Quantifying the Risks of Wellbore Failure during Drilling Operations Using Bayesian Algorithm

A thesis submitted to the faculty at African University of Science and Technology in partial fulfillment of the requirements for the degree of Master of Science in the Department of Petroleum Engineering

By

Emmanuel Adeyeye Adeyemi

Supervised by:

Dr. Xingru Wu

and

Prof. David Ogbe

African University of Science and Technology
www.aust.edu.ng
P.M.B 681, Garki, Abuja F.C.T, Nigeria

Abuja, Nigeria

June, 2019

CERTIFICATION

This is to certify that the thesis titled "**Quantifying the Risks of Wellbore Failure during Drilling Operations Using Bayesian Algorithm**" submitted to the school of postgraduate studies, African University of Science and Technology (AUST), Abuja, Nigeria for the award of the Master's degree is a record of original research carried out by Emmanuel Adeyeye Adeyemi in the Department of Petroleum Engineering.

Quantifying the Risks of Wellbore Failure during Drilling Operations Using Bayesian Algorithm

By

Emmanuel Adeyeye Adeyemi

A THESIS APPROVED BY THE PETROLEUM ENGINEERING DEPARTMENT

RECOMMENDED: .....

Supervisor, Dr. Xingru Wu

.....

Co-Supervisor and Committee Member, Prof. David Ogbe

.....

Head, Department of Petroleum Engineering

APPROVED: .....

Chief Academic Officer

.....

Date

ABSTRACT

The wellbore integrity plays an important role in petroleum operations like drilling, well completion and production. Caliper, Electrical image logs, Acoustic televiewers (ATV) and Optical televiewers (OTV) are some of the devices currently used in the industries to identify breakouts.

However, these techniques are restricted in applications. For instance; caliper sometimes indicates the effect of drill spiral grooves as borehole enlargement zones, poor resolution, and complicated processing procedure reduced the application of the electrical image logs. The ATV and OTV which give better outputs are not often used due to the high cost of installation and operation. These limitations necessitated a new approach to quantifying the wellbore instability.

This research work focuses on data analytics and development of Bayesian Algorithm (with code in Python to predict the wellbore failure probability using real-time pore pressure and fracture gradients data obtained from the wellbore using d-exponent.

Key Words: Bayesian, d-exponent, drilling, wellbore, pore pressure, breakouts

DEDICATION

To the Lord God Almighty, my Saviour.

ACKNOWLEDGEMENTS

I am overwhelmed by the kind disposition and unequalled support that I received from every staff and student of African University of Science and Technology, particularly from the Department of Petroleum Engineering during the entire period of my program.

Special thanks to my supervisors, Dr Xingru Wu and Prof. David Ogbe for their commitments to my academic development. The invaluable guidance they provided cannot be quantified. Their lifestyles have taught me that limits are self- imposed, and should always be challenged. My special thanks go to all my course mates for their cooperation and the spirited minds that brought uncommon unity. How will I thank the family of Babatunde and the entire BCC brethren for their contributions towards the success of my study? Kudos to my dad (Mr Adebiyi Francis) for accepting pain for my gain. Bravo to my mum (Mrs Dorcas Olayemi) for your labour of love. Prince Adeyemi Gbenga, you are a brother indeed. All my other siblings: Elizabeth, Esther, Janet and Joshua, I say thank you for always looking up unto me as a beckon of hope. Your eyes of expectation on me kept me through. Now I say to God be the glory for the success of this research work.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.0 Background

Wellbore stability significantly contributes to non-productive time by incurring high amounts of drilling costs and increasing safety risks. Early detection of relevant wellbore stability issues is therefore crucial. The continuous improvement of drilling and measuring technologies has resulted in a greater volume of higher quality data being available in real time for the identification and remediation of potential drilling hazards during the drilling process (Wessling et al., 2014).

It is not unusual that the difficulties encountered during a drilling operation can be tracked down to choices made during the planning stage. However, generating a robust drilling operation plan is not easy as there are often substantial uncertainties associated with the actual geological context. To address this problem, a method is proposed that evaluates quantitatively the risk levels of a drilling operational plan as a function of the underlying uncertainty associated with its description (Cayeux et al., 2016).

To improve the drilling process quality and efficiency one solution is to advance a specific computerized method for helping the petroleum industry in reducing unwanted downtime. In the thesis, this is being achieved by detecting process-deviations (symptoms) during the drilling process, and applies them to produce explanations and advices, supported by a general knowledge model (a drilling ontology) (Skalle et al 2016).

1.1 Problem Statement

(Udegbunam, 2014) indicated that there was an increasing concern in the oil and gas industry regarding wellbore stability problems. The need to improve well operations has become imperative as the operators move towards more challenging and harsher environments such as ultra-deep waters and high-pressure and high-temperature (HPHT) fields. Maintaining borehole stability is an important issue in the oil and gas industry and geothermal wells. In the drilling process, the economic loss caused by borehole instability reaches more than one billion dollars every year and the lost time accounts for over 40 % of all drilling-related nonproductive time (Yan et al., 2014). Wells with high down-hole temperatures have a higher risk of borehole instability resulting from the temperature change.

The well is balanced when the wellbore fluid gradient is equal to the formation pore pressure gradient. Many wells are drilled with a slight overbalanced condition where the loss of filtrate is accepted in order to develop a low permeability filter cake on the wellbore. Dropping below the pore pressure, an underbalanced condition, is likely to result in an influx of formation fluids into the well, which if not controlled can lead to a loss of the well section or a blowout. Exceeding the fracture pressure is equally hazardous as the loss of fluid to the formation can lower the annulus fluid level, lowering the downhole pressure, potentially enough to invite an influx into the well from another exposed formation. In either case, the excursion from the drilling margin can result in significant added costs to the well in the form of additional services, materials, and time to kill the well and perform remedial work. (Johnson et al., 2018).

In this research, an algorithm based on Bayesian analysis of real time data collected from the well while drilling is proposed with an aim to solving the age-long wellbore instability issues in oil and gas industries.

1.2 Objectives of the Research

The objectives of this research are to:

1. Develop a framework to predict the wellbore failure probability by using real-time pressure and temperature data from microchips retrieved from the wellbore.
2. Quantify and qualify the risks during drilling operations from the developed model
3. Build a real time model of spreadsheet to predict the needed mud weight for safe drill.

1.3 Outlay of the Thesis

- ✓ Literature survey for causes and remedies of wellbore failure during drilling operations.
- ✓ Identify existing methods of data collection and models in literature that can quantify and predict wellbore failure. Emphasis is laid on the temperature/pressure of the formation.
- ✓ Extract the data from proven technique used in Development and Field Evaluation of a Distributed Microchip Downhole Measurement System (Shi et al., 2015)
- ✓ Build the workflow, database and models for specific formations with varying properties using Naive Bayesian Classifier with codes in Python.
- ✓ Generate predictive model for wellbore failure under real-time condition.

CHAPTER TWO

LITERATURE REVIEW

This chapter discusses the basic concepts of the subject matter and the review of the existing works related to the study.

2.1 Wellbore Stability

Incorporating wellbore stability in the drilling program is essential as it highlights the likely sections in the well with a high probability of instability. Considering the geology of the field, diverse formations at different well depths provide a guide of the likely failure mechanisms and ways of avoiding or managing them. Loose hyaloclastite or tuff formations to hard granitic rocks are observed to occur in geothermal fields. The loose formation has a high likelihood of well collapse and zones of drilling fluid circulations while hard formations slow down the drilling progress. In addition to the geological aspect of the field, in-situ stresses, pore pressure, temperature, open hole and depth have an impact on geothermal drilling operations (Nganga & Company, 2018).

2.2 Risks in Drilling Operation

There are numerous sources of risk in a drilling operation and the quality of the evaluation made by the proposed risk evaluation method is directly related to treating as many threat sources as possible; however, it would be an overwhelming task to manage all potential sources of drilling problems at once. Therefore, in a first step, only a subset of possible causes of drilling incidents will be implemented. The causes from this initial list are associated with typical problems

encountered in drilling operations in the fields. Furthermore, this work focuses on conventional drilling, even though there is a potential to extend the method to other drilling methods like managed pressure drilling or dual-gradient drilling (Cayeux et al., 2016).

The first group of drilling problems is related to geo-pressure margins (see Figure 2.1). This includes the risk of the downhole pressure to be lower than the pore pressure anywhere along the open hole section during the evaluation of the drilling operation, therefore having the risk of causing a formation fluid influx (from now on this risk will be denoted R1-pore pressure). Similarly, the downhole pressure is compared to the collapse pressure of the open hole formation rocks to estimate the risk for borehole instabilities (denoted R2- collapse pressure). The risk of formation fracturing caused by a downhole pressure exceeding the fracturing pressure at any depth along the open hole section is also estimated (R3-fracturing pressure). Finally, a hydrostatic pressure being larger than the minimum horizontal stress of the exposed formation rocks leads to the evaluation of the risk of experiencing permanent mud losses in the case of a loss circulation incident (R4-min horizontal stress.)

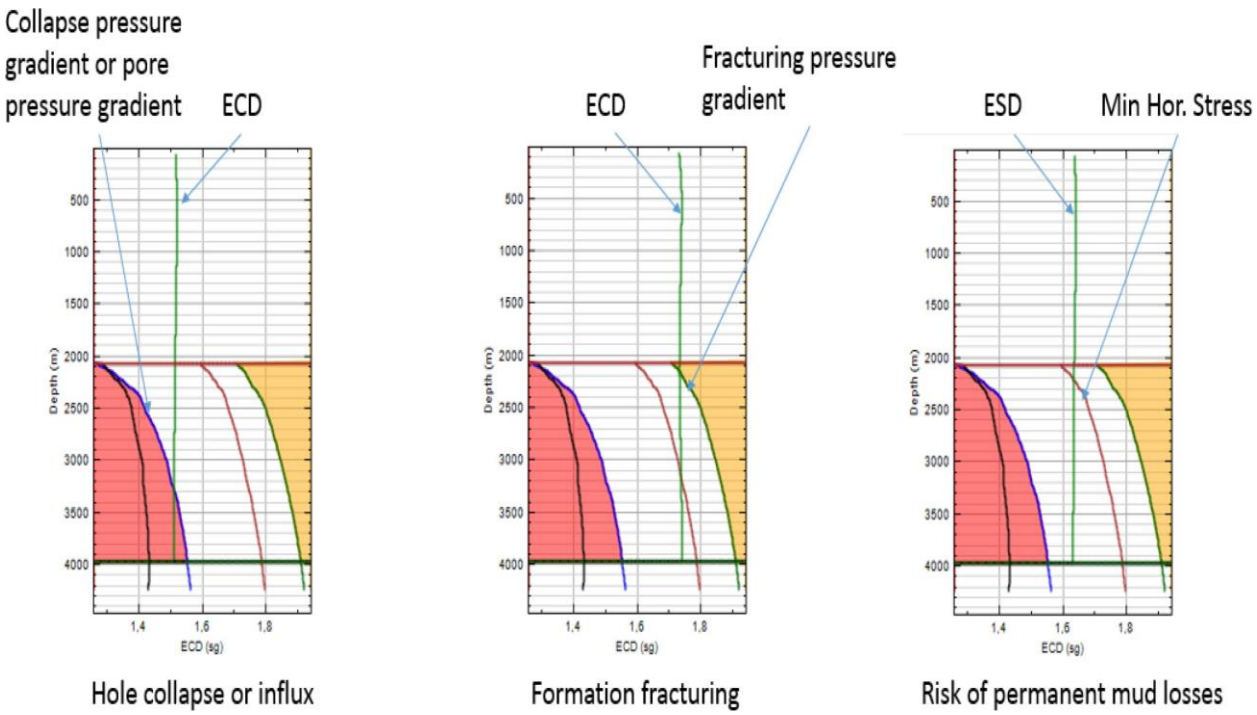


Figure 2. 1 Geopressure Margin (Cayeux et al., 2016)

Another group of risks is associated with cuttings transport. The high concentration of cuttings in suspension, anywhere along the annulus, may be the origin for a pack-off incident (R5-high cuttings concentration). Or in deviated wells, large accumulations of cuttings in beds may cause pack-offs, overpulls or set-down weights (R6-high cuttings bed) (see Figure. 2.2).

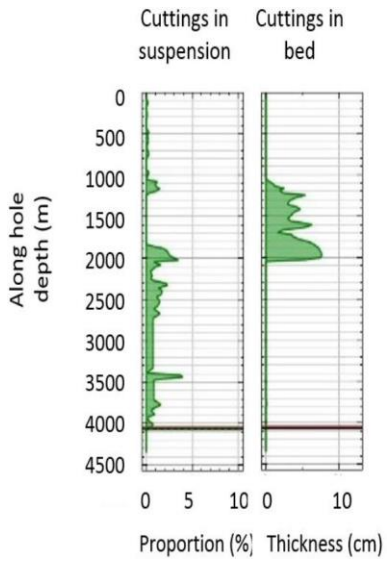


Figure 2. 2 example of a well with a long and thick cuttings bed (Cayeux et al., 2016)

Figure 2.2 above is an example of a well with a long and thick cuttings bed and relatively high concentrations of cuttings, both having the potential to create problems like pack-offs, overpulls and set-down weights. The percentage of cuttings are defined by the proportion of suspended particles in the hole on the scale of 0 to 10. The calipers are used to determine the thickness of the clogged cuttings in the hole.

The last group of sources of potential problems is connected to the limitations of the drilling equipment. For instance, the pump pressure shall not exceed the rating of the rig hydraulic system (R8-max pump pressure). The torque at the top-drive shall be lower than a maximum value that may depend on the selected gear for top-drives with multiple gear ratio and top-drive speed vs torque performance curves (R9-max top-drive torque)

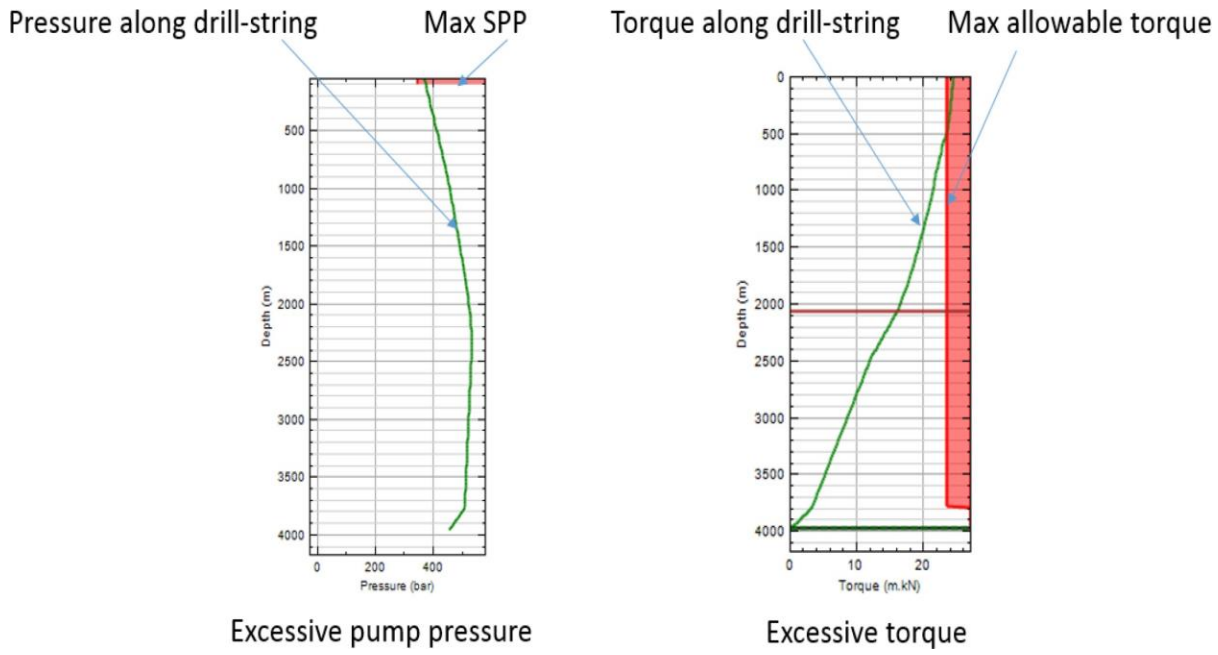


Figure 2. 3 maximum pump pressure being exceeded and drill-string torque above the maximum torsional strength of drillpipes (Cayeux et al., 2016)

Therefore, in this first initial stage of building the risk estimation framework, the following nine risks are evaluated: R1-pore pressure, R2-collapse pressure, R3-fracturing pressure, R4-min horizontal stress, R5-high cuttings concentration R6-high cuttings bed, R7-buckling, R8-max pump pressure, R9-max top-drive torque.

As mentioned earlier, there are many other sources of drilling dysfunctions like formation washout, stick-slip, bit bouncing, whirl, differential sticking, well collision, well path tortuosity, directional control problems, downhole telemetry shortage, downhole equipment failure.

Even for the aforementioned drilling problems, they may have origins that are not yet treated. For instance, formation fluid influx may be caused by gas influx even when the downhole pressure is larger than the pore pressure. Borehole collapse may result from chemical and thermal diffusion despite a downhole pressure greater than the collapse pressure (Zhai & S. Abou-Sayed, 2011). The intention is to expand the list of analyzed risks step by step.

Finally, the first version of the described risk evaluation tool is limited to the simulation of drilling operations. In a drilling context while using a conventional drilling method, there are three parameters that can be controlled: The weight on bit (WOB), the drill-string rotational speed (RPM) and flow-rate. The automatic risk analysis consists in varying these control variables throughout the simulation of the drilling operation. During the analysis, the maximum value used for the variation of the WOB is imposed by the bit manufacturer limit. The maximum bound of the flow-rate variation is imposed by the number of available pumps, their maximum speeds and stroke volumes. The maximum rotational speed is slightly more complicated when the top-drive has several gear ratios.

2.3 The Existing Works on Risks of Wellbore Stability

After studying the effect of thermoplastic on wellbore stability, (Chen & Ewy, 2006), concluded that pore pressure changes may might be independent of temperature change in high permeability formation such as sandstones. They further stressed that heating the wellbore decreases stability by increasing pore pressure hoop stress and then decreasing the collapse pressure failure index and

raising both the collapse and the fracturing mud weight for both permeable and impermeable boundary conditions.

In a similar perspective, wellbore inclination angle had been pointed out as one of the factors that affect wellbore stability. Based on numerical modeling and simulation, (Dosunmu & Bose, 2011) stated that for any rock type, an increase in borehole inclination angle increases rock borehole stability. Examining it from the solution point of view, breakouts in wellbore provide valuable information with respect to the evaluation of maximum horizontal stress magnitude and verification of the geomechanical model built for a field. After processing eight chosen petrophysical logs applied to the from five wells drilled in a carbonate reservoir, (Soroush & Tool, 2012) proposed that statistical classifiers (Parzenand Bayesian) could be applied to the de-noised petrophysical logs to classify depths into Breakouts (BO) and non-breakout (nBO) zones. He pointed out that the Bayesian algorithm in conjunction with wavelet de-noising is an excellent tool to detect borehole enlargement zones in carbonate formation and that it is a good replacement technique for image and caliper logs.

The necessity of real-time monitoring of drilling operation in the oil and gas industry called for an automated system capable of capturing relevant data required to combat the menace of wellbore instability. (Wessling et al., 2014) put forward automation of a single module of the geomechanical analysis workflow. The workflow included automation setting of a normal compaction trend and determination of the onset of the overpressure using statistical methods, image interpretation algorithms to identify depth and width of breakouts; and automatic shale discrimination at the bit with drilling technologies together with continuous, step-wise development of algorithms are promising ways forward towards automation beneficial.

Modern exploration and exploitation of petroleum reserves required efficiently managed drilling operations. (Yu et al., 2012) introduced the concept of microchip technology to acquire subsurface data such as pressure and temperature. They injected 13 tracers in an 11,500 ft well circulated with a 13 ppg mud at 400-500 gpm. After travelling about 6.6 km in the 8-3/8 inch wellbore to a depth of 11,050 ft, seven tracers returned to the surface successfully. It was found that the first tracer returned to surface after 51 minutes. The data collected on the tracer shows a bottom hole pressure of about 7,500 psi and flowing bottom hole temperature of 190 $°F$. They concluded that the microchip system could be used for subsurface measurement or as a diagnostic tool for identifying problematic zones.

As far back as 1991, (Adam T. Bourgoyne Jr et al., 1991) submitted that to safely drill a deep well for hydrocarbon exploration or production, it is necessary to prevent formation fluids from flowing into the well. This is typically done by adjusting the density of the drilling mud so that the wellbore pressure is at all depths above the pressure of formation fluids (the pore pressure). On the other hand, the mud density cannot be so high as to cause hydraulic fracturing of the formation (the fracture pressure). The pore pressure and the fracture pressure gradients thus provide minimum and maximum values for the mud weight that define a mud weight window.

In current practice, estimates of the pore and fracture pressures can be obtained from information on the variation in compressional wave velocity with depth (in turn obtained from surface seismic data and borehole measurements). These estimates can then be calibrated with pressure data acquired during drilling. Although these pore and fracture pressure estimates are recognized to be inaccurate, a major limitation of current practice is that there is no quantification of their uncertainty.

2.4 The Logging While Drilling (LWD) Telemetry Technology

2.4.1 Electromagnetic Telemetry

Each hydrocarbon bearing formation has unique properties that affect the transmission of electrical signals and which varies with the depth, spacing, and sequence of different types of geological layers, among others. The electrical properties of the drillstring also dynamically vary with length, the constitution of the drilling fluid and temperature.

Electromagnetic telemetry adapts to the electrical environment encountered in the well by means of a microprocessor in a monitoring tool, which continuously scans the frequency spectrum with electromagnetic signals to determine the optimum frequency for data transmission between the tools and processing units (computers) on the surface (Jr. et al., 2015). This transmission can be achieved by both through the drillstring body or through the formation being drilled. Figure 2.4 shows a general electromagnetic telemetry system.

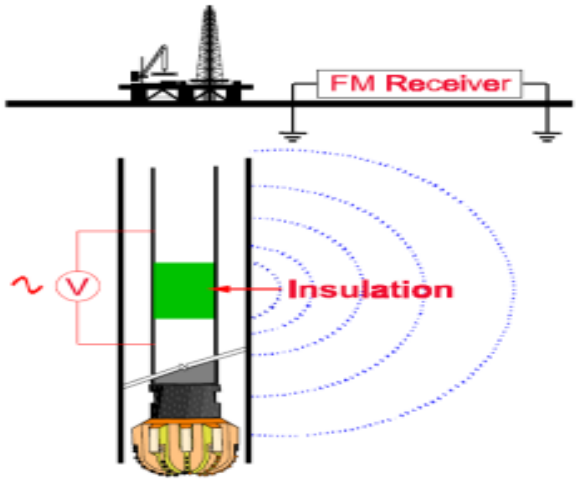


Figure 2. 4 Electromagnetic telemetry system (Hughes, 1997).

By determining the relative attenuation of the transmitted signal, the operator at the surface can send a command back to the tool (or vice-versa) to replace the frequency of the carrier wave with one that may suffer less attenuation and which may have a better signal to noise ratio. For example, highly resistive formations below the drill string cannot function as ground and the drill string itself must then be treated as a vertical dipole. In this case, the most effective frequency for the carrier wave would be one resonant with the drill-string, identified by the tool like the one drawing maximum input current. The most frequently used optimal selection technique is one in which the receiver periodically sends recognition signals to the transmitter, so that when this signal is not captured by the same transmitter, it changes the carrier wave frequency gradually decreasing it until it returns capturing the recognition signals from the receptor. However, by lowering the frequency of the carrier wave, the amount of transmitted data is consequently diminished. In this case, the tool in the well has to be programmed to transmit only the most critical information relevant to the operation. (Mwachaka, Wu, & Fu, 2019) wrote that the first papers on electromagnetic telemetry were quite promising, claiming that rates of up to 100 bits per second (bps) could be achieved with the use of signal repeaters. However, the development of mud pulse telemetry in parallel on the market brought more efficient results, with higher data transmission rates reaching greater depths, without the needs of repeaters usage. The huge attenuation suffered by the electromagnetic signals, caused by the formation properties, drilling mud and surface conditions led to a decline of this technology.

Another aspect of restricting this technology is water. Due to its high electrical conductivity, it limits the use of electromagnetic telemetry in offshore operations with large water depths, making it be relegated just to the oil & gas (O&G) market onshore and for shallow depths, considering the

low transmission rates (Jr. et al., 2015). On the other hand, in activities where usual drilling mud may not be used, alternatives are to use aerated.

2.4.2 Acoustic Telemetry

The acoustic telemetry system is operated with batteries and is distinguished from other systems since it works by generating acoustic waves capable of transmitting real-time data to the surface through the walls of pipes at distances up to 12,000 ft. A variety of data can be acquired and transmitted: pressure, temperature, time, command and information about the system status. This telemetry system is used mainly in exploratory wells and also for well-testing operations. Meters equipped with quartz crystal sensors located just above the tester valve generate accurate signals of temperature and pressure in the deep end, being stored in recording memory or directly transmitted to the surface. Each quartz sensor is able to store up to 440,000 readings in its memory, and the entire system is capable of storing more than 1.3 million scans, allowing great flexibility.

In real-time, the transmitter sends packets of data with all information acquired in every two minutes directly to the next repeater. A packet is a group 12 acoustically transmitted data sets, each set containing pressure information, temperature and time for sample intervals of 10 seconds. The repeater then forwards the packet to the surface to be decoded. The system is able to communicate in both directions, allowing the operator to send additional commands on the surface below to the system, for example, if any changes have to be done. Figure 6 exemplifies the acoustic telemetry system.

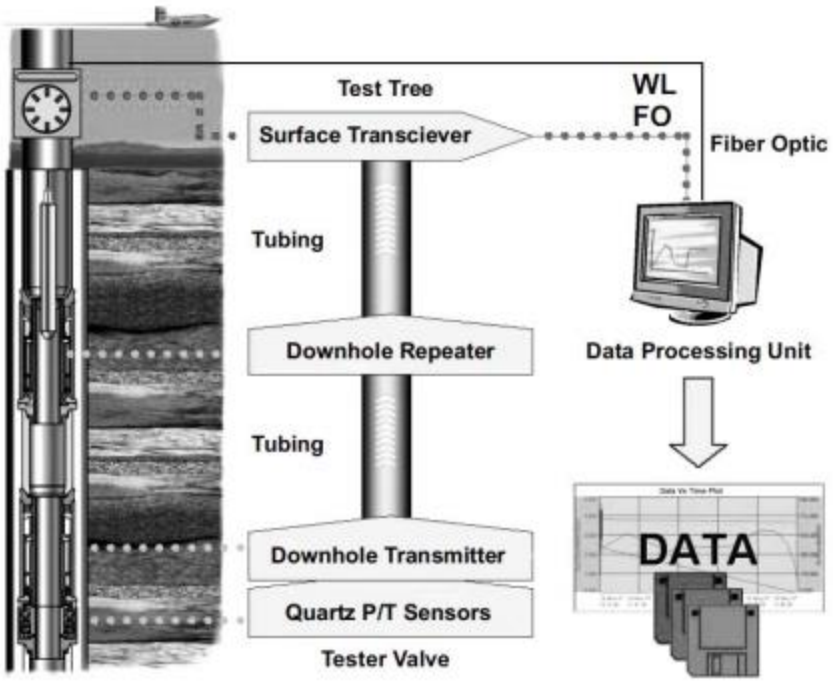


Figure 2. 5 Flowchart of an acoustic telemetry system operation. (Jr. et al., 2015).

In wells deeper than 12,000 ft, more repeaters are necessary but are not as straight forward as it looks like due to the "cross-talk" problem (unwanted interference of a transmission channel to another). Moreover, in such environment, the great difficulty lies in transmitting the acoustic signals to the surface through the docking column, which due to its size and components can cause large attenuation. Therefore, the acoustic signal is brought up to the sub transducer surface, and from there converted into an electrical signal and transmitted through a twisted pair cable to the surface. In 2011, XACT Downhole Telemetry Inc. published a study on its acoustic telemetry module capable of transmitting data to a 20 bps rate in wells up to 2500 m ((Jr. et al., 2015)).

2.4.3 Mud Pulse Telemetry

This technology employs a module that modulates resistance to the flow of the drilling fluid through the inside of the drill-string, generating an increase and decrease of the stand-pipe pressure. The mud pressure pulses are propagated approximately at the speed of sound to surface. Transducers located on the surface detect and convert the pressure signal into a digital electrical signal through analog/digital (A/D) converters. This signal is then sent to a computer that will process and decode the signal received through specific and developed software to recognize and treat these signals.

In an ideal situation, each pressure pulse created by the module would spread column up and would be easily detected by transducers. However, the pressure of the drilling fluid undergoes significant fluctuations and containing noises from various sources such as the drill noise, noise from the torque and from the mud pump itself, etc. The drill noise is caused by its vibration during the drilling operation, which partially restricts the output of drilling fluid causing a high-frequency noise. Torque noise is caused by the increase in drill string torque when the drill is in contact with the formation; after contact, the torque in the column is relieved generating a peak of large amplitude and low-frequency pressure. Finally, mud pump noises are due to the cyclical piston movement for displacing mud into the circulating system.

Some drilling systems use a buffer on the surface to reduce noise caused by the mud pump, and the pulsation dampener can also be adjusted for that. However, while they absorb some of the pressure fluctuations, they also act as a mirror, reflecting pressure pulses back to the telemetry module, sometimes destructively, creating interference and hindering or making the detection of

pulses by the transducer in surface more difficult. A basic telemetry module contains two sections: one for communication and others for control of the generated pulses.

Various logging tools send their data to a digital signal processor (DSP) located in the communication section. A compression module reduces the amount of data transmitted over techniques involving data filtration. It is also used differential encoding which allows a data string to be represented with fewer bits than usual. The multiplexing module selects the data of the different tools and assembles a single chain of data being transmitted, divided into blocks that can contain information about synchronization and error correction. A coding module then converts the digital signal channels to be transmitted on a set of timings that are communicated to the pulse control section for generating them. The pulse control section consists of a processor, memory, an opening coil, closing coil, two banks of capacitors and battery. The control pulses operate a valve through the opening and closing coils generating the pressure pulses. The coils drain relatively high amounts of current in operation, in some cases more than the battery can deliver. Its power though must be within the operating capacity of the battery. To solve the problem of the current supply, each reel is associated with a capacitor bank. The battery charges the capacitors between the operations of the reels, and when the processor activates the reels they unload supplying current making the valve operation possible.

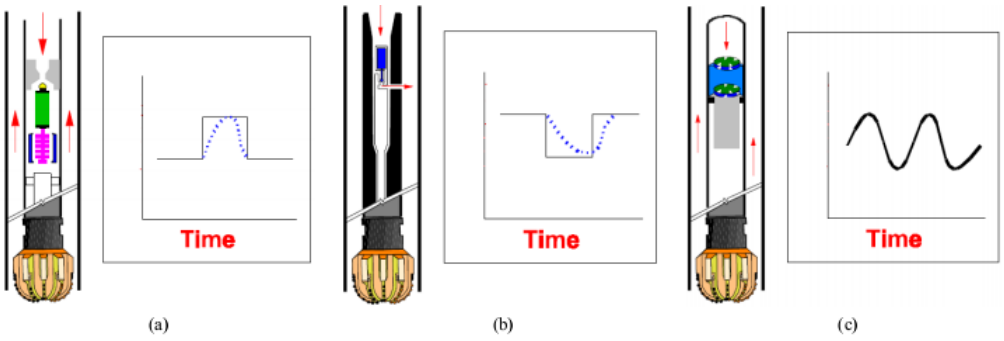


Figure 2. 6 Mud pulses: positive (a), negative (b), and continuous (c). (Hughes, 1997)

The valve that creates pressure pulses can have different shapes and constructions, being classified among different types in terms of the operating system. The three most commonly used types are: positive pulse (Figure 2.6 (a)), negative pulse (Figure 2.6 (b)) and continuous pulse (Figure 2.6(c)).

Any of these three types can be used provided that the valve can produce variations quickly enough (in the order of microsecond). The duration of the pulses may vary from 80 ms until approximately 400 ms, depending on the drilling system parameters. In the positive pulse, the equipment creates a restriction in which the stand-pipe pressure increases. In the negative one, the equipment allows part of the flow to leak to the annulus, creating a pressure decrease. Finally, in the continuous one, there is a modulator and a stator creating restrictions and reliefs continuously.

2.4.4 Wired-Drill-Pipe Telemetry

This technology uses individually modified drill-pipes to provide a two-way telemetry system for real-time transmission, speeding up to 57,000 bps, making possible to obtain large amounts of data from down-hole (Bybee, 2015). The technology used high strength coaxial cables and low-loss inductive coils built in connections on each tubular joint to transmit information. Signal repeaters are placed at specific locations along the drill-string to ensure an acceptable signal to noise ratio. These repeaters operate as individually accessible nodes within the telemetry network, and thus, being able to identify sites that can provide potentially valuable measuring data. Figure 2. 7 shows a schematic of the mentioned wired drill pipe, focusing on its threads. A bi-directional network architecture, which is this specific case, allows the transmission of downhole data to the surface at high speed while commands can still be sent from the surface to the equipment itself. By inserting a physical and electrical interface, existing logging-while-drilling (LWD) and measuring-while-drilling (MWD) tools become fully compatible. Although there are small variations between tools

from each manufacturer, the interfaces are generally consistent in the industry: it involves a network inductive connection at the top, a suitable electromechanical coupling on the bottom, a network card, modem and a power source.

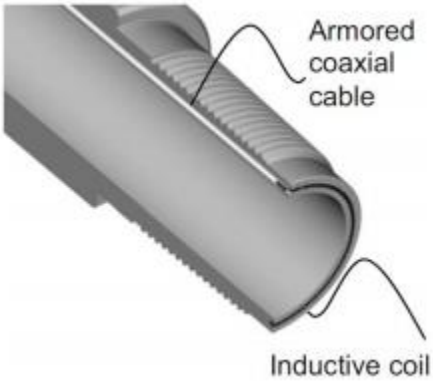


Figure 2. 7 Wired-drill pipe joint. High-Speed Wired-Drillstring Telemetry (Bybee, 2015)

Table 2. 1 Comparative performance among LWD telemetry technologies (Bybee 2015).

Features	LWD telemetry technology			
	Electromagnetic	Acoustics	Mud pulses	Wired drill pipe
Max. transmission data rate(bps)	10	20	20	57,600
Max. depth (meters)	5500	3700	12,200	Unlimited
Data quantity	Medium	Low	High	Very high
Signal attenuation	High	High	Medium	N/A
Signal interference	High	Medium	Medium	Low
Installation	Medium	Medium	Low	High

2.5 A Distributed Microchip Downhole Measurement System

A new generation of a near real-time, cost-effective distributed downhole temperature and pressure measurement system that utilizes improved microchip technology was developed in the laboratory and tested in the field.

The improved microchip technology splits temperature and pressure measurement into two individual systems. Design of the printed circuit board and the electrical components has minimized the measurement error and power consumption of the system. Significant improvements in the integrity of the drilling microchip have been achieved by using a new protective material. Experimental results show that the accuracy of the system measurement is within ±0.5°C for temperature measurements and ±0.05% for pressure measurements.

The drilling microchip is typically deployed by using a tracer injection system continuously (controlled by a computer) or by dropping it into the drillpipe while making a pipe connection during rig operation. It travels along the inner passage of the drillpipe, exits the bit nozzle, and returns toward the surface in the wellbore annulus because of dynamic circulation with drilling fluid. Then, it is usually captured at the shale shaker at the surface.

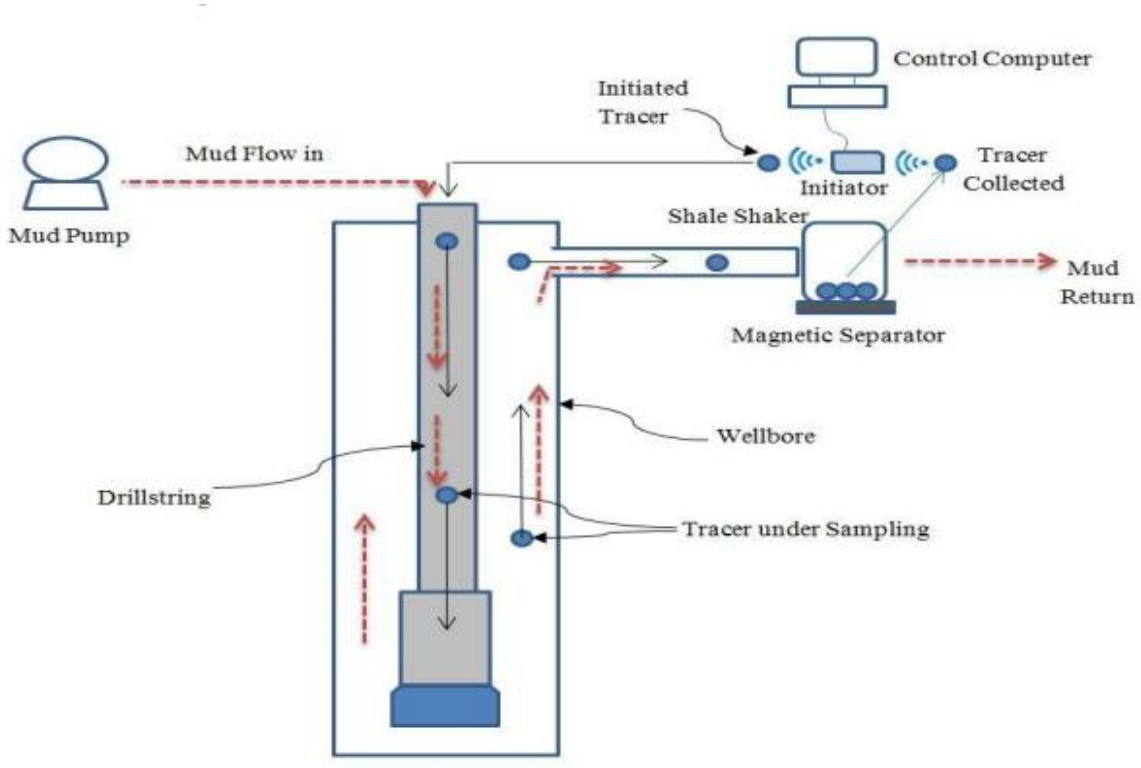


Figure 2. 8 A Distributed Microchip Downhole Measurement System (Shi et al., 2015)

A total of 14 new generation tracers were deployed into two different wells in two batches (4 pieces and 10 pieces) during the field evaluations in Saudi Arabia. Nine of the deployed 14 microchips were recovered at the surface with the measurement data stored in the on-chip memory. The results of the field tests proved that the microchip system is able to survive in a well with 13,800 ft vertical depth under more than 150°C bottomhole temperature and 10,000 psi bottomhole pressure conditions. Measurements retrieved from the recovered microchips show excellent consistency. (Shi et al., 2015)

The circulating mud temperature in the field test well has also been predicted by using the existing thermal model. The modeling result has been compared to the microchip measurements and it shows a strong agreement between two temperature profiles inside the drillpipe. However, the modeling result is probably not accurate in the annular section based on comparison to the microchip measurements.

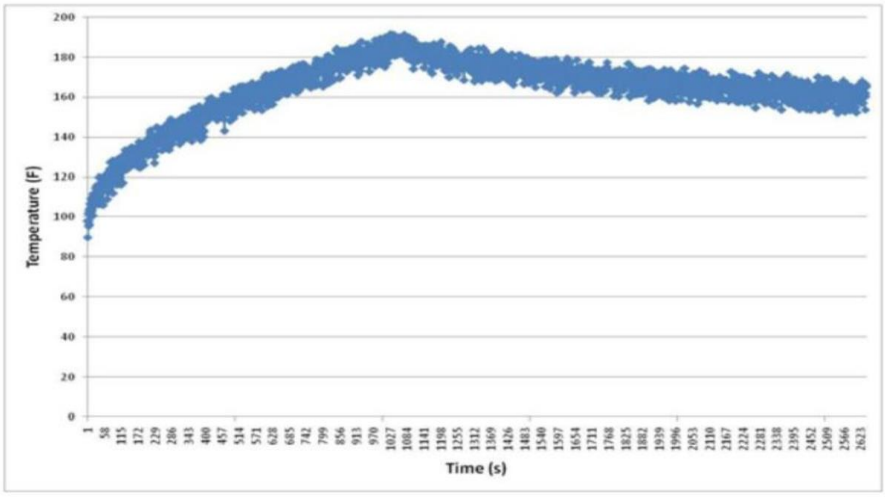


Figure 2. 9 Temperature data downloaded from the first generation microchip in the field test. (Shi et al., 2015)

The innovative distributed microchip downhole measurement is a unique data acquisition system designed to provide downhole in-situ circulating mud temperature and pressure (i.e., equivalent circulating density) profiles. It can function as a diagnostic tool for problems such as seepage, lost circulation zones, and open-hole restriction due to poor hole cleaning or an under-gauged hole encountered in the drilling and cementing process.

2.5.1 New Generation Microchip Development

The new generation distributed microchip system separates temperature and pressure measurement into two individual systems (Shi et al., 2015). Each system has an individual high precision sensor specially designed for temperature or pressure measurement. By splitting the system, each individual system has double spaces to store more data samples. The protective coating of the microchip is also improved in terms of microchip integrity. Developments of the two individual new generation systems are described in the following section.

2.5.2 New Generation Temperature Sensor

The new generation temperature sensor is designed to apply a precise and continuous temperature measurement. An innovative digital temperature sensor was used in the new design to replace the fiber optic temperature sensor in the first generation microchip system. Compared to the original sensor, the new sensor has several advantages. Firstly, the accuracy of the temperature sensor (±0.5°C) is much higher. Secondly, the size of the sensor is much smaller so that it saves space on the printed circuit board (PCB) layout. Another improvement is the low power consumption of the new temperature sensor. The operation current is low, and it increases the service life of the battery when taking measurements. All of these advantages make the new temperature sensor ideal for space-constrained, power sensitive applications such as the microchip application.

The new temperature sensor is in the temperature integrated circuit (IC). There is an I/O control interface to establish a connection between the temperature IC and the micro-controller unit of the system, as shown in Figure 2. 10. The sensor can implement read and write command from the host.

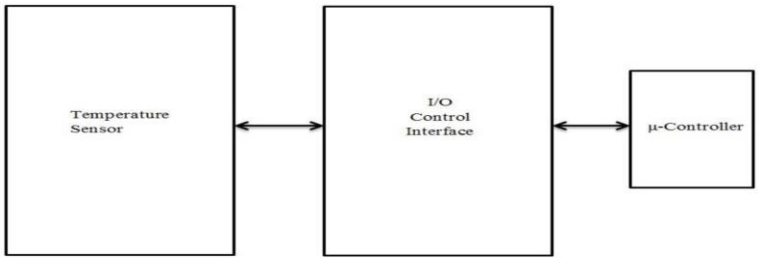


Figure 2. 10 Internal Structure of Microchip Temperature sensor. (Shi et al., 2015)

2.5.3 The Schematic of microchip Instrument System

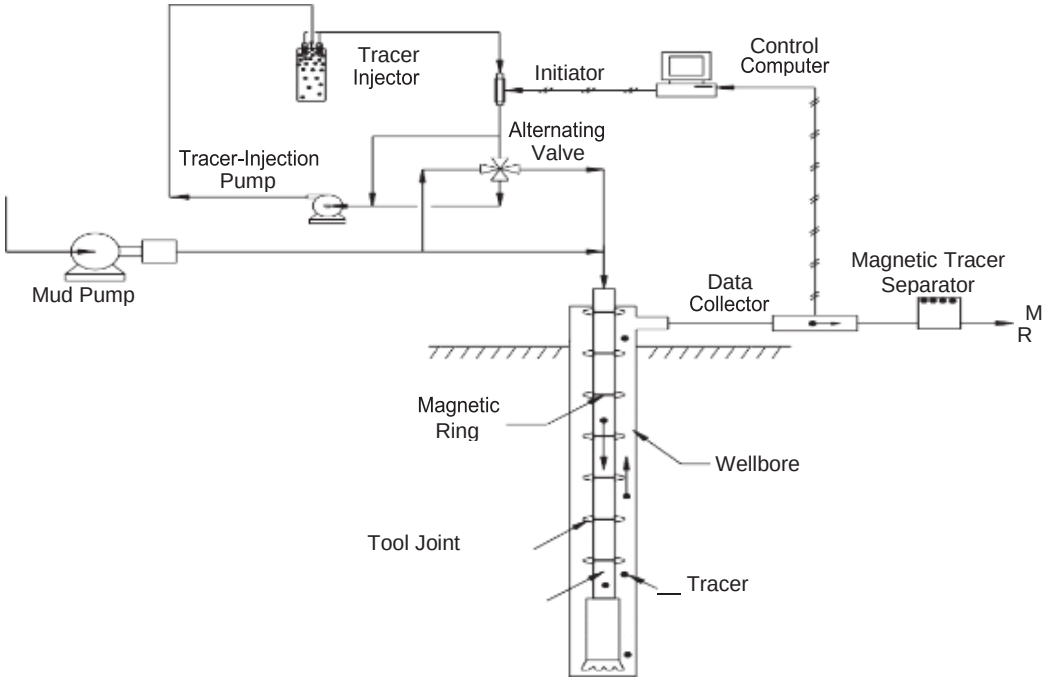


Figure 2. 11 The schematic diagram showing the deployment of microchips into the wellbore. (Yu et al., 2012).

The prototype of the instrument system is shown in Figure 2. 11 works as follow:

The surface devices include an initiator to reset the circuit on the tracer before the tracer is injected into the well-bore, and a data collector to retrieve data from the tracer's on-chip memory when the tracer is carried back to the surface by the drilling fluid. The initiator and data collector will use wireless communication to reset the circuit and download data from the tracer, respectively.

A magnetic tracer separator (see Figure 2. 11) will be installed after the data collector to recycle the tracers. A lithium cell (battery) used in the tracer contains stainless steel, which can enhance the separation of tracers from the drilling fluid in the magnetic tracer separator. When used in experiments or in the field, tracers will be injected together with the drilling fluid. Before a tracer is injected into the flowline, it passes through an initiator, which will reset the circuit for recording data. As the tracer travels through the wellbore, it will measure the temperature and pressure throughout the wellbore and store the data in the on-chip memory at a sampling rate set by the initiator. When the tracers are carried out of the borehole by the drilling fluid, they will pass through a data collector (controlled by a computer) through which the tracer will communicate with the surface devices to send the data stored in its on-chip memory.

Fiber-optic temperature and pressure sensors for harsh environments (up to 15,000 psi and 250°C) have been used in the field successfully during the last 10 years and will be integrated into the instrument system in this project. Other sensors can be integrated into the system when they are readily available. When used in the field, magnetic rings can be placed on the tool joints

(glued to the surface of the pin). An on-chip magnetic sensor can be fabricated on the SOC IC. As the tracer travels through the borehole, it could sense the weak magnetic field around each tool joint

along the drillstring. This information can be recorded in the on-chip memory and used to determine the real-time location of the tracer in the wellbore (Yu et al., 2012).

CHAPTER THREE

METHODOLOGY

This chapter presents approaches used in the work with a step-by-step analysis starting from the data preparation to the feeding of the data to Bayesian Algorithm.

3.0 Basic Techniques for Data Gathering and Preparations

3.1 Background of Bayesian Algorithm

Naive Bayes is a simple technique for constructing classifiers: models that assign class labels to problem instances, represented as vectors of feature values, where the class labels are drawn from some finite set. There is not a single algorithm for training such classifiers, but a family of algorithms based on a common principle: all naive Bayes classifiers assume that the value of a particular feature is independent of the value of any other feature, given the class variable. For example, a fruit may be considered to be an apple if it is red, round, and about 10 cm in diameter. A naive Bayes classifier considers each of these features to contribute independently to the probability that this fruit is an apple, regardless of any possible correlations between the color, roundness, and diameter features.

The Bayesian algorithm provides a probabilistic framework for a classification problem. It has a simple and sound foundation for modeling data and is quite robust to outliers and missing values (Belief, & Han, 2012)

3.2 The sequential steps taken in preparing the data used in the model

Step 1: Obtain the d-exponent parameter from Jorden and Shirley's correlation (Solano, Uribe, &

Frydman, 2007); $d_{exp} = \frac{\mathrm{Log}\left(\frac{R}{60N}\right)}{\mathrm{Log}\left(\frac{12W}{1000d_b}\right)}$ ……3.1

Where R is rate of penetration in ft/hr, N is rotary speed in rev/min, W is weight on bit in klbf and d_b is bit size in inches.

Step 2: Obtain modified d-exponent from Rehm and McClendon's correlation (Shajari & Najibi,

2015); $d_{mod} = d_{exp}\left(\frac{\rho_n}{\rho_e}\right)$ ……3.2

Where ρ_n is normal pore density and ρ_e is equivalent circulating density both in ppg.

Step 3: Plot modified d-exponent against depth and estimate the slope and intercept of the resulting straight line using exponential curve fitting equation: $d_{est} = ((d_{mod})e^{mD}$ ……3.3

Step 4: Obtain d-estimate from the resulting slope and intercept in 3.3 above. By inference from Zamora's correlation (Stunes, 2012):

$d_{est} = \left(\frac{1}{m}\right)[\ln D + \ln(d_{mod})]$ ……3.4

Where D is depth in ft, m is the slope of the normal pore pressure region in ft^{-1} and (d_{mod}) is the intercept on modified d exponent axis.

Step 5a: Estimate the overburden stress (σ_{ob}) from Mitchell's approximation of Eaton:

$$\sigma_{ob} = 0.84753 + 0.01494\left(\frac{D}{1000}\right) - 0.0006\left(\frac{D}{1000}\right)^2 + 1.199 * 10^{-5}\left(\frac{D}{1000}\right)^3 \ldots\ldots\ldots\ldots\ldots 3.5$$

Where D is in ft and σ_{ob} in psi/ft.

Step 5b: Accuracy check of overburden stress estimated in 5a above from Bourgoyne's correlation:

$$\sigma_{ob} = 0.052\left\{\rho_{sw}D_w + \rho_g D_s - \frac{(\rho_g - \rho_{fl})\phi_o}{K}(1 - e^{-kD_s})\right\} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 3.6$$

σ_{ob} is overburden pressure in psi/ft, D_w is the depth of the sea waterft, ρ_{sw} is the sea water density in ppg, ρ_g is the grain density in g/cm^3, ρ_{fl} is fluid density in g/cm^3, ϕ_o is surface porosity, k is porosity decline constant in ft^{-1} and D_s is the depth below the surface of the sediments in ft.

Step 6: Determine the formation pore pressure in psi/ft and ppg from:

$$F_P = \sigma_{ob} - (\sigma_{ob} - N)\left(\frac{d_{mod}}{d_{est}}\right)^m \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 3.7$$

Where N is Normal formation pore pressure in psi/ft and m is exponent constant.

Step 7: Obtain the minimum mud window from F_P using trip margin of 0.5 ppg

Step 8: Determine the Poisson ratio from:

$$\gamma = 0.23743 + 0.05945\left(\frac{D}{1000}\right) - 0.00668\left(\frac{D}{1000}\right)^2 + 0.00035\left(\frac{D}{1000}\right)^3 - 6.71 * 10^{-6}\left(\frac{D}{1000}\right)^4 \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 3.8$$

D is the depth in ft.

Step 9: Determine the fracture gradient from: $F = \frac{\sigma_{ob-F_P}}{D}\left(\frac{\gamma}{1-\gamma}\right) + \frac{F_P}{D}$ 3.9

Where F_P is the formation pore pressure in psi and σ_{ob} is overburden pressure in psi and D is depth in ft.

Step 10: Obtain the maximum mud window as Fracture gradient F_g less kick margin of 0.5 ppg.

Step 11: Plot the graph of mud windows defined by steps 7 and 10 above against depth.

3.2 Input Data for Bayesian Algorithm Model

Gaussian Naives Bayes (GaussianNB) is used to analyse the data for the prediction of probability of wellbore failure. This is because the data is continuous and normally distributed. Appendix B shows the plot of the input data.

The input parameters (depth, mod_dex, est_dex, sig_f, Poisson) were used to the mud windows.

Step 1: Convert the input data parameter from xlsx file into csv file.

Step 2: Launch the Jupiter notebook (What is Jupiter notebook?) and import the input parameter as csv file

Step 3: Define and import the class targets as 'N', 'K', and 'LC'

Step 4: Split the input data into 'training' and 'test' data

Step 5: Train the model using training data set

Step 6: Generate the confusion matrix using both trained and test data

Step 7: Obtain the prediction ability of the model from the values of Precision, Recall and Accuracy.

3.3 Confusion Matrix: Model for Validation of Prediction

A confusion matrix is a table that is often used to describe the performance of a classification model (or "classifier") on a set of test data for which the true values are known. The confusion matrix itself is relatively simple to understand, but the related terminology can be confusing.

The table below presented the basis for the confusion matrix used in the model.

Table 3. 1 Confusion matrix

ACTUAL	PREDICTION			
		n	k	l
	n	TP_n	e_{nk}	e_{nl}
	k	e_{kn}	TP_k	e_{kl}
	l	e_{nl}	e_{nl}	TP_l

Formulae for prediction of Precision, Recall and Accuracy.

$$FN_{\_k} = e_{\_kn} + e_{kl} \quad \text{3.10}$$

$$FN_{\_n} = e_{\_nk} + e_{\_nl} \quad \text{3.11}$$

$$FN_{\_l} = e_{\_ln} + e_{lk} \quad \text{3.12}$$

$$FP_{\_n} = e_{\_kn} + e_{\_ln} \quad \text{3.13}$$

$$FP_{\_k} = e_{\_nk} + e_{\_lk} \quad \text{3.14}$$

$$FP_{\_l} = e_{\_ln} + e_{\_kl} \quad \text{3.15}$$

$$P_{\_n} = e_{\_kn} + e_{\_ln} \quad \text{3.16}$$

$$TN_{\_n} = tp_{kick\_} + e_{\_lk} + e_{\_kl} + \text{tp}_{\text{lost\_}} \quad \text{3.17}$$

$$TN_{\_l} = tp_{\_norm\_} + e_{\_nk} + e_{kn} + tp_{kick\_} \quad \text{3.18}$$

$$TN_{\_k} = tp_{norm\_} + e_{\_nl} + e_{\_ln} + \text{tp}_{\text{lost\_}} \quad \text{3.19}$$

$$Precision = \frac{TP}{(TP+FP)} \quad \text{3.20}$$

$$Precision_n = \frac{tp_{norm\_}}{(tp_{norm} + FP_n)} \quad \text{3.21}$$

$$Precision_k = \frac{tp_{kick\_}}{(tp_{kick} + FP_k)} \quad \text{3.22}$$

$$Precision_l = \frac{tp_{lost\_}}{\left(tp_{lost_{+FP_l}}\right)} \quad \text{3.23}$$

$$Recall = \frac{TP}{(TP + FN)} \quad \ldots\ldots 3.24$$

$$recall_n = \frac{tp_{norm\_}}{(tp_{norm} + FN_n)} \quad \ldots\ldots 3.25$$

$$recall_k = \frac{tp_{kick\_}}{(tp_{kick} + FN_k)} \quad \ldots\ldots 3.26$$

$$recall_l = \frac{tp_{lost\_}}{(tp_{lost} + FN_l)} \quad \ldots\ldots 3.27$$

$$Accuracy_l = \frac{(TN + TP)}{(TN + TP + FN + FP)} \quad \ldots\ldots 3.28$$

$$accuracy_n = \frac{\left(TN_n + tp_{norm\_}\right)}{\left(TN_n + tp_{norm_{+FN_n}} + FP_n\right)} \quad \ldots\ldots 3.29$$

$$accuracy_k = \frac{\left(TN_k + tp_{kick\_}\right)}{\left(TN_k + tp_{kick_{+FN_k}} + FP_k\right)} \quad \ldots\ldots 3.30$$

$$accuracy_l = \frac{\left(TN_l + tp_{lost\_}\right)}{\left(TN_l + tp_{lost_{+FN_l}} + FP_l\right)} \quad \ldots\ldots 3.31$$

The equations above were used to develop the inputs for the Bayesian Algorithm (see appendix B) and nomenclature for the meaning of terms used.

CHAPTER FOUR

Results and Discussion

The following numerical results were obtained from the spreadsheet based on the methodology applied in the study.

4.1 Numerical Examples of the Model

The above step-by-step procedures were performed using spreadsheet to obtain the results below.

The input data begins with modified d-exponent. The coding of all correlations is done with the spreadsheet and the results can be found in appendix A.

Table 4. 1 Test data for the Bayesian Algorithm

Depth (ft)	mod d-ex	Est_dex	σ_{ob} (psi/ft)	Poisson	class
0	0.00	0.00	0.85	0.24	N
1000	0.00	0.00	0.87	0.29	N
2000	0.00	0.00	0.88	0.33	N
3000	0.00	0.00	0.89	0.36	N
4000	0.00	0.00	0.90	0.39	N
5000	0.00	0.00	0.92	0.41	N
6000	0.00	0.00	0.92	0.42	N
7000	0.00	0.00	0.93	0.43	N
8100	1.53	0.00	0.94	0.44	N
9000	1.56	0.00	0.95	0.44	N
9600	1.58	0.00	0.95	0.45	N
10100	1.50	0.00	0.96	0.45	N
10400	1.59	0.00	0.96	0.45	N
10700	1.61	0.00	0.96	0.45	N
10900	1.62	0.00	0.96	0.45	N
11100	1.58	0.00	0.96	0.45	N
11300	1.65	0.00	0.96	0.45	N
11500	1.49	0.00	0.96	0.45	N
11600	1.62	0.00	0.97	0.45	N
11800	1.55	0.00	0.97	0.45	N
12100	1.59	0.00	0.97	0.45	N
12200	1.68	0.00	0.97	0.46	N
12300	1.42	1.63	0.97	0.46	K
12700	1.28	1.64	0.97	0.46	K
12900	1.19	1.64	0.97	0.46	K
13000	1.14	1.64	0.97	0.46	K
13200	1.23	1.64	0.97	0.46	N
13400	1.13	1.65	0.98	0.46	N
13500	1.13	1.65	0.98	0.46	N
13600	1.08	1.65	0.98	0.46	LC
13700	1.01	1.65	0.98	0.46	LC
13800	0.98	1.65	0.98	0.46	N
13900	1.01	1.66	0.98	0.46	LC
14000	0.91	1.66	0.98	0.46	LC
14200	0.93	1.66	0.98	0.46	LC
14400	0.86	1.66	0.98	0.46	LC

Depth (ft)	mod d-ex	Est_dex	σ_{ob} (psi/ft)	Poisson	class
14600	0.80	1.67	0.98	0.47	LC
14800	0.86	1.67	0.98	0.47	LC
14900	0.80	1.67	0.98	0.47	LC
15000	0.90	1.67	0.98	0.47	LC
15200	0.82	1.68	0.98	0.47	LC
15300	0.87	1.68	0.99	0.47	LC
15400	0.92	1.68	0.99	0.47	LC
15500	0.87	1.68	0.99	0.47	K
15700	0.80	1.68	0.99	0.47	K
16200	0.80	1.69	0.99	0.47	N
16800	0.65	1.70	0.99	0.48	N

Table 4. 2 Training data for the Bayesian Algorithm

depth (ft)	mod d-ex	Est_dex	σ_{ob} (psi/ft)	Poisson	class
0	0.00	0.00	0.85	0.24	N
1000	0.00	0.00	0.86	0.29	N
2000	0.00	0.00	0.88	0.33	N
3000	0.00	0.00	0.89	0.36	N
4000	0.00	0.00	0.90	0.39	N
5000	0.00	0.00	0.91	0.41	N
6000	0.00	0.00	0.92	0.42	N
7000	0.00	0.00	0.93	0.43	N
8100	1.52	0.00	0.94	0.44	N
9000	1.55	0.00	0.94	0.44	N
9600	1.57	0.00	0.95	0.45	N
10100	1.49	0.00	0.95	0.45	N
10400	1.58	0.00	0.95	0.45	N
10700	1.60	0.00	0.95	0.45	N
10900	1.61	0.00	0.95	0.45	N
11100	1.57	0.00	0.96	0.45	N
11300	1.64	0.00	0.96	0.45	N
11500	1.48	0.00	0.96	0.45	N
11600	1.61	0.00	0.96	0.45	N

depth (ft)	mod d-ex	Est_dex	σ_{ob} (psi/ft)	Poisson	class
11800	1.54	0.00	0.96	0.45	N
12100	1.58	0.00	0.96	0.45	N
12200	1.67	0.00	0.96	0.46	N
12300	1.41	1.58	0.96	0.46	K
12700	1.27	1.59	0.97	0.46	K
12900	1.18	1.59	0.97	0.46	N
13000	1.13	1.59	0.97	0.46	K
13200	1.22	1.60	0.97	0.46	N
13400	1.12	1.60	0.97	0.46	N
13500	1.12	1.60	0.97	0.46	LC
13600	1.07	1.60	0.97	0.46	LC
13700	1.00	1.60	0.97	0.46	LC
13800	0.98	1.61	0.97	0.46	N
13900	1.00	1.61	0.97	0.46	LC
14000	0.91	1.61	0.97	0.46	LC
14200	0.93	1.61	0.97	0.46	LC
14400	0.86	1.61	0.97	0.46	LC
14600	0.80	1.62	0.98	0.47	LC
14800	0.86	1.62	0.98	0.47	LC
14900	0.80	1.62	0.98	0.47	LC
15000	0.90	1.62	0.98	0.47	LC
15200	0.82	1.63	0.98	0.47	LC
15300	0.87	1.63	0.98	0.47	LC
15400	0.92	1.63	0.98	0.47	LC
15500	0.87	1.63	0.98	0.47	N
15700	0.80	1.63	0.98	0.47	N
16200	0.80	1.64	0.98	0.47	N
16800	0.65	1.65	0.99	0.48	N

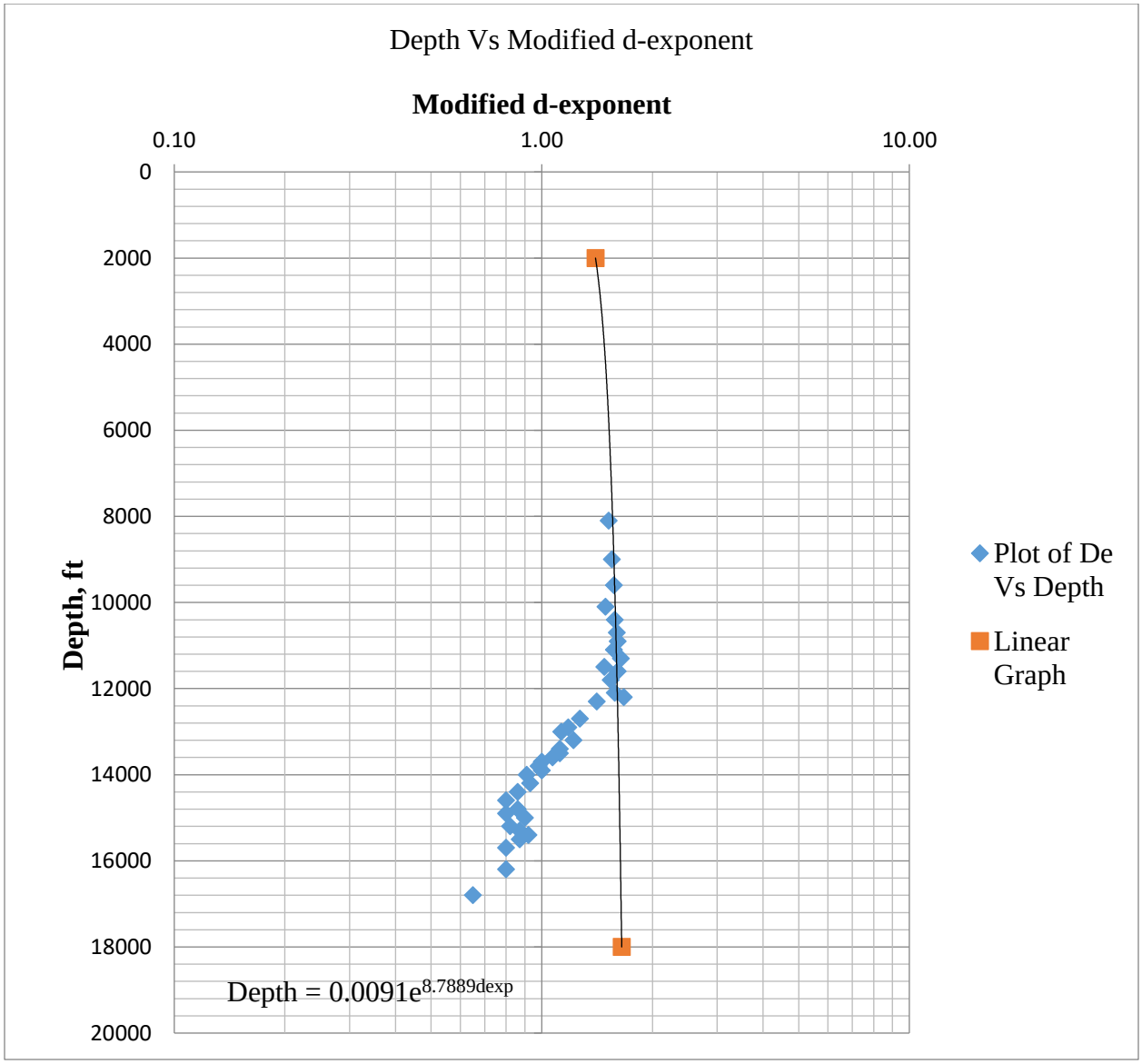


Figure 4. 1 Plot of Depth against d-exponent

It can be seen from the figure 4.1 above that d-exponent fairly increases with depth under normal pressure condition untill the depth of about 13,000ft where it decreases with depth due to abnormal pressure.

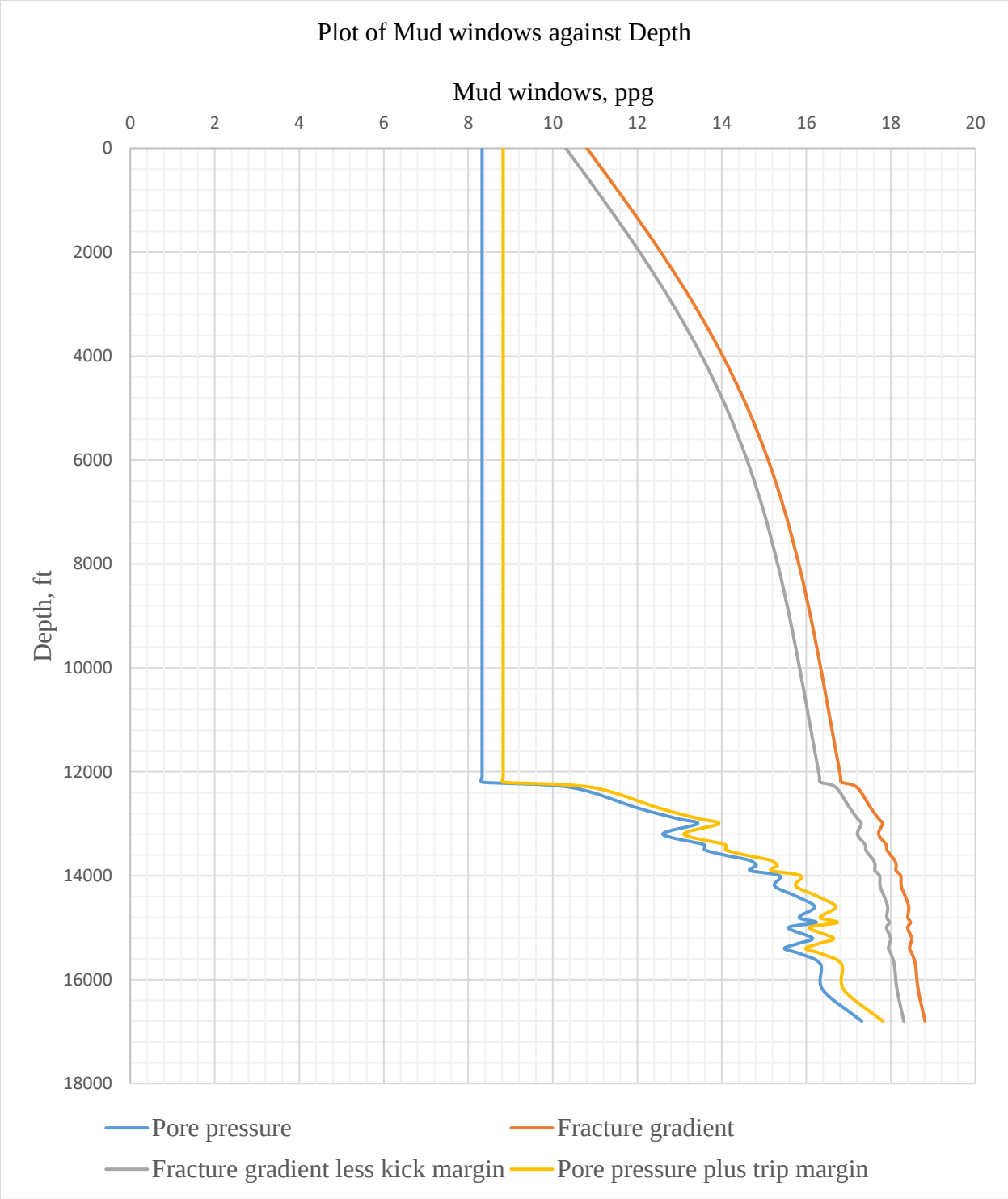


Figure 4. 2 Graph showing the variations of mud window against depth

4.2 Numerical Result of the Bayesian Model

Table 4. 3 Numerical result of Bayelsian Algorithm at a given depth

	Depth (ft)	Mod_dex	Est_dex	σ_{ob} (psi/ft)	Poisson	Class	
Training data	15700	0.80	1.63	0.98	0.47	N	
Testing data	15700	0.80	1.68	0.99	0.47	K	
Input data	15700	0.80	1.68	0.98	0.47	LC	Predicted

The confusion matrix is used to determine the validity of the model based on the values of Precision, Recall and accuracy of model.

The confusion matrix model is initialized using the zero values presented below.

Table 4. 4 Initial input values for confusion matrix model

$TP_{\_n} = 0$	$e_{\_nk} = 0$	$e_{\_nl} = 0$
$e_{\_kn} = 0$	$TP_{\_K} = 0$	$e_{\_kl} = 0$
$e_{\_ln} = 0$	$e_{\_lk} = 0$	$TP_{\_l} = 0$

After running the Algorithm with the initialized data in Table 4.4 above, following results are obtained.

Table 4. 5 Results of confusion matrix

Precision normal = 1.0	Precision Kick = 0.8	Precision lost = 0.68
Recall normal = 0.82	Recall kick = 0.67	Recall lost = 1.0
Accuracy normal = 0.89	Accuracy kick = 0.94	Accuracy lost = 0.87

The results of the confusion matrix show that the Bayesian algorithm model used for this work has capability of predicting the wellbore status with 89% accuracy for normal condition, 94% accuracy if kick is to be experienced and 87% accuracy if lost circulation is anticipated.

CHAPTER FIVE

Conclusion and Recommendation

This chapter presents the summary of conclusions inferred from the applied methodology and results obtained. It also gives the necessary suggestions for further work.

5.1 Conclusion

The following conclusions were made:

- The Bayesian algorithm developed has the ability to predict the state of the wellbore while drilling with the accuracy of 94% for the kick and 87% for lost circulation.
- The d-exponent is a better parameter for formation pressure prediction while drilling.
- The reliability of the model is dependent of the degree of normally distributed of the well data.
- Pore pressure and fracture gradient approximately define mud windows in the absence of stress data.
- The Gaussian Naives Bayes is an excellent tool to predict the probability of wellbore failure during drilling operations.
- The microchip technology for wellbore data collection will be a great tool if incorporated with electromagnetic waves for instantaneous wellbore data measurement as compared to telemetry system.
- The telemetry system is still an excellent tool for wellbore data collection.

5.2 Recommendation

- This study can further be enlarged in scope to incorporate temperature parameters in defining the mud windows so as to account for thermal effect on wellbore stability.
- Leak off test data can be used in place of fracture gradient predicted by Ben Eaton.
- Minimum horizontal stress or effective stress can be used together with the pore pressure to determine the minimum mud window.
- In the future, it can be extended to risk evaluation during tripping, cementing, etc.

Nomenclature

d-exp	d-exponent
mod-dex	modified d-exponent
est_dex	estimated d-exponent
$TP_{\_n}$	true positive of normal
$TP_{\_k}$	true positive of kick
$TP_{\_l}$	true positive of lost circulation
$FN_{\_n}$	false negative of normal
$FN_{\_k}$	false negative of kick
$FN_{\_}$	false negative of lost circulation
$e_{\_kl}$	error excluding true positive of normal
$e_{\_kn}$	error excluding true positive of lost circulation
$e_{\_ln}$	error excluding true positive of kick
K	kick
L or LC	lost circulation
N	normal state of the well
FP	formation pore pressure

Greek symbols

ϕ	porosity in fraction or %
ρ	density in ppg
γ	Poisson
$\boldsymbol{\sigma}$	overburden in psi/ft

Subscripts

exp	Exponent
mod	modified
est	estimate
ob	overburden
sw	sea water
w	water
g	gas
s	sea
fl	fluid density
o	surface

Reference

Jr. Adam T. Bourgoyne, Keith K. Millheim, Martin E. Chenevert, Jr. F. S. Young. (1991). *Applied Drilling Engineering*. Society of Petroleum Engineers (SPE textbook series 2)

Bybee, K. (2015). High-Speed Wired-Drillstring Telemetry. In *Journal of Petroleum Technology* (Vol. 60). https://doi.org/10.2118/1208-0076-JPT

Cayeux, E., Daireaux, B., Karimi Balov, M., Haavardstein, S., Magne Stokland, L., & Saasen, A. (2016). *Automatic Performance Analysis and Estimation of Risk Level Embedded in Drilling Operation Plans*. (1999). https://doi.org/10.2118/181018-ms

Chen, G., & Ewy, R. T. (2006). Thermoporoelastic Effect on Wellbore Stability. *SPE Journal*, *10*(02), 121–129. https://doi.org/10.2118/89039-pa

Dosunmu, A., & Bose, E. E. (2011). *Diagnosing Wellbore Failures What Are We Doing Wrong?* (August). https://doi.org/10.2118/140722-ms

Hughes, B. (1997). *Baker Huges INTEQ's guide to measurement while drilling information guide*. (September). Retrieved from http://elearning.utm.my/17182/pluginfile.php/427700/mod_resource/content/1/Guide to MWD - Baker Hughes.pdf

Johnson, A., Sundaramoorthy, S., Piccolo, B., & Fraczek, J. (2018). *Real Time Condition Monitoring of the Wellbore Seal through Hydraulic Fluid Analysis Using an Active Wellbore Sealing System during Managed Pressure Drilling*. https://doi.org/10.4043/28436-ms

Jr., I. N. de A., Antunes, P. D., Gonzalez, F. O. C., Yamachita, R. A., Nascimento, A., &

Goncalves, J. L. (2015). A Review of Telemetry Data Transmission in Unconventional Petroleum Environments Focused on Information Density and Reliability. *Journal of Software Engineering and Applications*, *08*(09), 455–462. https://doi.org/10.4236/jsea.2015.89043

Mining, T., Mining, D., Belief, B., & Han, J. (2012). *Naïve Bayesian Classifier Learn more about Naïve Bayesian Classifier Classification*.

Mwachaka, S. M., Wu, A., & Fu, Q. (2019). A review of mud pulse telemetry signal impairments modeling and suppression methods. *Journal of Petroleum Exploration and Production Technology*, *9*(1), 779–792. https://doi.org/10.1007/s13202-018-0483-y

Nganga, S. I., & Company, G. (2018). *Wellbore Stability Analysis in Geothermal Well Drilling*. 1–7.

Shajari, M., & Najibi, S. H. (2015). *Petroleum Science and Technology Application of the dc-Exponent Method for Abnormal Pressure Detection in Ahwaz Oil Field : A Comparative Study*. (February 2012). https://doi.org/10.1080/10916466.2010.483439

Shi, Z., Chen, Y., Yu, M., Zhou, S., & Al-Khanferi, N. (2015). *Development and Field Evaluation of a Distributed Microchip Downhole Measurement System*. https://doi.org/10.2118/173435-ms

Skalle, P., Aamodt, A., & Swahn, I. (2016). *Detection of Failures and Interpretation of Causes During Drilling Operations*. https://doi.org/10.2118/183022-ms

Solano, Y. P., Uribe, R., & Frydman, M. (2007). *Predict Pore Pressure Using the D Exponent Method : An Example From The Carbonera*. *3*, 103–111.

Soroush, H., & Tool, W. O. (2012). *IPTC 14141 Bayesian Algorithm Opens Way to Wellbore Stability*. (February), 7–9. https://doi.org/10.2523/IPTC-14141-MS

Stunes, S. (2012). *Methods of Pore Pressure Detection from Real-time Drilling Data*. (June).

Udegbunam, J. E. (2014). *Improved Well Design with Risk and Uncertainty Analysis by*.

Wessling, S., Bartetzko, A., Pei, J., & Dahl, T. (2014). *Automation in Wellbore Stability Workflows*. (March), 27–29. https://doi.org/10.2118/149766-ms

Yan, C., Deng, J., Yu, B., Li, W., Chen, Z., Hu, L., & Li, Y. (2014). Borehole Stability in High-Temperature Formations. *Rock Mechanics and Rock Engineering*, *47*(6), 2199–2209. https://doi.org/10.1007/s00603-013-0496-2

Yu, M., He, S., Chen, Y., Takach, N., LoPresti, P., Zhou, S., & Al-Khanferi, N. M. (2012). *A Distributed Microchip System for Subsurface Measurement*. (Figure 1), 1–22. https://doi.org/10.2118/159583-ms

Zhai, Z., & S. Abou-Sayed, A. (2011). *Thermal-Poro Plastic Stress Effect on Stress Reorientation in Production and Injection Wells*. https://doi.org/10.2118/140948-MS

Appendix A.

The table used to generate input data for the Bayesian algorithm

D* 100	Mod d-ex	Est d-ex	σ_{ob}	FP	FP	FP+s afety	γ	Frac	Frac	Frac_ less kick	Mud Weig ht	Cl as s
ft			psi/ft	psi/ft	ppg	ppg		psi/ft	ppg	ppg	ppg	
0	0.00	0.00	0.85	0.43	8.33	8.83	0.24	0.56	10.81	10.31	9.00	N
10	0.00	0.00	0.86	0.43	8.33	8.83	0.29	0.61	11.70	11.20	9.00	N
20	0.00	0.00	0.88	0.43	8.33	8.83	0.33	0.65	12.56	12.06	9.00	N
30	0.00	0.00	0.89	0.43	8.33	8.83	0.36	0.69	13.34	12.84	9.00	N
40	0.00	0.00	0.90	0.43	8.33	8.83	0.39	0.73	14.03	13.53	9.00	N
50	0.00	0.00	0.91	0.43	8.33	8.83	0.41	0.76	14.61	14.11	9.50	N
60	0.00	0.00	0.92	0.43	8.33	8.83	0.42	0.79	15.10	14.60	9.50	N
70	0.00	0.00	0.93	0.43	8.33	8.83	0.43	0.81	15.50	15.00	9.50	N
81	1.52	0.00	0.94	0.43	8.33	8.83	0.44	0.82	15.85	15.35	9.50	N
90	1.55	0.00	0.94	0.43	8.33	8.83	0.44	0.84	16.10	15.60	9.50	N
96	1.57	0.00	0.95	0.43	8.33	8.83	0.45	0.84	16.25	15.75	9.50	N
101	1.49	0.00	0.95	0.43	8.33	8.83	0.45	0.85	16.36	15.86	9.50	N
104	1.58	0.00	0.95	0.43	8.33	8.83	0.45	0.85	16.43	15.93	9.50	N
107	1.60	0.00	0.95	0.43	8.33	8.83	0.45	0.86	16.50	16.00	9.50	N
109	1.61	0.00	0.95	0.43	8.33	8.83	0.45	0.86	16.55	16.05	9.50	N
111	1.57	0.00	0.96	0.43	8.33	8.83	0.45	0.86	16.59	16.09	9.50	N
113	1.64	0.00	0.96	0.43	8.33	8.83	0.45	0.86	16.63	16.13	9.50	N
115	1.48	0.00	0.96	0.43	8.33	8.83	0.45	0.87	16.68	16.18	9.50	N
116	1.61	0.00	0.96	0.43	8.33	8.83	0.45	0.87	16.70	16.20	9.50	N
122	1.67	0.00	0.96	0.43	8.33	8.83	0.46	0.88	16.84	16.34	9.50	N
123	1.41	1.58	0.96	0.54	10.43	10.93	0.46	0.89	17.20	16.70	10.20	K
127	1.27	1.59	0.97	0.62	12.02	12.52	0.46	0.91	17.53	17.03	11.40	K

D* 100	Mod d-ex	Est d-ex	σ_{ob}	FP	FP	FP+safety	γ	Frac	Frac	Frac_less kick	Mud Weight	Class
ft			psi/ft	psi/ft	ppg	ppg		psi/ft	ppg	ppg	ppg	
129	1.18	1.59	0.97	0.67	12.94	13.44	0.46	0.92	17.71	17.21	13.50	N
130	1.13	1.59	0.97	0.70	13.43	13.93	0.46	0.93	17.80	17.30	13.00	K
132	1.22	1.60	0.97	0.66	12.60	13.10	0.46	0.92	17.71	17.21	15.40	N
134	1.12	1.60	0.97	0.71	13.58	14.08	0.46	0.93	17.89	17.39	17.00	N
135	1.12	1.60	0.97	0.71	13.59	14.09	0.46	0.93	17.91	17.41	17.50	L
136	1.07	1.60	0.97	0.73	14.05	14.55	0.46	0.94	17.99	17.49	18.00	L
137	1.00	1.60	0.97	0.76	14.64	15.14	0.46	0.94	18.09	17.59	18.90	L
138	0.98	1.61	0.97	0.77	14.82	15.32	0.46	0.94	18.13	17.63	16.00	N
139	1.00	1.61	0.97	0.76	14.67	15.17	0.46	0.94	18.12	17.62	19.00	L
140	0.91	1.61	0.97	0.80	15.37	15.87	0.46	0.95	18.24	17.74	18.50	L
142	0.93	1.61	0.97	0.79	15.25	15.75	0.46	0.95	18.25	17.75	18.50	L
144	0.86	1.61	0.97	0.82	15.78	16.28	0.46	0.95	18.34	17.84	18.50	L
146	0.80	1.62	0.98	0.84	16.20	16.70	0.47	0.96	18.43	17.93	18.50	L
148	0.86	1.62	0.98	0.82	15.83	16.33	0.47	0.96	18.40	17.90	18.50	L
149	0.80	1.62	0.98	0.84	16.24	16.74	0.47	0.96	18.47	17.97	18.50	L
150	0.90	1.62	0.98	0.81	15.57	16.07	0.47	0.96	18.40	17.90	18.50	L
152	0.82	1.63	0.98	0.84	16.15	16.65	0.47	0.96	18.50	18.00	18.50	L
155	0.87	1.63	0.98	0.82	15.85	16.35	0.47	0.96	18.50	18.00	16.50	N
157	0.80	1.63	0.98	0.85	16.33	16.83	0.47	0.97	18.58	18.08	17.00	N
162	0.80	1.64	0.98	0.85	16.39	16.89	0.47	0.97	18.65	18.15	18.00	N
168	0.65	1.65	0.99	0.90	17.31	17.81	0.48	0.97	18.8	18.31	18.20	N

Appendix B

The code pages of the Bayesian algorithm

In [3]:

```
from sklearn.naive_bayes import MultinomialNB
from sklearn.naive_bayes import GaussianNB
import csv
import pandas as pd
import numpy as np
from sklearn import preprocessing



Depth     = []
Mod_dex   = []
Est_dex   = []
Sig_f     = []
Poisson   = []
Targets   = [] #these are my classes

with open('training_data_2.csv') as csv_file:
    csv_reader = csv.reader(csv_file, delimiter=',')
    line_count = 0
    for row in csv_reader:
        if line_count == 0:#skip first row
            line_count += 1
        else:
            Depth.append(float(row[0]))#append depth
            Mod_dex.append(float(row[1]))#append mod weight
            Est_dex.append(float(row[2]))
            Sig_f.append(float(row[3]))
            Poisson.append(float(row[4]))
            Targets.append(row[5])#append the class/target
            line_count += 1
    print(f'Processed {line_count} lines.')



df           = np.array(list(zip(Depth,Mod_dex,Est_dex,Sig_f,Poi
sson)))
labelEncoder = preprocessing.LabelEncoder()
```

```
label        = labelEncoder.fit_transform(Targets)
df           = df.astype(np.float64)
```

```
Processed 48 lines.
```

In [4]:
```
#classifier = MultinomialNB()
classifier = GaussianNB()
classifier.fit(df,Targets)
```

Out[4]:
```
GaussianNB(priors=None)
```

In [9]:
```
predicted=classifier.predict([[15700,0.8,1.684170495,0.98,0.4710
22732]]) # Depth Mod_dex Est_dex Sig_f Poisson
predicted[0]
```

Out[9]:
```
'LC'
```

In [6]:

```
#confusion matrix 3 variables

tp_norm_ = 0
tp_kick_ = 0
tp_lost_ = 0

e_nk = 0
e_nl = 0

e_kn = 0
e_kl = 0

e_ln = 0
e_lk = 0


with open('test_data_2.csv') as csv_file:
    csv_reader = csv.reader(csv_file, delimiter=',')
    line_count = 0
    for row in csv_reader:
        if line_count == 0:#skip first row
            line_count += 1
        else:
            depth    = (float(row[0]))
            mod_dex  = (float(row[1]))
            est_dex  = (float(row[2]))
            sig_f    = (float(row[3]))
            poisson  = (float(row[4]))
            target   = (row[5])
            line_count += 1
            predicted = classifier.predict([[depth, mod_dex, est
_dex, sig_f, poisson]]) # Depth Mod_dex

            model_prediction =predicted[0]

            if model_prediction==target:
                if target == 'N':
                    tp_norm_=tp_norm_+1
```

```
                if target == 'K':
                    tp_kick_=tp_kick_+1
                if target == 'LC':
                    tp_lost_=tp_lost_+1

            if target == "N":
                if model_prediction == "K":
                    e_nk=e_nk+1
                if model_prediction == "LC":
                    e_nl=e_nl+1

            if target == "K":
                if model_prediction == "N":
                    e_kn=e_kn+1
                if model_prediction == "LC":
                    e_kl=e_kl+1


            if target == "LC":
                if model_prediction == "N":
                    e_ln=e_ln+1
                if model_prediction == "K":
                    e_lk=e_lk+1




FN_n = e_nk + e_nl
FN_k = e_kn + e_kl
FN_l = e_ln + e_lk

FP_n = e_kn + e_ln
FP_k = e_nk + e_lk
FP_l = e_nl + e_kl

TN_n = tp_kick_ + e_kl + e_lk + tp_lost_
TN_k = tp_norm_ + e_nl + e_ln + tp_lost_
TN_l = tp_norm_ + e_nk + e_kn + tp_kick_
```

```
#Precision = TP/(TP+ FP)
precision_n = tp_norm_/(tp_norm_ + FP_n)
precision_k = tp_kick_/(tp_kick_ + FP_k)
precision_l = tp_lost_/(tp_lost_ + FP_l)

#Recall=TP/(TP+FN)
recall_n = tp_norm_/(tp_norm_ + FN_n)
recall_k = tp_kick_/(tp_kick_ + FN_k)
recall_l = tp_lost_/(tp_lost_ + FN_l)

#Accurracy = (TN+TP)/(TN+TP + FN+FP)
accurracy_n = (TN_n + tp_norm_)/(TN_n + tp_norm_ + FN_n + FP_n)
accurracy_k = (TN_k + tp_kick_)/(TN_k + tp_kick_ + FN_k + FP_k)
accurracy_l = (TN_l + tp_lost_)/(TN_l + tp_lost_ + FN_l + FP_l)


print("Precision normal:",precision_n)
print("Precision kick:",precision_k)
print("Precision lost:",precision_l)
print("\n\n")

print("Recall normal:",recall_n)
print("Recall kick:",recall_k)
print("Recall lost:",recall_l)
print("\n\n")

print("Accurracy normal: ",accurracy_n)
print("Accurracy kick: ",accurracy_k)
print("Accurracy lost: ",accurracy_l)
```

```
Precision normal: 1.0
Precision kick: 0.8
Precision lost: 0.6842105263157895
```

```
Recall normal: 0.8214285714285714
Recall kick: 0.6666666666666666
Recall lost: 1.0



Accurracy normal:  0.8936170212765957
Accurracy kick:  0.9361702127659575
Accurracy lost:  0.8723404255319149
```

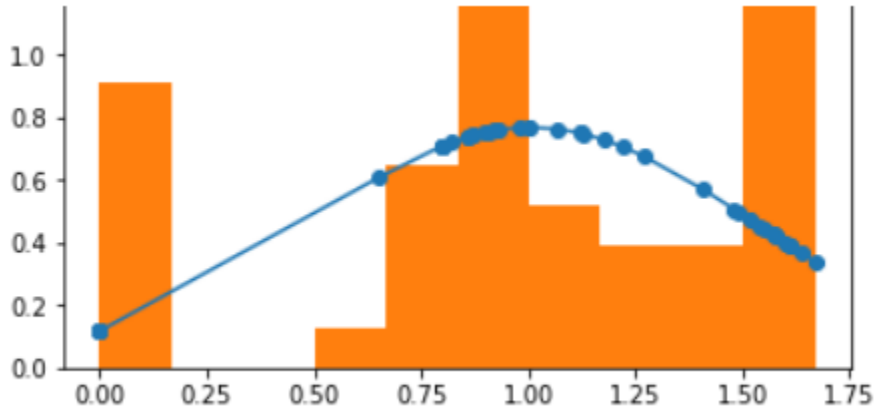
In [12]:

```
#plotting graphs

'''
Depth     = []
Mod_dex   = []
Est_dex   = []
Sig_f     = []
Poisson   = []
Targets   = []

'''
import matplotlib.pyplot as plt
```

In [48]:

```
import numpy as np
import scipy.stats as stats
import pylab as pl

Mod_dex = sorted([0,0,0,0,0,0,0,1.52,1.55,1.57,1.49,1.58,1.6,1.6
1,1.57,1.64,1.48,1.61,1.54,1.58,1.67,1.41,1.27,1.18,1.13,1.22,1.
12,1.12,1.07,1,0.98,1,0.91,0.93,0.86,0.8,0.86,0.8,0.9,0.82,0.87,
0.92,0.87,0.8,0.8,0.65
])  #sorted

fit = stats.norm.pdf(Mod_dex, np.mean(Mod_dex), np.std(Mod_dex))
#this is a fitting indeed

pl.plot(Mod_dex,fit,'-o')

pl.hist(Mod_dex,normed=True)      #use this to draw histogram of
your data

pl.show()                      #use may also need add this
```

```
C:\Users\Emmanuel\Anaconda3\lib\site-packages\matplotlib\axes
\_axes.py:6571: UserWarning: The 'normed' kwarg is deprecate
d, and has been replaced by the 'density' kwarg.
  warnings.warn("The 'normed' kwarg is deprecated, and has be
en "
```



In [4]:

```
import numpy as np
import scipy.stats as stats
import pylab as pl

Sig_f = sorted([0.85,0.86,0.88,0.89,0.90,0.91,0.92,0.93,0.94,0.9
4,0.95,0.95,0.95,0.95,0.95,
                0.96,0.96,0.96,0.96,0.96,0.96,0.96,0.96,0.97,0.97
,0.97,0.97,0.97,0.97,0.97,
                0.97,0.97,0.97,0.97,0.97,0.97,0.98,0.98,0.98,0.98
,0.98,0.98,0.98,0.98,0.98,
                0.98,0.99])  #sorted

fit = stats.norm.pdf(Sig_f , np.mean(Sig_f ), np.std(Sig_f ))  #
this is a fitting indeed

pl.plot(Sig_f ,fit,'-o')

pl.hist(Sig_f ,normed=True)      #use this to draw histogram of
 your data

pl.show()                    #use may also need add this
```

```
C:\Users\Emmanuel\Anaconda3\lib\site-packages\matplotlib\axes
\_axes.py:6571: UserWarning: The 'normed' kwarg is deprecate
d, and has been replaced by the 'density' kwarg.
  warnings.warn("The 'normed' kwarg is deprecated, and has be
en "
```

In [76]:

```
import numpy as np
import scipy.stats as stats
import pylab as pl

Poisson = sorted ([0.24,0.29,0.33,0.36,0.39,0.41,0.42,0.43,0.44,
0.44,0.45,0.45,0.45,0.45,0.45,0.45,
                   0.45,0.45,0.45,0.45,0.45,0.46,0.46,0.46,0.46,
0.46,0.46,0.46,0.46,0.46,0.46,0.46,
                   0.46,0.46,0.46,0.46,0.47,0.47,0.47,0.47,0.47,
0.47,0.47,0.47,0.47,0.47,0.48])  #sorted

fit = stats.norm.pdf(Poisson  , np.mean(Poisson  ), np.std(Poiss
on  ))  #this is a fitting indeed

pl.plot(Poisson  ,fit,'-o')

pl.hist(Poisson  ,normed=True) #use this to draw histogram of yo
ur data

pl.show()                      #use may also need add this
```