much more frequent, providing typically 20x the sampling rate compared to the depth-based data given from the operator. In time-based data, all data that is not on-bottom drilling data must be omitted.

To bypass this issue, the time-based data is converted into depth-based data. This is done by evaluating whether the bit is on bottom or not. In the data from the service company, the measured depth is reported in two different parameters, i.e. hole depth and bit depth. The hole depth represents how deep the well is and the bit depth represents the depth of the bit.

1. Hole and bit depths are extended to all data points.

TIME	HDEPT	BDEPT	ROP	RPM	TORQUE	TIME	HDEPT	BDEPT	ROP	RPM	TORQUE
41:55.3	-999.25	20833.27	11.47	58.9	7960	42:20.1	20833.27	20833.41	22.72	3.55	6500
41:55.7	20833.27	-999.25	-999.25	-999.25	-999.25	42:21.0	20833.27	20833.41	-999.25	-999.25	-999.25
41:56.0	-999.25	-999.25	-999.25	-999.25	-999.25	42:22.0	20833.27	20833.41	-999.25	-999.25	-999.25
41:56.3	-999.25	-999.25	-999.25	-999.25	-999.25	42:23.0	20833.27	20833.41	-999.25	-999.25	-999.25
41:57.0	-999.25	-999.25	-999.25	-999.25	-999.25	42:23.2	20833.27	20833.42	11.6	12.21	4950
41:57.3	-999.25	-999.25	-999.25	-999.25	-999.25	42:24.0	20833.27	20833.42	-999.25	-999.25	-999.25
41:58.0	-999.25	-999.25	-999.25	-999.25	-999.25	42:25.0	20833.27	20833.42	-999.25	-999.25	-999.25
41:58.4	-999.25	20833.3	34.72	43.26	8160	42:25.8	20833.42	20833.42	-999.25	-999.25	-999.25
						42:26.0	20833.42	20833.42	-999.25	-999.25	-999.25

Figure 3.9: First stage

2. Off-bottom data are omitted by checking whether hole depth is larger than bit depth.

1	TIME	HDEPT	BDEPT	ROP	RPM	TORQUE	1	TIME	HDEPT	BDEPT	ROP	RPM	TORQUE
45891	52:22.2	-999.25	11215.78	-999.25	0	0	153626	47:44.1	-999.25	21008.29	81	57.62	11500
45892	52:23.0	-999.25	-999.25	-999.25	-999.25	-999.25	153627	47:44.3	21008.29	-999.25	-999.25	-999.25	-999.25
45893	52:23.4	20654.12	-999.25	-999.25	-999.25	-999.25	153628	47:45.0	-999.25	-999.25	-999.25	-999.25	-999.25
45894	52:23.4	-999.25	-999.25	-999.25	-999.25	-999.25	153629	47:46.0	-999.25	-999.25	-999.25	-999.25	-999.25
45895	52:24.0	-999.25	-999.25	-999.25	-999.25	-999.25	153630	47:46.5	-999.25	-999.25	-999.25	-999.25	-999.25
45896	52:25.0	-999.25	-999.25	-999.25	-999.25	-999.25	153631	47:47.0	-999.25	-999.25	-999.25	-999.25	-999.25
45897	52:25.7	-999.25	11215.78	-999.25	0	0	153632	47:47.2	-999.25	21008.32	34.67	59.01	11530
45898	52:26.0	-999.25	-999.25	-999.25	-999.25	-999.25	153633	47:47.5	-999.25	-999.25	-999.25	-999.25	-999.25
45899	52:27.0	-999.25	-999.25	-999.25	-999.25	-999.25	153634	47:49.0	-999.25	-999.25	-999.25	-999.25	-999.25
45900	52:27.6	-999.25	-999.25	-999.25	-999.25	-999.25	153635	47:50.0	-999.25	-999.25	-999.25	-999.25	-999.25
45901	52:28.5	-999.25	11215.78	-999.25	0	0	153636	47:50.3	-999.25	21008.39	81.95	61.28	11320

Figure 3.10: Second stage.

3. MWD data are only recorded every so often and “waste”, or in between measurements read as -999.25, which must be removed. MWD (WOB, RPM, ROP, TOR) data are not recorded at the same time as the memory sub data. Memory sub data record more frequently and must be aggregated to fit the frequency of the logged MWD data. This is performed by maximizing logged data between each MWD data point.

Torque	Lat_RMS	Lat_Peak	Ax_RMS	Ax_Peak	Torsional_	Torque	Lat_RMS	Lat_Peak	Ax_RMS	Ax_Peak	Torsional_
-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25
-999.25	2.385175	4.297924	0	0	1.611722	-999.25	2.385175	4.297924	0	0	1.611722
-999.25	2.279319	4.688645	0	0	1.611722	-999.25	2.279319	4.688645	0	0	1.611722
-999.25	2.314111	4.200244	0	0	1.660562	-999.25	2.314111	4.200244	0	0	1.660562
-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25
-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25	-999.25
-999.25	2.082443	4.542125	0.04884	1	1.611722	-999.25	2.082443	4.542125	0.04884	1	1.611722
-999.25	2.211327	3.907204	0	0	1.611722	-999.25	2.211327	3.907204	0	0	1.611722
-999.25	2.109189	4.932845	0	1	1.562882	-999.25	2.109189	4.932845	0	1	1.562882
10830	-999.25	-999.25	-999.25	-999.25	-999.25	10830	2.385175	4.932845	0.04884	1	1.660562
						10830	2.385175	4.932845	0.04884	1	1.660562

Figure 3.11: Third stage.

- Once extracted as depth-based data, the data can be treated accordingly in the application.

3.5 Changepoint detection scheme

Through the package called Ruptures, the changepoint detection scheme has been derived. As presented in the theory, it relies on a negative log-likelihood method for detecting changepoints and is summarized big-picture in figure Figure 3.12 below.

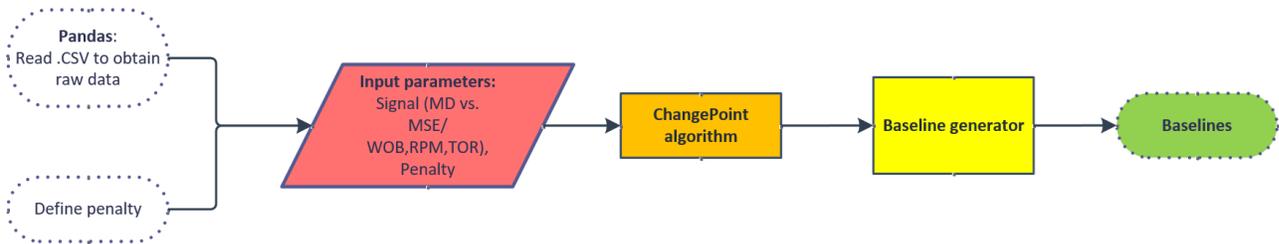


Figure 3.12: Flowchart overview of creating baselines.

The input data for the functions are primarily .csv-files read by the Pandas module and structured into data frames. To ensure the data type is correct, the input data are converted to NumPy arrays which are more efficient to perform computations with compared to conventional lists or other data types. As the data are processed towards baselines, the data goes through two sub-functions, one which is the changepoint algorithm that defines the number of changepoints and their position. The second function is the baseline generator which translates the changepoint positions and median value between them into baselines.

A more comprehensive and detailed procedure of the changepoint algorithm is outlined in the flowchart below, ref. Figure 3.13. The procedure consists of a first stage which initializes the data on a signal by fitting certain internal parameters and defines a cost function to calculate the cost of the given subset, defined as a changepoint interval. Using the fitted input signal and the cost function, the PELT algorithm from the Ruptures package calculates the optimal changepoint segmentation based on the criteria defined. The overall procedure goes over two stages which are integrated using the Ruptures package and supplied as output to the baseline generator.

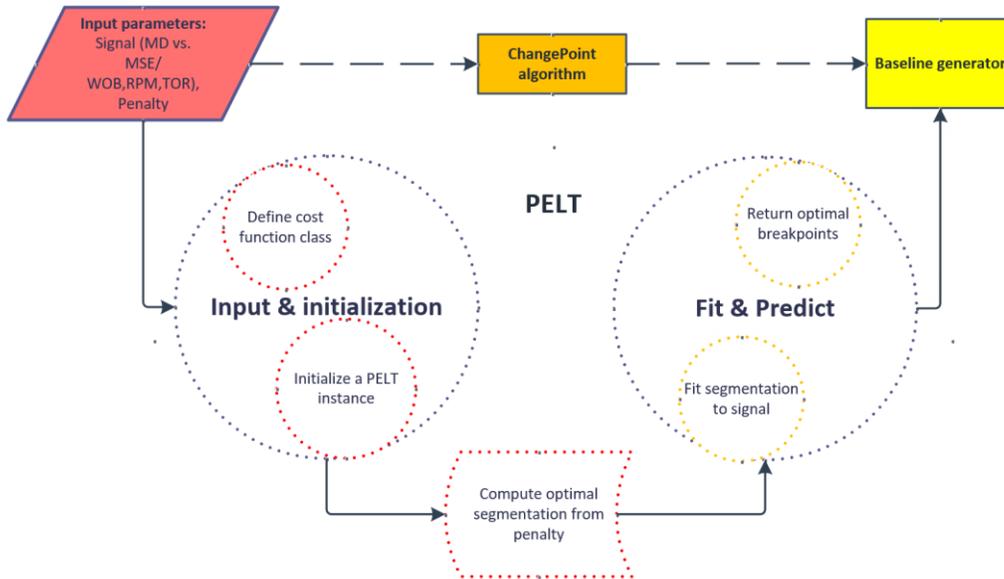


Figure 3.13: PELT sub-functions overview.

From the optimal changepoints and input signal, the developed baseline generator provides a simple solution to translate the breakpoints into values for plotting and computing. The median of the sub-signal is computed by the NumPy median function. By using the median, outliers with high value will have less impact compared to using the arithmetic mean. The output from this function is the end-product of the changepoint detection scheme, which is an array with the value of the baseline and the start and end of each segment, ref. Figure 3.14.

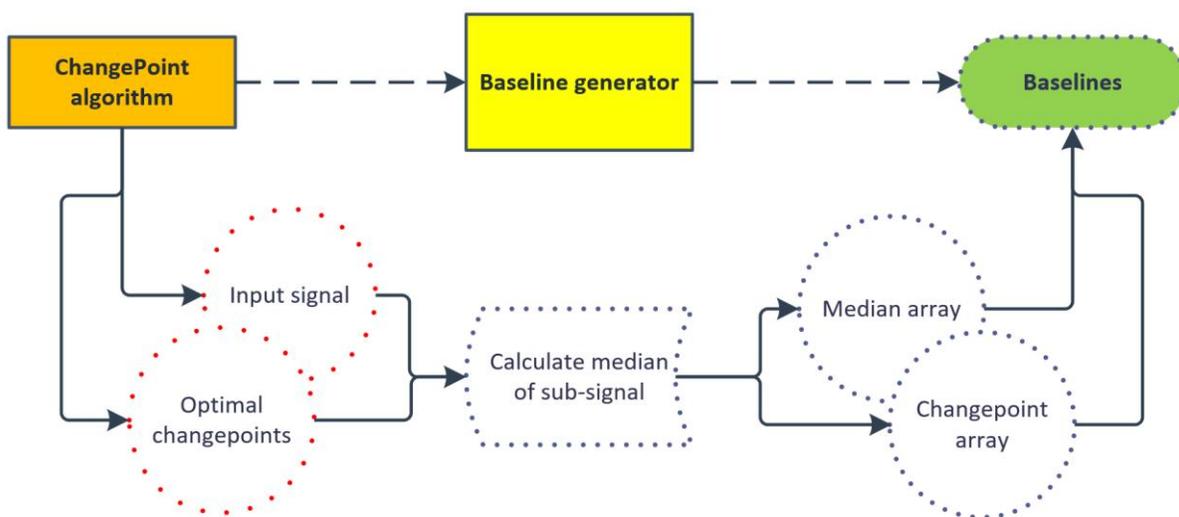


Figure 3.14: Baseline generator overview.

3.6 Picking out data points of interest

Having created the desired baselines for the different parameters, data has now been compressed in several orders of magnitude. In order to select representable parameter combinations or overlying baselines responses, a PickPoints function has been developed, illustrated in Figure 3.15.

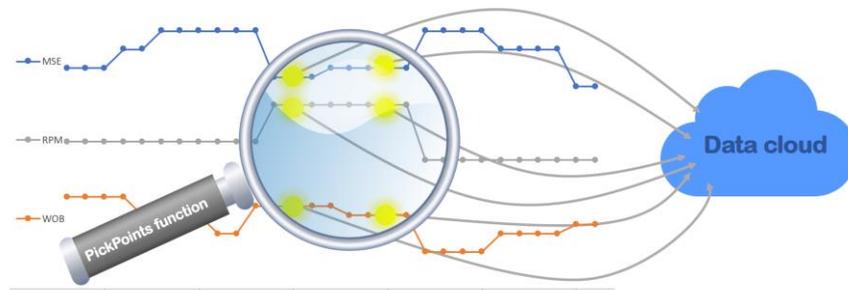


Figure 3.15: PickPoints function illustrated; every baseline combination is picked out and combined into representable data points.

First, the baselines of the desired parameter intervals are generated. Then, by discretizing the baselines, the PickPoints function can detect the overlapping values and pull them out as different parameter combinations. At last, duplicates are dropped to minimize memory usage and only provide unique combinations in the data storage. The resulting parameter combinations are now processed and readily prepared for the next step, regression analysis.

3.7 Establishing lithology specific road maps for optimal parameter selection

While gathering data points and reducing them to representable data points may provide a certain level of information, real value lies in integrated statistical analysis. Regression analysis has its value in predicting future behavior from established trends in historic well data.

Evaluation of a regression function is used to forecast and predict missing data to optimize processes, such as the drilling parameter optimization process presented in this thesis. By using historical reference well data to identify behavior of future operations, one overlaps the area of machine learning by using “trained data” to identify trend patterns (Draper & Smith, 1981). Several different methods for regression exist, and for the interested reader further material and an overview of regression models is presented in (Fox, 2015), i.e. a modern global perspective on regression analysis.

Well data from MWD logs may typically be extracted as simple comma-separated values, i.e. stored in csv-files. Csv-files exemplifies an effective way to store well data as they can be easily imported and exported across multiple table-oriented application platforms or relational databases.

Even though data files may contain tens of thousands of data points per well, they rarely exceed a tenfold of megabytes in size. For software and functions where data is read and processed in simple ways this does not pose a challenge. However, when introducing regression modelling and more sophisticated analyses, reducing this data from thousands to a magnitude of tens can prove to have a significant impact on computational time and even on the accuracy of the model by removing unwanted “noise”.

After selecting a desired sub-interval, e.g. a specific lithology or section, a cropped data subset is gathered from the original data set. Although the data set is cropped, several duplicates, outliers and other surplus information exist. As outlined in the changepoint algorithm and baseline generator, one may reduce large data sets in tens of magnitude by picking out representable baseline values for each sub-section. Hence, the cropped subset can be tuned and thereby further compressed using a set of tuning parameters, provided through the user input in the application module. With sufficient tuning of the changepoint algorithm, the representable parameter

selection is stored with the different “operating parameters” grouped in n data point. For n representable operating parameters and MSE responses, the data is grouped as,

$$\mathbb{P} = ((WOB_1, RPM_1, MSE_1), \dots, (WOB_n, RPM_n, MSE_n))$$

Notice how WOB, RPM and MSE are now decoupled from their associated depth, hence one may aggregate data from different wells and depths and compare them with each other. This is done on several basic assumptions and may involve potential errors, further discussed below and in the Discussion section. The basic assumptions for comparing parameters are given below.

- Lithology characteristics such as rock compressive strength is assumed to be constant and homogenous in lithology intervals and across compared data.
- Bit aggressiveness is assumed constant in compared wells or at least having a negligible effect on bit MSE.
- BHA configurations are assumed to be of similar character, meaning RSS data is not compared to bent motor data.
- Only downhole bit MSE should be compared between wells due to being decoupled from additional drill string friction from alternating well paths.

Note that Surface MSE may yet be utilized when aggregating lithology data between wells, in certain situations and if handled correctly. These situations include if well path, formation depth and compressive rock strength are nearly identical between wells of comparison, which is discussed in more detail in the *Discussion* chapter. When the tuned and compressed data are imported into the regression function, a meshed grid of the operating area is generated and characterized by the maximum and minimum operating parameters experienced in the input data set. The regression function is defined as the polynomial described in theory, ref. eq. (12). To assign the coefficients used in the equation, the SciPy Optimize package is utilized with the Curve Fit module. A set of initial coefficients may be chosen as a “first guess”, which for this approach are all set to a value of 1. Using the defined polynomial function and the fitted coefficients, a corresponding MSE value is forecasted.

In order to generate optimum recommended parameters, one must first define what optimal parameters are. For this optimization process, two criteria are used.

1. ***A predefined threshold (120% of minimum representable MSE) of “acceptable” MSE is used as the upper limit for efficient drilling.***

The reason for not using the minimum directly is to bypass potential errors and thereby missing out on potential performance gains. Desired operating range is therefore over-fitted.

2. ***To maximize the achievable performance, RPM and WOB must be maximized within acceptable limits.***

During efficient drilling, ROP responds proportionately to increases in both RPM and WOB. Maximizing the individual parameters is therefore equal to maximizing the product of them. Recommended operating parameters are then assigned where the product of RPM and WOB reaches its maximum.

Below is a summarized flow chart of how the application optimizes the operating parameter selection based on the representable input data.

Optimizing parameter selection from surface regression

1. Generate mesh-grid covering domain of data:
2. Create quadratic curve $y = Xa$
3. Define polynomial coefficients through method of least squares from the SciPy package
4. Evaluate polynomial coefficients with quadratic curve on the grid to obtain surface
5. Minimize surface within domain to obtain recommended operating area

$$(WOB_{ideal}, RPM_{ideal}, MSE_{ideal})$$

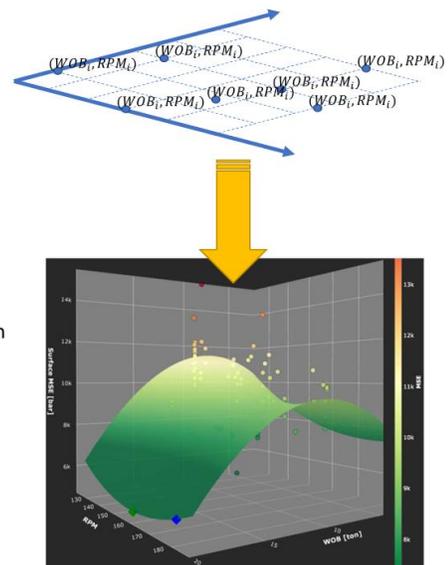


Figure 3.16 Parameter optimization scheme summarized.

3.8 Diagnostics

Being able to programmatically identify bit dysfunction is crucial in order to address drilling inefficiencies in an automatic drilling system. A comprehensive literature study about the physics of each bit dysfunction and detailed procedures of dysfunction diagnostics are presented in (Berge-Skillingstad & Anderssen, 2018). From the theory outlined in this project comes application in practice presented below, for some of the main dysfunctions. For the sake of the confidence in the diagnostics tools, only certain dysfunction-indicators have been included.

Below is a generalized pattern of the rudimentary workflow involved in drilling diagnostics, in a simplified procedure:

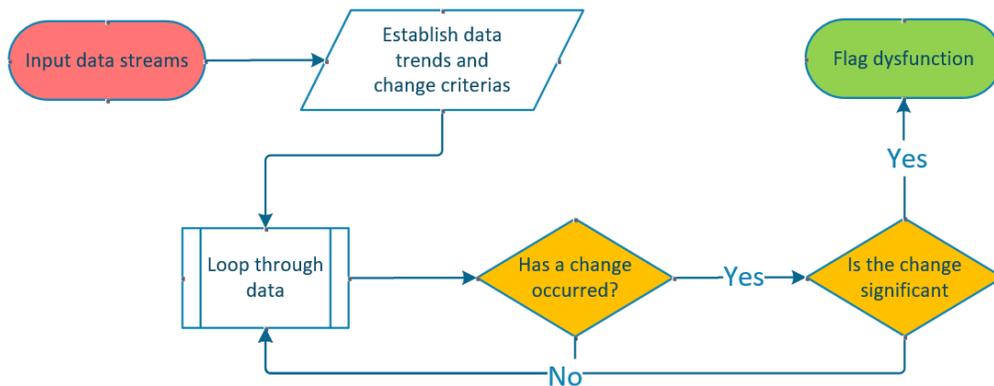


Figure 3.17: Generic diagnostics procedure strategy used in this thesis.

3.8.1 Establishing a linear trend of “expected” MSE

Recall that MSE is expected to change as confined compressive rock strength increases with greater hole depth. This is true for downhole bit MSE. This is also true for surface MSE, however, surface torque will increase considerably more from additional drill string friction in an elongated well path. It is important to keep this in mind when surveilling MSE to diagnose drilling inefficiency or bit dysfunction. For this reason, a linear regression trend line of MSE is established throughout every interval of investigation in order to attempt to capture the

“expected” increase in MSE with hole depth. The purpose of the regression line is to capture any increasing trend from confined compressive rock strength in downhole bit MSE, or additional increase from drill string friction in surface MSE (if no downhole bit MSE is available for diagnostics).

As such, any changes in MSE will always be evaluated in relative severity against the MSE regression trend line. This is a conservative approach, as any change in confined compressive rock strength or immediate change in drill string friction will pale in comparison to an onset of a bit dysfunction. Keep in mind that any drilling inefficiency will shift the regression line upwards in an increasing trend. Conclusively, comparing relative change in MSE versus the regression trend line will lessen the experienced severity. This will, in effect, underfit the severity and flagging of diagnostic results. If an MSE baseline is underlying the regression trend line, one can assume a relative “efficient” baseline, even though it may not be efficient. This is, however, favorable in an automated diagnostics algorithm, in order to increase the confidence level of the actual flagged dysfunctions and avoid overfitting results. A MSE baseline overlying the general trend line will indirectly dictate an inefficient drilling operation.

3.8.2 Whirl

In-depth whirl diagnostics and theory behind operational procedures and considerations are presented in (Berge-Skillingstad & Anderssen, 2018). Whirl is by far the most common bit dysfunction and according to (Dupriest et al., 2010), occurs in approximately 40 % of all world-wide footage drilled. As emphasized in (Berge-Skillingstad & Anderssen, 2018), it is important to separate between BHA whirl and bit whirl. BHA whirl will not affect performance or drilling efficiency, whereas bit whirl may severely affect drilling efficiency and cause unintended bit trips from premature damage¹. In other words, only bit whirl will cause a considerable response in MSE. A generalized pseudo-code of how bit whirl is programmatically diagnosed in the application is presented in the figure below, Figure 3.18.

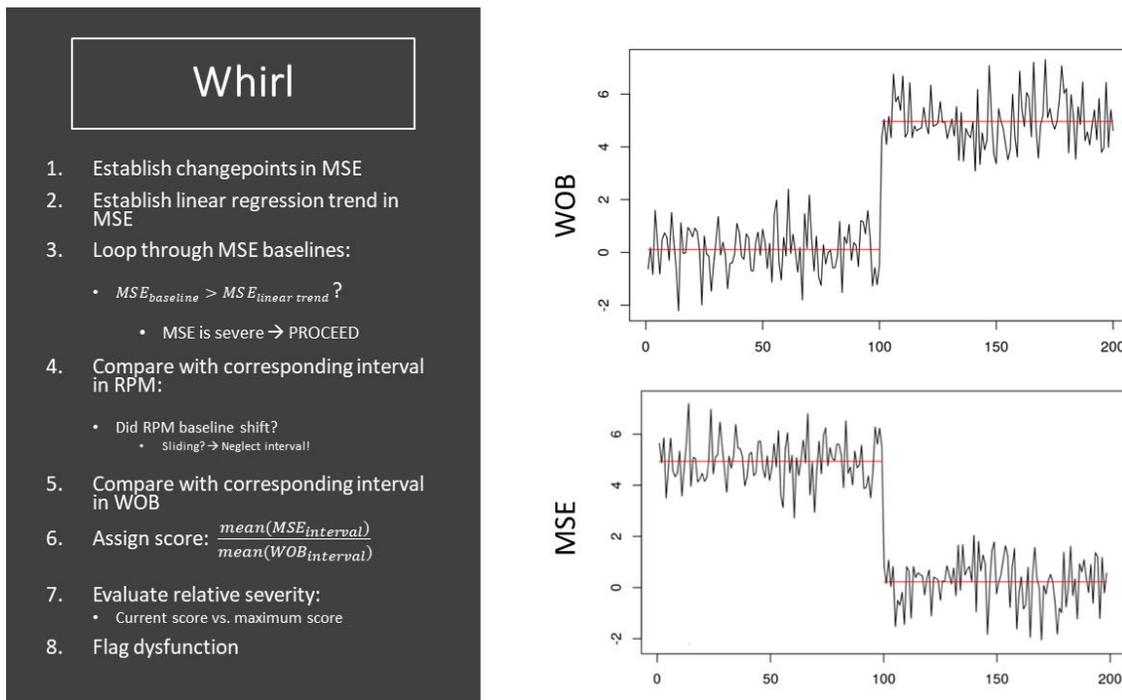


Figure 3.18: Whirl diagnostics scheme overview.

¹ BHA whirl may still be damaging to tools in the assembly, however, the severity of lateral vibrations in the BHA are heavily dependent on the node and anti-node positions of the sinusoidal wave in the whirling BHA. This means that stabilizer placement and BHA configuration (tool position relative to sinusoidal anti-nodes/nodes) is vital to avoid unnecessary tool failures from BHA whirl.

To assess the challenges of a confidence level, a calculated whirl score has been derived from the relationship between MSE and WOB. Bit whirl is uniquely characterized by a disproportionate increase in performance from an increase in WOB. In other words, a decreasing MSE from an increasing WOB uniquely identifies bit whirl and clears way for a primitive way of assessing a whirl score by simply exploiting this relationship, ref. eq. (20). Lastly, the whirl score is divided by the maximum experienced score in the interval to normalize the value between 0 and 1.

$$Score_{whirl} = \frac{\frac{mean(MSE_{baseline})}{mean(WOB_{corresponding\ interval})}}{maximum\ score} \quad (20)$$

Where,

$Score_{whirl}$	Whirl score on data point
$MSE_{baseline}$	Baseline value of MSE at current data point
$mean(WOB_{corresponding\ interval})$	Mean WOB at same interval as MSE baseline
$maximum\ score$	Maximum score on interval analyzed, used to normalize score

3.8.3 Stickslip

Similar to whirl, in-depth stick-slip diagnostics theory, operational procedures and considerations are presented in (Berge-Skillingstad & Anderssen, 2018). Bit speed oscillations do not affect performance significantly, as such MSE is not a good diagnostics tool for identifying stickslip in itself. Full-stick events, however, cause severe and erratic MSE spikes, and are typically what are most damaging to the bit. It is not clear whether an increase in MSE is only due to full-stick or due to a coupled whirl event from a temporary loss of DOC. Conclusively, there are times stickslip and torsional severity diagnostics will overlay whirl diagnostics. This is further discussed in the *Discussion* chapter.

Although MSE is not a good tool for this type of dysfunction, an approach to diagnose just this has been developed. The strategy is to combine two independent diagnostic tools to weigh up a solid assessment. A summarized pseudo-code of how stickslip is programmatically diagnosed in the application is presented in Figure 3.19 below.

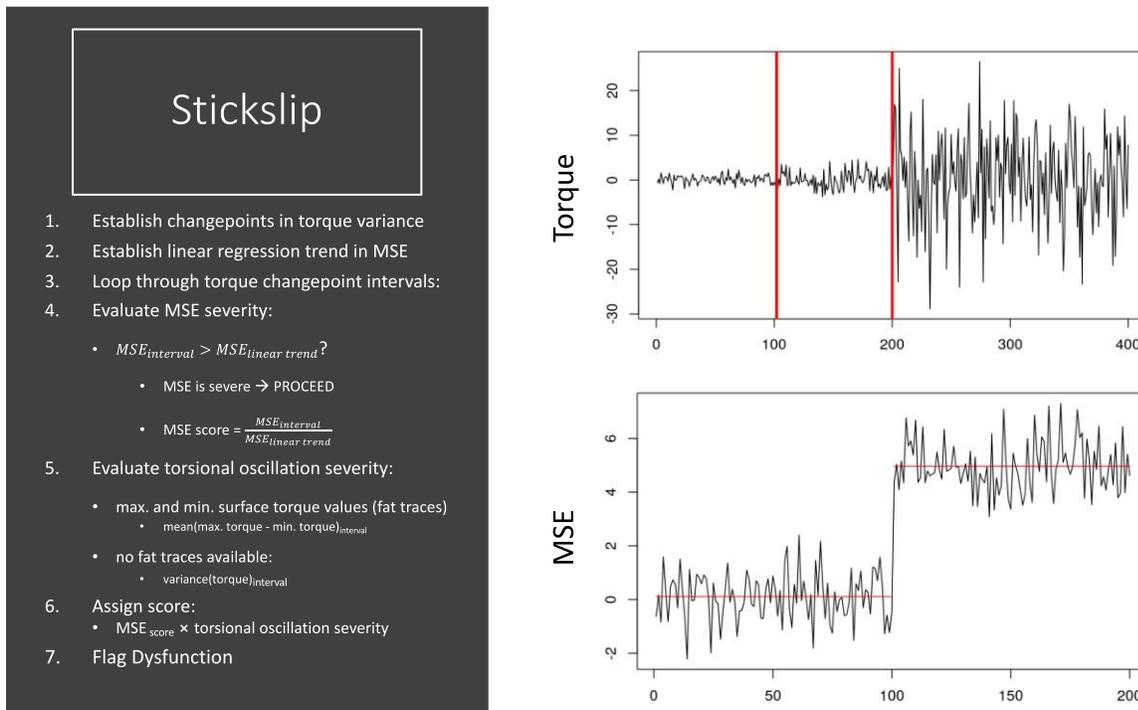


Figure 3.19: Stickslip diagnostics scheme overview

As with the whirl diagnostics, this tool also relies on changepoint detection. However, the changepoint detection scheme is built around changes in variance of a signal. By evaluating changepoints in torque variance, sub-intervals are obtained and form the basis for assigning scores.

3.8.4 Torsional severity

As mentioned, stickslip is in several ways significantly harder to diagnose than e.g. whirl. To bypass some of the diagnostics difficulties, a torsional severity confidence level system has been developed. Each data point is assigned with a simple scoring system based on the severity of oscillations in torque and MSE. The confidence level is defined as two separate scores for each data point, and a total aggregated score. The procedure for torsional severity diagnostics is explained and visualized in Figure 3.20 with a more detailed explanation of each step below.

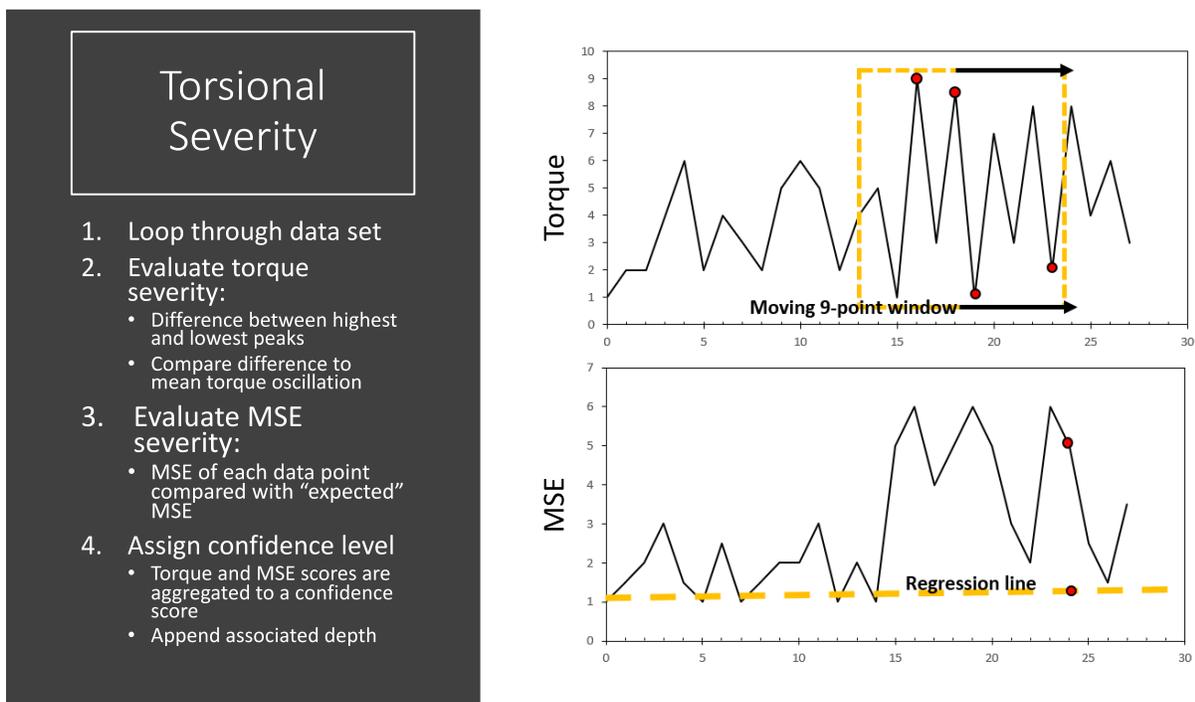


Figure 3.20: Torsional severity diagnostics overview

1. Loop through data set

- a. Every data point is now considered, unlike whirl or stickslip diagnostics where baseline values are considered.

2. Evaluate torque severity

- a. Last m data points of torque are considered:
 - i. k largest torque values of the maximum torque are picked out
 - ii. k smallest torque values of the minimum torque are picked out
- b. Torque oscillation is calculated as difference between the sum of the k largest and the sum of the k smallest torque values.
- c. Torque score, i.e. severity is calculated by dividing the torque oscillation of point i by the mean oscillation.

3. Evaluate MSE severity

- a. MSE score calculated by dividing MSE of point i by the reference MSE value.
 - i. If diagnosing the whole well or sub-interval, reference MSE is the linear regression MSE value of a straight-line prognosis.
 - ii. If diagnosing smaller intervals, mean MSE across the interval is used as a reference MSE. This is because the big-picture MSE trend will be clouded by local variations on smaller intervals.

4. Assign confidence level

- a. MSE and torque score are aggregated to a total score through simple summation
 - i. If torque score is less than T or MSE score is less than regression line, total score is set to 1. This is because the response in MSE is likely to have a different root cause.

Defined variables and their impact on the stickslip diagnostics function:

k	9	Number of extrema of torque used to measure oscillation
m	2	Number of data points to look back on to evaluate oscillation
i	N/A	The i -th number of the loop
T	1.2	Threshold for torque score before aggregating a confidence level

3.8.5 Stringer detection & damage assessment

Drilling into what is commonly known as stringers, i.e. interbedded layers with different compressive rock strength, may in certain cases be severely damaging to the bit. This is due to the erratic vibrations that occur when the bit transitions into layers with high contrast in compressive rock strength. The operational recommendations while drilling into formations with known stringers are well discussed within the industry, and several different “best-practices” are further outlined in (Berge-Skillingstad & Anderssen, 2018).

Motor stalling events are often associated with entering hard stringers, as the motor torque rating has a tendency to be exceeded when entering formations with sharp contrasts in compressive rock strength. The differential pressure across the motor will exceed the motor differential rating in a motor stalling event, indirectly indicating a stringer.

Surveilling MSE is an effective metric for assessing whether or not a stringer event has been damaging to the bit. This is because the minimum baseline value of MSE during efficient drilling is immediately affected by changes in drilling performance from bit damage. In sub-intervals, one may assume that MSE should stay constant, considering efficient drilling and a negligible increase in confined compressive rock strength. For a given BHA, a certain efficient baseline value of MSE can be identified, and monitored for changes in drilling performance. MSE should always fall back to the minimum efficient baseline value (+ a threshold to account for lithology changes) in a parameter exploration initiative (i.e. step-test to determine efficient drilling), if the bit is green. If the new minimum MSE baseline does not fall back to the previous minimum efficient value following a potential damaging event, one may assume damage to the bit. MSE will permanently shift following a decrease in performance ratio from a damaged bit.

Evaluating motor stalling events together with efficient baseline monitoring in MSE is what compromises the stringer detection and damage assessment diagnostics. The procedure and pseudo-code for stringer detection and damage assessment is depicted in Figure 3.21 below.

Stringer detection & Damage assessment

1. Check if differential pressure and bit torque exceed motor rating
2. If yes, motor stalling event:
 - Establish MSE baselines
 - Monitor future baselines after stalling event
3. Evaluate if future baselines do not fall back to previous efficient MSE + threshold
4. Flag damaging stringer event

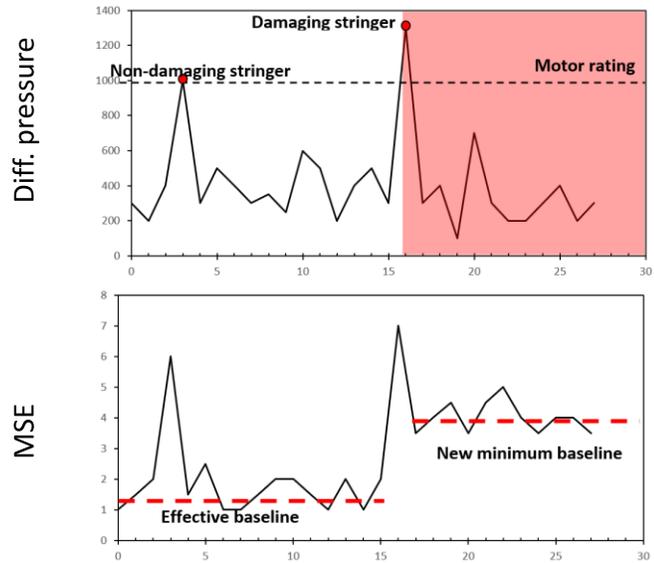


Figure 3.21: Stringer detection and damage assessment scheme overview

4 Results

The purpose of the results in this thesis is to attempt to add a level of confidence to the theory presented. This is done through simple demonstrations of separate results validating each concept in the drilling application and its functionalities. Furthermore, post-drill analysis and cases containing more complex problems will be presented for each module. This section will contain some level of discussion around results, however it can be differentiated by the *Discussion* section by the amount of detail included. In the *Discussion* section, a deeper debate on uncertainties, challenges, and further implementation will be outlined.

Well data from an anonymous operator is used to conduct the analyses presented below, together with well reports and other well technical data as supplements to the analysis. An important aspect which is discussed in this thesis is the level of confidence of the different types of analyses. Each confidence level is assigned with certain basic assumptions and is compared to “manual” interpretations and/or memory data from vibrational shock subs, if available. In some cases, the “manual” interpretations are considered the correct answer and is what the automated results are compared against.

Based on the analysis and conclusions made in this thesis, the goal is to prove why and how MSE should be the underlying metric for drilling automation, how it will facilitate automation of the performance management processes, and how this could have affected the observed problems in the presented wells. The conclusion is based on the ability to programmatically identify expected responses in MSE from different types of dysfunction.

Please keep in mind that the results in this chapter are generated from the user interface in the drillWiz application, such that several figures below may contain small fonts. The figures are provided in vector format (. SVG) should the reader need to zoom in for more details. In addition, the generic legend to the plots presented are depicted below.

Defining the following parameter curves and associated colors for plots

Table 1: Plot legends

Parameter	Unit	Color	Parameter	Unit	Color
WOB	tonnes		Surface MSE	bar	
Bit RPM	sec ⁻¹		Downhole MSE	bar	
Rotary RPM	sec ⁻¹		Surface torque	kNm	
Rate of penetration	m/hr		Bit torque	kNm	
Differential pressure	psi				

Additional parameter curves and associated colors for diagnostic plots

Table 2: Diagnostic Legend

Parameter	Color	Parameter	Color
Whirl severity		Stickslip severity	
Total torsional severity		MSE regression line	
MSE torsional severity		Torsional RMS (shock sub)	
Torque torsional severity		Lateral RMS (shock sub)	

4.1 Well information

To be able to fully understand the nature of bit dysfunctions and to correctly interpret MWD logs, one must have sufficient information about the well, i.e. BHA configurations, mud systems, lithology etc. The well data used in this thesis have been cherry picked as they all contain some level of dysfunction, and have all been drilled in the same area, meaning analogies between lithologies can be drawn. The wells have been provided by an anonymous operator from an onshore play in the U.S. For this reason, all data has been anonymized including well names and formation names. From a performance management point of view, the operator uses ROP as the primary metric for success.

Below is a tabulated visualization of the wells picked for analysis in this thesis, ref. Table 3: Well Information. The apparent goal of the operator from a directional steering point of view has been to use a RSS, however in several cases a bent motor has been utilized as a contingency due to several tool failures and steerability issues. An interesting circumstance to point out is the ratio between tool failure and number of stabilizers in the presented wells. As pointed out in (Berge-Skillingstad & Anderssen, 2018), stabilizer placement is one of the most important design factors when designing to prevent bit whirl. This is due to stabilizer placement determines how the sinusoidal wave from the BHA in compression arrives at the bit. Significant less tool failures are observed in the packed assemblies (three stabilizer BHA configuration) in comparison to the slick² assemblies in the presented wells. Note that stable (packed) assemblies do experience limited build angles, however slick assemblies are inherently prone to more bit whirl. Stabilizer placement does not eliminate whirl in the BHA, however it will affect the shape, node and anti-node positions of the sinusoidal wave in the BHA.

This means that the BHA configuration may severely affect how the sinusoidal wave from whirl arrives at the bit and BHA tools, potentially wrecking unwanted havoc to tools and causing unintended trips. Redesigning the BHA to mitigate whirl may not only increase drilling efficiency, but may also potentially increase tool longevity and prevent borehole quality issues

² A slick assembly is defined as a BHA with less stabilizers than a stable or packed assembly (i.e. without a “sleeve” or “sting” stabilizer)

such as spiral borehole patterns. Spiral borehole patterns are often associated with steerability issues, which is apparent in many of the presented wells.

If excessive bit whirl is present in the intervals affected by steerability issues, spiral borehole patterns may be the root cause, if not, other conditions may be the root cause such as i.e. inadequate MW and non-gauge hole. Without knowing what dysfunction is limiting steerability there is really no good way of redesigning the system to avoid running into the same mistake. In other words, there lies a substantial value to future operations in being able to identify which predominant dysfunctions are limiting your interval.

Table 3: Well Information

Well name	BHA config, ICP to TD³	Tool failures	Comment
Well A3	Bent motor	1 (Bit trip)	N/A
Well C3	RSS, then bent motor	6 (5 for RSS, 1 for motor)	RSS x2 stab, bent motor x1 stab
Well C4	RSS	3 (1 in curve, 2 in lateral)	3 stabs
Well C7	RSS, then bent motor	5 (2 for RSS, 3 for motor)	2 stabs
Well B2	RSS	2 (1 for RSS, 1 unknown)	3 stabs
Well B3	RSS	1 for MWD	3 stabs
Well B4	RSS	1 for RSS	3 stabs
Well B5	RSS	3 (1 for MWD, 2 for RSS)	2 stabs
Well B6	RSS, then bent motor	2 (1 for RSS, 1 for motor)	RSS x2 stab, bent motor x1 stab

³ Intermediate casing point to target depth of well.

Table 4: Formation List

The formation lithology list is depicted in Table 4: Formation List. Due to anonymity, new formation names representing various figures in astronomy and the space program have been given. What makes these wells ideal for analysis with the developed drilling application, is essentially that analogies may be drawn between wells due to similar lithologies and well design.

Figure 4.1 below illustrates a typical well design along with what the operator sees as main challenges and risks in every section. Potential for Geosteering issues are mentioned as a main risk in the 8 ½” section, although potentials root causes are not mentioned. A typical strategy for an operator experiencing challenges is to put focus on the consequence rather than on mitigating the root cause. This may again be a reason for the repeated steerability issues experienced in the 8 ½” section in the presented wells.

For the analyses performed in this thesis the sections in focus are the 17 ½”, 12 ¼” and 8 ½”. All sections are drilled with PDC bits. The 8 ½” section is drilled with OBM, which is important to keep in mind when looking at the generated torque plots, as OBM reduces the generated torque from friction significantly compared to WBM.

Formation name	Lithology
Gemini Sand	Sandstone
Nebula Sand	Sandstone
Orion Sand	Sandstone
Mars Sand	Sandstone
Meteor Sand	Sandstone
Saturn Sand	Sandstone
Jupiter Sand	Sandstone
Mercury Shale	Shale
Aldrin Shale	Shale
Shepard Shale	Shale
Gargarin Shale	Shale
Saturn Shale	Shale
Pluto Shale	Shale
Armstrong Shale	Shale
Lovell Limestone	Limestone
Venus Limestone	Limestone
Neptune Limestone	Limestone
Apollo Limestone	Limestone
Asteroid Salts	Salt and Anhydrite

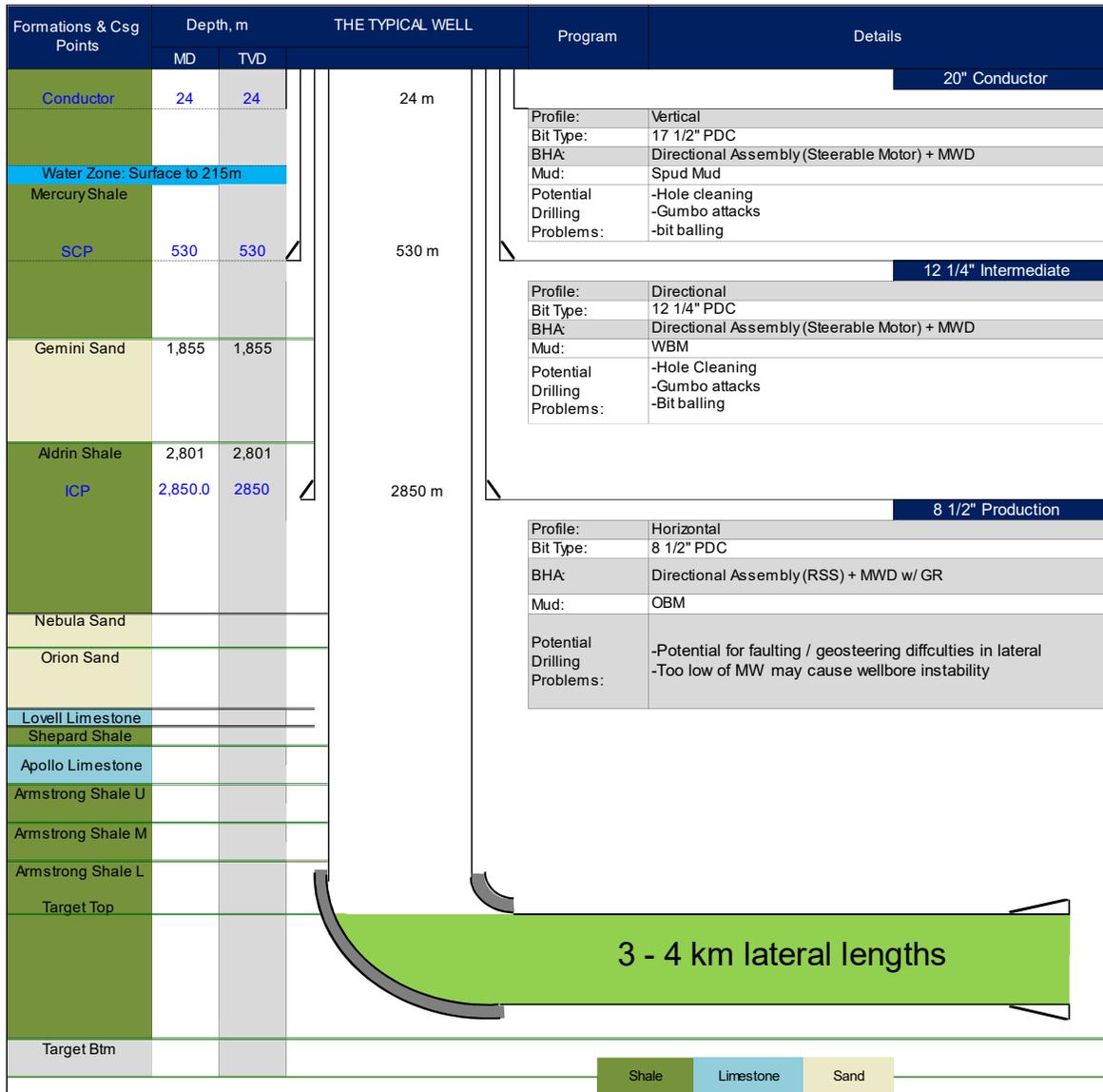


Figure 4.1: Typical well illustration and associated lithology depths. Important section specific data are displayed together with the operator's main concern in each section.

4.2 Creating baselines

Several of the drilling application's modules rely on the changepoint algorithm for detecting changes in mean shift. Hence, the results of using different input parameters for the penalty term is presented below in Figure 4.2. The input signal used for this demonstration is an arbitrarily selected interval of MSE versus depth in the A3 well. The changing variable in these results is the penalty term which represents the heuristic part of the method. A tradeoff between the number of representable baselines and error term may be observed in Figure 4.2 below. An increasing penalty term decreases the number of representable baselines, however it also increases the error or discrepancy between baselines and actual data points.

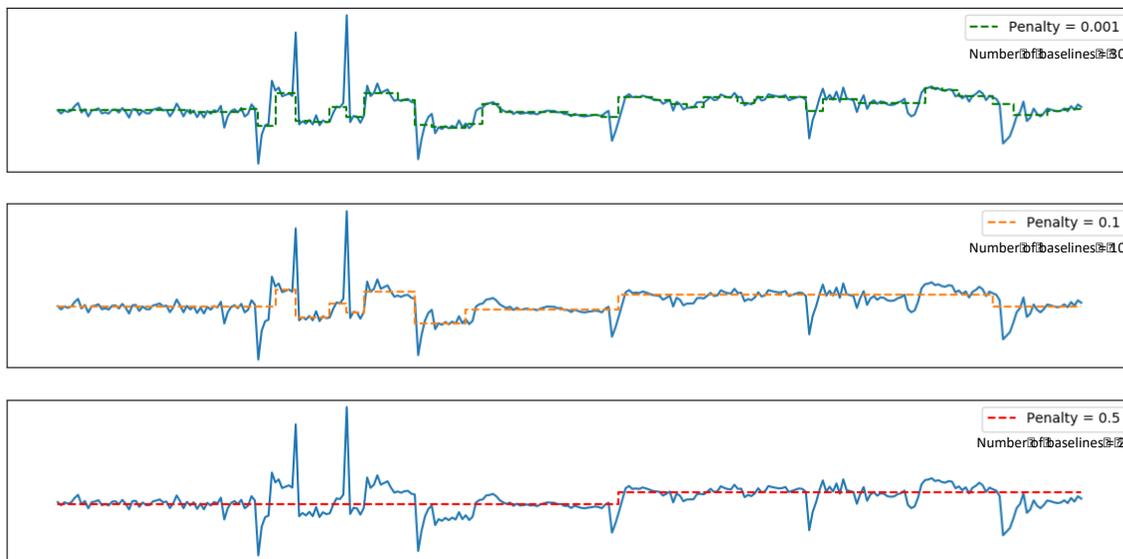


Figure 4.2: Baselines consequently reduces the data over several magnitudes by finding the trend of the data. Different penalties reveal different number of baselines.

Because of the way each baseline is returned and visualized, the total number of data points in each baseline array equals $2P$, where P represents the number of baselines. By looking at the graphs on Figure 4.2, one can clearly see how the baselines follow the general trend of the data, with a lower penalty resulting in more fine-tuned baselines. The obvious weakness (and strength)

of creating baselines like this, is that the effect of outliers will be omitted as the baseline is generated from the median of the interval.

4.3 Concept: establishing an overview

Initially, when opening the application an overview of the entire well data is presented in a landscape mode. Formation tops, lithology, and section depths are displayed by default, given that the user has provided the information in the input files. Figure 4.3 below is an example overview of well C4. The figure is split into five rows. The top row features downhole and surface MSE, respectively depicted as red and brown traces. The second and third row respectively depict WOB in green, and rotary and bit RPM in blue. The second last row (from top to bottom) depicts surface torque in light brown, and bit torque in dark green, if available. The last row contains motor differential pressure and ROP, respectively as blue and black traces. The plot background area is color-coded by formation lithology (green = shale, yellow = sandstone, blue = limestone) with light grey lines indicating formation tops. Vertical thick black lines represent section depths or casing shoe depths.

In default mode, the well 'Overview' plot provides an efficient initial view and visual holistic check of which formations or sections are prone to higher levels of drilling inefficiency (i.e. higher MSE). To illustrate, one may clearly see higher levels of surface MSE (brown trace, top plot) in the Apollo Limestone (blue shaded area around 4000m MD), from the overview of well C4 in Figure 4.3. This indicates higher levels of drilling inefficiency in the Apollo Limestone, relative to adjacent formations.

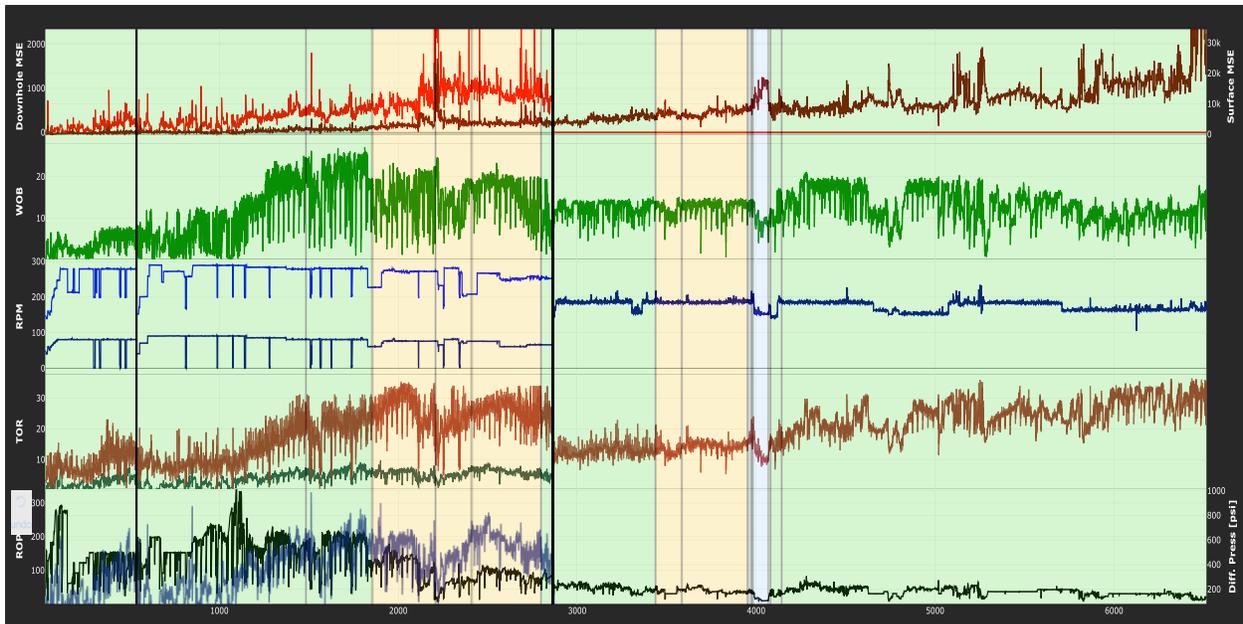


Figure 4.3: Well C4 Overview. Abrupt change in parameters when casing is set (Black vertical line).

One may also investigate data by section or formation lithology by selecting the desired interval from the range slider above the plot (presented in the User Interface chapter) or by simply selecting a range in the plot. Below, in Figure 4.4, is a snippet of the 8 ½”-section from well C4. Again, notice the clear response in MSE in the Apollo Limestone (blue shaded area at approximately 4000m MD). The increase in MSE in the Apollo Limestone indicates a higher amount of bit whirl from entering a harder formation and losing DOC. This effect is worsened when reducing WOB, causing even higher levels of bit whirl and loss of DOC. One may also notice other regions of drilling inefficiency in the Lower Armstrong Shale (green shaded area from approximately 4100m MD to TD). The increases in MSE neatly correspond with changes in operating parameters or lithology transitions, indicating several cases of distinct bit dysfunction, which will be discussed in more detail in the diagnostic section. The strength in analyzing well data like this, is the ability to get an initial idea of areas or specific lithologies which are prone to bit dysfunctions limiting performance in the well, right off the bat.

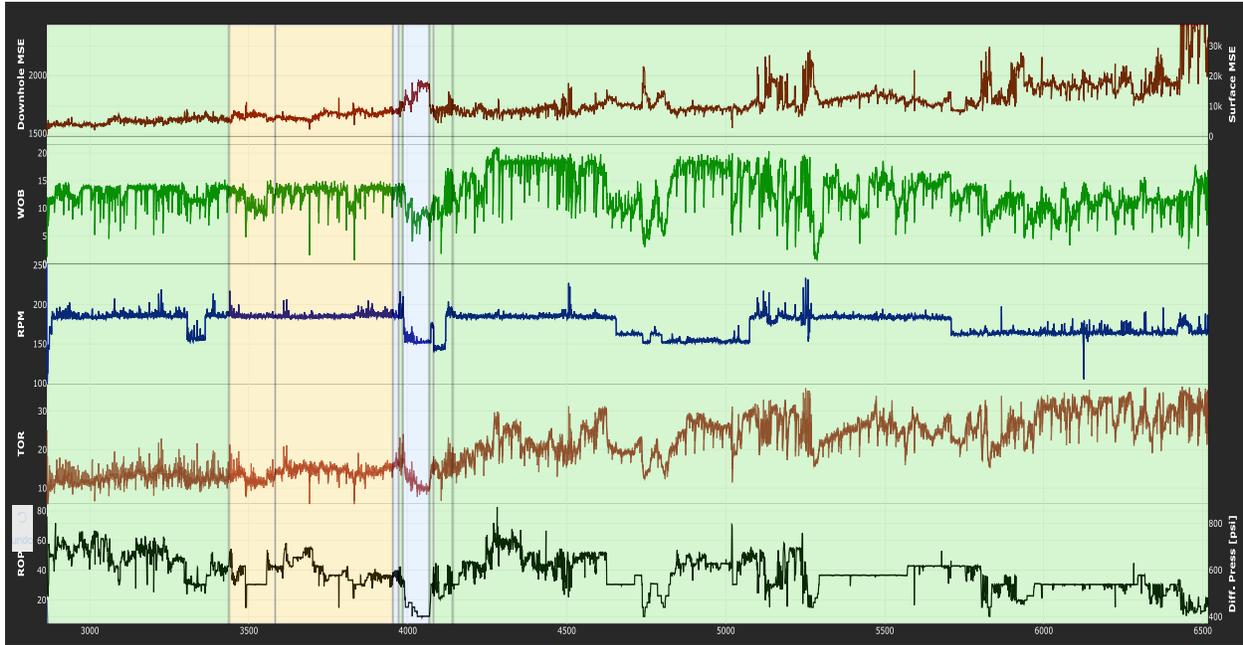


Figure 4.4: Well C4 8.5 Section, shale in green, sand in yellow and chalk in blue.

Once an overview and initial idea of potential drilling inefficiency in specific lithologies or sections have been established, one may further investigate lithology specific data by surface regression for determining optimal parameter ranges.

4.4 Concept: lithology specific road-maps

Being able to identify trends and predict future behavior is where regression analysis has its value. Understanding how data is organized in relation to each other is where analysis provides its value. By evaluating the relationship of a dependent and a set of independent variables, one can understand how the dependent variable varies with changes in one of the independent variables.

The theory and method of creating lithology specific “road-maps” have been presented in the theory and method chapter, respectively. Below is a field example to show the concept of establishing such a road-map for an arbitrary formation. Selecting a new data subset from a desired formation lithology may typically contain many data points.

Now imagine a new offset well is to be drilled near the C-pad. Well C3 and C4 are now our reference wells. One may establish a road map by aggregating and tuning data from both reference wells on a formation basis, using the drillWiz application. After selecting the desired well and formation, e.g. the Mercury Shale on well C3, and clicking the “Surface Road-Map” button in the application, the left part of Figure 4.5 is generated. The road-map specific data points from the interval are scattered with a color map from green to red based on the MSE severity. Now, sufficient tuning of changepoint penalties (input boxes in the application) is required to obtain representable baselines of the interval. A visual check of the baselines is presented to the right in the figure. The next step is to push the “Filter Baseline” button which yields the picture to the right of Figure 4.5.

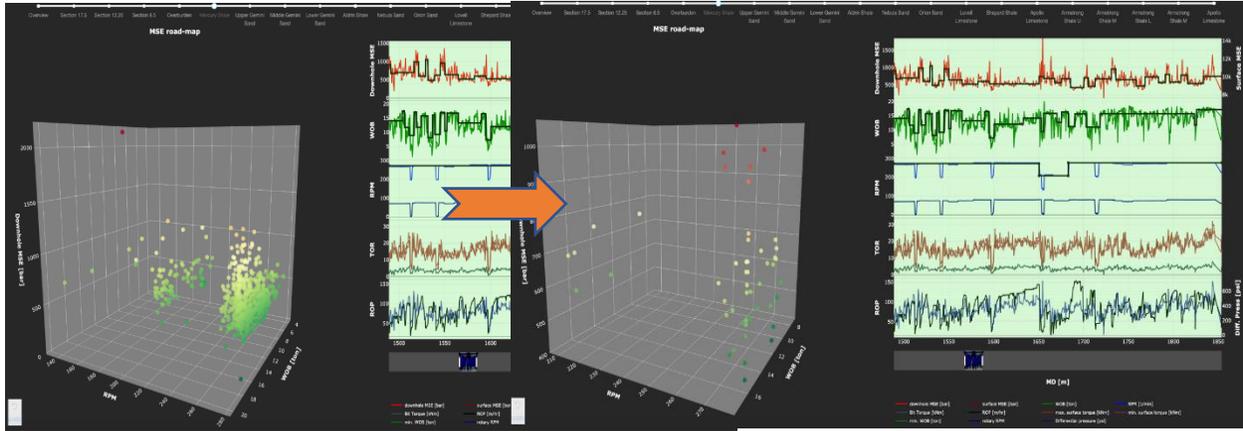


Figure 4.5: Filtering of baselines results in less data points for analysis.

Once the data has been compressed to a representable combination of parameters in the interval, one may create the surface road-map. Clicking the “Create Surface Regression” button, reveals the left stage of Figure 4.6. The Mercury Shale in well C3 is now regressed and optimized. However, one may think this was a bad or insufficient result. The next step is therefore to aggregate more data from other wells penetrating the same lithology. To do so, one must click the “Store Data” button, before choosing the next well. The “Store Data” function will temporarily store the selected data points in the browser memory. The next time the “Store Data” button is clicked, it will automatically check that the new data points are from the same lithology under investigation, before also storing them. After selecting the same lithology in well C4 and filtering new baseline values, one may store the new data points from both well C3 and C4 in the scatter plot. If the surface regression function is activated again, the middle part of Figure 4.6 is revealed, proving a different plot and regression trend. A new optimum parameter selection is identified for the aggregated data points, represented in the top right part of the application and to the right in the figure below, ref. Figure 4.6.

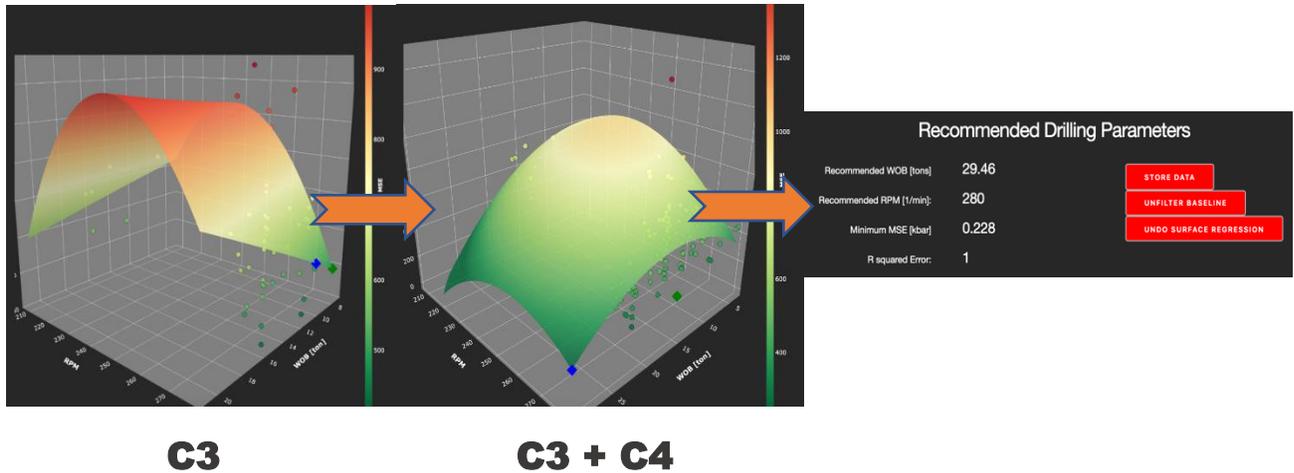


Figure 4.6: Introducing additional data points results in more accurate surface regressions, but also heavier computations. Output parameters are displayed together with recommended parameters.

4.5 Case study: lithology specific road-map for drilling the Apollo Limestone

Having been introduced to the general concept of lithology specific road-maps and how data goes from long columns of numbers and graphs to tangible outputs as figures and operational recommendations, the next step is to do a deeper case study for the “next well”. To do this, a formation with known issues is picked out for analysis to determine the dos and don’ts of the next operation.

Considering the issues reported in the 8 ½” section of well C7, no apparent issues were reported from the ICP of 2855 m MD to the Apollo Limestone transition at 4100 m MD. Considering that this section had 5 trips in total, and all of them originating in the Apollo Limestone, this is an interesting interval to investigate. The trips in this section were a combination of an inability to achieve desired build rates, and tool failures. Considering BHA configuration, one must differentiate the intervals with bent motor configurations and RSS. Conclusively, one may separate the results into two parts, one case when drilling with a RSS with no motor, and the other when drilling with a bent motor.

4.5.1 Drilling with a RSS

The following intervals, presented in Table 5: Case data for RSS, are used as the basis for the Apollo Limestone case when drilling with an RSS.

Table 5: Case data for RSS

Case data	
Formation	Apollo Limestone
BHA configuration	Rotary Steerable System
Section	8 ½” Build section
Wells analyzed	Well C3, Well C4, Well C7 ⁴

With data provided from the directional service company, surface MSE has been calculated for the wells C3, C4 and part of C7. In all three wells, the depths and inclinations entering the Apollo Limestone are quite similar, providing confidence in the assumption that well path and rock strength will be sufficiently similar for comparison of MSE between the wells. Remember, the exact value of MSE is of less significance, as we are more interested in the established trend.

To pick out the “representable” parameters desired, the following changepoint algorithm parameters have been used as an input in the application, ref. Table 6: Baseline parameters to achieve results.

Table 6: Baseline parameters to achieve results

Baseline parameters for changepoint algorithm	
WOB penalty	0.3
RPM penalty	0.1
MSE penalty	0.3
Minimum baseline length (unitless)	5

⁴ Only run 1 and 2 are analyzed as the remaining runs were drilled with a bent motor

Using the input from Table 6 above, Figure 4.7 below is generated, displaying the Apollo Limestone section in well C3. Overlapping baselines of surface MSE, WOB and RPM are picked out with the PickPoints function, described in the Method section. Notice that the parameters vary quite a lot compared to the generated baselines. The reason for not using a higher accuracy is divided in two. First, one reduces the data tremendously by using a higher penalty compared to the original data. Secondly, as mentioned above, we do not need to know exactly what happens in every increment of depth, as the overall trend is what we are looking for, being able to generate a road-map with “red and green” operating areas.

A quick look at the close correlation between WOB and surface torque indicates that weight is being delivered to the bit as expected (i.e. no weight transfer issues of significance). Due to the fact that inclination is only a few degrees and close to vertical in this interval, surface torque and thereby surface MSE is a viable option for analysis.

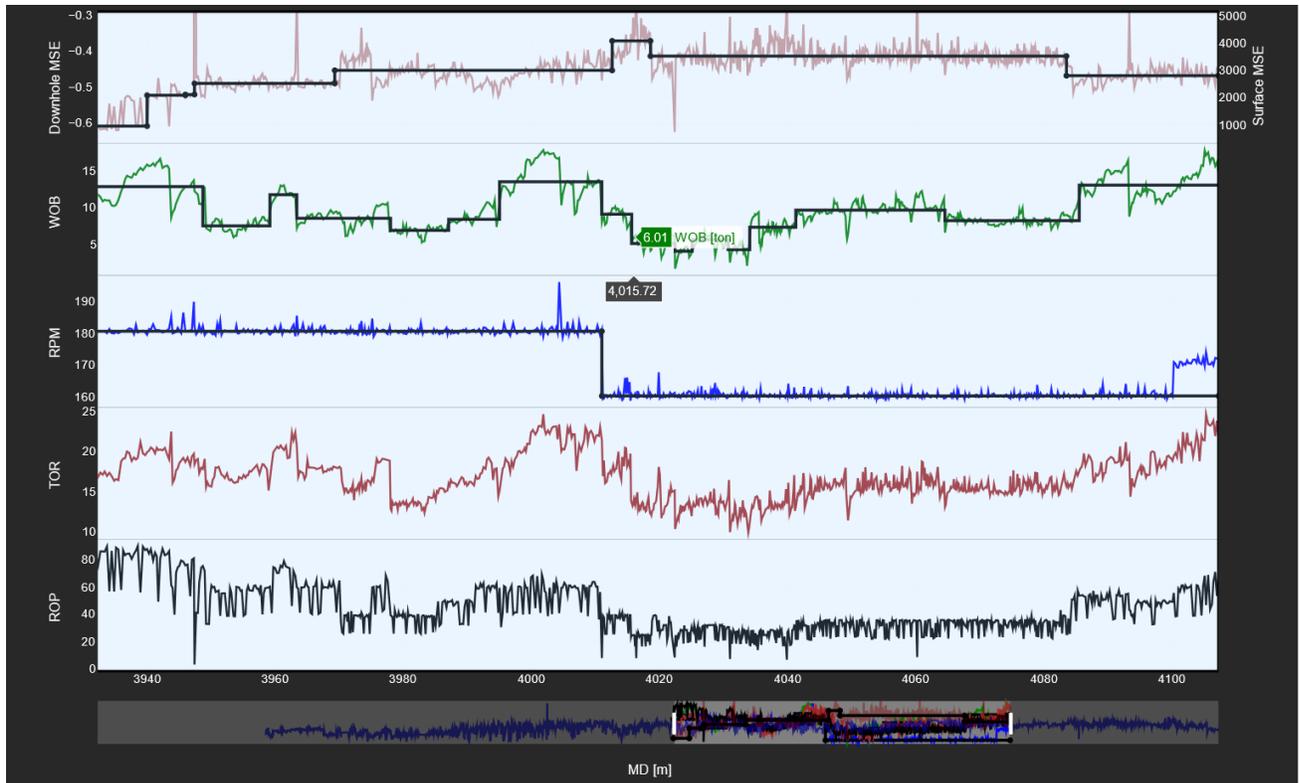


Figure 4.7: Baselines generated with sufficient accuracy.

Using the baselines generated in Figure 4.7 above, a resulting set of parameters from well C3 are obtained and saved in the application. Moving onto well C4, the next set of parameters are obtained in the same way, using the same changepoint parameters. Figure 4.8 depicts the developing 3D-scatter plot when the wells C3, C4 and the two runs from well C7 are aggregated in the same plot. Notice that for each aggregated well, more pieces are put into the puzzle filling the operating window, and painting a picture of where to operate and not.

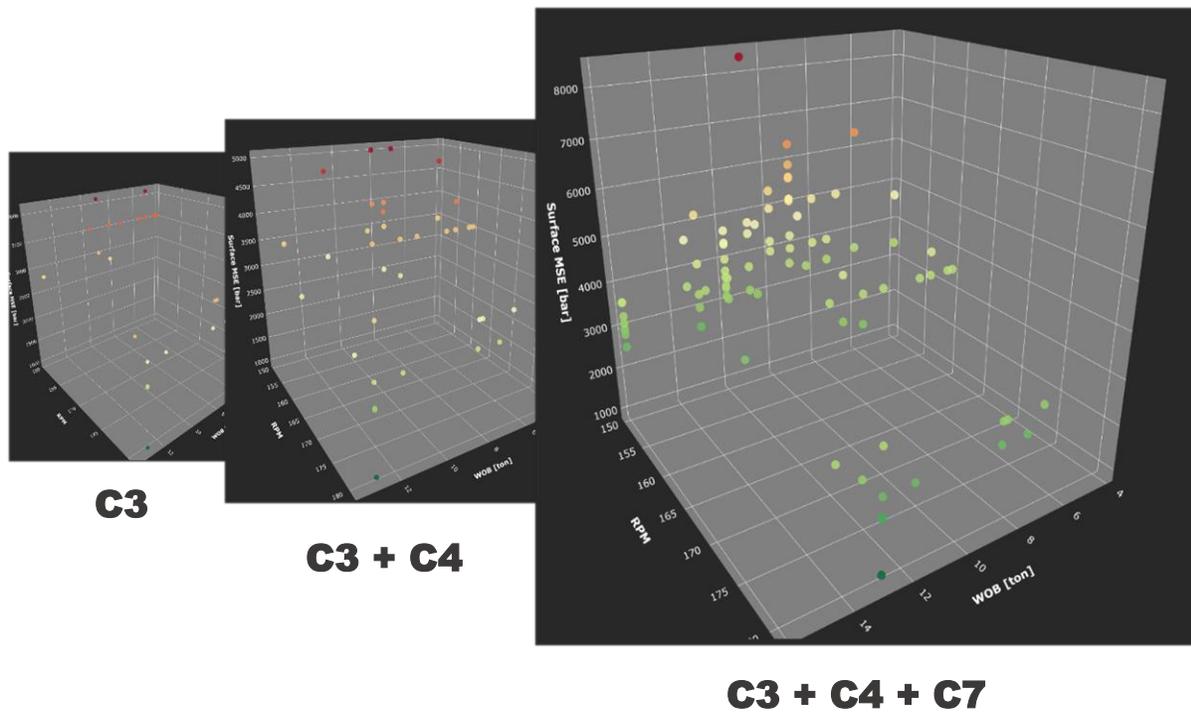


Figure 4.8: 3 wells are aggregated to fill the operating area with only representable data.

Having populated the operating domain sufficiently, we have now run out of pieces and must therefore pull out our regression tool to “forecast” the rest of the puzzle. When initiating the surface plot and the optimizer algorithm, a 3D-surface plot is generated as illustrated in Figure 4.9 to the right. The green marker represents the minimum of the surface, while the blue marker represents the optimal operating parameters within the acceptable MSE threshold.

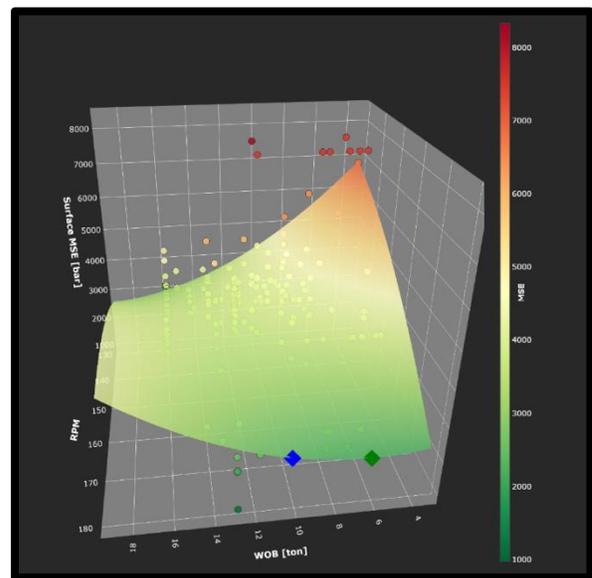


Figure 4.9: Surface generated from aggregated data points. Optimum located at blue point, minimum MSE at green point.

By further evaluating the generated 3D-surface through rotation, one can see how well the domain is fitted with data, ref. Figure 4.10 below. An obvious weakness with the data points picked for regression, is the overlapping MSE values for same input of WOB and RPM. However, as previously mentioned, we are only interested in the bigger trends. Knowing exactly the MSE value at each combination has little value, but knowing which direction of operating parameters propose a lower MSE is paramount.

From Figure 4.10 one can conclude that the desired operating RPM should be 180 and the WOB should preferably be about 10 tonnes. Although the algorithm produces a recommended parameter combination, one can clearly see that further exploration is necessary. This is because the operating area around the minimum MSE value is not very well populated and represented by datapoints compared to other areas, which may reduce the credibility of the trend. Another fact is that this recommended parameter selection should only be a preliminary goal for real-time step tests in the interval, in order to maximize performance gains by further exploring the operating window. Looking at the surface plot, one can clearly see that running 135-155 RPM together with low WOB results in high MSE, most likely due to bit whirl. This is clearly indicated as a region to avoid.

Recommendation from surface regression: Operate at 180 RPM and start step test at 10 tonnes.

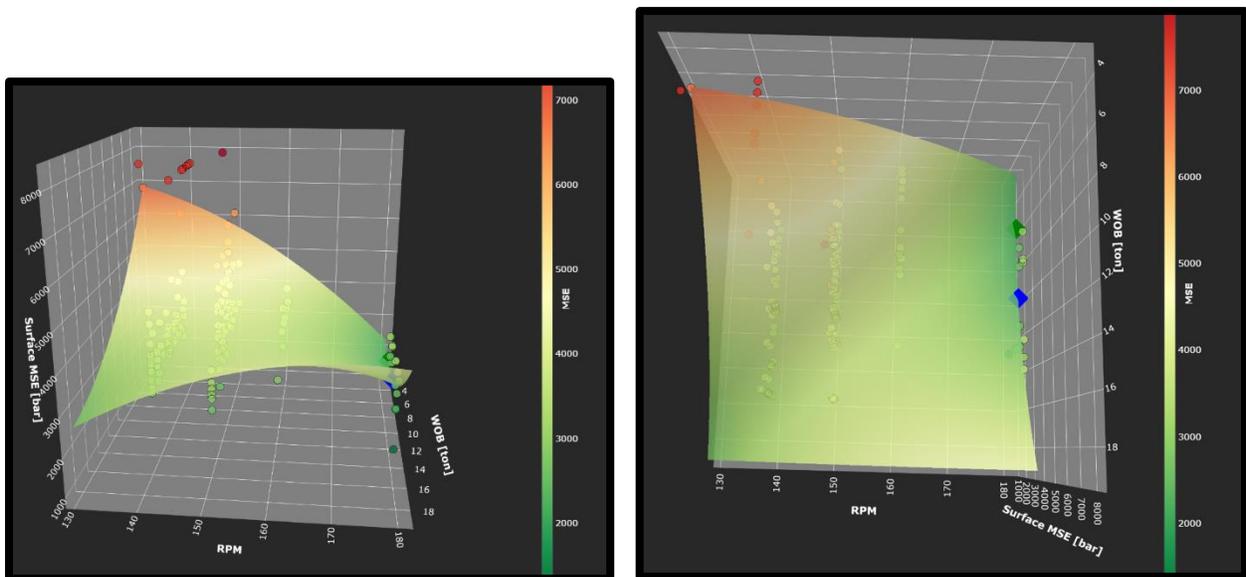


Figure 4.10: A second and third angle of the generated surface regression for the RSS case.

4.5.2 Drilling with a bent motor

When analyzing the Apollo Limestone, it can be interesting to look at how a case study varies when considering the runs with bent motors. The common denominator of the motor runs is that they all were a contingency after the initial RSS configurations failed to reach TD for different reasons. Switching to a bent sub with a motor facilitates the use of downhole MSE, which will provide a much more accurate picture of what is happening at the bit compared to the surface MSE. Remember, downhole MSE is only dependent on what's below the motor. The footage drilled in this case study is taken from the lateral part of the well, emphasizing the need for downhole MSE. A summarized table of some of important case data for wells analyzed in this case is presented below, ref. Table 7: Case data for the bent motor drilling scenario.

Table 7: Case data for the bent motor drilling scenario

Case data			Baseline parameters for changepoint algorithm	
Formation		Apollo Limestone	WOB penalty	1.5
BHA configuration		Bent motor	RPM penalty	1
Section		8 ½" Lateral section	MSE penalty	1.5
Wells analyzed		Well B6 ⁵ , Well C7 ⁶	Minimum baseline length	5
Motor RPM factor	[rev/gal]	0.28		
Motor torque factor	[ft-lbs/psi]	9.75		
Motor torque rating	[ft-lbs]	12,188		

⁵ Only run 1 and 2 are used as the rest of the well was drilled using a RSS with no motor.

⁶ Only run 5 and 7 are used as the rest of the well was drilled using a RSS with no motor.

The plots presented in Figure 4.11 below shows how the motor runs in well B6 and C7 are aggregated to fill the operating area. Because of the short length in one of the C7 runs, both runs were aggregated together in the figure. If we look at the aggregated data sets, one can notice the significantly higher RPM settings, due to the additional downhole motor RPM.

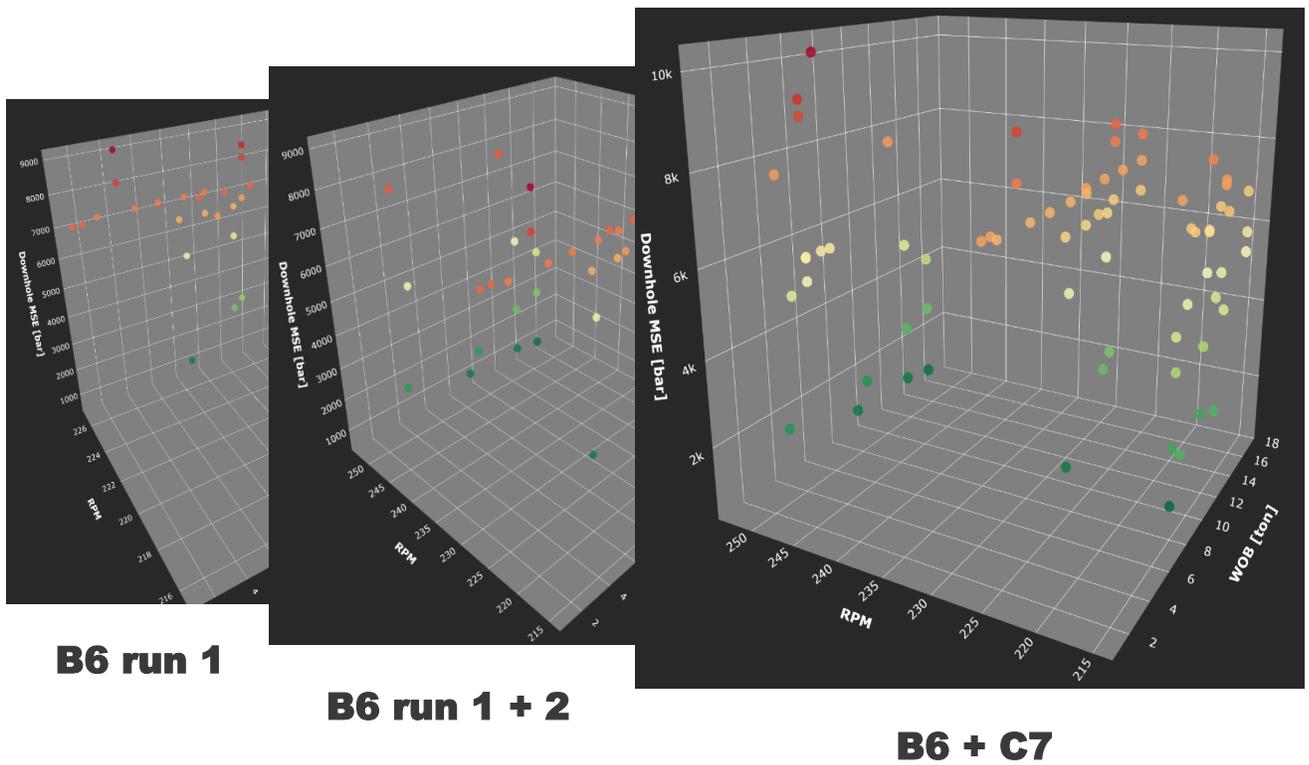


Figure 4.11: 3 wells are aggregated to fill the operating area with only representable data.

While the three-dimensional scatter plots might seem hard to interpret at first, the surface regression provides a convenient weighting of all the data points and picks out the representative regions and trends. Applying the surface regression to the data, yields Figure 4.12, with the optimal point colored in blue and the minimum MSE setting in dark green. The optimal area of the surface tilts towards the region of 17 tonnes WOB and 250 RPM. As seen on the scattered data, operating at about 230 RPM most likely results in resonant vibrations and yields high MSE values. This plot does in fact make sense when considering that low WOB initiates bit whirl, and too low RPM with high WOB may result in stickslip or simply be a resonant frequency.

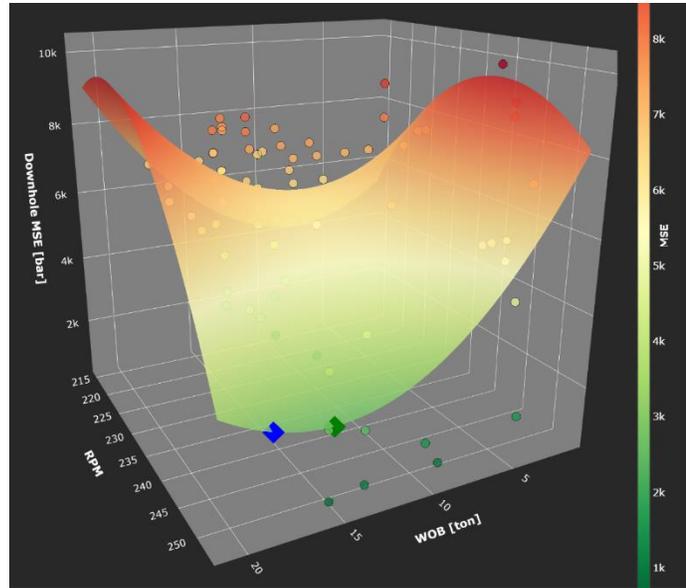


Figure 4.12: Generated surface from bent motor drilling. Optimal point in blue, minimum in green.

4.6 What results will you show your manager?

When considering new methods and techniques like the application and methods presented in this thesis, one may easily forget about the most important aspect of new technology. How does this benefit my company, and in this case, how does this make me money? That’s what your manager will ask for, and this is what you will show him/her.

Introducing a drilling performance factor, assuming efficient drilling operations:

$$\beta = \frac{WOB_{After}}{WOB_{Before}} \times \frac{RPM_{After}}{RPM_{Before}} \quad (21)$$

Table 8: Summarized results from the two road-map cases.

Case	Parameters at current regime (WOB,RPM)	Suggested parameters w/ drillWiz (WOB, RPM)	β , Drilling performance factor
C3: RSS, no motor	(9,160)	(11, 183)	1.40
B6: Bent motor	(10,215)	(17,250)	1.98

Table 9: Economic potential for well B6 and well C3.

Assuming a rig rate and using the data from well B6 and C3, the calculation in Table 8 and Table 9 shows the potential in performance gains, excluding gains from avoided tool failures and additional rig time from tripping pipe.

One might think this is an overly confident estimate, i.e. saving over \$40,000 in one formation, just from parameter adjustments on

B6. This is from assuming an efficient drilling operation. If indeed inefficient drilling were

	Well B6	Well C3
Rig rate ⁷	\$20,000	
Formation	Apollo Limestone	
BHA config	Bent motor	RSS, no motor
ROP at current settings	23 m/hr	31 m/hr
Section length [m]	2,591 m	247 m
Performance gains	\$46,464	\$2,940

⁷ Rig rate is assumed to be \$20, 000 as commercial numbers will not be used, despite the anonymity.

occurring before applying these parameter changes, even more gains (disproportionate) would have been seen.

Considering the amount of tool failures experienced in this formation with the current operating parameters, one may assume that drilling has been inefficient and severe vibrations may potentially be the root cause for tool failures. Moving towards more efficient operating parameters will consequently mitigate excessive vibrations and increase tool longevity considerably, avoiding unnecessary trips. This is not because operating parameters damage your tools (i.e. high WOB), it is the onset of a dysfunction and severe impact shocks associated with it, which is typically the root cause for damaging downhole equipment.

Moving on to the next part of the application, diagnostics, it will be interesting to see if in fact severe vibrations and bit dysfunction have been limiting performance and causing tool damage in the Apollo Limestone.

4.7 Proof of MSE with simple diagnostics

4.7.1 Bit whirl

After stating that MSE should be the underlying metric for drilling efficiency, it is paramount to show proof of MSE as a diagnostics tool in simple scenarios. The following two subchapters will show basic concepts in diagnosing whirl and stick-slip from well data. The figure below depicts a small data subset from the 8 ½’’ section of well A3 with constant RPM, ref. Figure 4.13. As discussed in (Berge-Skillingstad & Anderssen, 2018), bit whirl is a dysfunction that can be characterized and diagnosed with reasonable uniqueness by a disproportionate increase in performance from an increase in WOB. This is clearly depicted in the figure below.

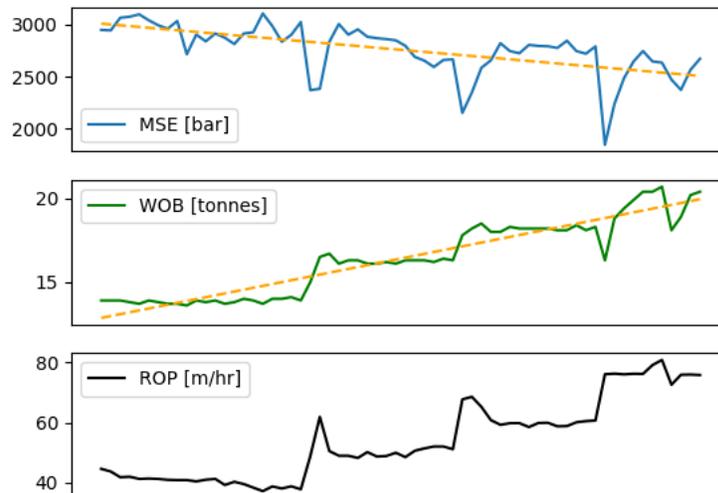


Figure 4.13: Whirl scenario depicted. Notice how MSE decreases with increasing steps of WOB.

The figure depicts a so-called step test or parameter exploration initiative, by slacking off weight in steps while observing for a response in MSE. As depicted by the yellow trend lines in Figure 4.13, there is a clear decrease in MSE from a disproportionate response in performance (ROP) when increasing WOB. This is a clear indication of bit whirl, which is characteristically the only bit dysfunction which gets better with increasing WOB.

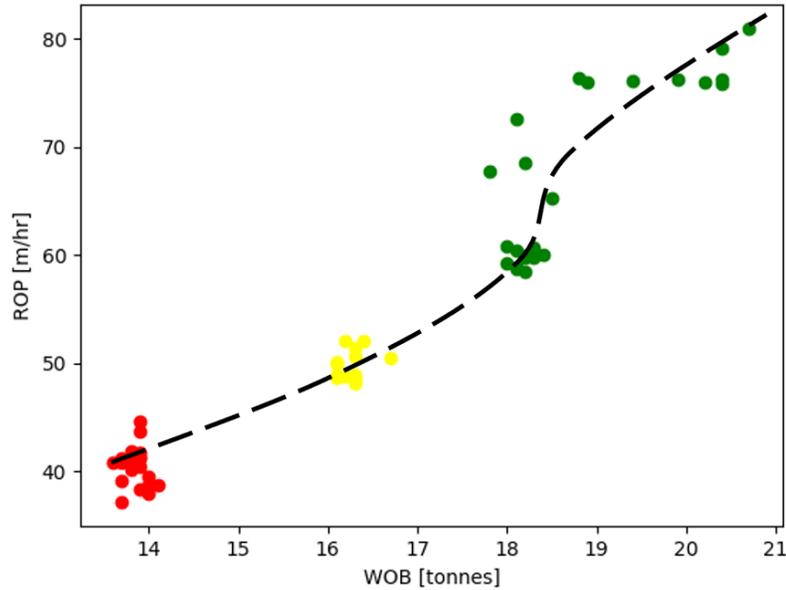


Figure 4.14: ROP vs. WOB proportionality shift at higher WOB. Indicates whirl.

Recall that analyzing trends in MSE will conveniently identify which region of the drill-off curve the bit is being operated in, ref. Figure 2.1. The figure above depicts the operating parameters and response in ROP from the step test, ref. Figure 4.14 Note the non-linear response at approximately 18 tonnes WOB from sufficiently burying the cutters, constraining lateral bit movement and preventing bit whirl. The figure above neatly depicts the transition from region 1 to region 2 in the traditional drill-off curve in Figure 2.1.

From the generic diagnostics plot presented in (Berge-Skillingstad & Anderssen, 2018), Figure 4.14 is a typical response in bit whirl. Thus, it can with confidence be diagnosed accordingly. From an operational point of view and based on the response in MSE, the drillers action to increase WOB in this situation was correct. This emphasizes the need for parameter exploration as a continuous process by always looking for the “optimal” set of parameters.

Below is another figure from an arbitrary data subset affected by bit whirl, ref. Figure 4.15. A difference from the previous figure is that the depth of cut is now displayed to illustrate the physics of the phenomena. MSE spikes radically when WOB is reduced and bit whirl is initiated. Following the decrease in WOB is a substantial loss of DOC which reduces the lateral resistance of the bit, inducing excessive energy losses from bit whirl.

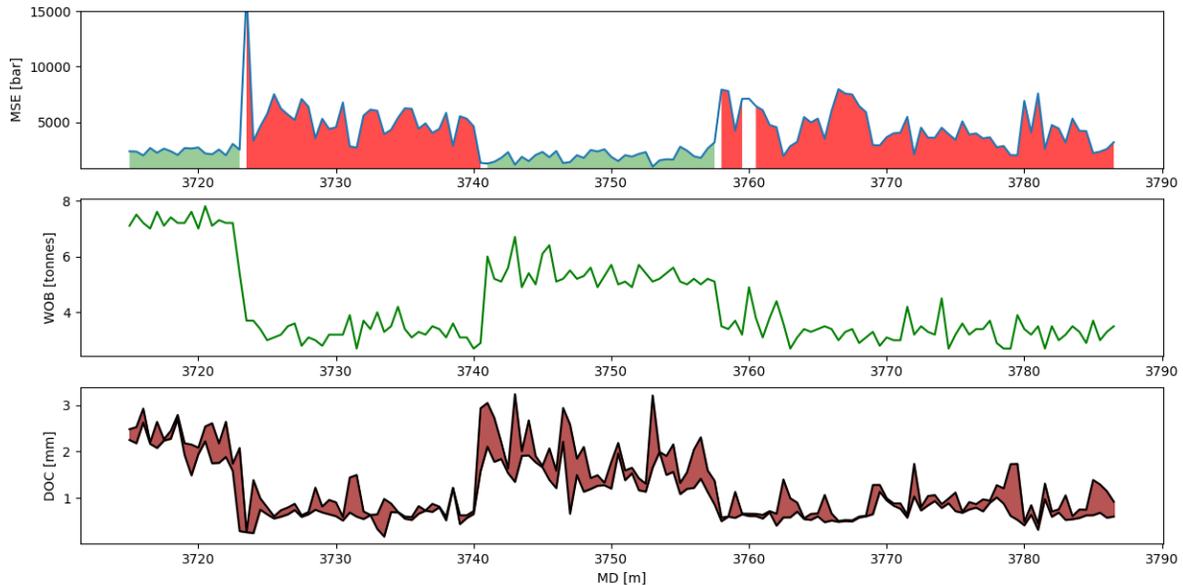


Figure 4.15: Whirl scenario taken from an arbitrary well. Whirl intervals colored in red.

4.7.2 Stickslip

Recall that MSE alone might not be the best diagnostics tool for identifying stickslip as bit speed oscillations from torsional vibrations do not affect performance significantly. However, full-stick events cause erratic and severe MSE spikes and are what are typically most damaging to the bit. Figure 4.16 below depicts an interval with high amounts of shale drilled at about 2500m MD in well A3. In the figure, large torque oscillations between 30 and 5 kNm can be observed in the interval. Even with large torsional oscillations from rapid changes in maximum and minimum torque present, only certain smaller intervals are severely affected by large MSE spikes. These are the most severe areas of interest, and where we can diagnose full-stick.

By setting an arbitrarily chosen “Survival MSE” of 4100 bar, one may observe a corresponding survival WOB at approximately 18 tonnes. Every time the weight is raised above 18 tonnes, MSE spikes and the risk of damaging the bit increases drastically. A red color has been filled under the MSE graph wherever WOB was raised above 18 tonnes, providing a good match between the high WOB and the proposed full-stick response in torque and MSE.

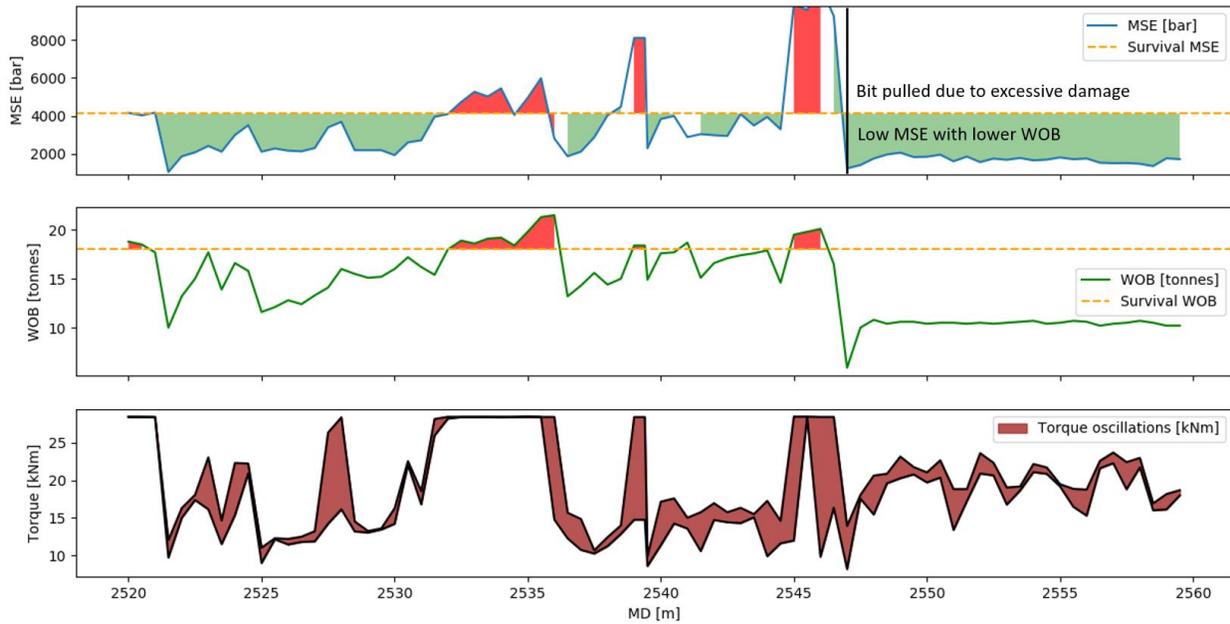


Figure 4.16: Stickslip scenario from well A3 with a bit trip on the black line. Stickslip events colored in red.

According to the well reports, the bit was eventually pulled due to an abrupt change in performance, which could be explained by full-stick as the bit was described with nose damage. Notice how the WOB after the bit trip was considerably lower and did not trigger any major torque oscillations or MSE spikes. Casing was set not very much deeper in the section, illuminating the potential savings if a lower “survival” WOB had been run from identifying stickslip earlier in the section, and potentially avoiding the bit trip.

4.8 drillWiz Diagnostics Scenarios

Having been introduced to the concept of identifying bit dysfunction with distinct responses in MSE and operating parameters, comes application in practice with several case scenarios below. The cases below have all been run through the diagnostic algorithms presented in the *Methods* chapter. Specifically, well C3 and C4 have been run through the drilling application drillWiz, and the automated diagnostic results are presented below. Keep in mind that the plots are in vector format (.SVG).

4.8.1 Case: Well C4 diagnostics study

Recall the 8 ½' section in well C4 from *Establishing an Overview*. The diagnostics overview of the same section is presented below, in Figure 4.17. When selecting the diagnostics module in the drillWiz application, two additional rows of traces are added to the standard 'Overview' plot. The top row now features two traces depicting whirl and stickslip severity in black and red traces, respectively. The whirl severity trace (in black) is estimated by the pseudo code presented in the whirl diagnostics section under the *Methods* chapter. Similarly, the red stickslip severity trace is estimated by the torque changepoint pseudo code for indicating stickslip severity, presented under the stickslip diagnostic section in the *Methods* chapter. Note that the stickslip severity trace depends on variance changepoints in surface torque and is considered a diagnostic with relatively low confidence. This is due to changepoints in torque variance and MSE alone is considered a generally weak diagnostics tool for stickslip, compared to more superior alternatives such as calculating or directly measuring bit speed RPM. For this reason, and from a computational efficiency standpoint, the top red stickslip severity trace is only generated upon selecting the additional 'torque changepoint' radio point requirement in the drillWiz application. The second top row contains a torsional severity score, which supplements and acts as a stickslip severity estimate. The row contains a torque severity score in blue, MSE severity score in green, and total severity score in red. The scoring system and pseudo-code is presented under the torsional severity section in the *Methods* chapter. The following five sub rows contain the same traces as depicted and described in the 'Overview' plot under *Establishing an Overview*. New to

these rows is, however, a ‘dashed’ linear regression line of MSE overlying the current MSE trace, representing the overall trend increase in MSE. Additionally, the generated MSE baselines are overlying the current MSE trace (black MSE baselines overlying the brown surface MSE trace). Upon selecting the ‘torque changepoint’ alternative in the drillWiz application, changepoints in torque variance are displayed as vertical red lines overlying the second to bottom torque row of the figure plot. These vertical lines represent changes in surface torque variance or frequency, indicating a changepoint in oscillation.

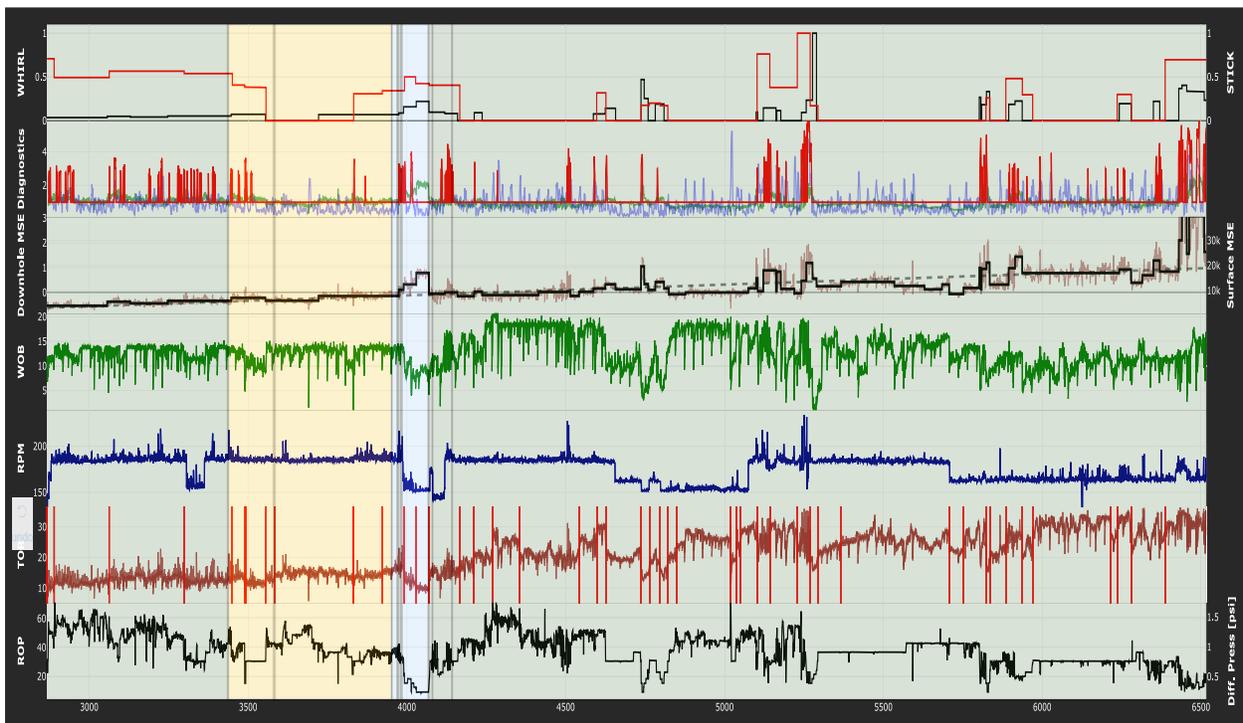


Figure 4.17: Well C4: 8 ½’’ section diagnostics (unitless penalty parameters: MSE penalty = 0.3, Torque penalty = 1, minimum baseline length = 5)

Right off the bat, one may notice several distinct impressions of bit dysfunction in the previously described regions. Take the Apollo Limestone (blue shaded area at 4000m MD), there is a relatively large response in whirl from the combined effect of increasing MSE and decreasing WOB. This is evident in the whirl severity trace (black) in the top plot, which we will come back to. There also seems to be a large response in stickslip that transitions into whirl in the Lower Armstrong Shale at approximately 5200m MD. Let’s take a closer look.

Figure 4.18 below contains a zoomed-in snippet around the inefficiencies occurring shortly after 5000m MD from the diagnostics overview above. Notice that specifically two regions seem to have erratic MSE responses, well above the general linear regression trend line (dashed black line in the MSE plot). Concurrently, at approximately 5100m - 5140m MD and 5240m - 5265m MD, both the torque changepoint algorithm and torsional severity algorithm generates large responses in total torsional severity score with a large torque variance and MSE response. This indicates high levels of stickslip severity (depicted by the red traces in the two top rows). This diagnostics confidence in stickslip severity is further solidified by the oscillation response from the top drive SoftSpeed software tuning RPM in the same interval. A natural response by the operator is to reduce WOB to lessen the severity of stickslip, however, the reduction in WOB is coupled by a large whirl event depicted by the top black trace in the diagnostics plot. The large response in whirl severity comes from reducing WOB after both stick scenarios, but not alleviating MSE. At this low WOB level, MSE continues to remain severe, well above the general interval trend line. A reduction of WOB from 18 to 2 tonnes at approximately 5275m MD has alleviated the stickslip response, but has now initiated a severe whirl response. Further increasing WOB to first 6, then 15 tonnes seems to suppress bit whirl as MSE moves towards a more stable efficient level in the remaining interval.

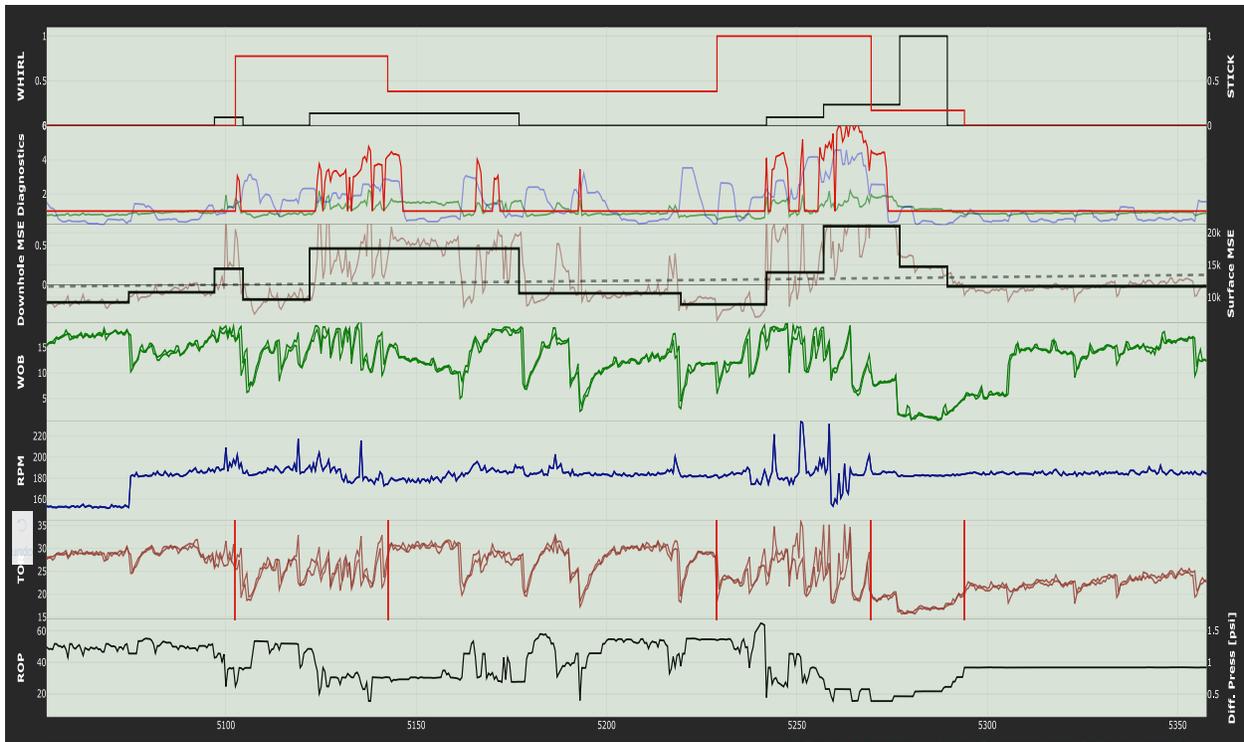


Figure 4.18: Well C4: 8 1/2'' section diagnostics: Zoomed-in (unitless penalty parameters: MSE penalty = 0.3, Torque penalty = 1, minimum baseline length = 5)

The general preferred outcome of the application in this situation is to identify the relative severity of limiters from offset reference wells in distinct sections or formations beforehand, enabling an engineering redesign process to extend the limiters in the next well. Apparent in this interval, is that the operator is limited by stickslip in the Lower Armstrong Shale. The natural response is to lower WOB, which transitions the bit dysfunction into bit whirl. A desired outcome of the application in this scenario, is to alert the operator to redesign the BHA to limit coupled stickslip and whirl responses in the Lower Armstrong Shale, for the next well. Such a proposal could have included e.g. thicker pipe for less torsional vibrations or depth of cut control to allow higher WOB levels to suppress bit whirl, but at the same time limit torsional oscillation from DOC oscillation in an aggressive bit.

Another desired outcome is to identify optimal operating settings. For instance, in the interval above one can notice severe stickslip responses seem to onset every time WOB exceeds 18 tonnes. In addition, bit whirl seems to onset at WOB levels lower than 10 tonnes. MSE moves

towards a more stable and efficient level at approximately 15 tonnes WOB from 5300m MD and onwards. This information indicates a desired operating setting between 15-18 tonnes WOB. Note that the Lower Armstrong Shale is very heterogeneous meaning efficient operating settings may adaptively change, further emphasizing the value of recurrent step-tests to explore the operating area.

Given that the diagnostics determine only a relative severity, based on the maximum severity experienced in the selected interval, it may be interesting to look at certain intervals one at a time.

Let's take a closer look at the Apollo Limestone in well C4 again. The operator has provided us with memory sub data of vibrational shock sensors. To this reason, if available, the diagnostics module of the drillWiz application will overly the plot with an additional top row containing shock sub data. The figure below contains such memory sub data, which serves as an additional level of confidence to compare the automated diagnostics results against, ref. Figure 4.19. The top row now contains lateral RMS values in black, and torsional RMS values in red. Interesting to note here is that, data is more high frequent, showing more trends more clearly, in a diagnostics point-of-view. Upon transitioning into the Apollo Limestone, quite severe lateral and torsional RMS vibration levels are recorded. These torsional RMS vibrations occur concurrently with responses in the diagnostic torsional severity score and torque changepoint stickslip severity. However, no noticeable responses in MSE are recorded, relative to the general trend line or the rest of the interval. This indicates that indeed torsional and lateral vibrations are occurring in the BHA, but these do not seem to translate to the bit, affecting the drilling efficiency.

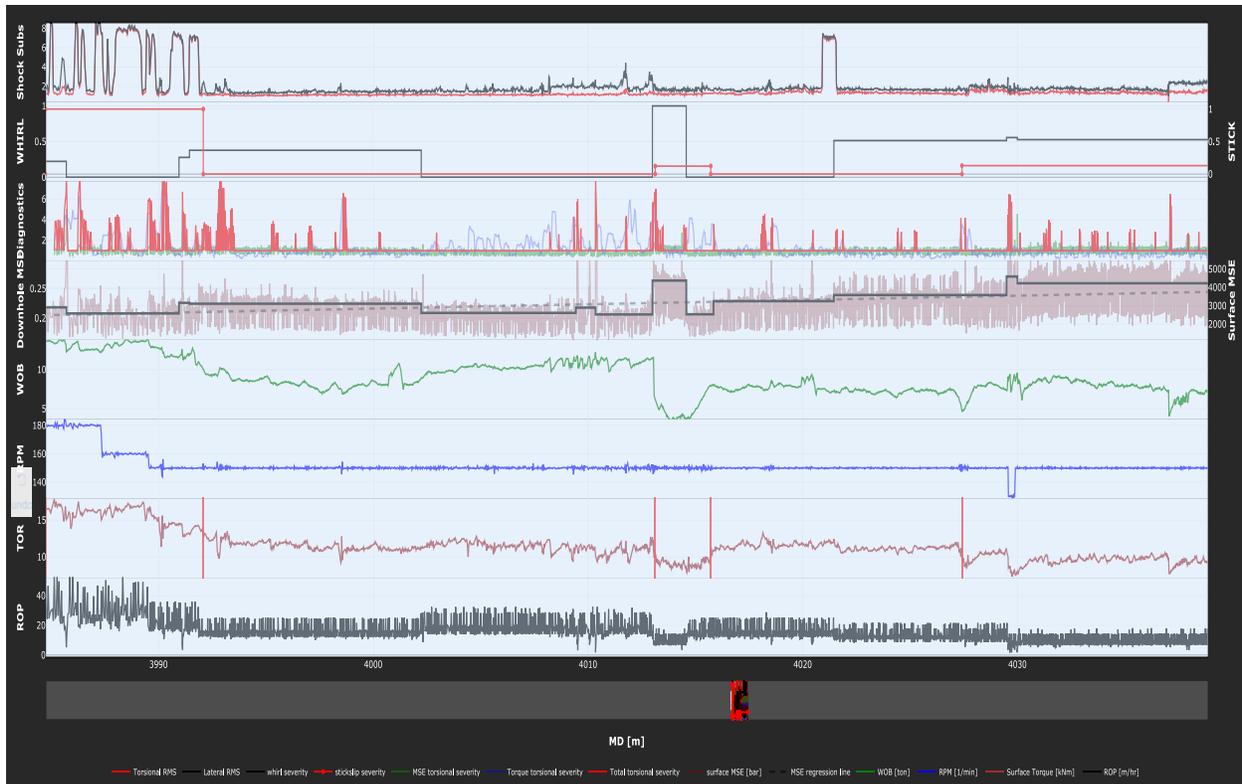


Figure 4.19: Well C4: Apollo Limestone diagnostics (unitless penalty parameters: MSE penalty = 0.5, Torque penalty = 5, minimum baseline length = 20)

Currently, the bit is drilling efficiently at approximately 180 RPM and 14 tonnes WOB in the Apollo Limestone. Due to relatively high shock levels recorded in the BHA, the operator decides to decrease WOB. Lateral and torsional vibrational levels in the shock sub recede when reducing the WOB, however the MSE baseline increases. Note that due to the increasing severity in the interval, the changes in MSE may seem insignificant, however they do amount for many 100 bars. Reducing WOB from approximately 14 tonnes to 7 tonnes has now resulted in a substantial loss of DOC, decreasing the bit's resistance to lateral movement in the hard Apollo Limestone, initiating bit whirl, which is depicted by the whirl severity trace (black) in the second top plot. Throughout the interval, the operator is steadily increasing WOB, observing a quiet response in vibrational levels, and decreasing trend in MSE.

At approximately 4010m MD, vibrational shock sub levels start to increase slightly (which is also depicted by the torsional severity score). The operator responds by decreasing WOB. WOB is drastically reduced from approximately 11 tonnes to 3.5 tonnes. This action initiates the largest response in whirl severity in the interval, which is clearly depicted by the black whirl severity

trace, which is responding to the large increase in MSE and reduction in WOB. Upon increasing WOB, MSE alleviates to some degree restricting higher levels of bit whirl. WOB is remained approximately constant with a slight decreasing trend deeper into the interval, which is clearly depicted by a larger response in MSE. This may be due to the confined compressive rock strength in the Apollo Limestone is increasing deeper throughout the interval. Conclusively, a loss of DOC will occur with an increasing confined compressive rock strength without a proportionately increase in WOB. The relative larger increase in whirl severity compared to stickslip severity indicates that the rest of the interval is affected by mostly bit whirl.

4.8.2 Case: Well C3 diagnostics study

Having identified that the Lower Armstrong Shale is generally afflicted by stickslip (transitioning into bit whirl at low WOB settings), and the Apollo Limestone being afflicted by large responses in bit whirl, one can move onto a different reference well, well C3, which is on the same well padding as well C4.

Naturally, one could expect the same bit dysfunctions limiting well C4 to also limit well C3, given that the well path, formation lithology, and BHA configuration is nearly identical. The difference, in fact, is that well C3 had much more problems than well C4. Well C3 had in total 6 trips in the 8 ½ ‘’ section due to several tool failures and steerability issues, while C4 had 3 such trips. An immediate difference between the two wells is that well C3 had a slick assembly, compared to C4 which had a packed assembly configuration, perhaps the operator got more cautious from all those tool failures in C3.

Below, in Figure 4.20, is the diagnostics overview of the Lower Armstrong Shale in well C3. Right off the bat one can see a lot of things happening. An initial impression is that there seems to be many distinct bit whirl events (depicted by the black trace in the second top row). The increased responses in MSE neatly correspond with a reduction in WOB. In addition, the torsional severity score seems to tramline the most severe peaks from the recorded BHA shock sub measurements. Conclusively, the overall first impression is that well C3 also seems to be

limited by stickslip responses (in torsional severity) which transitions into bit whirl at lower WOB settings, in the Lower Armstrong Shale.

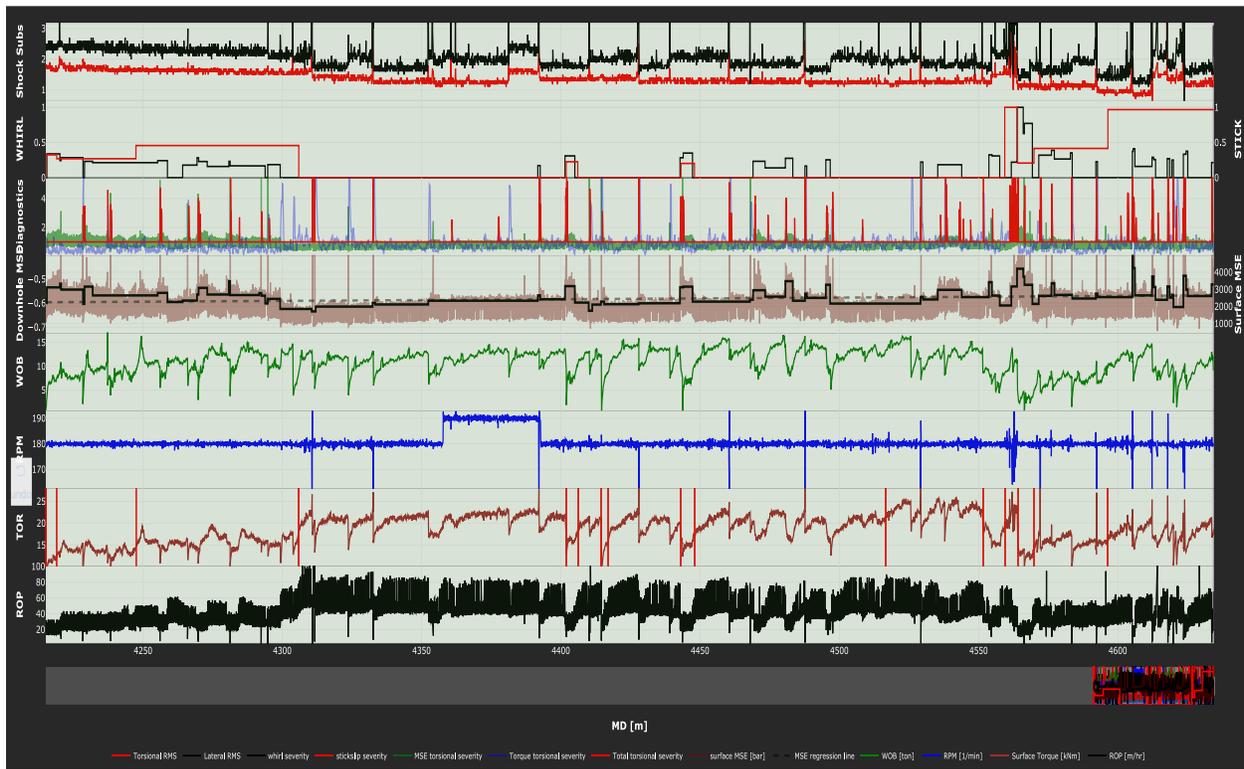


Figure 4.20: Well C3: 8 1/2" section diagnostics (unitless penalty parameters: MSE penalty = 0.3, Torque penalty = 3, minimum baseline length = 15)

Let's take a close look at the events occurring at approximately 4550m MD. Below in Figure 4.21 is a snippet of the zoomed-in interval (4550m – 4635m MD) from the diagnostics plot above. Again, one may immediately see several large responses in stickslip from the stickslip and torsional severity trace which transition into bit whirl events by excessively reducing WOB. Nearly every significant spike in vibrational shock sub measurements are tramlined by spikes in the torsional severity algorithm. Following most of these spikes are onset of bit whirl events from lowering WOB. One may notice quite significant torsional and lateral vibrational levels at approximately 4560m MD, which correspond with an increasing WOB and large oscillations in surface torque and SoftSpeed RPM tuning. Following this response, MSE starts to increase and the stickslip severity algorithm responds with a high severity. Immediately after reducing WOB, BHA vibrational levels calm down but MSE continues at an increased level due to an onset of bit

whirl. This stickslip response and transition into bit whirl is surprisingly similar to that one of approximately 5250m MD in well C4, in Figure 4.18 above.

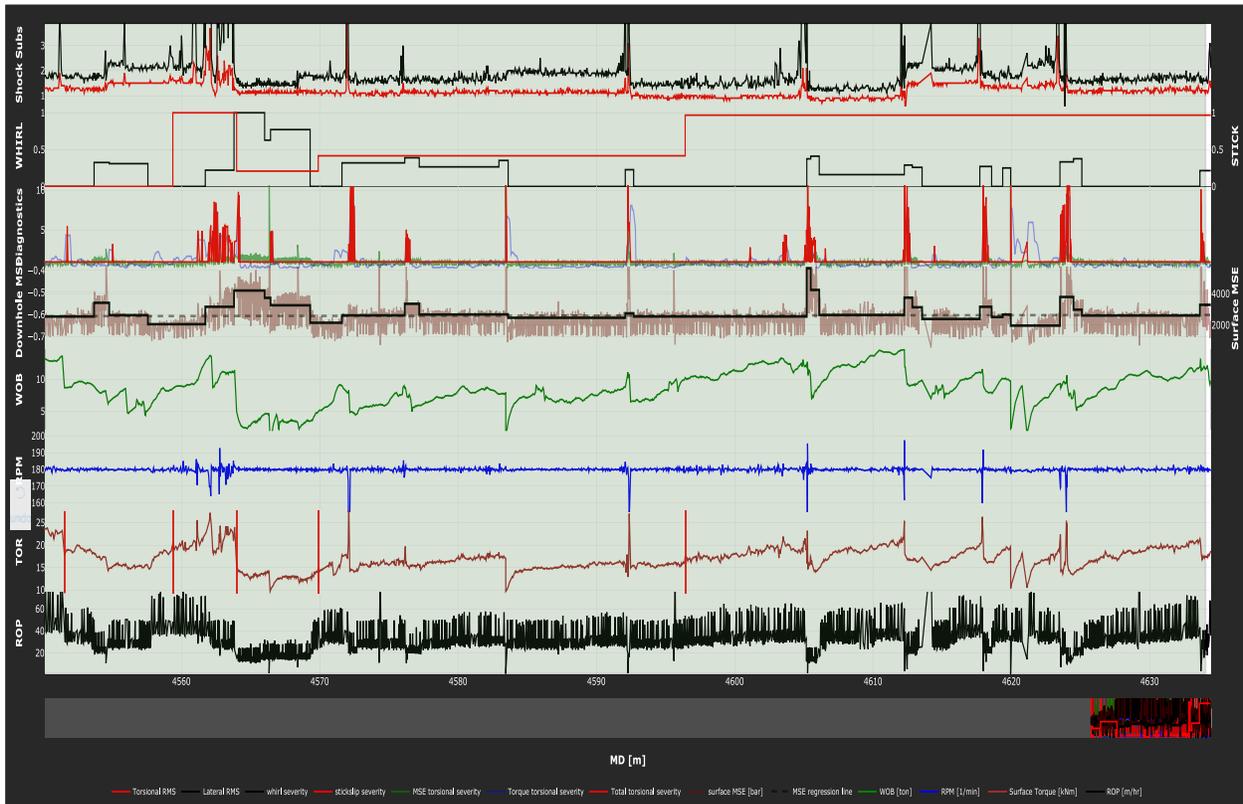


Figure 4.21: Well C3: 8 1/2'' section diagnostics: Zoomed-in (unitless penalty parameters: MSE penalty = 0.3, Torque penalty = 3, minimum baseline length = 15)

Conclusively, it is apparent that both well C3 and C4 are limited by severe bit whirl in the Apollo Limestone, and laminar interfacial severity in the Lower Armstrong Shale (large stickslip responses occurring infrequently). This diagnostics conclusion is also in agreement with the service company incident report of motor failures in well C3. Where the service company concluded that high vibrational levels were responsible for some of the tool failures. Sadly, bit pictures were not included in the well reports which could have further increased the confidence level in the diagnostics, through bit forensics. This is due to each bit dysfunction having distinct failure mechanisms observed on the bit, and emphasizes the value in properly documenting bit pictures after every trip.

4.9 Stringer detection & bit forensics

There will always be some degree of non-uniqueness in accurately determining bit dysfunction. To further increase confidence in diagnostics, bit pictures will always present compelling evidence of what is actually happening downhole. The below example is taken from the 12 ¼” section in well C3, and includes pictures of the bit after reaching TD. Even though the bit reached TD in one run, the bit was considered unable to be rerun due to severe bit damage.

Table 10: Case data for stringer detection

Case data		
Well analyzed		Well C3
Section		12 ¼”
BHA configuration		Bent motor
Bit post-run		Not reusable
Motor RPM factor	[rev/gal]	0.28
Motor diff. rating	[psi]	900

Below are the pictures taken by the rig crew immediately after tripping out the 12 ¼” BHA. The first thing to notice about these pictures other than the damage itself, is the location of it. Note that the inner cutters seem to be in good shape, while the shoulder cutters have rendered quite some damage. Especially the blade on the bottom picture which has suffered a full shear, i.e. a hinge failure, ref. Figure 4.22. Now that we know that the bit is damaged, the next step is to identify the origin.



Figure 4.22: Bit pictures with clear IFS damage.

Initiating the drillWiz diagnostics function with the above well parameters yields Figure 4.23 to the right when viewing the 12 ¼” section.

Notice the large area filled in red. This interval indicates permanent damage to the bit, which has resulted in decreased performance. This interval is flagged by the Stringer detection & damage assessment scheme, colored in red to signal a damaged bit (red bit). At this depth, the differential pressure and bit torque exceeds the motor rating, indicating a stalling event and potential damaging

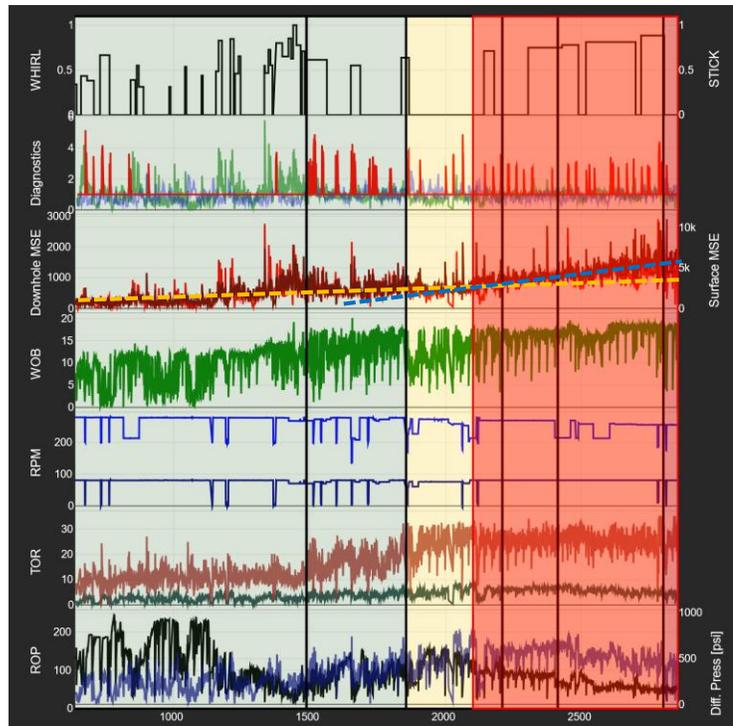


Figure 4.23: Damaging stringer detected, indicated by red fill.

stringer. Immediately following the event, the trend in MSE shifted indicating damage to the bit. The following future MSE baselines never reached the previous efficient baseline value (+ a threshold of 20%).

To illustrate the shift in MSE, two lines have been drawn on top of the curve. The orange line represents the MSE trend before the stalling event, and the blue line represents the trend after the stalling event. A shift like that should not occur when drilling in the same lithology with the same parameters, and are consequently clear indications of bit damage.

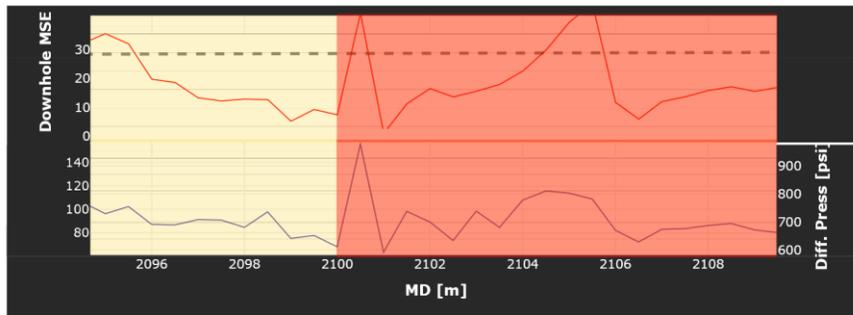


Figure 4.24: Differential pressure spiking along with MSE, i.e. typical stringer response.

If you zoom in on the event, notice the differential pressure exceeding the motor rating of 900 psi, and the increase in downhole MSE, ref. Figure 4.24. Although the MSE immediately goes down following the stalling event, it does not go down to the previous minimum baseline value (+20% threshold), nor does it until the bit was pulled. Now that we know the location of the damaging event, and have documented pictures of the



Figure 4.25: High load on bit shoulder in transition between laminated layers when exiting stringer.

bit damage, the only thing left is to identify in detail how this damage occurred. Considering the theory presented by Dupriest et al. in (Berge-Skillingstad & Anderssen, 2018), the origin of the damage must be from interfacial severity, i.e. a stringer. As the damage occurred on the shoulder of the bit, and not on the nose cutters. This indicates a high point load at the shoulders, and not at the nose when the damage occurred. Hence the damage most likely occurred when exiting a hard stringer, ref. Figure 4.22.

5 Discussion

5.1 Uncertainties and the impact of this

5.1.1 Corruption of MSE data through parameter uncertainties

Although mechanical specific energy poses a physically correct view of what happens downhole, one must keep in mind that several different parameters, all with different individual errors, are used to calculate MSE. Being aware of the dependency between different drilling parameters, and how they are measured or obtained is imperative in order to do proper MSE analysis.

For instance, consider a long lateral well where drill pipe buckling may get less severe when reducing WOB and may result in a considerable decrease in drill string friction and surface torque. This may have a significant impact on surface MSE. As a result, surface MSE may decrease considerably when reducing WOB by less drill string friction from buckling. This may potentially be erroneously interpreted as an increase in drilling efficiency. The significant weakness in not having downhole bit MSE available, is that there is simply not always a certain way of knowing whether the decrease in WOB in the above scenario was a good choice or not. It is hard to know if the decrease in surface MSE was due to less drill string friction from alleviating buckling, or by actually increasing drilling efficiency from alleviating a bit dysfunction. What may well as likely be happening, is an increase in bit whirl and decrease in drill string friction, or a decrease in bit dysfunction with constant drill string friction. There is simply no way of decoupling what is happening at the bit or in the drill string. Hence, emphasis is put on the importance of including downhole torque measurements for downhole bit MSE. Another important aspect of including downhole bit MSE, is the illumination of weight transfer issues. By actually decoupling events occurring in the bit and the drill string, the operator may with increasing confidence observe if weight transfer issues are limiting performance from observing any large discrepancy between surface and downhole MSE responses.

Bit torque from differential pressure across mud motors

As a part of the downhole motor MSE equation (eq. (3)), differential pressure across the motor is multiplied with a motor differential factor to estimate the torque at the bit. Differential pressure is calculated by subtracting the off-bottom standpipe pressure, measured on each connection, from the on-bottom standpipe pressure which is measured while drilling. As the differential pressure is dependent on standpipe pressure, other factors than changing resistance, i.e. rock strength and depth of cut at the bit, can also affect the calculated bit torque. This means drilling with losses, pack-offs or other factors may cause the differential pressure to react and corrupt the bit torque value.

In order to compare optimal parameters in lithologies from multiple wells, BHA configuration plays a vital part. For instance, running with a RSS and mud motor will involve about 60+ ft of extra stabilizers and BHA below the motor. This additional BHA will add additional drill string friction compared to a bent motor BHA configuration, and consequently increase the downhole MSE. The amount of additional torque the mud motor must deliver is primarily dependent on the degree of whirl below the motor to the bit, which may be considerable in some cases. An ideal case is of course a downhole sub with the ability to measure downhole torque directly, however being an additional expenditure.

RSS with mud motors are a rare sight in unconventional wells due to the extra expenditure. This means that bit torque and downhole bit MSE is unavailable in intervals run with only a RSS and no motor, which is apparent for many of the wells in this thesis.

Another source of error in calculating bit torque from differential pressure is weight transfer issues. Recall that the bit torque required to turn the bit should only depend on DOC and compressive rock strength. Thus, a weight transfer issue may be identified any time the driller slacks of weight (increases WOB) and the differential pressure does not respond. The additional WOB is not getting to the bit. If WOB reaches the bit, DOC goes up and differential pressure must increase. Differential pressure is an important parameter to monitor to see if weight is transferring properly. Any weight transfer issue must be a geometric interference problem as the surface friction factor does not change as we change WOB. The dominant root causes for weight transfer issues are typically borehole spiral patterns, which is a consequence of severe bit whirl.

Bit RPM

The most common way to obtain bit RPM is quite simple. Although bit RPM represents a quite accurate overview of what the mean value of rotation is, it does come with significant limitations when attempting to analyze transient events. When drilling with a motor, the bit RPM is calculated by adding the additional motor RPM to the rotary RPM. The rotary RPM is measured with decent accuracy, however, motor RPM accuracy is another challenge. Through a motor RPM factor, given as the rotation obtained when rotating with no resistance per gallon and the total pump output volume per minute, one can calculate the motor RPM.

Imagine a stickslip scenario where the bit RPM fluctuates. Given the length of the drill string, the rotary RPM will be affected based on the torsional wave propagating to the surface, and will not provide any instantaneous accuracy at all. The apparent motor RPM will be constant as long as the pump rate is constant. This does not mean it is impossible to obtain a somewhat accurate bit RPM. One example of this can be read in (Ertas, Bailey, Wang, & Pastusek, 2013), who presents a method to use the peak-to-peak values of surface torque, a pre-drill estimate of wellbore trajectory, and rotary RPM to obtain a measure of bit RPM with 10-15% accuracy.

Weight on bit

WOB is a crucial parameter in both diagnostics and the surface road map part of the application, and one may easily forget that this parameter involves several uncertainties. Fundamentally, hook load is measured at the deadline going from the crown block sheave to the anchor. This measurement in itself involves uncertainties, but the big uncertainties comes from estimating the WOB from the hook load. To do this, the hook load on-bottom is subtracted from the hook load off-bottom, which is an initialized value that must be reset at certain times throughout the drilling process. In a short vertical section, the weight applied at the bit may be quite accurate, however when evaluating whether the reduced hook load is applied to the bit in a long lateral is another question, due to a long list of possible weight transfer problems that may occur.

Rate of penetration

ROP measurements in this thesis are given as instantaneous on-bottom rate of penetration estimates and is measured on surface by pipe velocity going into the hole. Measuring ROP on surface compared to downhole measurements will not be very different averaged up on a big-picture basis, however, when analyzing transient events or oscillations, i.e. from shortening and lengthening of drill pipe from torsional oscillations, it may prove different.

Surface road map discussion (downhole vs. surface MSE measurements)

Creating a lithology specific road-map from reference well studies may be beneficial in well planning. Such a road-map may identify desired parameter ranges in pre-drill stages and identify which parameters to avoid (e.g. resonant RPM settings). Downhole bit MSE will provide a unique reference point for identifying these optimal parameters as confined compressive rock strength can be assumed constant within the lithology in a horizontal plane. In addition, downhole bit MSE is unaffected by drill string friction, such that in an undulating well path downhole bit MSE should be approximately comparable and independent of MD length.

The exception is of course if downhole bit MSE is unavailable, and the only available torque measurements are at surface (containing drill string friction). This is apparent in several inspected wells in this thesis, as downhole torque readings are unavailable in wells without mud motors or downhole torque subs. Surface MSE may yet be utilized, in certain situations and if handled correctly. Normalizing surface MSE with respect to depth may be a suitable option. This can be done by depth correcting surface MSE by the established linear regression trend throughout the well. Assuming that the slope of the linear MSE regression trend accounts for the additional drill string friction from an elongated well path, this correction may be tuned for surface MSE in lithology interceptions and at different measured depths along the drill string. Another situation is when well path, formation depth, confined compressive rock strength, and BHA configuration are nearly identical, such as in several wells on the same well padding in this thesis. The exact value of surface MSE will still deviate from well to well by discrepancies between them, however, remember that the exact value of MSE is of less significance as long as it is in the same ball park. In surface regression we are only interested in the established trend,

and not so much in the exact MSE value. This is due to fluctuations in MSE from discrepancies in drill string friction among near “similar” wells will pale in comparison to MSE fluctuations from onset of drilling inefficiency.

Depth-Of-Cut Control will fool your MSE response

As discussed in detail in (Berge-Skillingstad & Anderssen, 2018), DOCC is a useful tool to constrain bit whirl by allowing a high WOB, while at the same time reducing bit aggressiveness when cutting excessively deep in the rock to avoid large torsional oscillations and possible stick-slip events. Most PDC bits today have some degree of DOCC inserts, however knowing when the rubbing inserts engage with the rock can easily be overlooked. Consider a case where the WOB is not sufficient to engage the DOCC. Assuming torsional oscillations are not severe at the applied DOC, increasing WOB will eventually trigger the DOCC to be activated. As the DOCC is activated, the bit becomes less aggressive, resulting in a change in the proportionality between WOB and ROP. Consequently, MSE will increase and one could faultily interpret this change as transitioning into some kind of bit dysfunction. It is important to keep this effect in mind before hastily making decisions about drilling inefficiency from parameter responses.

Time-based vs. depth-based data

A fundamental assumption in this thesis is that the extracted depth-based drilling data is representable. Depth-based data is assigned as a single data point for each depth, regardless of the time used on that depth. Consider at a certain point, the driller decides to pull up and ream, or if a hard streak is encountered and different drilling parameters are used on approximately the same depth. For these and many more different scenarios, non-uniqueness in data sets may be present, which provides an additional dimension of uncertainty to the analysis.

Even though depth-based data and time-based data has its discrepancies, depth-based data is one-to-one with time-based “on-bottom” data. As long as time-based data is on bottom progressive drilling, depth-based is exact. In other words, algorithms working on depth-based data will conclusively work for on-bottom time-based data. Choosing maximum depth-based data values

will extract the most severe time-based transient events. This means that maximum depth-based data will be equivalent to time-based data when analyzing severe events.

Note that depth-based on bottom ROP will decrease on every connection due to connection procedures such as reciprocating the string, rotating on/off bottom, surveys, breaking-in the bit etc. Accordingly, depth-based ROP and WOB will decline on every connection, resulting in a considerable MSE spike. This signature may look like a stringer, however it is important to look at the frequency (is it occurring every stand length etc.) and differential pressure spikes (if available) to separate MSE spikes from connection procedures or actual transient events.

5.1.2 Uncertainties when determining order of surface regression

A fundamental part of the automatic parameter selection process is the method chosen to perform the surface regression. The results presented in this thesis are based on the 2nd-degree polynomial function, eq. (22). A set of polynomial parameters are fitted using the representative data points of WOB, RPM and MSE to obtain a surface equation to forecast the MSE on different parameters outside the current operating area.

2nd order polynomial

$$MSE_i = a_0 + a_1WOB_i + a_2RPM_i + a_3WOB_i \times WOB_i + a_4WOB_i \times RPM_i + a_5RPM_i \times RPM_i \quad (22)$$

Where,

(WOB_i, RPM_i) Input data points

a_0, \dots, a_k Polynomial coefficients

n Total number of data points

k Degree of polynomial

MSE_i Surface regression output

Although the 2nd order regression provides a somewhat flexible picture of the data trends, one can argue that other mathematical approaches would be a better solution. Consider a case where multiple extrema exist for a given set of data points. If applying a 2nd degree solution, a surface

regression will smooth over some of the extrema and may risk losing information about possible resonant frequencies in operating parameters. This might have a crucial effect in the parameter determination process. A possible solution can be found by Taylor expanding the 2nd order polynomial to higher orders of regression, ref. equation (23).

Expanding to Nth order polynomial with n data points:

$$MSE_i = a_0 + a_1WOB_i + \dots + a_{n-2}(WOB_{n-2})^2 \times (RPM_i)^{N-2} + a_{n-1}WOB_i \times (RPM_i)^{N-1} + a_n(RPM_i)^N \quad (23)$$

The below pictures, Figure 5.1, visualizes how higher degree polynomials may provide different solutions in the surface road-map. In this example, the Orion Sand has been used as the lithology for analysis with the changepoint algorithm values for WOB, RPM and MSE of 0.01, 0.001 and 0.01, respectively. The scatter plot in the upper left corner shows the filtered vales before applying surface regression. In the next picture to the right, one can observe how the 2nd degree polynomial solution forecasts the surface in this sand, with the optimal solution being at about 160-170 RPM and 9-12 tonnes WOB.

In the next picture on the bottom left, a solution with a 3rd degree polynomial has been utilized to approximate the surface road-map, yielding a local minimum of the operating area similar to the 2nd degree solution. However, now an even lower minimum exists at a lower WOB (stretching towards 2 tonnes) and 210 RPM due to the surface trying to fit a more accurate trend. After a quick look at this solution, one can easily say that this minimum is not valid due to it would most likely provide high levels of bit whirl from an insufficient DOC. The last solution presented is the 4th degree polynomial solution. This solution provides even more extreme changes in the surface and several local minima. One can see the solution yielded from the 2nd degree is being approached, however, the problem comes when extending the solution further away from the data points.

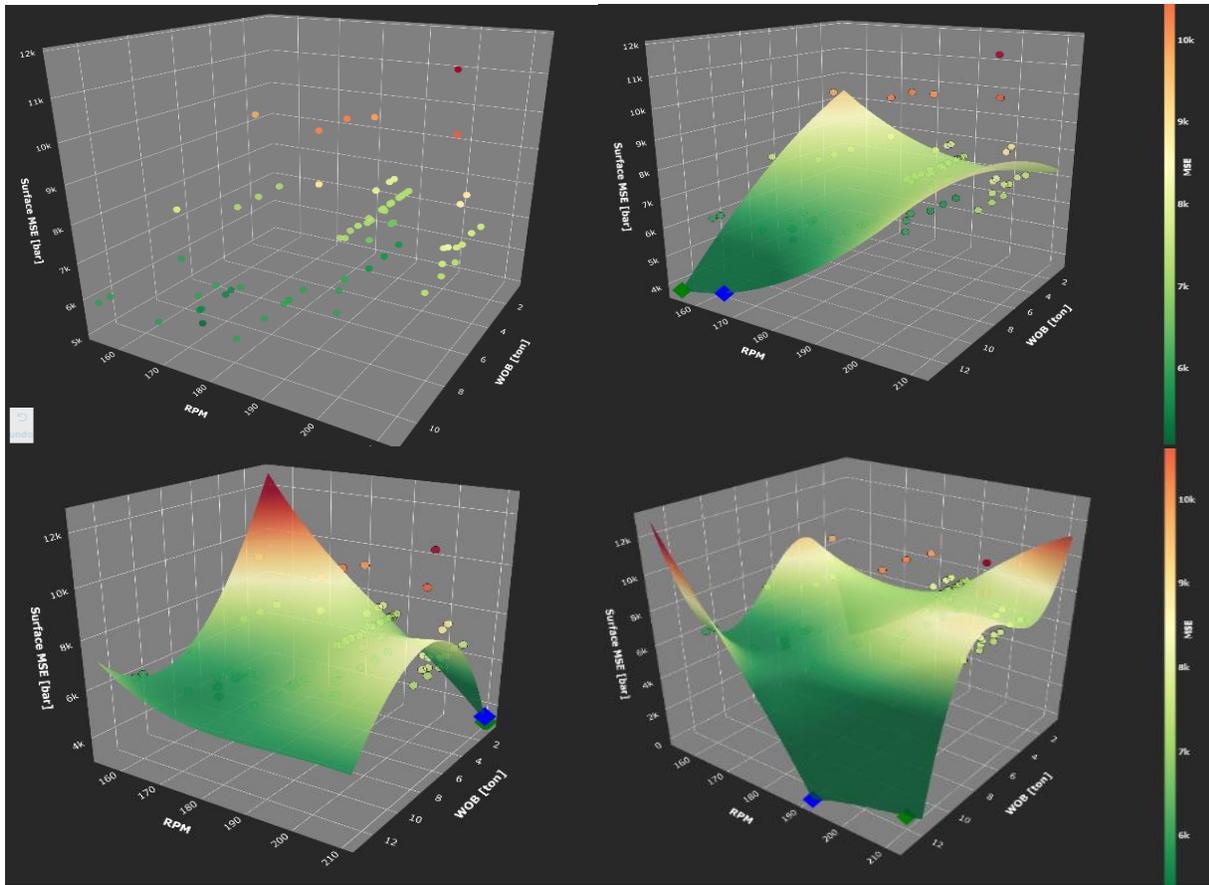


Figure 5.1: Different orders of surface regression yield different solutions.

Due to its higher polynomial degree, a 4th degree solution can prove to be quite accurate close to existing data points due to its additional degrees of freedom, but when forecasting a solution in areas with undersaturated data points, extreme changes take place as no existing data constrains the solution. This can be seen in the lower right corner of Figure 5.1, where the solution drops towards a minimum which cannot be considered a physically correct solution (severely low MSE, approaching zero).

To sum up the results from increasing degrees of freedom, i.e. higher degree polynomials, one can see that the surface provides a more accurate fit close to existing data points (data saturated areas). But when extending the surface outside the domain of the input data, extreme and physically incorrect solutions are obtained. This illustrates the principle that a weighted solution must be obtained to be able to balance accuracy with a physically correct trend approximation.

Considering the current level of parameter exploration, the operating area is not explored sufficiently. Conclusively, certain lithologies are “underexplored” for a proper regression analysis. This forces a trade between accuracy for better trend forecasting, and is why a 2nd degree solution is preferred.

Another approach may be to include a logarithm-based solution, to avoid extreme changes in MSE forecasts from changes in WOB and RPM. Adding a higher degree logarithmic term is something that was considered implemented, but was left out for further development as a higher fundamental understanding of mathematics and statistics should be in place before attempting something like this.

5.1.3 Confidence-level and uncertainty in dysfunction analysis

Having been introduced to several comprehensive cases of automated bit dysfunction diagnostics, its applications and desired utilities, it is important to discuss the current limitations in confidence level. The primary purpose of the drillWiz diagnostics tool is to aid the drilling engineer in manual interpretation of reference wells. The level of confidence in the current state of affairs is only given in a relative confidence level. In other words, dysfunctions are given a relative severity score compared to the maximum severity experienced in the well or interval under investigation.

Recall that dysfunction diagnostics will not flag an event unless the investigated MSE baseline is overlying the regression trend line, which indirectly and conservatively dictates an inefficient drilling operation. This means that even if dysfunctions are compared in relative severity, the flagged events are still characterized as inefficient drilling operations. By such, the engineer can allocate his or her scarce resources to only focus on the most pressing events or predominant dysfunctions limiting performance in the well.

Even though dysfunctions are not flagged unless the MSE score is severe, there are still cases where the diagnostics algorithm fares poorly.

By definition, the whirl diagnostics algorithm will respond any time the MSE baseline scores higher than the general regression trend. This means that whirl diagnostics will always allocate a

score to intervals with “inefficient” MSE baselines present. As a result, the whirl diagnostics algorithm may indicate bit whirl, even if there is no bit whirl at all in the interval. If there is no bit whirl in the interval, the whirl diagnostics algorithm will still assign a maximum score of 1 to the most severe MSE response. Conclusively, the scoring system will only work effectively if bit whirl is present in the interval. This is because the combination of a high MSE and low WOB will always generate the highest relative score. In intervals with bit whirl, other severe MSE responses will pale in comparison to the score of the actual bit whirl events. Note that bit whirl may also be initiated from a substantial loss of DOC upon entering a harder formation without proportionately increasing WOB, further emphasizing the need of lithology specific diagnostics to observe changes from lithology transitions.

The same discussion above is also valid for the stickslip diagnostics algorithm. Given a certain amount of changepoints in torque variance, a relative maximum score of 1 will always be set between the most severe changepoints. This, in combination with a severe response in MSE, will always be flagged by the stickslip changepoint algorithm, even if there is no stickslip present in the interval.

This weakness is apparent in several wells, and is illustrated in the figure below, ref. Figure 5.2. Figure 5.3 depicts a snippet of diagnostics from the Gemini Sands in Well C4. The Gemini Sands are severely limited by bit balling, which is illustrated in the figure below. There seems to be a lithology transition to an interbedded shale at approximately 2200m MD. ROP and torque drops drastically at this depth, initiating a severe MSE response at a relatively high WOB. ROP and torque recovers when lowering WOB, reducing the accumulated ribbon height and recovering some DOC, which is a characteristic response in bit balling diagnostics. However, in the figure below, both the stickslip and whirl diagnostic algorithms reach their maximum severity at what appears to be bit balling. This is also true for the torsional severity diagnostics. This is one of the primary reasons why the diagnostics algorithms should always support a manual interpretation, as the confidence level in dysfunction uniqueness is still relatively low.

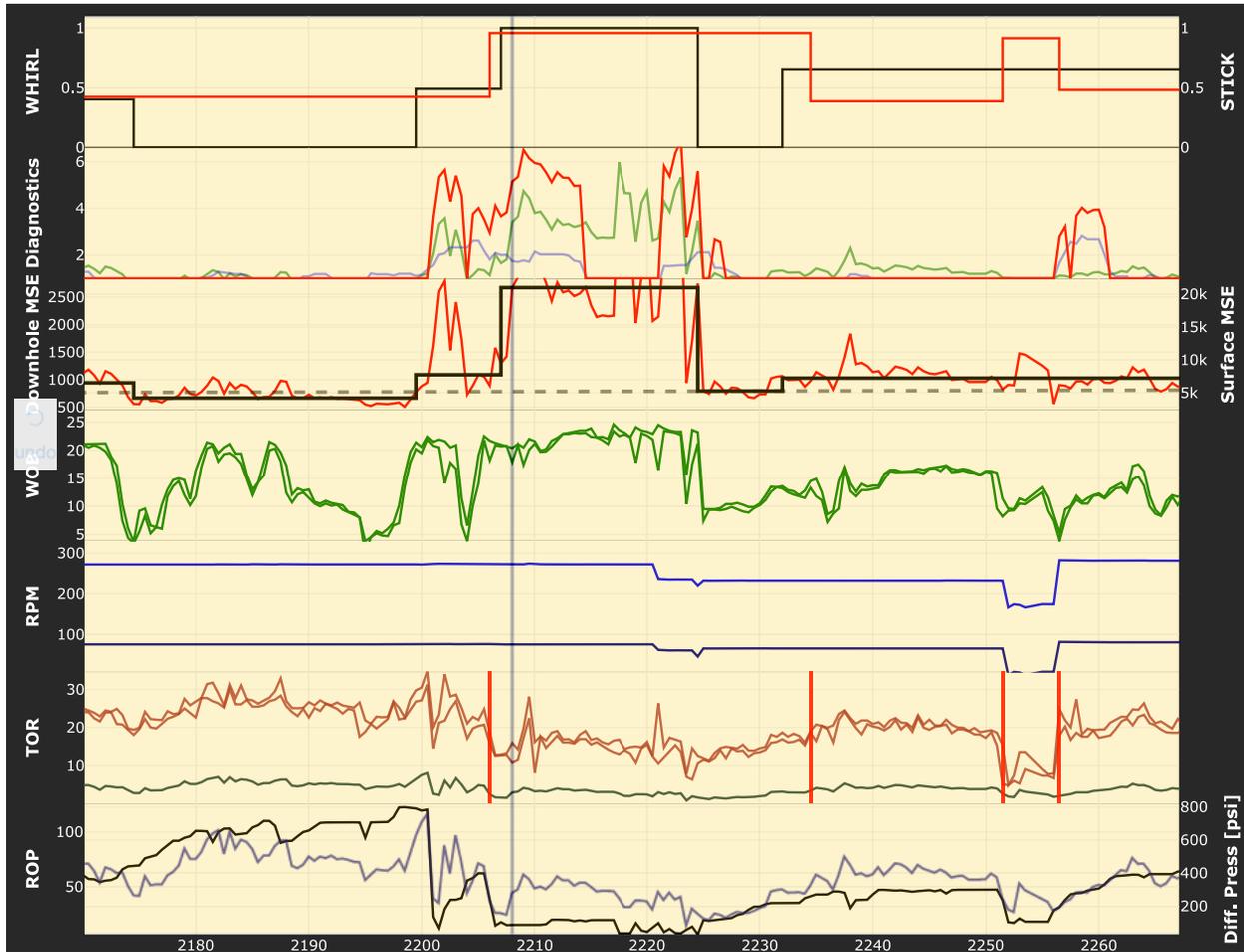


Figure 5.2: *Weakness in Diagnostics: Bit balling event in the Gemini Sands, well C4*

As previously discussed, MSE is generally considered a weak diagnostics tool for stickslip. This is due to bit speed oscillations not severely affecting performance, and more superior tools for accurately determining bit speed oscillations exist. This may be directly measuring bit RPM, or by calculating expected bit RPM from surface torque measurements and an expected torque response along the drill string from a T&D model.

Furthermore, as discussed in (Berge-Skillingstad & Anderssen, 2018), it is not clear whether an increase in MSE is only due to a full-stick event or due to a coupled stickslip and whirl event. This is especially true for laminar interfacial severity, which often occurs in interbedded formations with high contrasts in compressive rock strength (i.e. cemented calcite stringers), faulted formations, or other discontinuities. A full-stick event from entering a hard stringer is

often accompanied by a high impact whirl event from a temporary sudden loss of DOC and lateral bit acceleration. Conclusively, there are many times a complex mix of torsional and lateral vibrations occur together, such as coupled stickslip and whirl.

In several of the figures presented in the *Results* chapter, there seems to be an overlapping of flagged stickslip and whirl events. For instance, Figure 4.20 of Well C3 depicts an almost tramlining between stickslip and whirl severity in certain intervals. This may be due to the combination of coupled stickslip and whirl events causing similar responses in both diagnostics.

All in all, there are strengths and weaknesses in comparing relative severity in diagnostics. The strength in the approach is effectively pointing out the most severe intervals of drilling inefficiency, aiding as a utility tool to suggest the potential dysfunctions. Creating a tool which programmatically helps the engineer to diagnose potential bit dysfunction may ultimately ease the workflow, and provide support in conclusions, when suggesting to redesign the next well to extend the current performance limiters. Having said that, the diagnostics tool should only serve as a support to manual interpretations. In the current state of affairs, confidence is not yet tested properly enough to achieve an acceptable confidence level in diagnostics. Further work in increasing confidence should be done by comparing and calibrating the diagnostic results with more bit pictures. Bit forensics serves as a strong support in diagnosing bit dysfunction, as every dysfunction creates unique bit damage. Unfortunately, the diagnostic results could not be compared to bit pictures in every run in this thesis, due to documenting bit pictures were not a standard operating procedure at the time of these wells.

5.2 Further work

The application presented in this thesis serves as a stand-alone software for post-drilling analysis of MWD data. Although the most obvious improvement for this application is to enhance the current level of confidence in the different analyses, other functions which have not been presented in *Results* have been in a concept-only phase. This includes bit balling diagnostics, which has a presented pseudo-code below. Additionally, a strategy for further validation of diagnostic confidence level with machine learning is proposed, and a suggested data parsing filter for further implementation and automation of data preparation. Finally, potential markets for commercialization and real-time implementation have been looked into with a generalized discussion and potential strategy outlined.

5.2.1 Strategy for validation of diagnostics

An overall strategy for increasing diagnostic confidence by machine learning can be developed. This could be done by establishing an expert committee to set up a training and test database. The database would include diagnostic solutions from manual interpretations and bit picture forensics conducted by the expert committee. The diagnostic algorithm may then be run with varying input sensitivities such as, e.g. different changepoint penalties, and compared against the training database. The goal of such a machine learning scheme would be to calibrate the diagnostics algorithm to develop a generic method or optimal set of generic input penalties which best match the training database. This could potentially fully automate the diagnostic algorithm by removing any dependency on manual input sensitivities.

Another outcome of machine learning would be to increase the level of uniqueness in distinguishing between different diagnostic results, a challenge described in page 104, *Confidence-level and uncertainty in diagnostic analysis*. By assigning a gratification score every time a specific diagnostic agrees with the expert committee, and negative gratification score every time it disagrees, the machine learning algorithm would then be able to calibrate the model to increase its uniqueness.

5.2.2 Automated data parsing to match all files and databases

Considering the time-consuming work load in data preparation, to make each well match the design of the application, great potential lies ahead to automate the process of data parsing. The current regime requires the input .csv files to be organized in a strict manner for all functionalities to work. The next step to improve data preparation would be to create a parsing filter in the application to parse all data automatically into the required structure, regardless of its file type, length, units, etc. Doing so could easily be done using external competency if investors would like to move the project into commercialization or further development.

5.2.3 Proposed bit balling diagnostics scheme for next level implementation

By incorporating lithology into the process, the diagnostics of whether a bit is balled up or not is far from complicated. Especially in shallow, young shales where bit balling is a common phenomenon. A typical bit balling response can be seen in Figure 5.2 where torque decreases and ROP collapses at a high WOB setting. MSE shows a clear response of drilling inefficiency and spikes accordingly. An important aspect of including bit balling diagnostics in the dysfunction analysis is to prevent other dysfunction tools to flag this incorrectly.

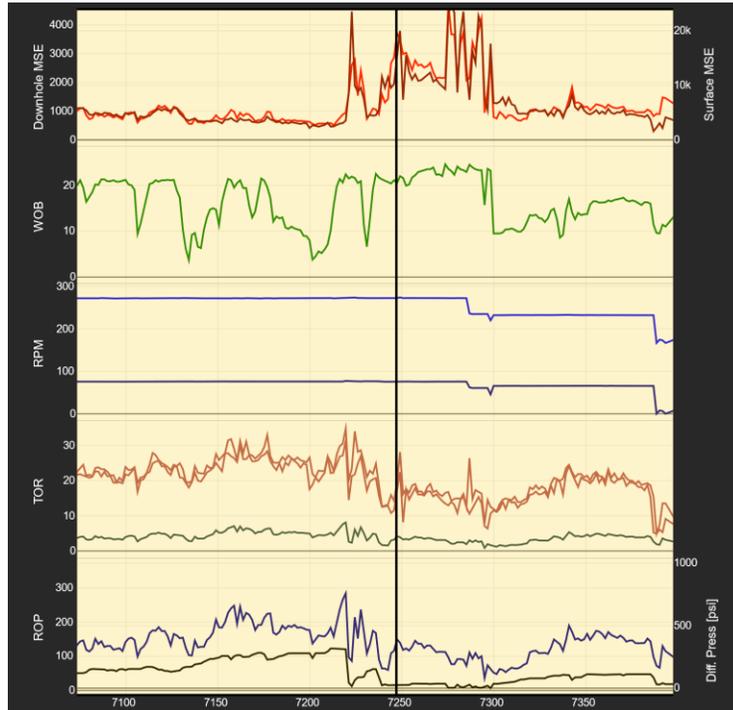


Figure 5.3: Bit balling response seen in what is most likely a shaly part of the Gemini Sand in the transition from the upper to middle zone. The interval is taken from the 12 1/4" section of well C4.

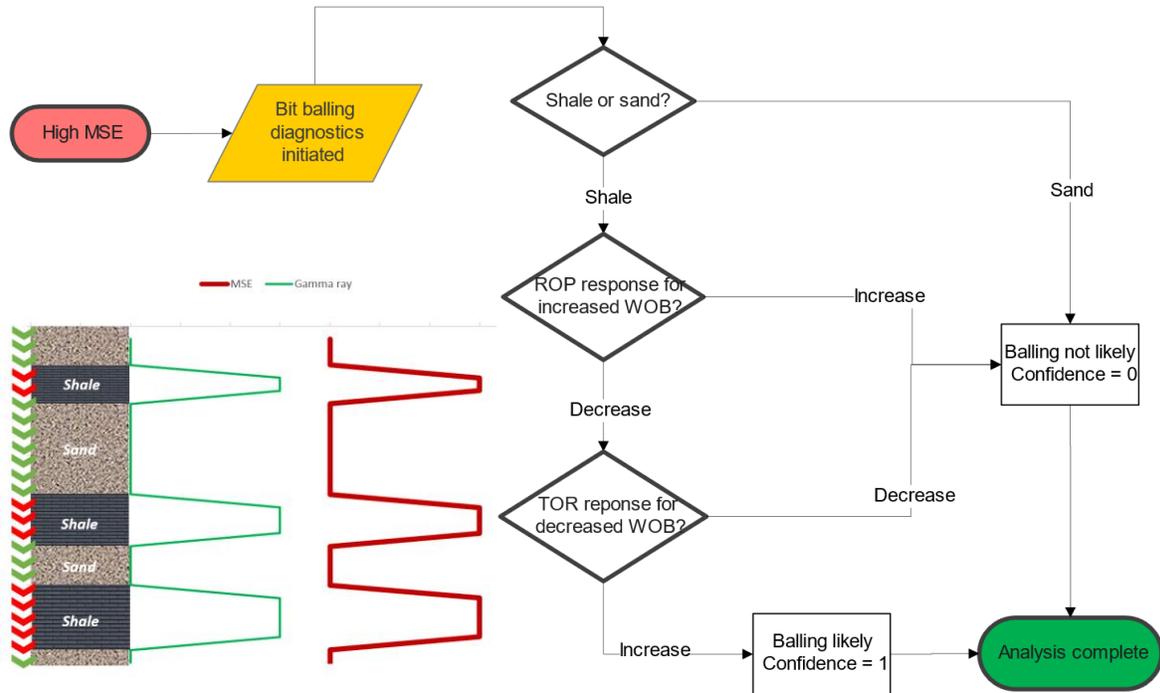


Figure 5.4: Proposed bit balling diagnostics procedure.

From the discussion surrounding bit balling presented in (Berge-Skillingstad & Anderssen, 2018), the above pseudo code for diagnostics has been derived, ref. Figure 5.4.

Programmatically, the procedure is simple enough to implement. However, to be able to exploit the lithology aspect fully, a more detailed lithology identification scheme must be in place. This can be done, for instance by using the gamma ray data to identify shale content. As bit balling is impacted by hydraulics, incorporating pump output and its effect on bit balling can also be an effective measure to assess the confidence level of the dysfunction.

5.2.4 Potential markets for commercialization: NOVOS application platform – The Reflexive Drilling System

In 2016, National Oilwell Varco introduced NOVOS, a software operating system to facilitate drilling systems automation through improved third-party integration. External users may be granted access to the software automation platform to upload and develop applications and smart algorithms for rig implementation. NOVOS applications provide the ability to integrate equipment and process controls on the rig with smart algorithms that may increase drilling performance through automating reflexive drilling tasks or taking corrective responses to changes in drilling efficiency in a consistent matter ("National Oilwell Varco," 2019).

Conventionally, third-party control of rig operations requires broad collaboration with rig owners and detailed knowledge of rig control systems and equipment. Third-party developed NOVOS applications are scalable across entire rig fleets with various equipment capabilities and NOV control systems, and do not require a complete control system overhaul. External users may now develop smart applications to specific procedures and scale it across entire rig fleets to provide consistency and standardization of operations through automation.

Introducing third-party involvement for developing applications in drilling systems automation processes is a huge step towards collectively developing a holistic system for drilling automation. The drilling industry is beginning to take a step into the digital ecosphere.

Digitalization is beginning to change how we operate with an increasing connectivity between humans and machines. Improving sensor technology, data acquisition tools, data processing power and high-speed telemetry is moving the industry towards a data driven architecture.

The drillWiz application is stand-alone as it is, however such an automated post-drill evaluation tool may easily be rewritten in C, and ultimately support the integration of a real-time surveillance tool, based on the same principles, and adapted for the NOVOS application platform. Potential hurdles, benefits, and the way forward for transitioning a post-drill diagnostics software into a real-time decision-making software is discussed to some degree in the below chapter of real-time implementation.

5.2.5 Real-time implementation

Having established the concept that MSE can effectively forecast parameter operating trends and autonomously identify bit dysfunction in post-drill analysis, the next step lies in real-time implementation. Recall that the underlying metric for quantifying drilling efficiency is MSE and a linear borehole is defined as a borehole with linear responses to changes in operating conditions (efficient drilling). Due to linear responses being predictable responses, MSE should be at the heart of every drilling automation operation.

Transitioning the offline changepoint detection tool in the drillWiz application to an online detection scheme (discussed on page 10, *detecting a change*), will essentially unlock many of the developed tools and diagnostics for real-time implementation.

Utilizing a high sensitivity online changepoint detection scheme effectively identifies any changes in drilling efficiency/MSE through continuous monitoring and assessment. As discussed in this thesis, a change in MSE may have a number of different root causes, e.g. lithology transitions or dysfunction onset from a change in operating parameters. Many of these scenarios are already tested and developed for post-drill analysis in the drillWiz application.

Further on, conventional performance management is preferably done by parameter exploration. WOB and/or RPM is manually and incrementally increased in steps in order to achieve maximum efficient performance. During every parameter “step-test”, MSE is recorded. If MSE starts to increase, the founder WOB or RPM set point has been identified, and maximum efficient operating setting has been identified. By continuously conducting such “step-tests” and automatically detecting founder changepoints, one may effectively increase the efficiency of the system in a physics-based approach. This is a reflexive system, which may be automated in an efficient (predictable) drilling operation, and a process which may eventually be transitioned from human decision-making to computer decision-making through smart changepoint algorithms and diagnostics.

To connect this feature with another innovative and continuously growing concept, is to introduce wired pipe to the rig. Doing so will allow for tremendous increase in data quality and frequency and will unfold another dimension of the real-time potential of the drillWiz concept.

Note that this is only at a conceptual stage, and a substantial increase in confidence and diagnostic uniqueness is required before attempting anything like this. However, the potential is certainly there and transitioning human decision-making from a manual experience driven culture to a holistic data driven culture will not only provide consistency and standardization of processes, but may also increase effectiveness of operations from instantaneous corrective decision making.

6 Conclusion

- Smart changepoint algorithms for reducing data and identifying trends prove very efficient when analyzing drilling performance.
- Data acquisition and preparation is time consuming as the necessary data are spread out in different formats and locations, e.g. data silos. An improved parsing filter for data of different origin and structure would solve this.
- MSE proves a great tool for identifying trends and optimal operating areas for offset well analysis. Results show clear indications of high and low MSE for certain operating parameters.
- As higher, but efficient operating parameters may be identified by mapping responses in MSE from surface road-maps, significant performance gains can be achieved from increased ROP.
- Current parameter selection schemes (parameter ranges in detailed operating procedures) are often empirical-based and do not explore the operating area sufficiently. For an improved road-map and optimized physics-based parameter selection scheme, a fundamental change in operating parameter exploration strategy must be done, e.g. more comprehensive step-tests.
- MSE proves a powerful tool to evaluate drilling performance when assessing downhole efficiency and damaging events. Bit pictures and incident reports add to the confidence level of these interpretations.
- Diagnostic results developed in this thesis correlate well with downhole memory sub data. Torsional severity diagnostics tramline the most severe downhole vibrations from memory subs, by only utilizing surface measurements. Commercial potential is seen as expensive downhole memory subs may be eventually be replaced by smart diagnostics.
- In addition, diagnostic results correlate well with manual interpretations and external information such as bit pictures and incident reports. This is especially true for intervals with confirmed dysfunction.

- Uniqueness in dysfunction diagnostics is challenging to achieve in some intervals. This is apparent in intervals with MSE responses, but no unique dysfunction characteristics. The diagnostics will overfit a dysfunction, even if no clear response is present. Further strategy for dysfunction validation is proposed.
- Evaluating dysfunction diagnostics post-drill will allow for effectively identifying performance limiters in reference well studies and initiate a cost justification process to redesign the next well and increase performance.
- Monitoring dysfunction diagnostics real-time during drilling can mitigate problems before they cause premature tool failure and bit damage. This will consequently increase tool longevity (avoid unnecessary trips) and allow for efficient drilling at optimal operating settings.
- The stringer detection & damage assessment tool shows good results when assessing whether the drill bit can achieve the same efficiency after a potentially damaging event.
- Difficulties in diagnostics occur when lithology data is insufficient or wrong, invoking the need for more detailed lithology data. Heterogeneous formations provide challenging interpretations.
- With improvements in data quality through more accurate lithology information and high frequency downhole measurements (wired pipe), comes the ability to do next-level diagnostics, e.g. bit balling, more accurate stickslip analysis and better stringer detection.

Link to application:

<https://drillwiz.herokuapp.com/>

Link to video walk-through:

<https://www.youtube.com/watch?v=PrTRPb06Ank>

Bibliography

- Aminikhanghahi, S., & Cook, D. J. (2017). A Survey of Methods for Time Series Change Point Detection. *Knowledge and information systems*, 51(2), 339-367. Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/28603327>
<https://www.ncbi.nlm.nih.gov/pmc/PMC5464762/>. doi:10.1007/s10115-016-0987-z
- Auger, I. E., & Lawrence, C. E. (1989). Algorithms for the optimal identification of segment neighborhoods. *Bulletin of Mathematical Biology*, 51(1), 39-54. Retrieved from <https://doi.org/10.1007/BF02458835>. doi:10.1007/BF02458835
- Basseville, M. (1981). Edge detection using sequential methods for change in level--Part II: Sequential detection of change in mean. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 29(1), 32-50. doi:10.1109/TASSP.1981.1163508
- Basseville, M., Espiau, B., & Gasnier, J. (1981). Edge detection using sequential methods for change in level--Part I: A sequential edge detection algorithm. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 29(1), 24-31. doi:10.1109/TASSP.1981.1163523
- Baum, M., & Sylvester, R. (1990). Statistics with confidence: Confidence intervals and statistical guidelines. Martin J. Gardner and Douglas G. Altman (eds), British medical journal, London, 1989. no. of pages: 140. price: £7.95 (\$19.00). *Statistics in Medicine*, 9(3), 344-345. Retrieved from <https://doi.org/10.1002/sim.4780090319>. doi:10.1002/sim.4780090319
- Berge-Skillingstad, M., & Anderssen, V. H. (2018). *MSE: Next Generation Digital Drilling Optimization*. Department of Geoscience and Petroleum. NTNU. Trondheim.
- Brun, A., Aerts, G., & Jerkø, M. (2015). How to achieve 50% reduction in offshore drilling costs. A *McKinsey and Company Oil & Gas report*.
- Draper, N. R., & Smith, H. (1981). *Applied regression analysis*. New York: Wiley.
- Dupriest, F. E., Elks, W. C., Ottesen, S., Pastusek, P. E., Zook, J. R., & Aphale, C. (2010). *Borehole Quality Design and Practices To Maximize Drill Rate Performance*. Paper presented at the SPE Annual Technical Conference and Exhibition, Florence, Italy. <https://doi.org/10.2118/134580-MS>
- Dupriest, F. E., & Koederitz, W. L. (2005). *Maximizing Drill Rates with Real-Time Surveillance of Mechanical Specific Energy*. Paper presented at the SPE/IADC Drilling Conference, Amsterdam, Netherlands. <https://doi.org/10.2118/92194-MS>
- Ertas, D., Bailey, J. R., Wang, L., & Pastusek, P. E. (2013). *Drillstring Mechanics Model for Surveillance, Root Cause Analysis, and Mitigation of Torsional and Axial Vibrations*. Paper presented at the SPE/IADC Drilling Conference, Amsterdam, The Netherlands. <https://doi.org/10.2118/163420-MS>
- Foundation, P. S. Retrieved from <https://www.python.org/community/logos/>
- Foundation, P. S. (2019). Python 3.7.3 documentation. Retrieved from <https://docs.python.org/3/>
- Fox, J., Jr. (2015). *Applied Regression Analysis and Generalized Linear Models*. Retrieved from <http://public.eblib.com/choice/publicfullrecord.aspx?p=1996011>.

- Hinkley, D. V. (1969). Inference about the Intersection in Two-Phase Regression. *Biometrika*, 56(3), 495-504. Retrieved from <http://www.jstor.org/stable/2334655>. doi:10.2307/2334655
- Jackson, B., Scargle, J., Barnes, D., Arabhi, S., Alt, A., Gioumousis, P., . . . Tao Tsai, T. (2005). *An algorithm for optimal partitioning of data on an interval* (Vol. 12).
- Killick, R., & Eckley, I. A. (2014). changepoint: An R Package for Change-point Analysis. *Journal of Statistical Software; Vol 1, Issue 3 (2014)*. Retrieved from <https://www.jstatsoft.org/v058/i03>
<http://dx.doi.org/10.18637/jss.v058.i03>.
- Killick, R., Fearnhead, P., & Eckley, I. A. (2012). *Optimal Detection of Change-points With a Linear Computational Cost* (Vol. 107).
- Kuhlman, D. (2011). *A Python Book: Beginning Python, Advanced Python, and Python Exercises*: PLATYPUS GLOBAL MEDIA.
- Kunche, P. (2016). *Metaheuristic Applications to Speech Enhancement*.
- Mau, M., & Edmundson, H. (2015). *Groundbreakers: The Story of Oilfield Technology and the People who Made it Happen*: Fast-Print Publishing.
- National Oilwell Varco. (2019). Retrieved from <https://www.nov.com/NOVOS.aspx>
- Parmer, C., Johnson, A., Parmer, J., & Sundquist, M. (2018). Retrieved from <https://plot.ly/products/dash/>
- Pearl, J. (1984). *Heuristics: intelligent search strategies for computer problem solving*: Addison-Wesley Longman Publishing Co., Inc.
- Rafatian, N., Miska, S. Z., Ledgerwood, L. W., Yu, M., & Ahmed, R. (2009). *Experimental Study of MSE of a Single PDC Cutter Under Simulated Pressurized Conditions*. Paper presented at the SPE/IADC Drilling Conference and Exhibition, Amsterdam, The Netherlands.
<https://doi.org/10.2118/119302-MS>
- Scott, A. J., & Knott, M. (1974). A Cluster Analysis Method for Grouping Means in the Analysis of Variance. *Biometrics*, 30(3), 507-512. Retrieved from <http://www.jstor.org/stable/2529204>. doi:10.2307/2529204
- Smith, A. F. M. (1975). A Bayesian approach to inference about a change-point in a sequence of random variables. *Biometrika*, 62(2), 407-416. Retrieved from <https://dx.doi.org/10.1093/biomet/62.2.407>. doi:10.1093/biomet/62.2.407
- Teale, R. (1965). The concept of specific energy in rock drilling. *International Journal of Rock Mechanics and Mining Sciences and Geomechanics Abstracts*, 2(1), 57-73. doi:10.1016/0148-9062(65)90022-7
- Theil, H. (1961). *Economic forecasts and policy*. Amsterdam: North-Holland Pub. Co.
- Truong, C. (2018, 3/1/2018). Retrieved from <http://ctruong.perso.math.cnrs.fr/ruptures-docs/build/html/general-info.html> - user-guide
- Weisstein, E. W. (2019, 26/3/2019). Least Squares Fitting--Polynomial. Retrieved from <http://mathworld.wolfram.com/LeastSquaresFittingPolynomial.html>

