AN AUTOMATED DRILLING GEOMECHANICS SIMULATOR USING

MACHINE-LEARNING ASSISTED ELASTO-PLASTIC FINITE ELEMENT MODEL

A Dissertation

by

HUSSAIN IBRAHIM H ALBAHRANI

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Nobuo Morita
Committee Members,	Theofanis Strouboulis
	Fred Dupriest
	Sam Noynaert
Head of Department,	Jeff Spath

December 2020

Major Subject: Petroleum Engineering

ABSTRACT

The petroleum industry has long relied on pre-drilling geomechanics models to generate static representation of the mud weight limits. These models rely on simplifying assumptions such as linear elasticity, a uniform wellbore shape, and generalized failure criteria to predict failure and determine a safe mud weight, which lead to inaccurate results. Thus, this study's main objective is to improve the process for predicting the wellbore rock failure while drilling.

Wellbore failure prediction is improved through the use of a novel modelling scheme that involves an elastoplastic finite-element model (FEM), machine learning (ML) algorithms, as well as imaging data that accurately describes rock failure. Geomechanics data are modelled in the FEM code, which are then used to train the ML algorithms in conjunction with imaging data. The produced integrated model of FEM and ML is used to predict failure limits for new wells. This improved failure prediction can be updated with the occurrence of different drilling events such as induced fractures and wellbore enlargements.

The integrated modelling scheme was first applied to lab experimental results to provide a proof-of-concept and validation. This application showed improvement in rock-failure prediction when compared to conventional failure criteria such as Mohr-Coulomb. Also, offset-well data from wireline logging and drilling records are used to train and build a field-based integrated model, which is then used to showcase different field applications of the model.

The field applications presented exhibit the advantages of this modelling scheme where integrating a physics model such as the FEM with a ML algorithm can significantly improve failure prediction while requiring significantly smaller datasets for training than those required for purely data driven ML models. The applications also show that the integration process lead to a better understanding of how failure takes place, which can be used to re-assess the mathematical formulation of the failure surface. From a practical perspective, the improvement in failure prediction can help avoid non-productive time (NPT) events such as wellbore enlargements, hole cleaning issues, pack-offs, stuck-pipe, and lost circulation. This efficiency is to be achieved by a real-time implementation of the model and by taking advantage of available data that are not routinely utilized by drilling.

DEDICATION

I dedicate this work to my parents who taught me the value of hard work and honesty. I also dedicate this to Sara and Ward as you're my driving force.

ACKNOWLEDGEMENTS

I would like to thank my advisor, Professor Nobuo Morita for all his support and dedication in helping me complete this work and for serving as the chair of my committee. He gave me the space to try out new ideas no matter how outlandish they seemed at the beginning. He was also always generous with his time and available to guide me through the rough patches I encountered throughout the course of this study, and he did all of this with kindness and a smile. I will always be proud, grateful, and honored having worked under him. I also would like to thank my committee members, Professor Dupriest, whose class "High Performance Drilling Practices" ignited my interest in this area of research seven years prior and inspired many of the new ideas used in this work, Professor Noynaert, who was the chairman of my MS thesis committee who supported me to launch my career in research, and Professor Strouboulis, whose class "Introduction to the Finite Element Method" gave me the tools to embark on this journey with much more confidence. I also would like to thank my committee members for their guidance, positivity, and support throughout the course of this research.

I would like to take this chance to thank my employer and sponsoring company Saudi Aramco, especially the management of EXPEC ARC for giving me this invaluable opportunity to learn and improve.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supervised by a dissertation committee consisting of Professor Nobuo Morita, Professor Fred Dupriest, and Professor Sam Noynaert of the Department of Petroleum Engineering and Professor Theofanis Strouboulis of the Department of Aerospace Engineering.

Experimental work was completed and previously partially published by the Koki Tomozawa of Waseda University, Japan and by L.E. Walle, A.N. Berntsen, E. Papamichos, and D. Szewczyk of SINTEF, Norway.

All the coding, data processing, and modelling work conducted for the dissertation was completed by the student Independently.

NOMENCLATURE

AUROC	Area under the receiver operating characteristic curve
a_0, a_1	Constants describing linear formulation of the Drucker-Prager failure criterion
B	Strain-displacement matrix
B/O	Wellbore breakouts or enlargement
b_0, b_1	Constants describing linear formulation of the Mogi failure criterion
CAL_{max}, CAL_{min}	Maximum and minimum caliper log measurements in an interval
CFO	Curve face orientation angle with respect to minimum stress direction
CV	Cross validation
D	Consistent tangent matrix
E	Young's modulus
f	Function of the failure surface in a failure criterion
F	Body force vector
FEM	Finite element method or model
FEM-ML	Integrated finite element method and machine learning model
FMI/FMS	Formation micro-imager or micro-scanner tool
F^e	Elemental body force vector
FPR	False positive rate

ft	Feet
g	Function of the failure surface in the Lade failure criterion
$h_{W,b}(x)$	Output perceptron function for failure classification in neural networks
in.	Inches
J	Jacobian matrix
J_2	Second deviatoric stress invariant
J_3	Third deviatoric stress invariant
K	Global stiffness matrix
k	A constant in the Lade criterion
K^e	Elemental stiffness matrix
K_r	Ratio of minimum horizontal load to maximum horizontal load in true-axial experiments
K_z	Ratio of vertical load to maximum horizontal load in true-axial experiments
LOT	Leakoff test
LWD	Logging while drilling
MDT	Modular formation dynamic tester
ML	Machine Learning
N	Shape function
$\hat{p}_{mk}$	Probability that a mesh node from the training dataset is in the m region of the feature space from the k failure class

ppg	Pounds per gallon
PPV	Positive predictive value
P_w	Wellbore pressure
r	A constant in the Lade criterion
RC	Radius of curvature
r_d	Magnitude of radial failure extension around the wellbore
RGB	Red, Green, Blue color number for a pixel
r_{Node}	Mesh node radial distance from the borehole center
r_w'	First derivative of the borehole curve radius
s	Sigmoid activation function for neural networks
S_{ij}	Deviatoric stress tensor
S_{11}	First component of the deviatoric stress tensor
t	Adjustment of shear stress for the Lade failure criterion
T	Traction force vector for boundary stress
T^e	Elemental traction force vector
TP, TN, FP, FN	True positive, true negative, false positive, and false negative respectively
TPR	True positive rate
U	Internally stored strain potential
u	Displacement
u^e	Nodal displacement
U_p	Potential energy

u_x, u_y, u_z	Directional displacements
v	Poisson's ratio
W	External work
x, y, z	Global coordinates in the FEM mesh
$W_{ij}^{(1)}$	Connection weights for each input cell in neural networks
z_j	Hidden layer cell values in neural networks
3D	Three dimensional

Greek

ε^e	Elastic strain
ε^p	Plastic strain
$\alpha_j^{(2)}$	Connection weights for each hidden layer cell in neural networks
$\theta_{critical}$	Critical wellbore breakout angle for risk-controlled stability
$\sigma_1, \sigma_2, \sigma_3$	Minimum, intermediate, and maximum principle stresses
$\sigma_1^e, \sigma_2^e, \sigma_3^e$	Effective principle stresses
$\sigma_V, \sigma_H, \sigma_h$	Vertical overburden, maximum horizontal, and minimum horizontal in-situ stresses
σ_{ij}	Stress tensor
σ_{ij}^e	Effective stress tensor
σ_m	Mean normal stress
σ_n	Normal stress for the Mohr-Coulomb failure criterion

$\sigma_x, \sigma_y, \sigma_z$	Normal in-situ stresses transformed into the wellbore coordinates
$\sigma_z, \sigma_r, \sigma_R$	Vertical, minimum horizontal, and maximum horizontal loads applied in true-triaxial experiments
$\bar{\tau}$	Average shear stress
τ_0	Cohesion for the Mohr-Coulomb failure criterion
$\tau_{yz}, \tau_{xz}, \tau_{xy}$	Shear in-situ stresses transformed into the wellbore coordinates
β	Angle between the maximum horizontal stress and the wellbore projection on the two horizontal stresses plane
β_{Node}	Mesh node circumferential distance to minimum stress direction in degrees
Γ, S_e	Surface integration variables
Ω, V_e	Volume integration variables
γ	Inclination of the wellbore as a deviation from the maximum in-situ stress
ε	Strain tensor
θ	Mesh node circumferential angle
λ	Plastic strain multiplier
ξ, η, ς	Local coordinates in the FEM mesh
τ	Shear stress
ψ	Yield surface
ϵ	Cross-entropy function
ϕ	Internal friction angle for the Mohr-Coulomb failure criterion

TABLE OF CONTENTS

LIST OF FIGURES

xviii

xxvi

LIST OF TABLES

CHAPTER I

INTRODUCTION

1.1 General Introduction

The drilling window can be defined as the safe mud weight range that can prevent drilling troubles. This range is basically defined by four parameters and those are the pore pressure, the threshold for wellbore rock instability or shear failure, the threshold for initiating a wellbore fracture or a tensile failure, and the threshold for initiating differential sticking problems. In routine overbalance drilling operations, the mud weights are calculated at values above the pore pressure to prevent reservoir fluids from flowing into the wellbore. After accounting for the pore pressure, consideration is then turned towards determining the mud weight range that will prevent rock failure. The mud weight for preventing wellbore instability can be either below or above the pore pressure depending on the in-situ stress regime, the rock properties, and the wellbore profile among other things. This means that in some cases, the instability mud weight can be the parameter that sets the minimum safe mud weight, which is basically the lower limit of the drilling window, while in other cases that limit is set by the pore pressure. In either case, the rock failure mud weights are usually calculated based on analytical solutions and industry popular failure criteria, such as the Mohr-Coulomb criterion, and the drilling engineer ends up with a single or a static representation of the safe mud weight range as in **Figure 1**.
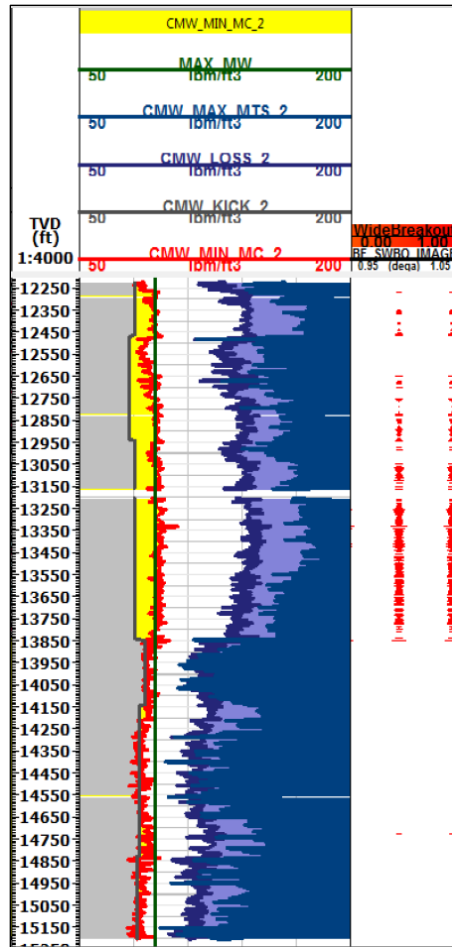
Figure 1. A pre-drilling geomechanics model showing the safe mud weight window as determined based on the popular analytical solution of the Mohr-Coulomb failure criteria (reprinted from AlBahrani and Al-Yami 2018).

The industry has long relied on pre-drilling geomechanics models to generate the static representation of the allowable mud weight limits. These models rely on simplifying assumptions such as: linear elasticity, uniform or circular wellbore shape, and generalized failure criteria such as the Mohr-Coulomb, Mogi, Drucker-Prager, and Lade failure criterion to predict failure and determine a safe mud weight. These assumptions can lead to inaccurate results, which in turn lead to the wrong recommendation regarding the safe

mud weight. Considering that the safe mud weight determines the magnitude of downhole hydrostatic pressure necessary to avoid shear and tensile failure of the wellbore rock, inaccurate determination of this value can lead to a multitude of drilling problems.

Another shortcoming in the field of drilling geomechanics modelling is the failure to take advantage of new information that becomes available as drilling progresses. Real-time or while-drilling produced data such as logging while drilling (LWD) wellbore images, shale shakers caving reports, and reports of lost circulation can be used to produce an updated description of the realistic wellbore shape along with other structural considerations. These updates can then be incorporated into the geomechanics model as they are received to produce new output and recommendation. Being able to reflect the effect of different drilling events, which are routinely occurring while drilling, on the geomechanics model will help create a dynamic representation of the allowable mud weight limits.

Failure to accurately estimate the wellbore strength and the safe mud weight, which is needed to avoid shear and tensile rock failures, routinely lead to a multitude of drilling problems. These problems include induced drilling fractures, loss of circulation, wellbore instability and enlargements, hole cleaning issues, pack-offs and tight spots, and stuck-pipe incidents. These problems are some of the major causes of non-productive time (NPT) in drilling. Also, incorporating the use of real-time data or while-drilling information into the modelling scheme makes way for generating drilling recommendations in an automated environment by taking advantage of available data that are not previously routinely utilized by drilling engineers.

1.2 Literature Review

1.2.1 The Use of Finite Element Modelling in Drilling Geomechanics

The use of the finite element method for a well-based analysis is common especially through the use of commercial softwares such as ABAQUS and ANSYS. There are numerous studies published which explore similar topics to the one considered here. Feng et al. (2015) employs the FEM to study the effectiveness of different wellbore strengthening techniques by introducing a static fracture to a symmetric wellbore. Salehi and Nygaard (2010) performed similar work with similar objective where they also introduced a static fracture to a symmetric wellbore for wellbore strengthening evaluation. The use of elasto-plastic FEM is well established in drilling geomechanics as seen in several sources (Zervos et al. 1998, Meng et al. 2010, Huang 2016).

1.2.2 Real-time Applications of Drilling Geomechanics

Application of geomechanics analysis while drilling or in real-time has become a popular topic in recent years. The overwhelming majority rely on the use of Mechanical Earth Models (MEM) to achieve this application. The output from MEM is commonly based on correlations that estimate the linear elasticity parameters and analytical solutions for stress determination and failure criteria. An example of this is where the sonic measurement from an LWD tool is used to update the linear elasticity parameters and stress estimation in MEM and provide new estimation of wellbore strength and mud weights based on analytical solution (Hamid et al. 2016). Similar approach is followed in several other sources (Magini et al. 2012, Bagala et al. 2010, Mylnikov et al. 2018). There

are no references discussing the application of geomechanics surveillance while drilling using numerical models (as opposed to analytical solutions). Most published work in this area discuss the use of a numerical models such as an elasto-plastic FEM to calculate the safe mud weight value in the pre-drilling planning phase. One example of this is done by Meng et al. (2010) where an elasto-plastic FEM is used to calculate the mud weight recommendation in the pre-drilling phase and then it uses the drilling experience to update the model input to produce better results for future wells. One more limitation in the application of the FEM in this reference is that, even in the pre-drilling phase, the model output doesn't consider structural issues such as the presence of micro-fractures. There are no references that discuss updating the numerical model and producing an updated output based on drilling events and structural considerations in real-time.

1.2.3 Modelling A Non-uniform and Realistic Wellbore Shape

There are several publications that present different approaches to modelling a non-circular wellbore. Cecílio et al. (2019) provides a finite element formulation for wellbore stability analysis using the DiMaggio-Sandler constitutive elastoplastic model. In their work, they employ an adaptive meshing approach to continuously define progressive failure while transferring elastoplastic deformation history. Zhang et al. (2011) model non-circular, but uniform, wellbore shapes to examine their effect on hydraulic fractures development. The wellbore shapes considered in Zhang's work vary from circular, elliptical, to circular with two identical and uniform enlargements on opposite directions of the wellbore. The same work previously by Feng et al. (2015) also show the results of an extended finite element model (XFEM) on fracture progression in

an imperfect wellbore. It illustrates the effect of the rugosity of an imperfect wellbore on the hoop stress and on the location of the developed fractures. Kanfar et al. (2015) use a 2D generalized plane strain FEM to model the time dependent shear failure and predict the resulting wellbore shape after failure while avoiding re-meshing for the resulting shape or progressive failure as was done by Cecílio et al. (2019) to save on computation time. Similar approach is also presented by Gaede et al. (2011) where re-meshing for the resulting wellbore in the 3D elastic FEM is avoided through the use of a hybrid analytical-numerical workflow. Cheng et al. (2019) presents the results of simulating progressive wellbore failure using the boundary element method where, in similar fashion to Cecílio et al. (2019), the progressive failure is represented by the removal of failed elements resulting in a circular wellbore with two identical and uniform enlargements on opposite directions. Morita and Ross (1993) present the results of a FEM based analysis of wellbore stability in horizontal elliptical uniformly-shaped wellbores. The effect of both fractures and instability is studied by Chen, Tan, and Detournay in three separate publications (Chen, Tan, and Detournay 2001, 2002, and 2003) where they analyzed the effect of drilling fluid filtration on wellbore stability in a naturally fractured rock rather than drilling induced fractures. As per the knowledge gained from this literature review, there's no prior published work that considers a realistic, completely non-uniform, and field image-based wellbore shape for drilling window modelling and calculations.

1.2.4 The Use of Image-based Analysis

The use of field image logs for deducing information relevant to geomechanical modelling, defining a realistic wellbore shape, and providing real-time type of analysis is

discussed in several publications. Barton et al. (1997) present an approach to use image logs to estimate the components of the stress tensor and the constraint of in-situ rock strength where this data can then be used to create a model for evaluating wellbore stability. Wessling et al. (2011) illustrated an automated algorithm to analyzing image logs to provide a real-time based update on the state of the wellbore. The image log analysis in this reference identifies alteration to the wellbore status such as the occurrence of enlargements. Other forms of image analysis and processing can also be used to identify and classify cavings in the shale shakers as shown by Skea et al. (2018). This reference uses a post-drilling image analysis technique to identify the presence of cavings from samples from the shale shakers, then proposes a cavings classification method, which is then used to determine the prevailing mode of failure.

1.2.5 The Incorporation of Machine Learning Algorithms in Rock Failure Prediction

Applications of machine learning in both drilling and geomechanics have recently become a vastly expanding area of research. Zhang et al. (2018) uses the output from a FEM to train Artificial Neural Networks (ANN) on the prediction of the maximum horizontal in-situ stress. Alternatively, the use of ANN is expanded to be trained based on the output of analytical solutions to produce the drilling allowable mud weight limits in (Phan et al. 2020) where the objective is to improve computational speed. Other references use drilling information along with LWD or wireline logs in a data-driven approach to predict rock properties and location of wellbore enlargements respectively (Kumar et al. 2012, Soroush et al. 2010). To summarize, the general objectives of employing machine learning algorithms in geomechanics applications are to:

1. Improve the estimation of parameters that are conventionally estimated from analytical solutions, numerical solution, lab measurements, or field measurements

2. Create the ability to estimate parameters that can't be estimated using available solutions or measurements

3. Facilitate the application of a model or a solution in a practical setting

4. Improve the computational speed of the available solutions

As per the literature review performed for this work, there's no published work that integrates the use of the FEM with a machine learning algorithm for the purpose of improved rock failure prediction. As mentioned previously, published work in this area is either purely data driven machine learning models or surrogate machine learning models. No source discusses the process of integrating a physics model such as the FEM with machine learning algorithms to improve rock failure prediction. However, the concept of integrating physics models with machine learning is an emerging area of research in other fields, where it is referred to as physics-based, theory guided data science, or hybrid modelling (Karpatne et al. 2017). There are two main advantages that the physics-based approach has over the purely data driven approach. First, in purely data driven models, the algorithm has to be trained on a large dataset that must contain all possible combinations. This means that if the machine learning algorithm was not trained or has not seen a specific combination of data, it will not be able to predict it. The second advantage, as illustrated by Karpatne et al. (2017), the physics-based approach allows for gaining more insight and even for scientific discovery by highlighting the underlying factors that led to the accurate

predictions made by the machine learning algorithm. The integration process and the illustration of their advantages are discussed more thoroughly in the upcoming chapters.

1.2.6 Conventional Failure Criteria

In rock mechanics and geotechnical engineering there are several published models that set a criterion for rock and soil failure. In general terms, these models rely on lab testing to define the failure envelope of rocks. The failure envelope is usually defined at what is known as strength parameters, which are specific to each developed model, and at certain limits of shear and normal stresses as observed in lab testing. These models suggest that rock failure in compression will take place if the stress state of the rock at the specific strength parameters falls above the defined failure envelope.

In petroleum engineering, rock core samples from a specific geological formation are excavated from the wellbore while drilling for the purpose of performing lab tests. Different lab tests, such as the triaxial loading, are performed on the core sample to determine the best fitting failure criterion and the values of strength parameters pertaining to the said criterion. Having determined the best suitable failure criterion and the failure envelop of this particular rock type, this information can then be used in different phases of the engineering design. The most common use of this is in drilling engineering where a failure criterion is used to calculate the minimum mud weight required to prevent wellbore instability due to shear failure of the exposed wellbore wall rock while drilling. Other failure criteria are also used to determine the limit of tensile failure of the rock, which defines the fracture gradient of the wellbore. Putting all of this together produces what is known as the drilling window, which can be defined as the limits of the safe and

allowable hydrostatic pressure exerted by the drilling fluid on the wellbore. The instability mud weight and the pore pressure are placed at the minimum limit of the drilling window and the fracture gradient is placed at the maximum limit. This procedure has been the conventional method to pre-determine the allowable mud weight in the pre-drilling design phase. Below is a review of the most common models used in the industry. It should be noted that in all of the following expressions, the convention for compressive stress is to be expressed as negative.

1.2.6.1 Mohr-Coulomb Failure Criterion

This criterion proposes that compressive failure takes place when the stress state exceeds a certain limit which is beyond the strength of the rock. The stress state is defined by the shear stress τ on a certain plane and the strength of the rock in Mohr-Coulomb is defined by the internal friction angle ϕ and cohesive forces τ_0 (Al-Ajmi and Zimmerman 2005). The criterion was initially defined by Coulomb as:

$$\tau = \tau_0 + \sigma_n \tan(\phi) \tag{1}$$

If compressive stresses are expressed as negative stresses, the criterion can also be expressed in terms of the principle stresses σ_1 and σ_3 instead of the normal stress σ_n as follows as shown in (Morita and Nagano et al. 2016):

$$\bar{\tau} = \frac{1}{2} \frac{\sigma_1 - \sigma_3}{\cos(\phi)} \tag{2}$$

$$\sigma_m = \frac{\sigma_1 + \sigma_3}{2} \tag{3}$$

$$Failure\ occurs\ if : \bar{\tau} > (\tau_0 - \sigma_m \tan(\phi)) \tag{4}$$

As can be seen from Equations (3), the value of σ_2 has no role in influencing the failure envelop in the Mohr-Coulomb criterion because it is satisfied on a specific plane that is in the direction of the intermediate principle stress σ_2. This means that this criterion does not consider the strengthening effect of σ_2, which leads to a conservatively higher estimates of the mud weight required for preventing wellbore instability when compared to other failure criteria. In other words, the Mohr-Coulomb failure criterion generally predict failure to take place earlier than other failure criteria that do consider the strengthening effect of σ_2.

1.2.6.2 Mogi Failure Criterion

To be able to examine the effect of the intermediate principle stress σ_2, true-triaxial tests were conducted by Mogi. He concluded that σ_2 does indeed have a strengthening effect, however, he also concluded that since the brittle fracture takes place in a plane that is in the σ_2 direction, the mean normal stress σ_m that works against the formation of the fracture is similar to the one expressed in the Mohr-Coulomb criterion in Equation (3), which does not consider σ_2 (Mogi 1971). Based on this, the Mogi failure criterion is expressed as follows:

$$\tau = \sqrt{\frac{1}{2} S_{ij} S_{ij}} \tag{5}$$

$$S_{ij} = \sigma_{ij} - \sigma_m \tag{6}$$

$$\sigma_m = \frac{\sigma_1 + \sigma_3}{2} \tag{7}$$

$$\text{Failure occurs if}: \tau > f\left(\frac{\sigma_1 + \sigma_3}{2}\right) \tag{8}$$

Where: σ_{ij} is the stress tensor, S_{ij} is the deviator stress tensor, and σ_m is the mean stress. The function in Equation (8) can be determined experimentally. For a linear formulation of this failure criteria, the following function can be used with the constants b_0 and b_1 being determined experimentally as material properties:

$$f = b_0 + b_1 \left(\frac{\sigma_1 + \sigma_3}{2}\right) \tag{9}$$

This f function defining the failure envelope of the Mogi failure is generally a monotonically increasing function. Using this type of function in the failure criterion emphasizes that failure takes place when the distortional strain energy is at a certain value that is increasing monotonically with $\left(\frac{\sigma_1 + \sigma_3}{2}\right)$ (Jaeger et al. 2007). An example of a common formulation of the f function in the Mogi failure criterion is the use of power law functions as shown in (Al-Ajmi and Zimmerman 2005).

Equation (5) is also known as the octahedral shear stress and can be expressed more conveniently in terms of the principle stresses as follows:

$$\tau = \frac{1}{3}\sqrt{(\sigma_1 - \sigma_2)^2 + (\sigma_2 - \sigma_3)^2 + (\sigma_3 - \sigma_1)^2} \tag{10}$$

1.2.6.3 Drucker-Prager and Lade Failure Criteria

Although this failure criterion was originally developed to evaluate the strength of soils, it is able to examine the strengthening effect of the intermediate principle stress σ_2 more prominently (Drucker and Prager 1952). The main difference here is that in this criterion, σ_2 is considered when calculating the mean normal stress σ_m:

12

$$\sigma_m = \frac{\sigma_1 + \sigma_2 + \sigma_3}{3} \tag{11}$$

This criterion uses the same expression for the octahedral shear stress τ as the one in the Mogi failure criterion in Equations (5) and (10). It also uses the same expression for the deviatoric stress tensor S_{ij} in Equation (6). As for the function of the failure criteria f, it is now a function of $\left(\frac{\sigma_1 + \sigma_2 + \sigma_3}{3}\right)$:

$$Failure\ occurs\ if : \ \tau > f\left(\frac{\sigma_1 + \sigma_2 + \sigma_3}{3}\right) \tag{12}$$

This function can be linearly expressed as:

$$f = a_0 + a_1 \left(\frac{\sigma_1 + \sigma_2 + \sigma_3}{3}\right) \tag{13}$$

Where constants a_0 and a_1 are determined experimentally as material properties.

In an opposite fashion to the Mohr-Coulomb failure criterion, the Drucker-Prager criterion is less conservative in the sense that it can over-estimate the strengthening effect of σ_2, which might lead to it predicting an over-optimistically low mud weight for preventing wellbore failure. This means that the Drucker-Prager criterion will generally predict failure to take place later than most other criteria, especially the Mohr-Coulomb one.

There's another form of the Drucker-Prager failure criterion where it was modified by Lade as detailed in (Morita and Nagano 2016). This Lade-modified Drucker-Prager failure criterion is expressed by the following relationships:

$$J_3 = S_{11}S_{22}S_{33} + 2S_{12}S_{23}S_{31} - S_{11}S_{23}^2 - S_{22}S_{13}^2 - S_{33}S_{12}^2 \tag{14}$$

$$r = \left(\frac{27}{2}|J_3|\right)^{\frac{1}{3}} \tag{15}$$

$$t = \frac{1}{2}\tau[1 + \frac{1}{k} - \left(1 - \frac{1}{k}\right)\left(\frac{r}{\sqrt{3}\tau}\right)^3] \tag{16}$$

Where J_3 is the third deviatoric stress invariant and S are the components of the deviator stress tensor. This criterion uses the same expression for the octahedral shear stress τ as the one in the Mogi failure criterion in Equations (5) and (10). It also uses the same expression for the deviatoric stress tensor S_{ij} in Equation (6). The failure surface t is a function of the mean stress σ_m:

$$t = g(\frac{\sigma_1 + \sigma_2 + \sigma_3}{3}) \tag{17}$$

As discussed in (Morita and Nagano 2016), while the original Lade model was developed for weak sands, when it is modified in the manner shown above and k is set at ≥ 0.9, it can be used to model strong and consolidated sands. It should be noted that if the k value is set at 1, this formulation becomes equivalent to the original Drucker-Prager formulation.

1.2.6.4 Implementations of the Conventional Failure Criteria in Finite Element

Modelling

The failure criteria discussed here can be implemented with any mechanical numerical model to determine the onset of failure at each location in a structure. For drilling geomechanics and finite element modelling, this is done by first calculating the principle stresses at each integration point in the defined mesh of the structure of the wellbore and formation. The principle stresses here are calculated based on the

displacement values obtained as output from the finite element model at the integration points. The role of the failure criteria is to evaluate the stress state at each point against the strength parameters assigned to the said point to decide on the possibility of failure. For example, if the Mogi failure criterion were to be used, the principle stresses calculated from the finite element model will be used to determine the value of the octahedral shear stress τ as described in Equations (5) or (10). Next, the failure criteria function f is calculated based on the strength parameters values assigned to each point. Finally, the value obtained from the failure criteria function f is subtracted from the calculated value of the octahedral shear stress τ. If the result of this subtraction is a positive value, this means that this point lies above the failure envelope, which means this point will be predicted to fail. An illustration of this process is shown in the flow diagram in **Figure 2**.
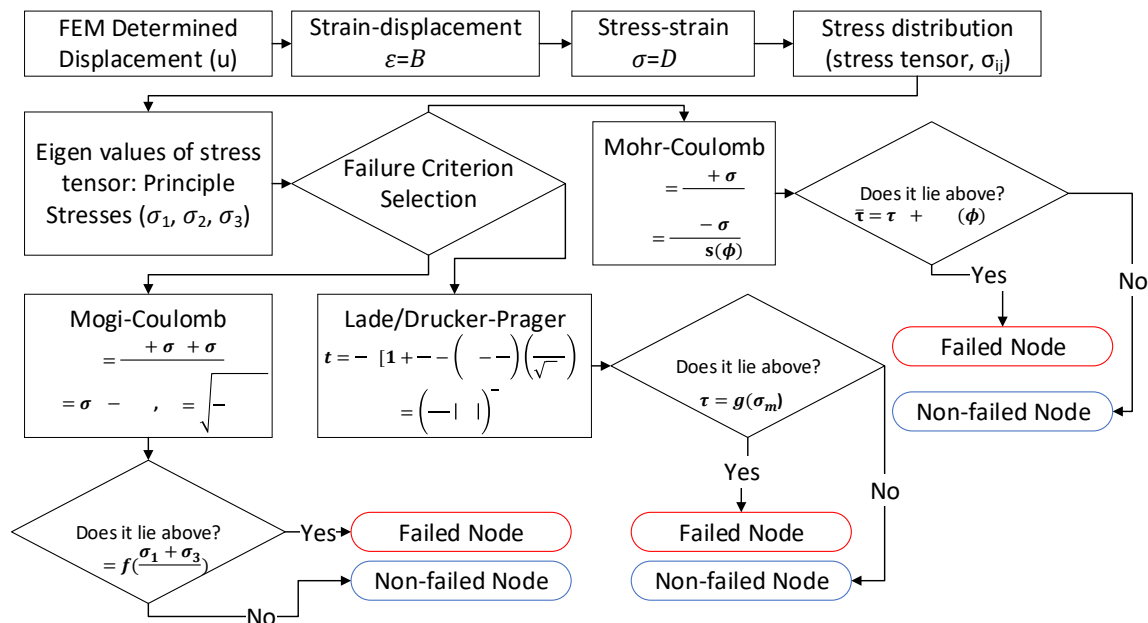
Figure 2: A flow diagram describing the process for implementing the conventional failure criteria into finite element modelling.

1.3 Research Objectives

The main objective is to improve the process for predicting the wellbore rock failure while drilling. This improvement is to be achieved by overcoming the aforementioned simplifying assumptions so that the plastic behavior of the rock, the realistic non-uniform wellbore shape, and an improved criterion for failure are all reflected by the model. Ultimately, this modelling effort should be able to produce an accurate three-dimensional (3D) representation of the wellbore structure. It should also produce a dynamic representation of the wellbore strength and the safe mud weight that can be updated as new drilling events take place.

It is proposed to improve wellbore failure prediction through the use of an integrated modelling scheme that involves an elasto-plastic 3D finite element model (FEM), machine learning (ML) algorithms for classification training and image recognition, and real-time drilling data such as image logs from a logging while drilling (LWD) tool that accurately describes the current shape of the wellbore. The available offset wells data are to be modelled in the FEM code, which are then used to train the ML algorithms. The produced integrated model of FEM and ML is to be used to predict failure limits for new wells. This improved failure prediction is to be updated with the occurrence of different drilling events. **Figure 3** depicts the proposed workflow for in-practice application of the integrated modelling scheme between the FEM and the ML.
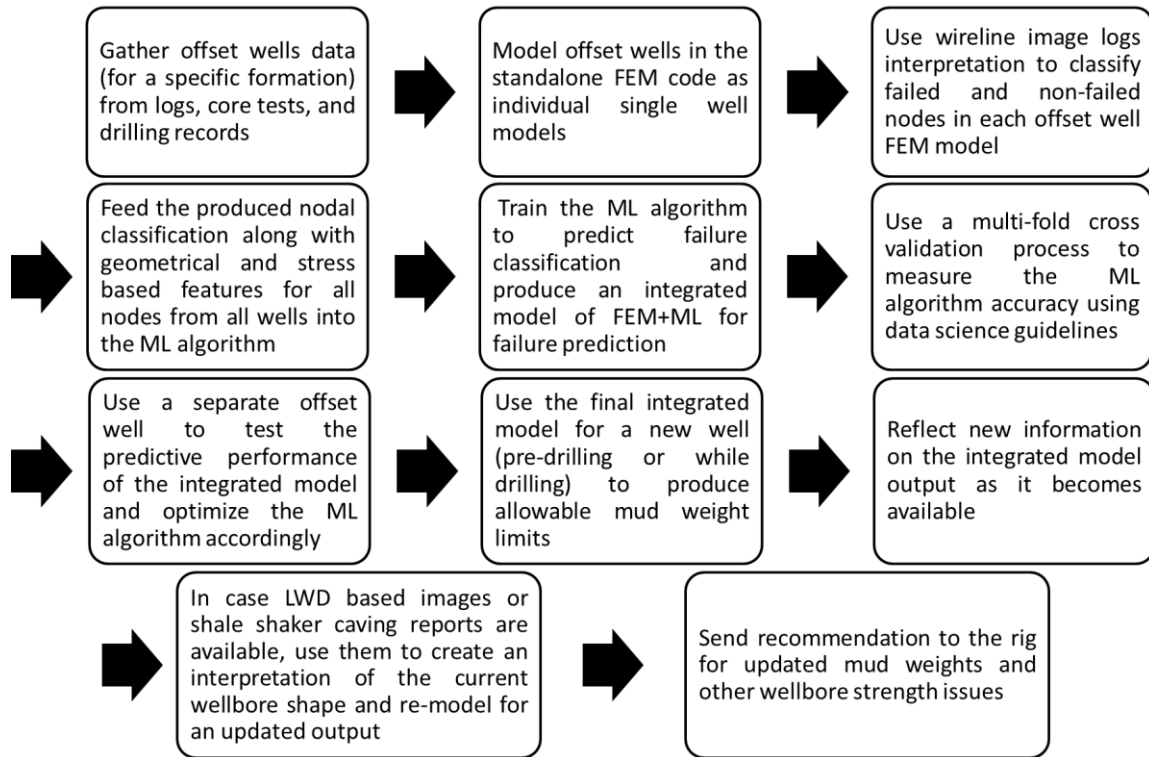
Figure 3. The targeted workflow for in-practice application of the integrated modelling scheme between the FEM and the ML.

Another goal is to provide a proof of concept and validation through applying the integrated modelling scheme to available lab experimental results. For the ML algorithms standalone predictive performance, a multi-fold cross validation process using the guidelines outlined by reputable data science references is to be used for testing. Finally, available actual field data are to be used to build an integrated model, which is then to be used to show that the model output for a test well (a well that will not be used to build and train the integrated model) reasonably matches the drilling events experienced by the well.

1.4 Potential Limitations

The most potentially limiting assumption in the proposed approach is the availability of data that describes the structural changes of the wellbore, for both the offset wells used for training the ML algorithm and for new wells. The ideal data for describing the structure and geometry of a wellbore is an image log. These logs are usually only available for reservoir intervals. Another useful type of data for describing the wellbore structure is a shale shaker caving report. However, while these reports can practically be gathered for all intervals, including overburden formations, producing detailed caving reports is not generally a prevailing practice in the industry.

Another possibly limiting assumption is that the deviation from a perfectly circular wellbore shape that is observed in an image log is due to wellbore failure. In some instances, the non-circular or non-uniform wellbore shape observed in an image log could be the result of drilling through naturally fractured, faulted, or rubblized zones. If this phenomenon is not identified correctly while interpreting the image logs using the ML image analysis code, this can introduce errors into the trained ML classification model.

One final potential limitation in the proposed approach is the possible loss of nuanced description of failure observed from image logs. Because the FEM code to be used mainly models the mechanical behavior (elasto-plastic model) of the rock, in some instances, this modelling effort might not capture the variation of rock failure that are related to the type of drilling fluid being used for example.

CHAPTER II

MODEL DEVELOPMENT

The model developed for this work relies on a new modelling scheme that integrates the use of a physics model, such as the finite element model, with a machine learning algorithm. As discussed in the first chapter, this modelling scheme can fall under the category of physics-based or theory-guided data science. Through this modelling scheme, two main advantages can be achieved:

1. Improve the prediction of rock failure through the use of machine learning algorithms in place of the conventional failure criteria.

2. Enable the machine learning algorithm to be trained on datasets that are smaller than those required for purely data-driven models.

Also, through the use of machine learning, it is possible to capture new forms of data while drilling and reflect them on the model to produce an updated representation of the drilling window and wellbore strength. These forms of data are said to be new here because, while they have been available in the industry for the past decades, they have not been conventionally exploited for the purpose of drilling engineering and design, especially in real-time settings. An example of these forms of data is image logs produced by logging while drilling tools or by wireline tools. These images are routinely used for reservoir engineering as they can provide valuable information about the structure of formations. However, while they also provide valuable information for drilling as they describe the extent and mode of rock failure experienced by the wellbore rock, they have

not been taken advantage of sufficiently to improve the process of drilling engineering design.

The main novel idea used to build the model and integrate the physics finite element model with the machine learning algorithm is the use of an automated image analysis and interpretations techniques. The use of these techniques plays three important roles in this modelling scheme, which are:

1. Enables taking advantage of the aforementioned forms of data, which are not conventionally incorporated into modelling for drilling engineering purposes

2. Provides a mean for rock failure classification for the purpose of a supervised machine learning training

3. Provides a mean for reflecting the effect of drilling events encountered while drilling in the model input to produce a timely updated output

Therefore, image analysis techniques essentially work to bridge the gap between the physics model and the machine learning algorithm while also facilitating for real-time applications of the integrated model.

A summary of the integrated process undertaken by the model developed in this work is illustrated in the flow chart in **Figure 4**. The same flow chart is broken down into different phases in **Figure 5** to highlight the functionalities of the integrated model. As can be seen in **Figure 5**, the first phase is training the integrated model using offset wells data, the second phase is producing results for a new well before drilling commences to allow for drilling engineering planning, and the final phase is the while drilling implementation that allows for a reactive modelling that responds to drilling events.
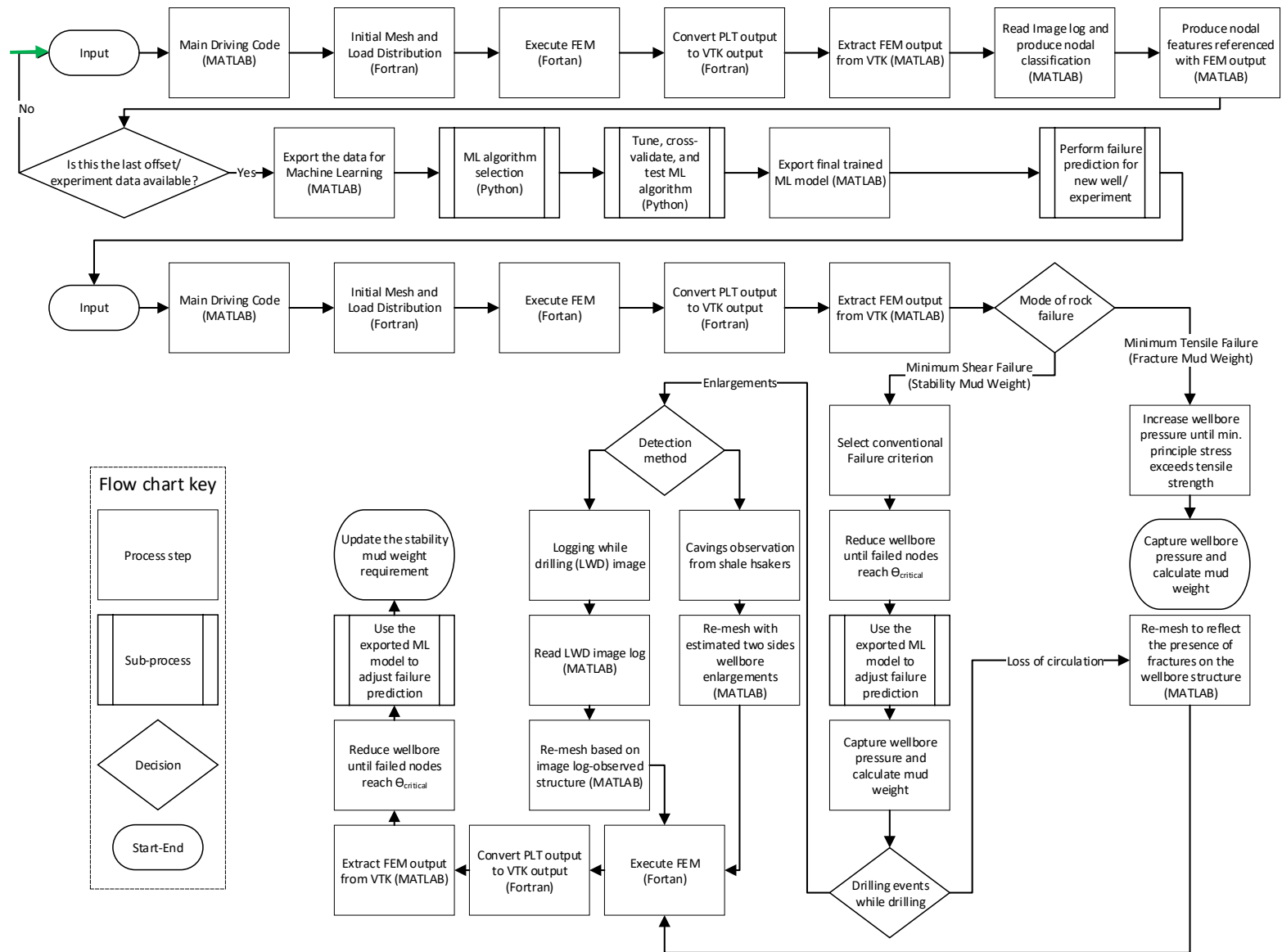
Figure 4. Flow chart detailing the integrated process undertaken by the model developed in this work.
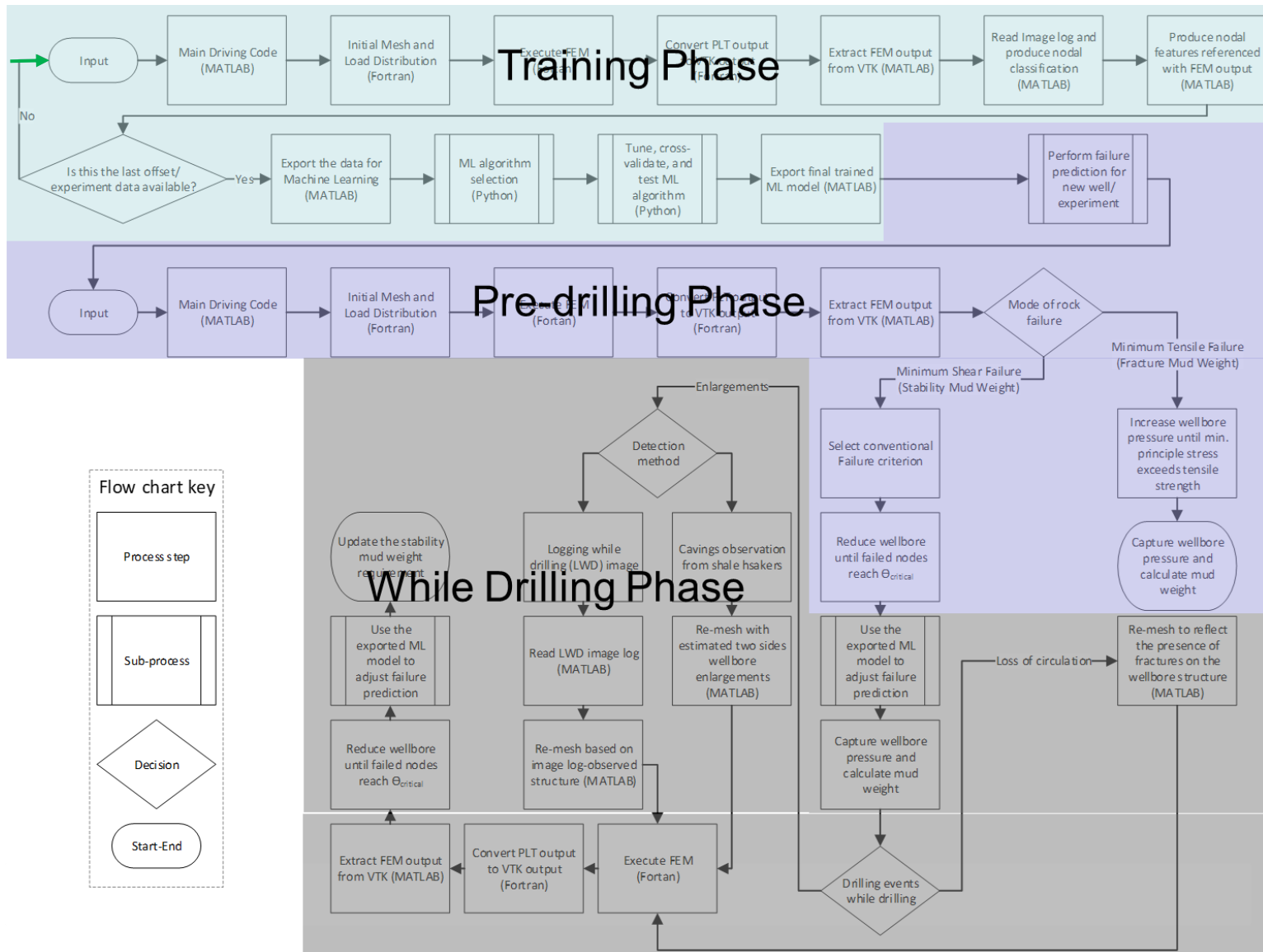
Figure 5. The same flow chart in Figure 4 broken down in different phases to highlight the model functionalities.

2.1 Elasto-plastic 3D Finite Element Method (FEM)

The physics model used for this work is an in-house three-dimensional elasto-plastic finite element model for solids. It follows a similar implementation to the one discussed by Owen and Hinton (1980). This model offers several advantages over what is currently used in drilling engineering for geomechanics modelling. As discussed in the first chapter, geomechanics modelling in drilling engineering mostly relies on analytical solutions or simplified numerical solutions, both of which are commonly linear elastic. The use of plasticity, in FEM or otherwise, is quite rare in drilling engineering as shown in the literature review section. Another advantage offered by this model is calculations speed, which allows for a real-time or a while-drilling implementation of the model. While-drilling implementation of a geomechanics model that considers the plastic behavior of the rock is non-existent in the literature as per the review conducted.

2.1.1 Mathematical Formulation

The constitutive relationships for the finite element code are derived based on the work shown by Zienkiewicz (1977). The discretization process undertaken here is quite common for geomechanics applications and it has been in wide spread use in different engineering fields for the past decades.

2.1.1.1 Discretization Using the Minimization of Total Potential Energy

To apply discretization using the minimization of the total potential energy, a body in equilibrium has to be considered. If there are body forces and applied boundary forces, the change of the potential energy with virtual displacement is zero. Now, in case the potential energy is minimum, the equilibrium is said to be stationary. If the external work

is W and the internally stored potential is U, then the following is true only if the body is in equilibrium:

$$\delta U_p = \delta (W + U) = 0 \tag{18}$$

This concept of the minimum potential energy can be used to discretized the fundamental elasticity equations shown below, where u is the displacement of the body in equilibrium with its body forces and applied boundary forces. This displacement u is estimated using the nodal displacement using the shape functions N:

$$u = Nu^e \tag{19}$$

The displacement-strain relationship with B as the strain-displacement matrix is:

$$\varepsilon = Bu^e \tag{20}$$

The stress-strain relationship with D as the consistent tangent matrix is:

$$\sigma = D\varepsilon \tag{21}$$

T_i is the boundary stress and the energy applied by the external stress is:

$$W = \int_{S_e} (u)^T T \, d\Gamma \tag{22}$$

Strain energy stored in the body:

$$U = \int_{V_e} (\frac{1}{2} \varepsilon^T \sigma - u^T F) d\Omega \tag{23}$$

By substituting Equations (22) and (23) into Equation (18),The potential energy becomes:

$$U_p = U + W = \int_{V_e} (\frac{1}{2} \varepsilon^T \sigma - u^T F) d\Omega + \int_{S_e} (u)^T T \, d\Gamma \tag{24}$$

Substituting the nodal displacement approximation ($u = Nu^e$):

$$U_p = (u^e)^T \int_{S_e} N^T T \, d\Gamma + \frac{1}{2}(u^e)^T [\int_{V_e} (B^T D B \, u^e)]u^e d\Omega - \int_{V_e} N^T F \, d\Omega \qquad (25)$$

If the body is at stationary equilibrium, the displacement that minimizes the potential is found by:

$$\frac{\partial U_p}{\partial u^e} = 0 \qquad (26)$$

Hence:

$$u^e \int_{V^e} \left((B^T)D \, B\right)d\Omega = \int_{V^e} N^T F \, d\Omega - \int_{S^e} N^T T \, d\Gamma \qquad (27)$$

$$K^e u^e = F^e + T^e \qquad (28)$$

Where:

$$K^e = \int_{V^e} \left((B^T)D \, B\right)d\Omega \; : Stiffness \; matrix \qquad (29)$$

$$F^e = \int_{V^e} N^T F \, d\Omega \; : Body \; force \qquad (30)$$

$$T^e = -\int_{S^e} N^T T \, d\Gamma \; : Traction \; Force \qquad (31)$$

Loadings from in-situ stresses and from the drilling fluid hydrostatic pressure are applied in increments to account for non-linear behavior and represented by the body force vector F and traction force vector T. The K stiffness matrix is assembled based on the consistent tangent matrix D and the strain-displacement matrix B. Equation (28) is solved numerically, through an iteration loop which takes place inside the load increment loop. The final output of the FEM model is the displacement u. The displacement output can

then be used in elasticity and plasticity equations to produce stress-strain at each numerical integration node.

2.1.1.2 Iso-Parametric Elements and Shape Functions

For this application, a 20-node iso-parametric Serendipity brick element is used. The use of the iso-parametric elements ensures three main advantages, which are enabling the use of higher order polynomials for shape functions, enabling the use of curvilinear coordinates, and reducing complexity by having the variable transformation performed with the same shape functions. The last advantage mentioned means that both the curvilinear coordinate transformation in Equation (32) and the variable transformation in Equation (33) are performed using the same shape functions N_i as follows:

$$x = \sum_{i=1}^{20} N_i(\xi, \eta, \varsigma) x_i \tag{32}$$

$$u = \sum_{i=1}^{20} N_i(\xi, \eta, \varsigma) u_i \tag{33}$$

The same curvilinear coordinate transformation in Equation (32) is also performed for the y and z coordinate in the 3D FEM. **Figure 6** visualizes the 20-node brick element, the local coordinates of (ξ, η, ς) and the global coordinates (x,y,z).
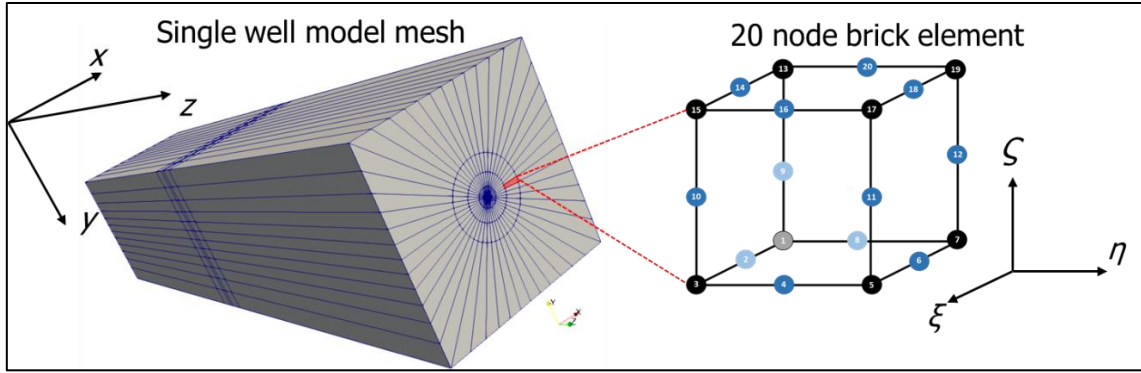
Figure 6. Visualization of the local and global coordinates and the 20-node brick element with numbered nodes.

As for the shape functions used for this work, they are quadratic Serendipity polynomials. They can be expressed based on the following general formulas from (Zienkiewicz 1977). For corner nodes and for side nodes respectively:

$$N_i(\xi, \eta, \varsigma) = \frac{1}{8}(1 + \xi\xi_i)(1 + \eta\eta_i)(1 + \varsigma\varsigma_i)(\xi\xi_i + \eta\eta_i + \varsigma\varsigma_i - 2) \qquad (34)$$

$$N_i(\xi, \eta, \varsigma) = \frac{1}{4}(1 - \xi^2)(1 + \eta\eta_i)(1 + \varsigma\varsigma_i) \qquad (35)$$

In these formulas, $(\xi_i, \eta_i, \varsigma_i)$ are the values of the local coordinates (ξ, η, ς), which can be read at each node number as depicted in **Figure 7**, where node number 1 is placed at the (-1 ,-1,-1) point.
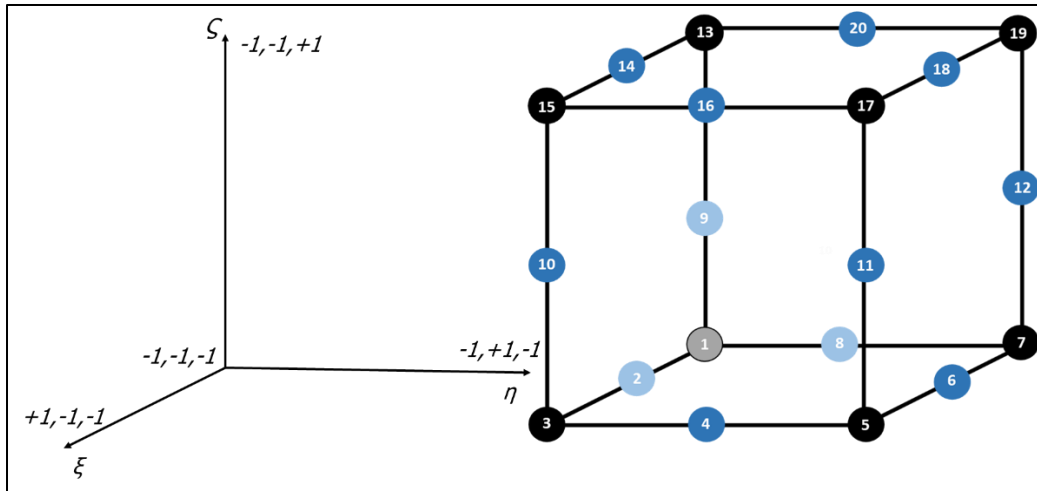
Figure 7. Values of the local coordinates for each numbered node in the 20-node brick element for the derivation of the shape functions.

2.1.1.3 Pre-processing Steps

There are some pre-processing steps that are required to prepare the well data to be used into the FEM code. These steps include meshing, defining the boundary conditions, setting the values for the loads (i.e. in-situ stresses and wellbore pressures), and transforming the loads into the wellbore coordinates.

For the boundary conditions, in all single well models shown in this work, both the top and the bottom layer of nodes are fixed with no allowed displacement. It should be noted that these fixed layers of nodes are part of the overburden and the underburden formations and they are not part of the interval being analyzed for rock failure. Another boundary condition is the application of the in-situ stresses after having them transformed into the wellbore coordinate. The transformed in-situ stresses $(\sigma_x, \sigma_y, \sigma_z)$ are applied to the nodes exposed at the top, bottom, and the four sides of the structure. Also, the wellbore pressure P_w exerted by the drilling fluid is applied to the nodes exposed to the wellbore

28

wall surface. In all cases, there a total of 8 exposed nodes for each 20-node element at these surfaces. **Figure 8** visualizes the boundary conditions on a single well mesh.
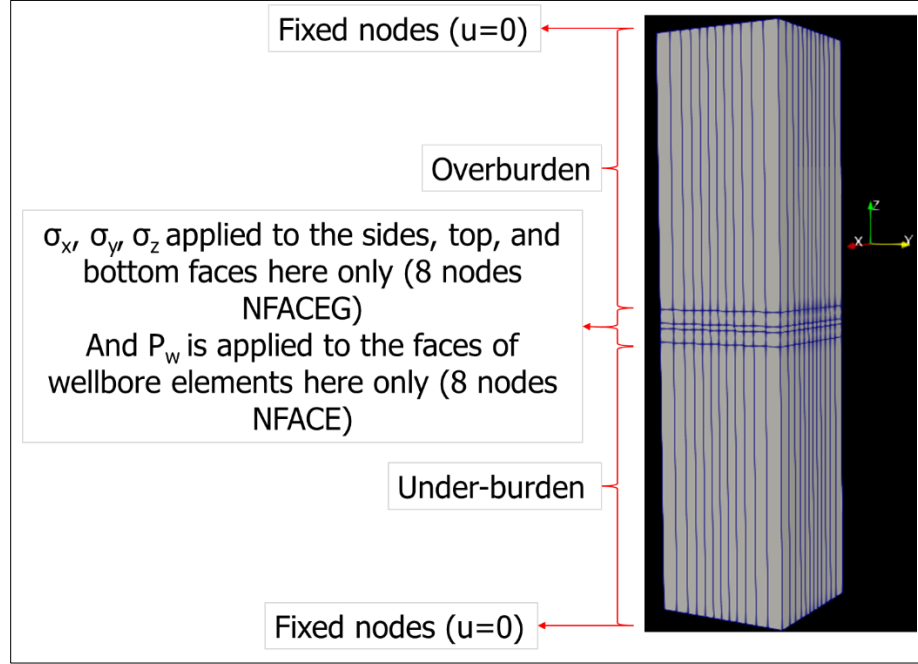
Figure 8. A visualization of typical boundary conditions on a single well mesh model.

The transformation of the in-situ stresses is carried out using the Euler angles for matrix rotation (Daneshy 1973, Richardson 1983). Based on this, the in-situ stresses $(\sigma_v, \sigma_H, \sigma_h)$ are transformed in the wellbore coordinate $(\sigma_x, \sigma_y, \sigma_z)$ using the following relationships:

$$\sigma_x = \sigma_v \, sin^2\gamma + (\sigma_H \, cos^2\beta + \sigma_h sin^2\beta)cos^2\gamma \tag{36}$$

$$\sigma_y = \sigma_H sin^2\beta + \sigma_h \, cos^2\beta \tag{37}$$

$$\sigma_z = \sigma_v \, cos^2\gamma + (\sigma_H \, cos^2\beta + \sigma_h sin^2\beta)sin^2\gamma \tag{38}$$

$$\tau_{yz} = \frac{\sigma_H - \sigma_h}{2} sin(2\beta) \, sin\, \gamma \qquad (39)$$

$$\tau_{xz} = \frac{1}{2}(\sigma_H \, cos^2 \, \beta + \sigma_h sin^2 \, \beta) \, sin(2\gamma) \qquad (40)$$

$$\tau_{xy} = \frac{\sigma_H - \sigma_h}{2} sin(2\beta) \, cos^2 \gamma \qquad (41)$$

Where the angle γ describes the inclination of the wellbore as a deviation from the maximum in-situ stress and the angle β specifies the relationship between the maximum horizontal stress and the wellbore projection on the two horizontal stresses plane. **Figure 9** provides an illustration of the coordinates systems and angles required for transformation.
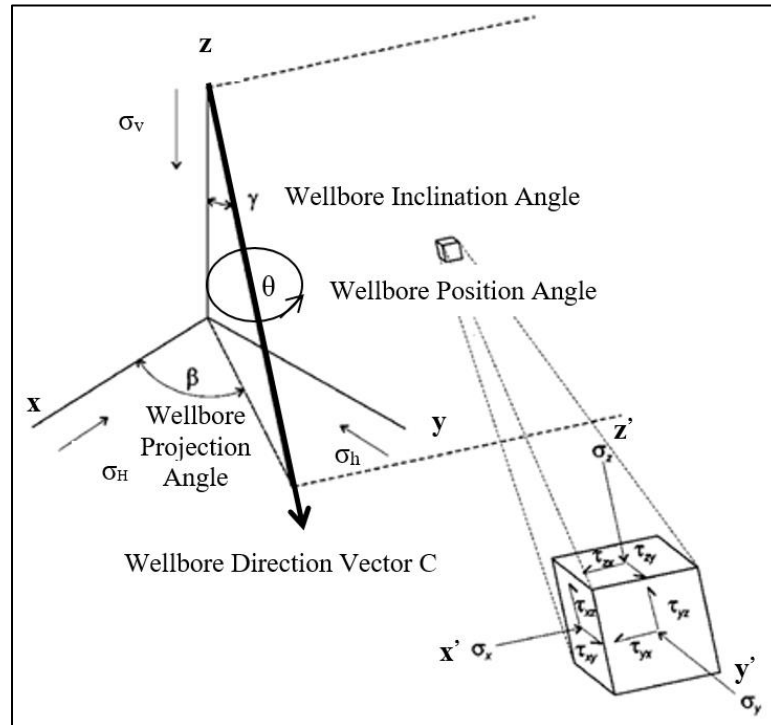
Figure 9. The relationship between the in-situ stresses reference frame and the defined wellbore section reference frame (reprinted from Watson et al. 2003).

2.1.1.4 Construction of the Stiffness Matrix

The construction of the elemental stiffness matrix K^e shown in Equation (29) requires the determination of consistent tangent matrix D, the strain-displacement matrix B, and the Jacobian matrix J. The determinant of the Jacobian matrix is used for the purpose of transforming the integral from the global coordinate system (x,y,z) to the local coordinate system (ξ, η, ς) as follows:

$$dxdydz = |J|d\xi d\eta d\varsigma \tag{42}$$

Based on this, the elemental stiffness matrix K^e in Equation (29) becomes:

$$K^e = \int_{V^e} B^T DB \, d\Omega = \int_{V^e} B^T DB \, |J| \, d\xi d\eta d\varsigma \tag{43}$$

Where the Jacobian matrix for 20-node brick element with the quadratic Serendipity shape functions from Equations (34) and (35) is defined as:

$$J = \begin{vmatrix} \sum_{i=1}^{20} \dfrac{N_i(\xi,\eta,\varsigma)}{\partial \xi} x_i & \sum_{i=1}^{20} \dfrac{N_i(\xi,\eta,\varsigma)}{\partial \xi} y_i & \sum_{i=1}^{20} \dfrac{N_i(\xi,\eta,\varsigma)}{\partial \xi} z_i \\ \sum_{i=1}^{20} \dfrac{N_i(\xi,\eta,\varsigma)}{\partial \eta} x_i & \sum_{i=1}^{20} \dfrac{N_i(\xi,\eta,\varsigma)}{\partial \eta} y_i & \sum_{i=1}^{20} \dfrac{N_i(\xi,\eta,\varsigma)}{\partial \eta} z_i \\ \sum_{i=1}^{20} \dfrac{N_i(\xi,\eta,\varsigma)}{\partial \varsigma} x_i & \sum_{i=1}^{20} \dfrac{N_i(\xi,\eta,\varsigma)}{\partial \varsigma} y_i & \sum_{i=1}^{20} \dfrac{N_i(\xi,\eta,\varsigma)}{\partial \varsigma} z_i \end{vmatrix} \tag{44}$$

The strain-displacement matrix B for the 20-node brick element is defined as:

$$B = (B_1 \quad B_2 \quad B_3 \quad \cdots \quad B_{19} \quad B_{20}) \tag{45}$$

The B_i elements of this matrix are defined as:

$$B_i = \begin{bmatrix} \dfrac{\partial N_i(x,y,z)}{\partial x} & 0 & 0 \\[2mm] 0 & \dfrac{\partial N_i(x,y,z)}{\partial y} & 0 \\[2mm] 0 & 0 & \dfrac{\partial N_i(x,y,z)}{\partial z} \\[2mm] \dfrac{\partial N_i(x,y,z)}{\partial y} & \dfrac{\partial N_i(x,y,z)}{\partial x} & 0 \\[2mm] 0 & \dfrac{\partial N_i(x,y,z)}{\partial z} & \dfrac{\partial N_i(x,y,z)}{\partial y} \\[2mm] \dfrac{\partial N_i(x,y,z)}{\partial z} & 0 & \dfrac{\partial N_i(x,y,z)}{\partial x} \end{bmatrix} \tag{46}$$

As can be seen, the dimensions of the B_i matrix are 6x3. This means that the B matrix in Equation (45) has a dimension of 6x60. The derivatives of the shape functions $\left(\dfrac{\partial N_i(x,y,z)}{\partial x}, \dfrac{\partial N_i(x,y,z)}{\partial y}, \dfrac{\partial N_i(x,y,z)}{\partial z}\right)$ are determined as follows using the inverse of the Jacobian matrix J:

$$\begin{bmatrix} \dfrac{\partial N_i}{\partial x} \\[2mm] \dfrac{\partial N_i}{\partial y} \\[2mm] \dfrac{\partial N_i}{\partial z} \end{bmatrix} = J^{-1} \begin{bmatrix} \dfrac{\partial N_i}{\partial \xi} \\[2mm] \dfrac{\partial N_i}{\partial \eta} \\[2mm] \dfrac{\partial N_i}{\partial \varsigma} \end{bmatrix} \tag{47}$$

The final part concerning the construction of the elemental stiffness matrix is the consistent tangent matrix D. The definition of this matrix differs for linear elasticity and plasticity. The linear elastic definition is as follows where it is based on the values of the linear elasticity constants, which are the Poisson's ratio v and the Young's modulus E:

$$D = \frac{E(1-v)}{(1+v)(1-2v)} \begin{bmatrix} 1 & \dfrac{v}{1-v} & \dfrac{v}{1-v} & 0 & 0 & 0 \\ & 1 & \dfrac{v}{1-v} & 0 & 0 & 0 \\ & & 1 & 0 & 0 & 0 \\ & & & \dfrac{1-2v}{2(1-v)} & 0 & 0 \\ & Symmetry & & & \dfrac{1-2v}{2(1-v)} & 0 \\ & & & & & \dfrac{1-2v}{2(1-v)} \end{bmatrix} \quad (48)$$

However, as yielding occurs and the transition from the linear elastic behavior to the plastic behavior takes place, the nonlinear pseudo consistent tangent matrix is determined using incremental equilibrium. This is done by reformulating the stiffness matrix at each iteration and solving the new system of equations each time. This is known as the Direct Iteration Algorithm and it is detailed in (Owen and Hinton 1980).

The final result for the elemental stiffness matrix from Equation (43) will be a 60x6 matrix. It can be shown that this dimension agrees with the dimension of the matrices to the right-hand side of Equation (28), where the body force and the traction forces matrices are as follow:

$$F^e = \int_{V^e} N^T F \, d\Omega = \int_{V^e}(N^T N \, |J| d\xi d\eta d\varsigma) \, F, \; where: F = \begin{bmatrix} F_x \\ F_y \\ F_z \end{bmatrix} \quad (49)$$

$$T^e = -\int_{S^e} N^T T \, d\Gamma \;, where \; T = \begin{bmatrix} T_1 \\ T_2 \\ T_3 \\ \vdots \\ T_{19} \\ T_{20} \end{bmatrix} \quad (50)$$

Considering that the shape function matrix N is 3x60 and its transpose N^T is 60x3, the resulting F^e matrix will be 60x1 and the resulting T^e matrix will be 60x1. This is consistent with the dimensions of the left-hand side in Equation (28), which contains the stiffness matrix with a 60x6 dimension and the unknown displacement matrix with 60x1 dimension.

2.1.1.5 Plasticity Flow Rule and Strain Hardening

As loading increases, the deformation behavior of the wellbore rock will transition from linear elasticity to plasticity by passing through what is known as the yield point. This transition is illustrated in **Figure 10** which shows a typical stress strain curve for a high strength sandstone.
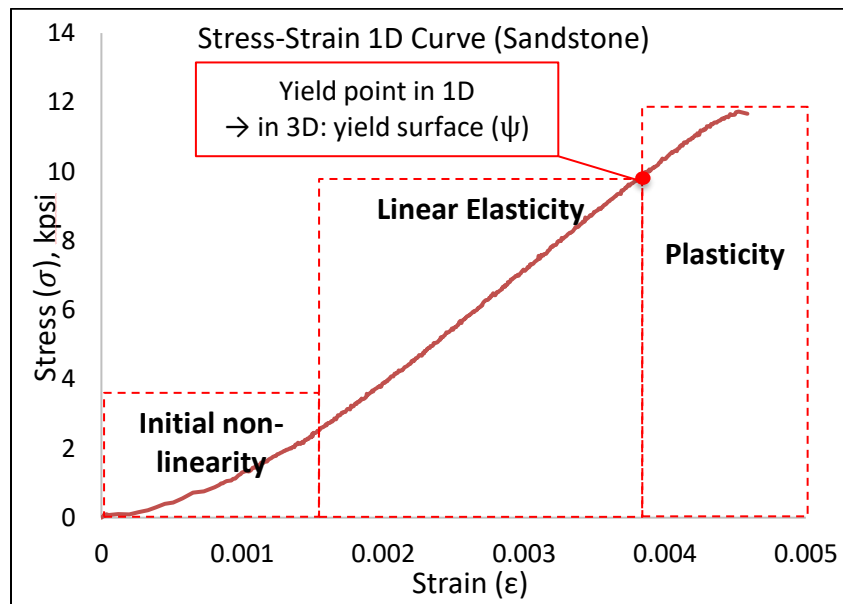
Figure 10. A typical stress-strain curve for a high strength sandstone shown the different behaviors exhibited by the rock as loading increases.

34

This model relies on the plastic flow rule for strain hardening to reflect the plastic behavior of the rock once it passes the yield point as shown by Morita (2004). This means that the strain is the result of two components, which are elastic strain ε^e and a plastic strain ε^p as follows:

$$\varepsilon = \varepsilon^e + \varepsilon^p \tag{51}$$

The plastic flow rule assumes that the flow direction is perpendicular to the yield surface ψ and it is defined as:

$$\Delta\varepsilon_{ij}^p = \lambda \frac{\partial\psi\left(\sigma_{ij}\right)}{\partial\sigma_{ij}} \tag{52}$$

This is known as the associative flow rule because it is assumed that the plastic potential surface is the same as the yield surface ψ. It also assumes the yield surface expands without changing the flow direction. This assumption means that the plastic strain multiplier λ is a constant. Intuitively, the plastic flow rule is the rate of change of the yield surface with respect to the stress tensor.

The yield criterion used in this work is the Drucker-Prager criterion. In a similar formulation the one shown by Huang et al. (2018), the associative Drucker-Prager yield criterion stipulates that yielding will take place when the deviatoric stress tensor S_{ij} and the mean stress σ_m satisfy the following relationship:

$$\psi\left(\sigma_{ij}\right) = \sqrt{\frac{1}{2}S_{ij}S_{ij}} - a_0 + a_1\sigma_m = 0 \tag{53}$$

Where constants a_0 and a_1 are determined experimentally as material properties and are used to correlate Drucker-Prager criterion to the Mohr-Coulomb criterion.

The following expression for strain hardening is used to calculate the scalar plastic strain ε^p from the plastic strain tensor determined by the flow rule:

$$\varepsilon^p = \int \sqrt{\frac{2}{3} d\varepsilon_{ij}^p d\varepsilon_{ij}^p} \tag{54}$$

Based on this, the scalar plastic strain is calculated as follows:

$$\varepsilon^p = \sqrt{\frac{2}{3} \varepsilon_{ij}^p \varepsilon_{ij}^p} \tag{55}$$

2.1.2 Description of the FEM Implementation

The elasto-plastic FEM is implemented through 33 subroutines and a driver code. The driver code calls twelve main subroutines and these perform several functions including receiving the input file, applying loads to construct and assemble the global stiffness matrix, and solving the system of equations. A flow chart of the driver and the twelve main subroutines is shown in **Figure 11**. For the construction of elemental stiffness matrices, there are two separate subroutines that calculate and inverse the Jacobian matrix based on Equation (44) and for Equation (47) and two more subroutines to evaluate and map the shape functions and their derivatives. These shape functions calculations are mainly performed for Equation (46) for the strain-displacement matrix B and for Equation (47). The flow chart with the Jacobian matrix and shape functions subroutines is shown in **Figure 12**.
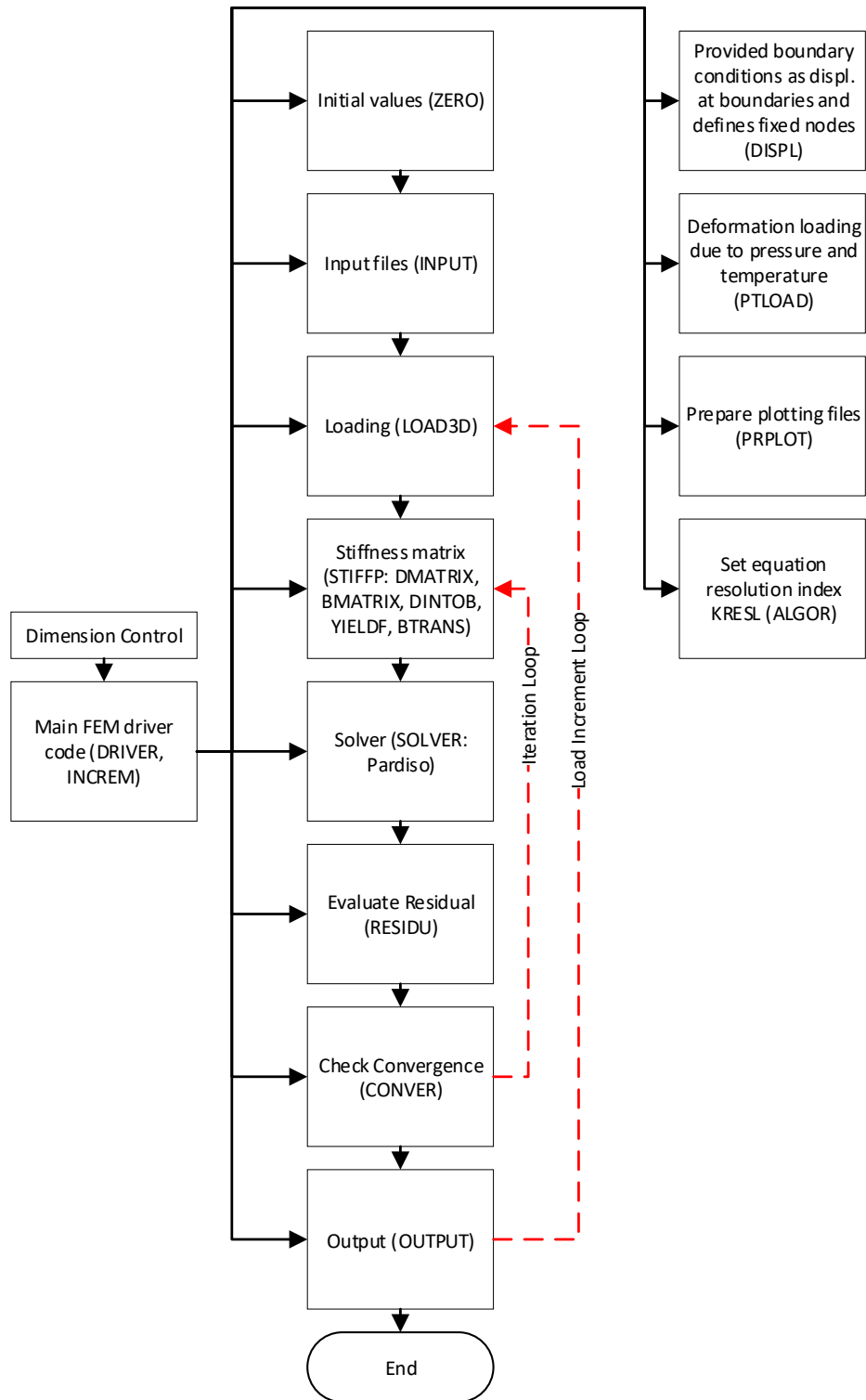
Figure 11. Flow chart of the driver code of the FEM and the 12 main subroutines called by it.
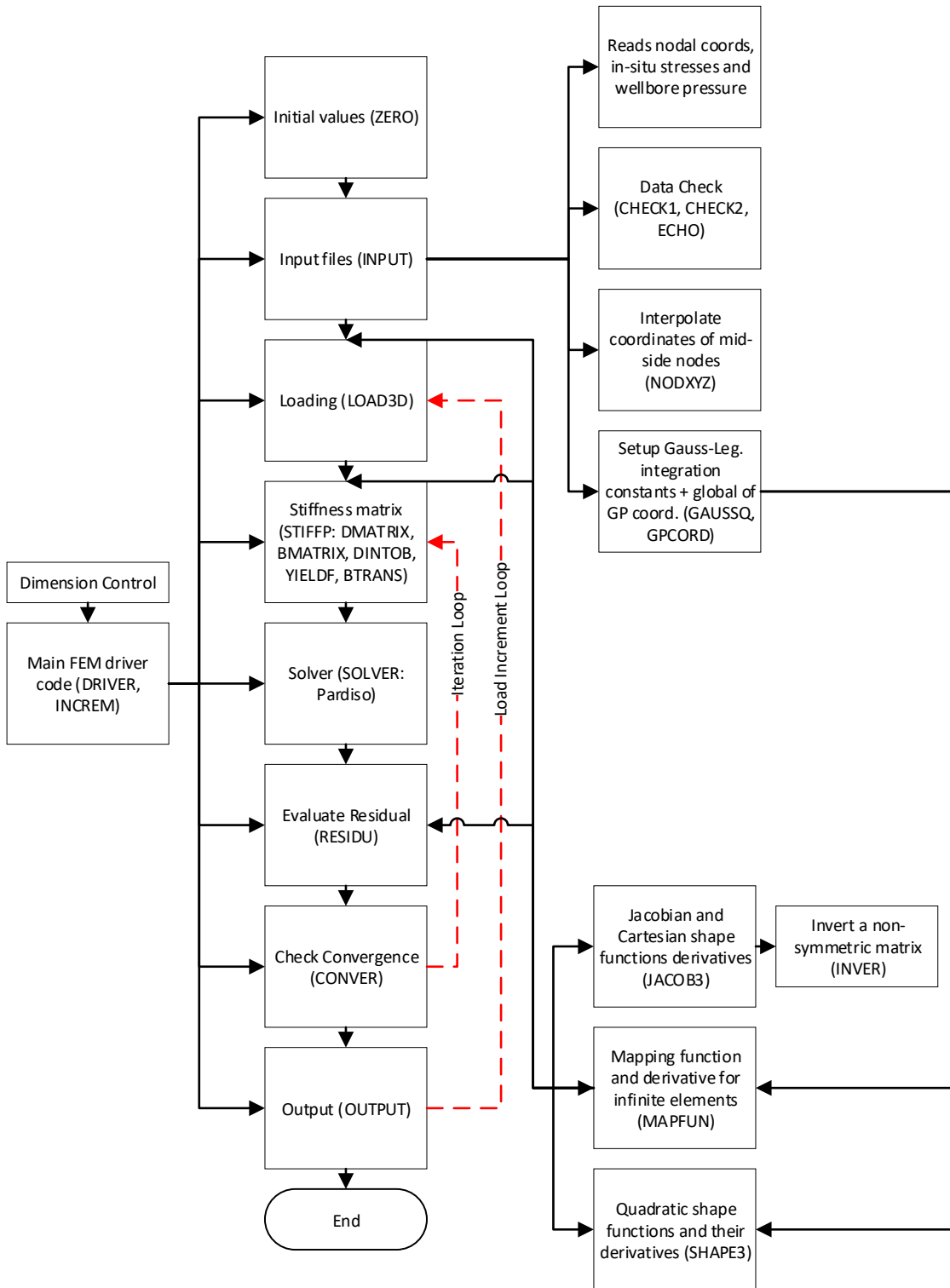
Figure 12. A flow chart detailing the process for the construction of the stiffness matrix using shape functions and Jacobian matrix calculations.

Carrying on with the construction of the stiffness matrix, the consistent tangent matrix D is calculated separately either based on its linear elastic form as in Equation (48) or based on its post-yield plastic pseudo nonlinear form. Moreover, the yield function ψ derivative and the plastic strain multiplier λ are calculated for the plastic flow strain hardening. The stiffness matrix is finally concluded by formulating the full displacement-strain matrix B and multiplying it by the consistent tangent matrix D. These steps that detail the process of the final steps in the construction of the stiffness matrix are shown highlighted in a red dashed box in **Figure 13**.

Upon solving the system of equations and determining the displacements u, the residual forces are calculated to check for convergence and equilibrium by subtracting the left-hand side from the right-hand side in the global form as follows:

$$\int_{V^e} \left((B^T) D\ B \right) u\ d\Omega\ = \int_{V^e} N^T F\ d\Omega - \int_{S^e} N^T T\ d\Gamma \qquad (56)$$

Where the left-hand side is the global stiffness matrix multiplied by displacement and the right-hand side is the body and traction forces. The value obtained from the subtraction of these two quantities should be equal to zero if the equilibrium condition is fully satisfied. However, that is not always achievable, therefore, a tolerance value is set to check for convergence. The tolerance value is usually set to be close but not equal to zero, which in this work and implementation is set at 10^{-6}. Once the residual forces are calculated and found to be less than the set tolerance value, convergence is said to be achieved, otherwise, the residual forces are carried to the next iteration. The same process is repeated for each separate load increment, where the load increments are defined in the input file manually.

These processes are carried out in two loops with the convergence loop nested in the load increment loop as shown in **Figure 14**. Finally, a full comprehensive flow chart for the full FEM code is shown in **Figure 15**.
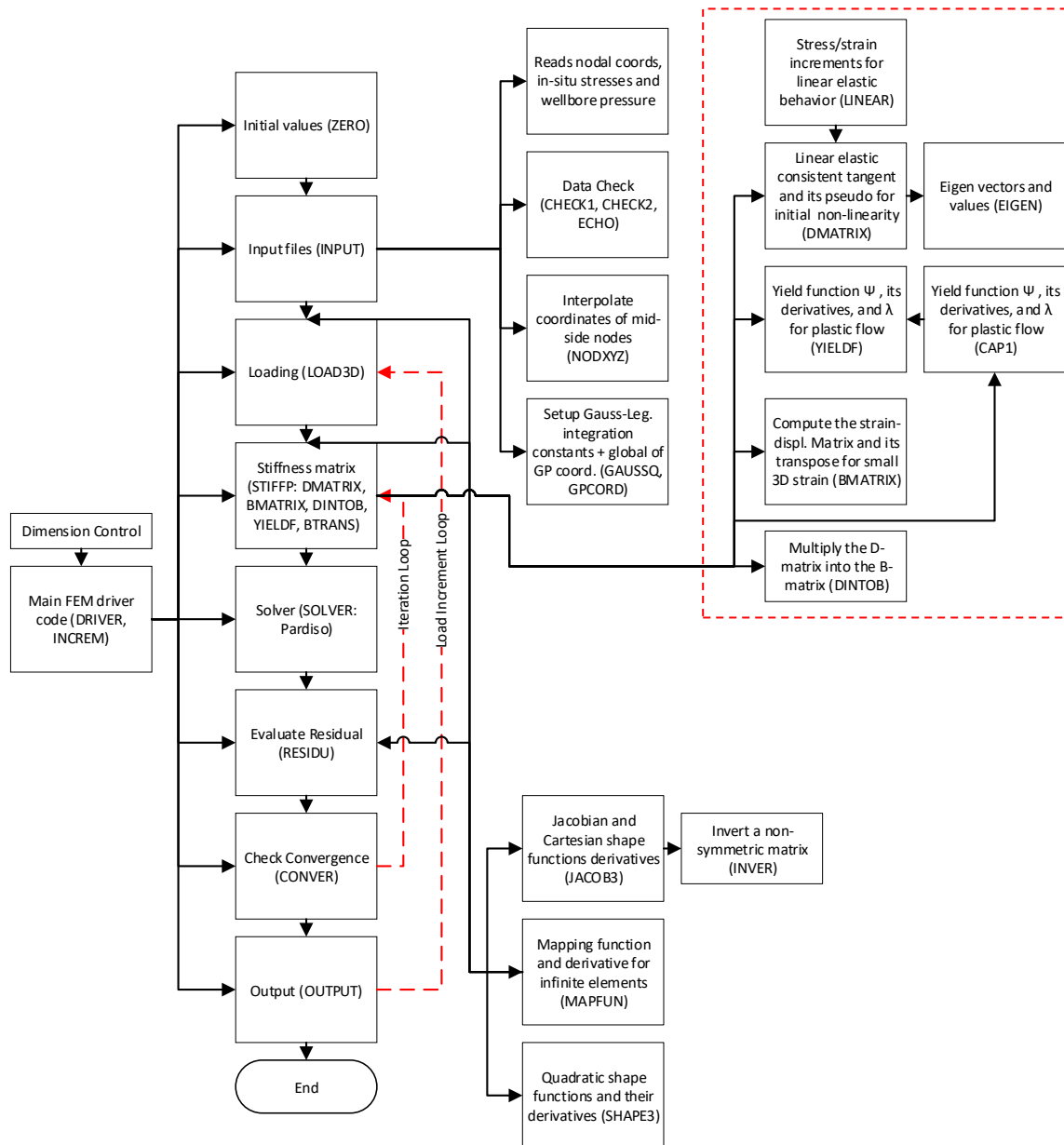
Figure 13. A flow chart highlighting the final steps taken to construct the stiffness matrix.
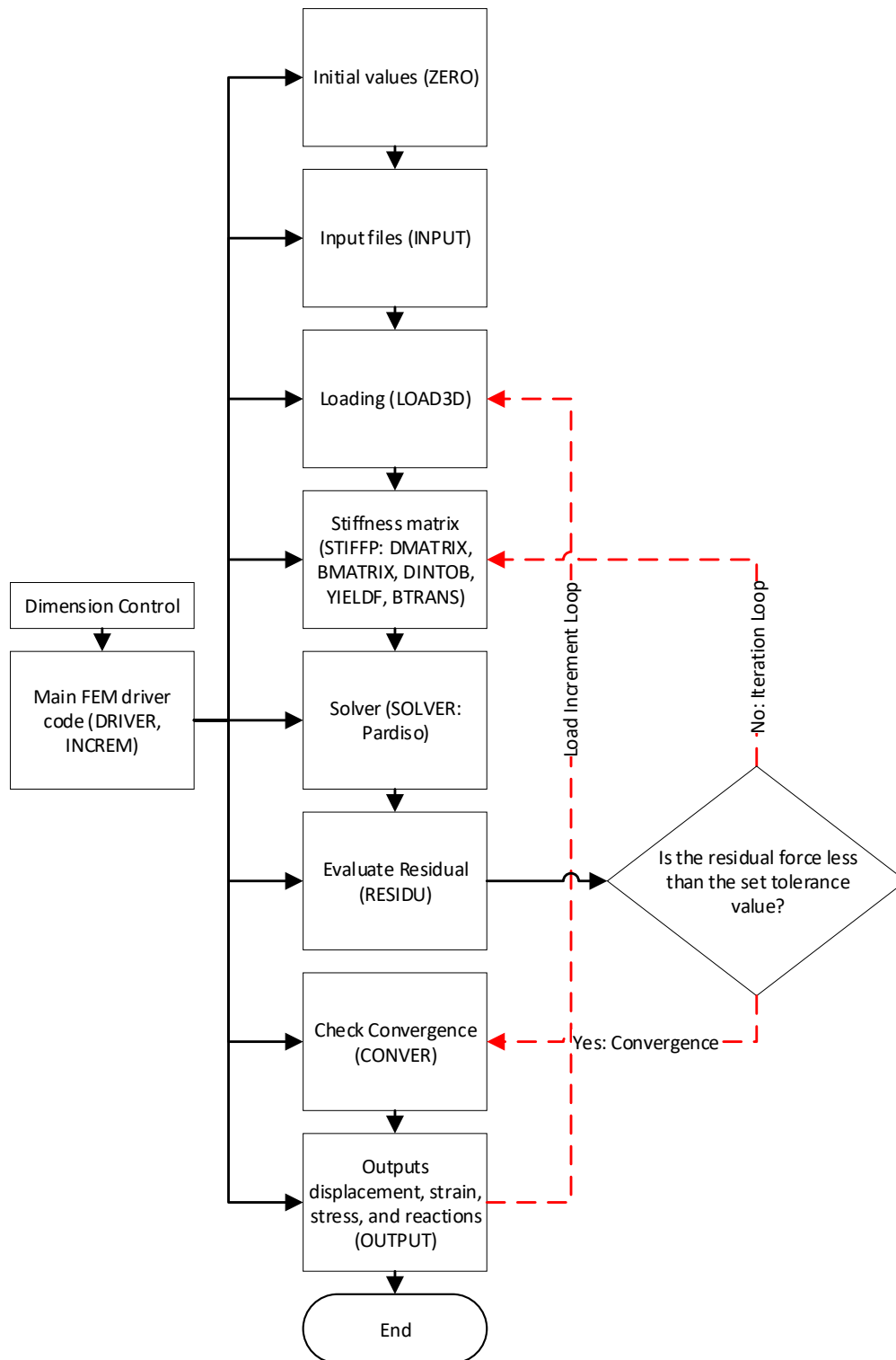
Figure 14. A flow chart detailing the process for residual forces calculations, convergence checks, and load the increment loop.
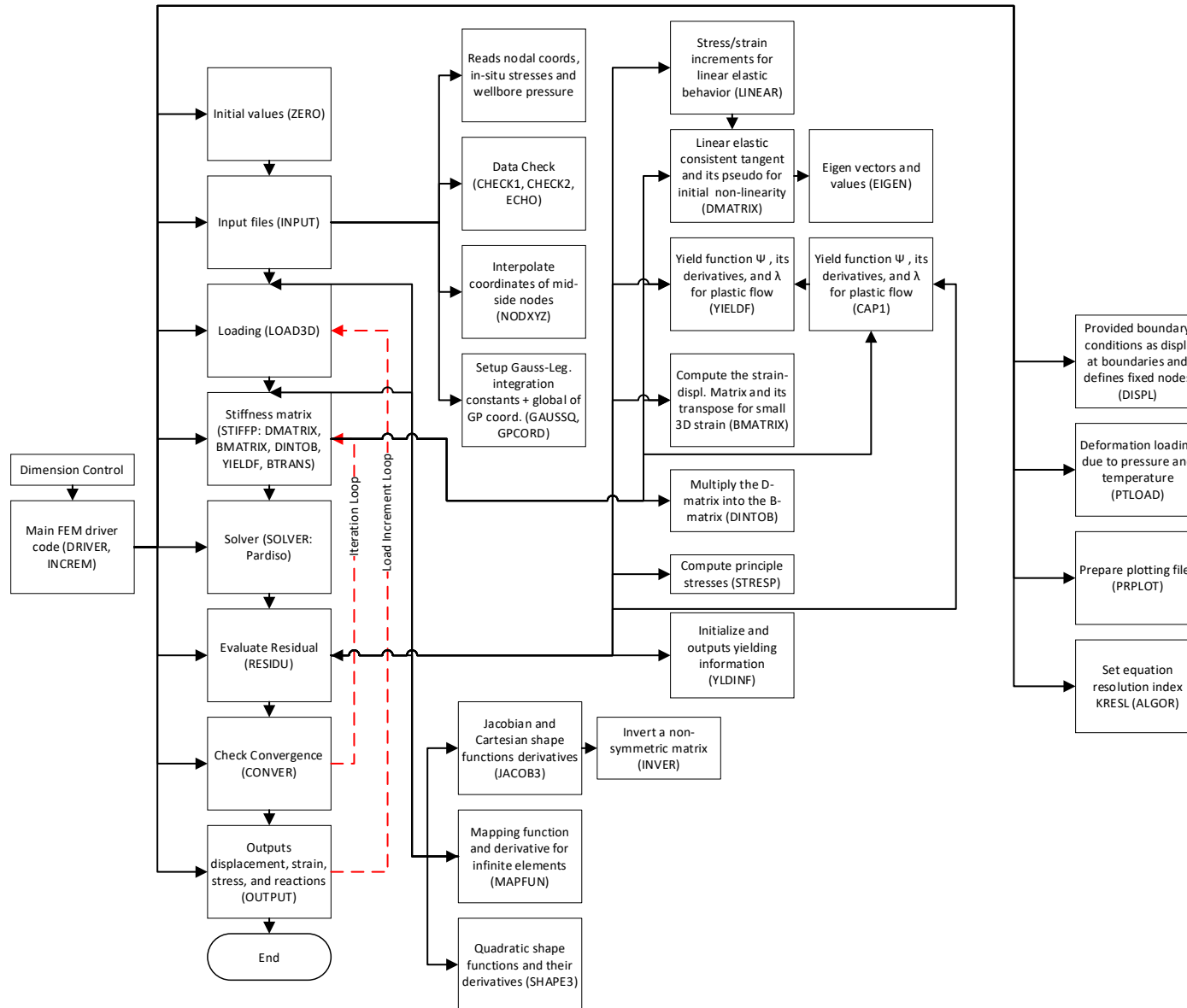
Figure 15. A full comprehensive flow chart of the elasto-plastic FEM code.

2.2 Development of Image Analysis Function for Both Lab Experiments and Field Use

As mentioned in the introduction of this chapter, image analysis techniques are proposed in this modelling scheme to provide a link between the finite element geomechanics model and the machine learning algorithms. Image analysis techniques serve this function by providing the ability to quantitatively determine the extent of wellbore wall rock failure from images. Conventionally, quantitative description of wellbore failure is provided using mechanical multi-arm caliper logs, however, these don't provide a comprehensive and full description of the wellbore shape as well as images can. The quantitative description of failure from images is used to provide a node-based classification of rock failure, which can be used to correlate the classification to each separate node in the finite element model mesh. By providing this link to the finite element model, a supervised training process through machine learning algorithms for the purpose of failure classification is now possible. The end result is an integrated FEM and ML model, which is trained based on offset wells data or previous lab experiments to predict the limits for rock failure in new wells or experiments.

Another function served by image analysis is enabling the model to reflect the effect of different drilling events as they occur through images produced from logging while drilling (LWD) tools. This ensures that events that led to any kind of change to the wellbore geometry and shape are now accurately reflected in the FEM mesh. This while-drilling consideration of changes to the wellbore shape and structure means that the drilling window can be re-evaluated and updated as drilling progresses.

Different image analyzers are developed for this model to be applied to lab-based images such as CT-scans and to field-based images such as resistivity image logs and ultrasonic image logs.

2.2.1 Bored Core CT-Scans Analyzer

The objective of developing an image analyzer for lab-based images is to provide a mean for producing a proof of concept and validation of the integrated FEM-ML model. The reason lab-based images, especially CT-scans, are useful for providing validation is that they are more easily interpreted than field-based images. Field-based images such as resistivity image logs are taken downhole in challenging environments due to high pressure, temperature, and drilling fluids noise, which means that the visual clarity of these images is not high. Therefore, field-based images cannot be assessed by the naked eye alone to assess the extent of failure, which is why it was deemed that relying on them to provide the initial validation is not the preferred route. This is not the case for lab-based images. In CT-scans, which are the form of lab images used in this work as seen in **Figure 16**, one can clearly determine the location and the extent of rock failure around a borehole in a core sample by simply visually examining the image. Based on this, the role of the bored core CT-scan analyzer is to identify the location of failure and determine the magnitude of failure in terms of change to the borehole radius.
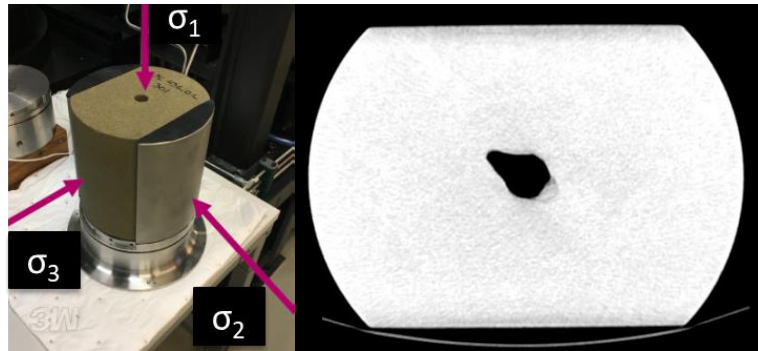
Figure 16. Left: bored core sample. Right: an example of a post-test CT-scan showing the failed borehole

Since the analyzer deals with images, the points in the image are actually pixels and the value for each point is actually the red-green-blue (RGB) color number. The CT-scan image analyzer relies on the concept of triangulation to assess failure as illustrated in **Figure 17 (a)**. Through this concept, a pixel that is centralized in the borehole is picked, which is then used as the first vertex of a triangle. From this vertex, either a horizontal or a vertical straight line is drawn by picking the pixels from the first vertex to the borehole wall. Then, a hypotenuse for a right triangle is formed at a specified angle. Next, the length of the hypotenuse is calculated using the Pythagorean theory. The process is repeated both vertically and horizontally to provide full readings of the borehole radii as shown in **Figure 17 (c)** and **(d)**. The number of the radii readings produced is dependent on the number of pixels available, i.e. image resolution, in the CT-scan. Finally, the vertical and the horizontal radii readings are combined and a spline smoothing is applied to the curve to eliminate any irregularities resulting from the pixels corners. This final result of this is shown in **Figure 18** .
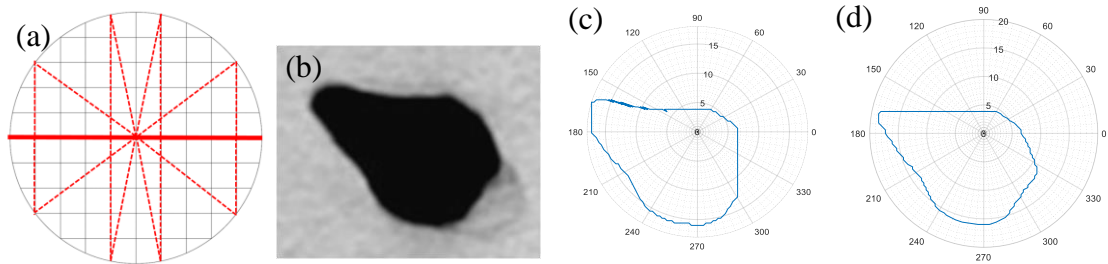
Figure 17. (a) The concept of triangulation used to calculate the borehole radii from images. This is a simple example of a circular borehole with horizontal triangulation lines. (b) Post-test CT-scan showing a failed borehole. (c) The radii measurements of the failed borehole based on horizontal triangulation. (d) The radii measurements of the failed borehole based on vertical triangulation.
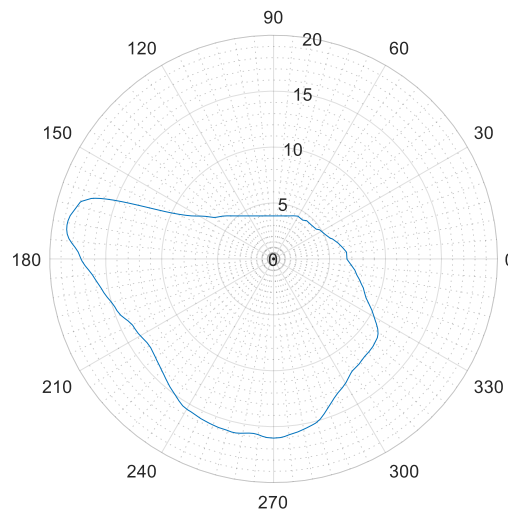
Figure 18. The full radial measurements of a CT-scan of a failed borehole as a result of combining the horizontal and the vertical triangulation readings and the spline smoothing step. The radial measurements in this chart are in mm.

The same process can be repeated for two separate CT-scans to determine the extension of the failure zones from one CT-scan to the other. **Figure 19** shows two CT-scans and the superimposed radii measurements. **Figure 20** shows the re-positioning of the two CT-scans from **Figure 19** to ensure the area of failure extension is accurately

46

identified. **Figure 20** also shows a mapping of the failure extension zone, where the changes of radial measurements from one CT-scans to the other is quantified at its specific circumferential location.
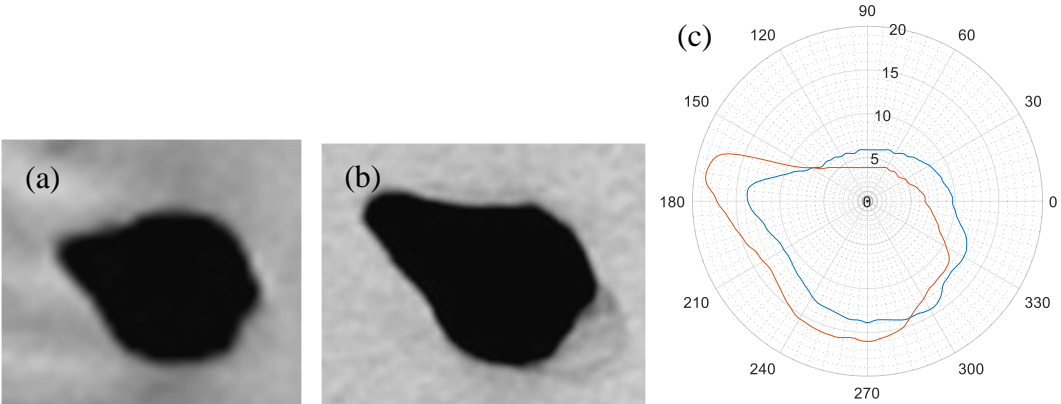
Figure 19. (a) CT-scan-1. (b) CT-scan-2. (c) superimposing the radii measurements of both CT-scan-1 and CT-scan-2.
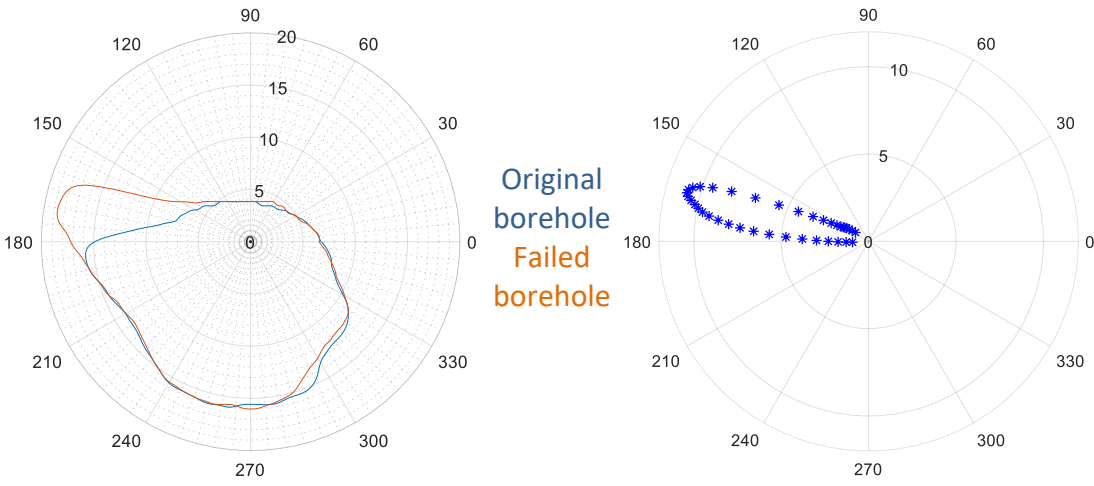
Figure 20. Left: re-positioning the two CT-scan to ensure centralization of definition of failure extension zones. Right: mapping the location and magnitude of borehole failure points

2.2.2 Ultrasonic Based Image Logs Analyzer

The goal of this analyzer is to convert this type of image logs into caliper readings that cover the 360° of the wellbore. This means that instead of relying on conventional mechanical caliper logs, which provide radial measurements at only four circumferential angles (0°, 90°, 180°, 270°) in their most common designs, it is now possible to rely on images to produce radial measurement at all circumferential angles. As in the case with CT-scans, the resolution of radial measurements is dependent on the image log resolution.

The ultrasonic image logs are the preferred form of image logs for this type of application. This is due to two main reasons. First, unlike the resistivity-based image logs, ultrasonic images have no gaps, and therefore, can provide a full and comprehensive description of the wellbore state. Second, they are available in LWD tools in the form of density-based images, which means they can be used in a real-time setting or a while-drilling setting to feed new information into the integrated model.

The main difference in the interpretation process between lab-based images and field-based image logs is that field-based image logs cannot be solely relied on to produce the interpretation. This is mainly due to the aforementioned challenging environment in which these images are taken and due to variations in images from one commercial logging tool to the other. This means that field image logs require more support information and validation than lab-based images. The support information used in this application is mechanical caliper log readings. For each image logging tool, the analysis of its images can be calibrated and validated against mechanical caliper measurements taken from the same interval where the images were produced. This ensures that the

quantitative assessment of enlargements and tight spots from an image is validated using the mechanical caliper measurements of the same enlargements and tight spots. Once this process is performed, the calibrated analyzer can be used to interpret images from the same logging tool with more confidence without the need to run a mechanical caliper log every time an image log is produced.

This process is used for interpreting ultrasonic image logs by converting the RGB color number of each pixel to radial measurements. An example of an ultrasonic image log is shown in **Figure 21**. The area highlighted with a red box is the interval that is used to provide an example of an image log interpretation.
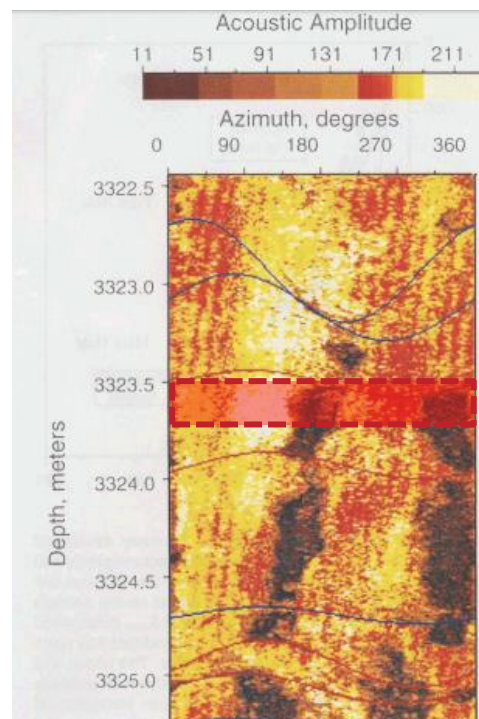
Figure 21. An example of an ultrasonic image log from 8.375" wellbore showing stress-induced wellbore enlargements (reprinted from Barton et al. 1997). The highlighted zone is the interval used for providing an example of image log interpretation.

The results of the interpretation of the ultrasonic image log in **Figure 21** are shown

in **Figure 22** and **Figure 23**. **Figure 22** illustrates the clear correlation between the RGB

number and the placement of enlargements and tight spots. This correlation is then used

along with calibrating values from a mechanical multi-arm caliper log to calculate the

radial measurements of the wellbore at all directions. As mentioned previously, the

number of radial measurements produced is dependent on the number of pixels available

in the image log (i.e. image resolution). In this particular example, the image log has

resolution of 221 pixels per the 360° degrees around the wellbore, which equates to around

one radius measurement for each 1.63° degrees around the circumference of the wellbore.

Figure 23 shows the final results of the image log interpretations that are calibrated with

mechanical caliper log readings in both cartesian and polar coordinates. When comparing

the ultrasonic image log interval in **Figure 21** to the interpreted radial measurements in

Figure 23, it can be clearly seen that the dark regions in the image that signify the presence

of wellbore enlargements are also reflected as radial measurements that are higher than

the bit radius.
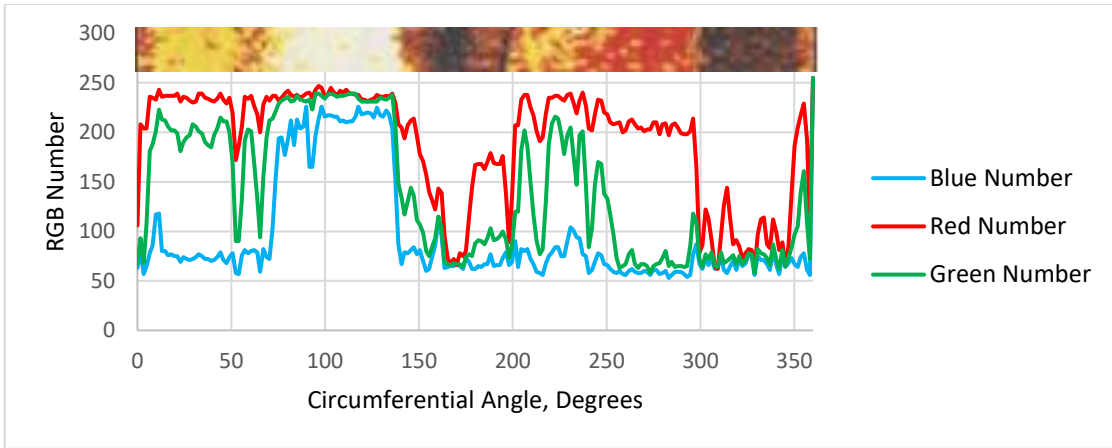
Figure 22. A plot of the RGB numbers from the ultrasonic image log showing a clear correlation between the location of enlargements in the image and the trends in the RGB numbers.
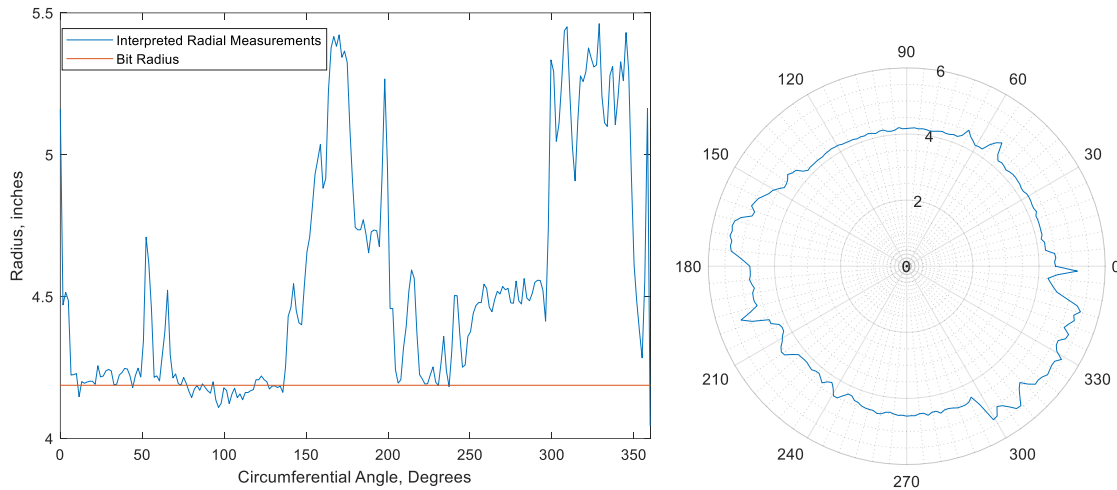
Figure 23. Left: the results of the ultrasonic image log interpretation showing radial measurements in cartesian coordinates. Right: the results of the ultrasonic image log interpretation showing radial measurements in polar coordinates to reflect the actual shape of the wellbore cross-section.

2.2.3 Resistivity Based Image Logs Analyzer

Just as in the case of the ultrasonic image log analyzer, the goal of this analyzer is

also to convert this type of image logs into caliper readings that cover the 360º of the

wellbore. The main difference between sonic-based images and resistivity-based images is that resistivity images normally contain several gaps due to the space between the electrical pads of the resistivity image logging tool. Resistivity based image logs also differ from sonic-based images in the way a breakout can be identified. In sonic images, changes in color are sufficient, however, in resistivity images, changes in color alone can be misleading as they can signify laminations, faults, natural, and induced fractures along with breakouts. This means that resistivity logs are efficient indicators of 'physical features' within the wellbore, but they are poor indicators of the spatial (specifically, radial) extension of these physical features. Fortunately, there's one distinguishing feature of breakouts in resistivity logs which is the lowered resolution of the image at that breakout orientation. The lower resolution at breakouts is due to the electrical pad of the logging tool being further away from the wellbore wall when the resistivity image measurements are taken. This change in color resolution is exploited in the image interpretation analyzer to identify the orientation of breakouts, still, it's not sufficient to assess the extent of the enlargement. For this issue, provided multi-arm caliper readings are used to correlate the radial extent at every location in the wellbore. The pixels RGB readings are feature scaled using the following formula to yield the needed radii:

$$Scaled\ RGB\ = CAL_{min} + \left(\frac{RGB_{adj} - RGB_{min}}{RGB_{max} - RGB_{min}} (CAL_{max} - CAL_{min}) \right) \qquad (57)$$

Where CAL_{min} and CAL_{max} are minimum and maximum multi-arm mechanical caliper radius readings respectively, which are used for calibration. RGB_{adj} is the adjusted RGB reading for filtering out noise and white gaps in the image. Using this method in

interpreting image logs means that the availability of multi-arm caliper data is a necessity when resistivity images are used, which also means this analysis can't be done in real-time. This is to be expected since conventional resistivity-based image logs are not available through LWD in the same manner as density-based images.

An example of a resistivity image log along with mechanical caliper log readings are shown in **Figure 24**. The regions highlighted in red boxes in the image log to the right are the low-resolution areas that signify the presence of wellbore enlargements. The area highlighted with a red box in the mechanical caliper log to the left is the interval from which the image log to the right was taken.
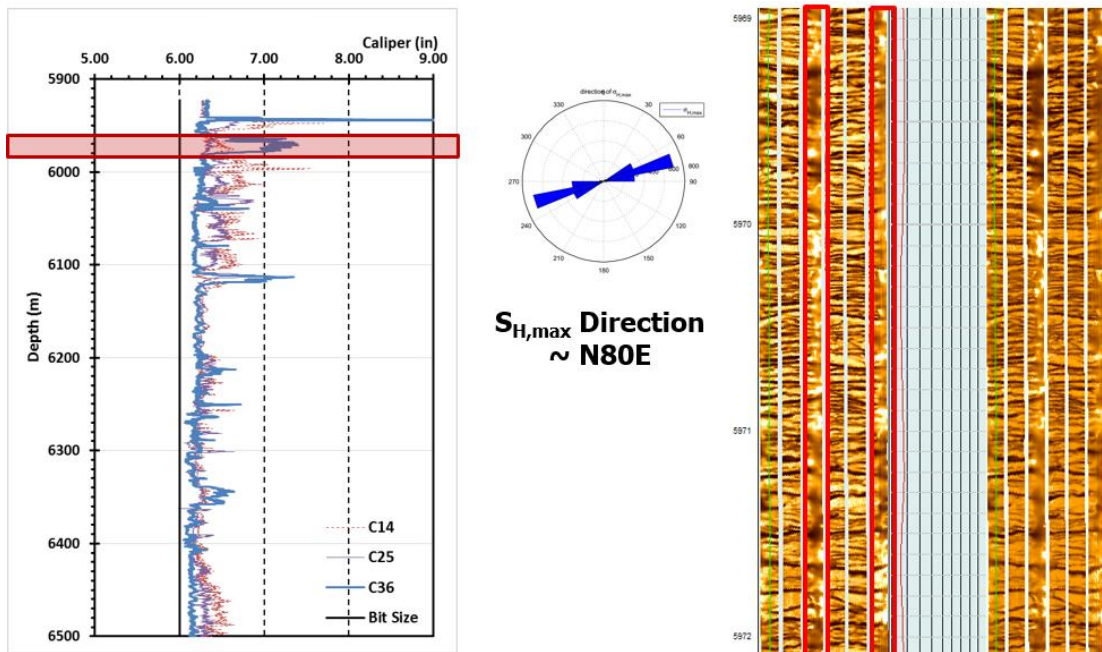
Figure 24. Left: mechanical multi-arm caliper log readings. Right: resistivity-based image log, which is also known as formation micro-imager log (FMI).

Figure 25 shows the original RGB number of the resistivity image log from **Figure 24** and the adjusted RGB values to account for the image gaps and other image noise. **Figure 26** shows the final results of the image log interpretations that are calibrated with mechanical caliper log readings in both cartesian and polar coordinates. When comparing the resistivity image log interval in **Figure 24** to the interpreted radial measurements in **Figure 26**, it can be clearly seen that the low resolution regions in the image that signify the presence of wellbore enlargements are also reflected as radial measurements that are higher than the bit radius. Also, when examining the mechanical caliper readings in the interval highlighted in red to the left in **Figure 24**, it can be clearly seen that there is an agreement that the wellbore is fully enlarged at all directions.
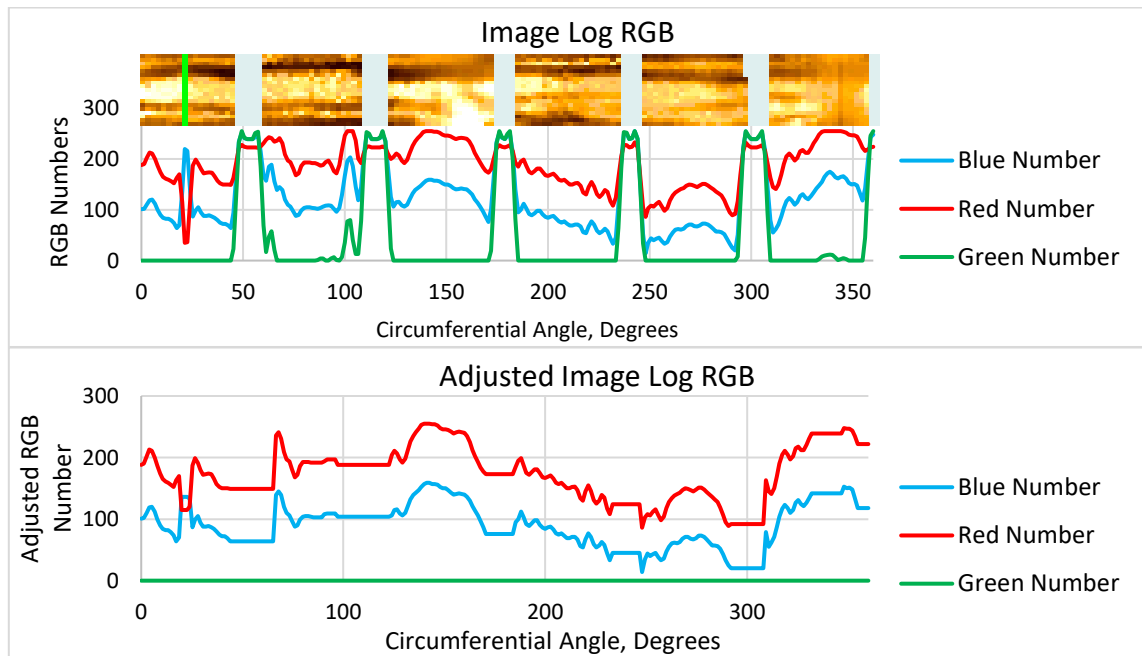
Figure 25. Top: A plot of the RGB numbers from the resistivity image log showing a clear correlation between the location of the image gaps, the low-resolution areas, and the trends in the RGB numbers. Bottom: the adjusted RGB numbers of the resistivity image log to remove noise.
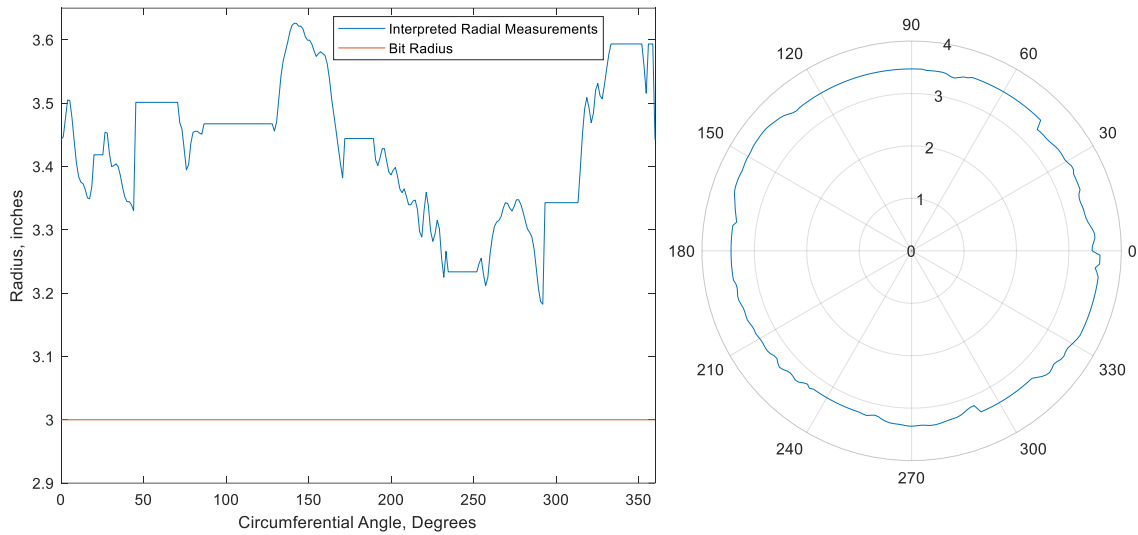
Figure 26. Left: the results of the resistivity image log interpretation showing radial measurements in cartesian coordinates. Right: the results of the resistivity image log interpretation showing radial measurements in polar coordinates to reflect the actual shape of the wellbore cross-section.

2.2.4 Geometric Shape Indicators for Mapping Progressive Failure

To conclude the image analysis section, there are other forms of information that can be extracted from the images, which need to addressed. This once again is concerned with interpreting images of boreholes, from either lab experiments or wells, that are not perfectly circular or uniformly shaped. For these types of irregular borehole shapes, the rules of symmetry no longer apply. And since the assumption of borehole symmetry is extremely common in the field of drilling geomechanics wellbore modelling as seen in (Feng et al. 2015, Salehi and Nygaard 2010) due to its ability to simplify models and significantly reduce computational cost, eliminating this assumption requires new considerations. The wellbore symmetry assumption stipulates that considering only a half or a quarter of the wellbore in the model, as seen in **Figure 27**, is sufficient because the

55

load distribution and failure occurrence in the second half or the remaining three quarters will be identical to those of the modelled portion of the wellbore. Once this assumption is removed by modelling a fully non-uniform wellbore shape, the distribution and response to the applied loads become skewed by the irregularities present around the wellbore wall. Intuitively, this in turn will affect the prediction of failure in general while ensuring that the failure distribution will vary around the wellbore. This issue is tackled in this work by extracting new parameters from images of non-uniform boreholes, which are to be used to assess and quantify the distribution of loads and failure around the non-uniform boreholes. These new parameters are designated in this work as the geometric shape indicators. Considering these parameters deal with modelling non-uniform borehole shapes, they are useful to provide a framework for progressive failure modelling. This is because single step failure models usually start from uniform (either perfectly circular or elliptical) borehole shapes, while modelling boreholes that had already experienced a form of failure requires the consideration of non-uniform shapes, hence, considering non-uniform borehole shapes is useful in progressive failure models.
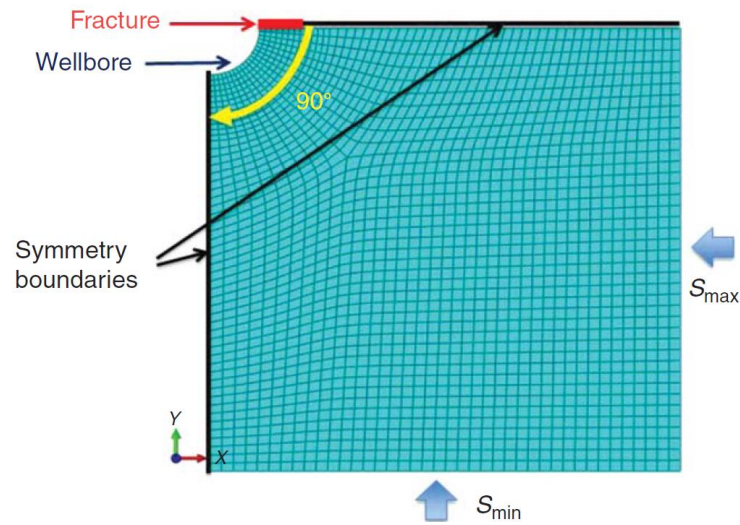
Figure 27. Invoking the symmetry assumption by modelling a quarter of the wellbore to reduce model complexity and computational cost (reprinted from Feng et al. 2015).

To illustrate the extraction of the proposed geometric shape indicators from images, the CT-scans in **Figure 28** are used. For illustration purposes, the first CT-scan will be considered as the initial borehole and the second CT-scan will be considered as the final borehole that resulted from a progressive failure step. As can be seen, the borehole in the starting CT-scan is non-uniformly shaped and the mapping of failure highlights a concentration of failure in certain non-symmetric regions of the borehole. What follows is a list of the geometric shape indicators as extracted from the first CT-scan. The main objective of extracting these parameters is to eventually use them as precursors to failure where they can be used as features in the machine learning algorithm for predicting progressive failure steps.
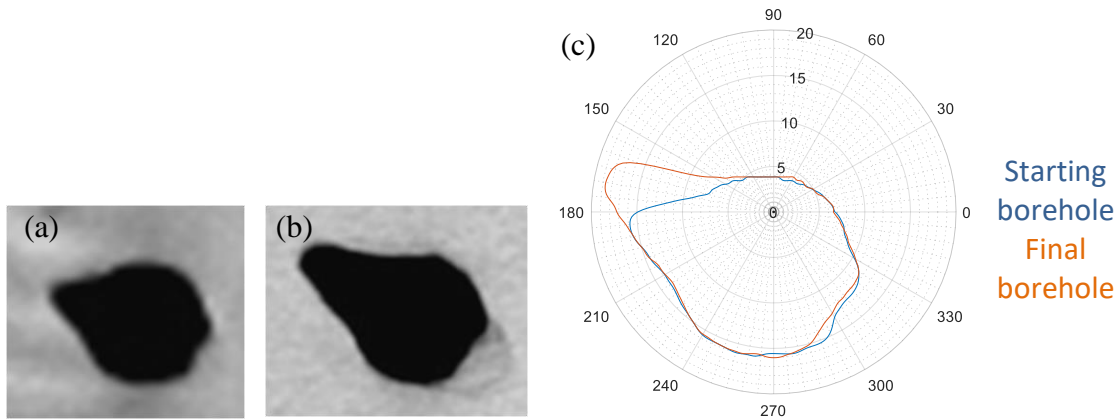
Figure 28. (a) CT-scan-1. (b) CT-scan-2. (c) superimposing the radii measurements of both CT-scan-1 and CT-scan-2 to highlight the region of borehole failure.

2.2.4.1 Angle to minimum stress direction (β_{Node})

As expected, shear failure of the wellbore becomes more likely close to the minimum stress direction. Each point or 'node' on the radius curve is assigned a value that describes its distance to the 'nearest' minimum stress direction in degrees. For example, if the minimum stress direction is $0°$-$180°$, then at the circumferential location of $150°$ Beta is equal to $30°$. The smaller this value of Beta, the more likely for shear failure to occur at that location. While this parameter is not necessarily extracted from the image of the borehole, it is used for normalizing the other geometric shape indicators that are actually extracted from the borehole image.

2.2.4.2 Radius of curvature (RC)

It was observed from the true-triaxial experiments that are detailed in the next chapter that failure is more likely to occur on concaved locations of a non-uniform borehole. To provide a quantification of this convexity, the radius of curvature is calculated. The radius of curvature is calculated at each point of the borehole curve based

on the geometrical theorem that the radius of a circumscribing circle of a triangle is equal to the product of its three sides divided by four times the triangle's area. Therefore, at each point on the borehole curve as interpreted from an image, a preceding and a succeeding point are picked to define the curve fragment of the borehole. Using these points, the radius of curvature (RC) is calculated using the following formula:

$$RC = \frac{sum\ of\ triangle\ sides\ lengths}{4 \times triangle\ area} \tag{58}$$

The radius of curvature is calculated either once for every available point on the borehole curve or only for a pre-determined set of discretized curves for further smoothing as shown to the left side in **Figure 29**. Also, this parameter is normalized with respect to its proximity to the minimum stress direction using the Beta angle. The normalized radius of curvature is shown to the right side in **Figure 29** where it is plotted along with r_d, which is the magnitude of radial wellbore failure at each specific circumferential location in the borehole. An illustration of the r_d variable is shown in **Figure 30**.
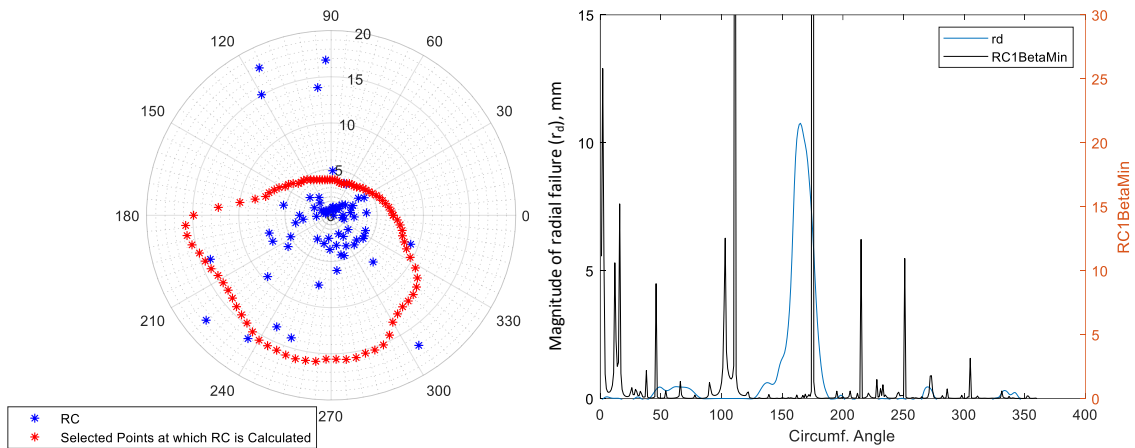
Figure 29. Left: the red points are the points of the borehole curve used to calculate the radius of curvature (RC) and the blue points are the calculated radius of curvature. Right: the normalized radius of curvature plotted along with r$_d$.
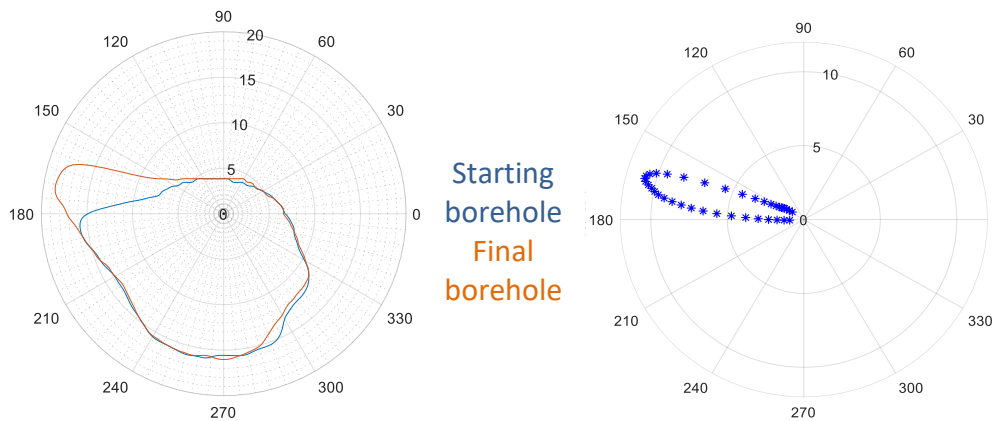
Figure 30. The definition of the r$_d$ variable, which is the magnitude of radial borehole failure at each specific circumferential location in the borehole. Left: superimposing the radii measurements two CT-scans to highlight the region of borehole failure. Right: the r$_d$ variable showing a mapping the location and magnitude of borehole failure points.

2.2.4.3 First derivative of the borehole curve radius (r$_w$')

This simply is the rate of change of the borehole radius around the circumferential angle. This definition means that regions that have a radius above that of the gauged

60

borehole will also have a high value of the first derivative. Observations made in the true-triaxial experiments that are detailed in the next chapter showed that failure, generally, favors regions of the borehole that has a higher radius than its two neighboring regions.

$$r'_w = \frac{\Delta r}{\Delta \theta} \tag{59}$$

The first derivative of the radius of the example borehole curve from the CT-scan is shown in **Figure 31** where it is plotted along with r_d. As can be seen in this plot, the regions of the non-uniform borehole that has a high first derivative value also experienced more failure. This shows the statistical value of using this particular geometric shape indicator as a precursor to borehole failure within the machine learning algorithms.
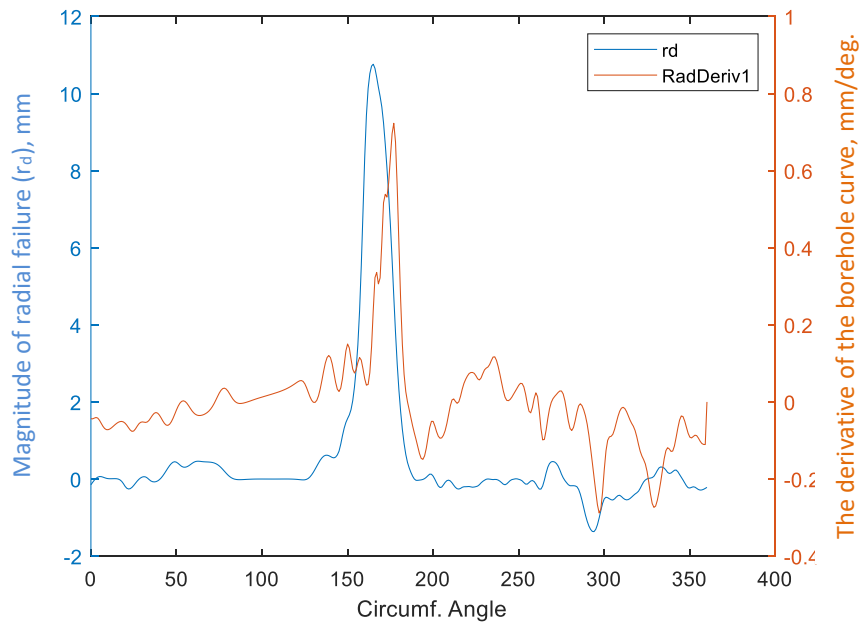
Figure 31. The first derivative of the borehole curve radius is plotted on the y-axis to the right, along with the magnitude of radial borehole failure r_d, which is plotted on the y-axis to the left.

2.2.4.4 Curve face orientation angle with respect to min stress direction (CFO)

This geometric shape indicator describes the angle between each line segment $\vec{L}$ of the borehole curve, which is defined by two consecutive points on the curve, and the direction of the minimum stress. This angle is always measured clockwise. This is done by calculating a normal vector $\vec{n}$ to the line segment and then calculating the angle between the normal vector and the minimum stress direction vector $\vec{s}_h$ as shown in **Figure 32**.
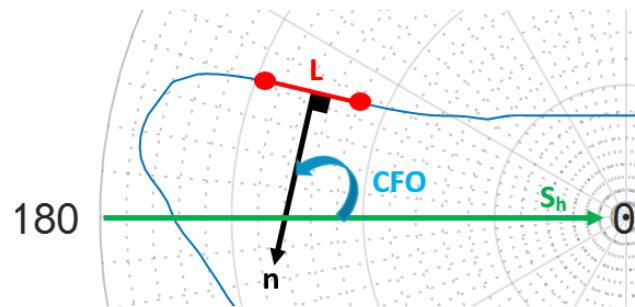
Figure 32. Procedure used to calculate the curve face orientation (CFO) feature around an irregular wellbore

The curve face orientation angle with respect to the minimum stress of the example borehole curve from the CT-scan is shown to the left side in **Figure 33** where it is plotted along with r_d. The right side of this figure shows the plot of the CFO in polar coordinates to be able to correlate its values with the actual shape of the borehole in the CT-scan. The cartesian plot of CFO in **Figure 33** illustrates the manner with which the CFO curve behaves around regions of the non-uniform borehole that experienced more failure. Capturing this correlation in the change of the curve trends between the CFO and the r_d variables is the role of the machine learning algorithms.
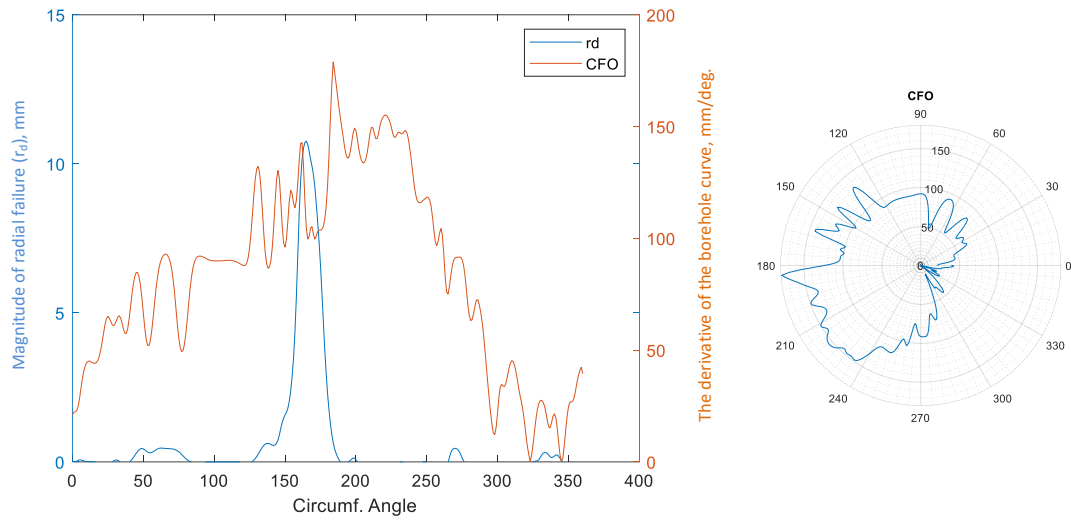
Figure 33. Left: curve face orientation (CFO) is plotted on the y-axis to the right, along with the magnitude of radial borehole failure r_d, which is plotted on the y-axis to the left. Right: a plot of the CFO variable in polar coordinates.

2.3 Development of the Meshing Function for 3D Structures and Non-uniform Boreholes

There are numerous open source and commercial tools for creating meshes for finite element analysis. These tools, such as SolidWorks, Abaqus, and ANSYS among others, allow the user to create meshes of complicated structures by relying on drawing tools. For the specialized purpose of drilling geomechanics modelling, it is common for researchers to build their own in-house codes for creating a mesh of a wellbore and the formations it penetrates. These in-house meshing tools are routinely simplified by relying on the wellbore symmetry assumption discussed in the previous section. The symmetry assumption provides simplicity in these meshing tools either by creating a mesh of a

quarter of the wellbore only or by using the coordinate of the quarter of the wellbore and manipulating the signs to produce a mesh and coordinates for a full wellbore.

None of the above-mentioned meshing methods provide the ability to create a mesh of a fully non-uniform boreholes. Also, as one of the objectives of this work is to enable the final integrated model to run automatically in while-drilling settings, it is necessary for the meshing and re-meshing tool to be able to create the mesh of the wellbore based on the most recent image log received without the need for user intervention. These new requirements mean that a new function that creates comprehensive meshes of non-uniform wellbores as interpreted from images is necessary. The development of this function is detailed in this section.

2.3.1 Creating Meshes with Non-uniform Borehole Coordinates and Other Wellbore

Features Based on Image Analysis

Once the radii measurements of a borehole are produced by the image analyzer described in the previous section, the meshing function receives these measurements and calculates the nodal coordinates of the mesh based on them. This is done by assigning a separate borehole radius measurement to each circumferential angle around the borehole. Then, the coordinates of each single line of nodes associated with each circumferential angle is calculated with a special care paid to whether the line of nodes is placed at the center or at the boundary of an element in the mesh. This process is repeated for each radial line of nodes at all the specified vertical layers until the full coordinates describing the borehole or the wellbore and the formations it penetrates is fully calculated. The

meshing function can perform this process based on lab images, CT-scans, ultrasonic image logs, and resistivity image logs as shown in **Figure 34**.
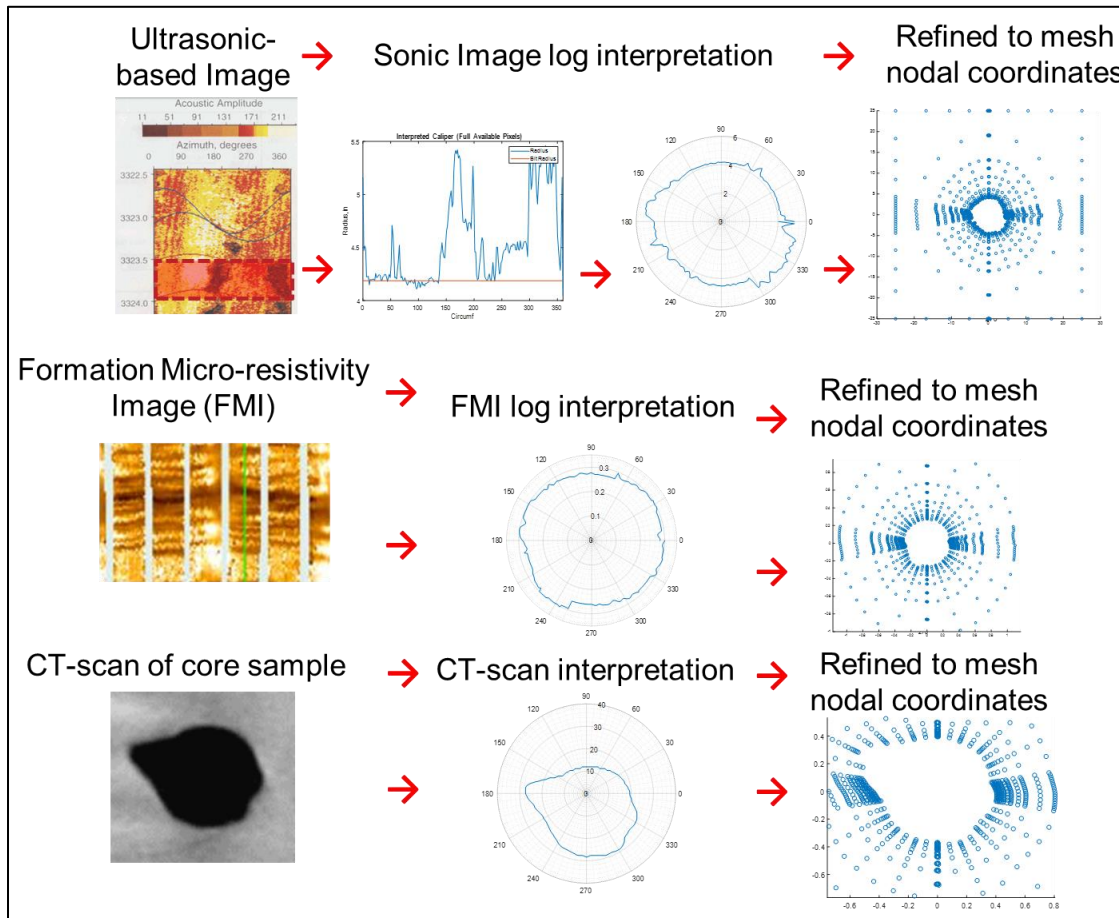
Figure 34. An illustration of the process of converting the borehole shape interpretations from field and lab images into a mesh nodal coordinate.

This meshing function also works to reflect the variation of the borehole shape along the axis of the borehole. For core samples, the mesh function reflects the different borehole shapes as described by CT-scans that are taken at different cross-sections of the core. The same goes for field-based images as the mesh can reflect the different shapes of

the wellbore as described by the image log across different depths. An illustration of this function based on CT-scans is shown in **Figure 35**.
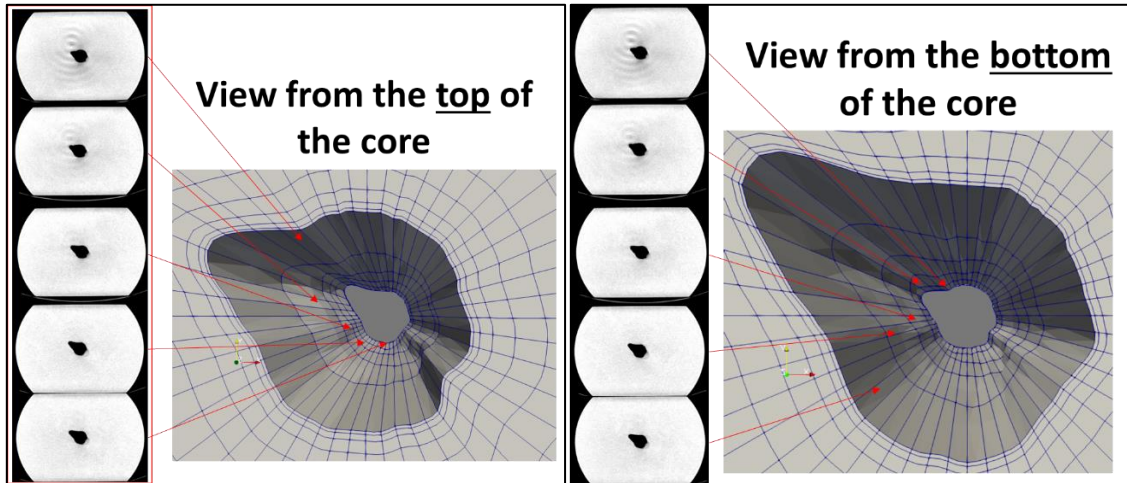
Figure 35. An illustration of the meshing function ability to reflect different borehole shapes based on CT-scan taken across different cross-sections of the core sample.

Another function of this meshing code is the reflection of different wellbore features, which is specific to field-based meshes. One wellbore feature considered is key-seating, which is the enlargement of one side of the wellbore due to having the drillstring rest on that side. The meshing code can reflect the dimensions of the key-seat enlargement on the coordinate of the wellbore. An illustration of the definition of key-seating and the results of the meshing code are shown in **Figure 36**.
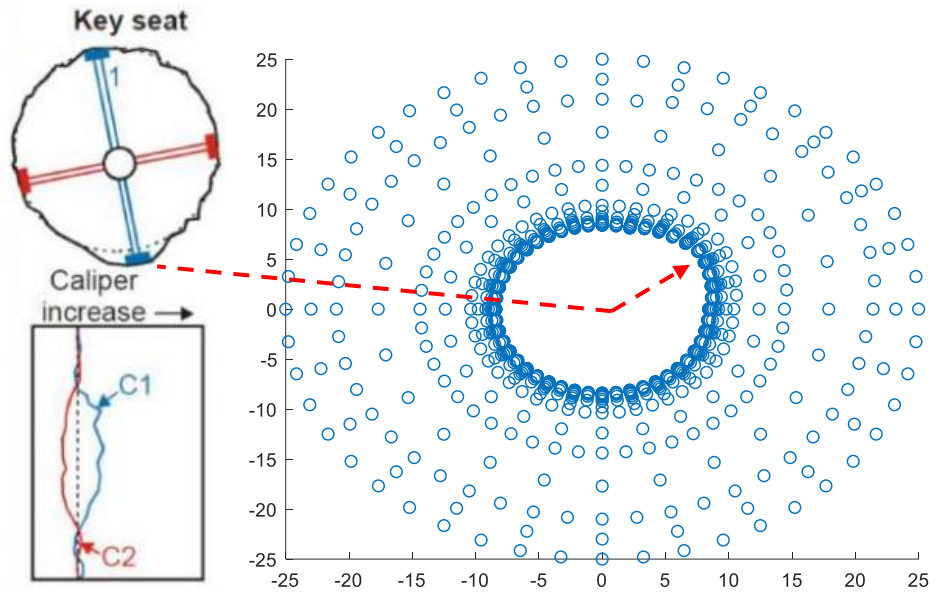
Figure 36. Left: the definition of key-seating. Right: the results of the meshing code showing the one side enlargement in the mesh nodal coordinates. The illustration of the key-seat to the left side is from (reprinted from Plumb and Hickman 1985).

One more wellbore feature that can be reflected by the meshing function is a fracture. The mesh function requires that the fracture aperture, length, tortuosity, and placement within the wellbore to be provided so that the geometry of the fracture is reflected. Based on this, the mesh will consider the fracture a part of the wellbore, which makes it exposed to the wellbore pressure. This way of meshing a fracture essentially means that the fracture is assumed to be a static feature of the wellbore. While this assumption is not useful for fracture propagation problems, it is useful for this work as it can be used to reflect the effect of drilling induced fractures, natural fractures, lost circulation, and wellbore strengthening on the drilling window (Feng et al. 2015). An illustration of the meshing a static fracture is shown in **Figure 37**.

67
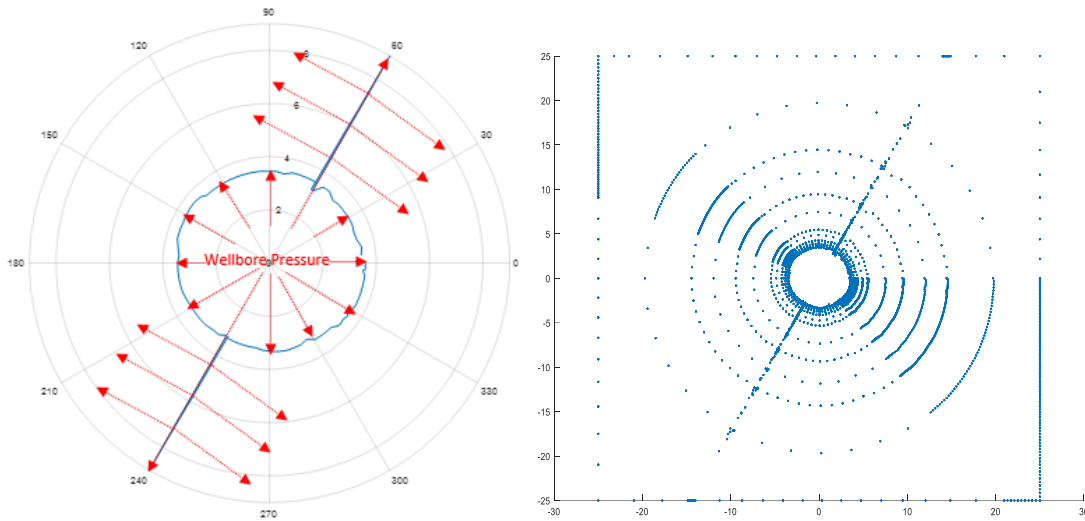
Figure 37. Left: the introduction of fractures to a non-uniform wellbore and a depiction of the wellbore pressure acting on both the wellbore wall and the fracture face. Right: the placement of nodal coordinate for the mesh of fractured non-uniform wellbore.

One final consideration for the mesh function is the shape of the structure boundaries. When creating a mesh for the geomechanics finite element model described in this work, it is necessary to specify the type of boundaries being used and reflect that on the nodal coordinates calculated. The mesh function can consider a finite and an infinite mesh boundary. The finite boundary means that the mesh elements at boundary of the structure, which will be subjected to the in-situ stresses as described previously, are all fully formed 20-node brick elements. As for the infinite boundary, the mesh nodal coordinates created for those consider a 12-node element, which are basically a 20-node brick element with one of the cube faces removed. It is worth to be noted that for these infinite elements, different shape functions will be used in the finite element model. The difference between these two boundaries as produced by the mesh function is shown in **Figure 38**.

68
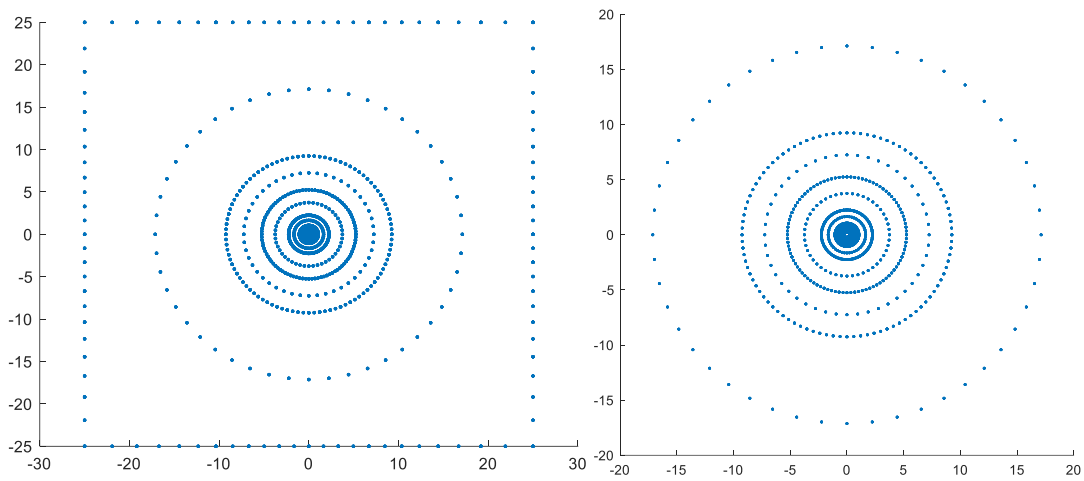
Figure 38. Left: plot of nodal coordinates of a mesh with finite boundary elements (20-node brick elements at the boundary). Right: plot of nodal coordinates of a mesh with infinite boundary elements (12-node infinite elements at the boundary).

The setup for the structure boundaries within the mesh needs to be altered when modelling core samples in experiments instead of wells in the field. A specialized meshing function is created to model the true-triaxial experiments that are discussed in the next chapter. As these core samples have round boundaries on two side and straight slabbed boundaries on the other two sides, the mesh function places finite element in all sides but in manner that mimics the core boundary shape. An illustration of the mesh created for these core samples is shown in **Figure 39**.
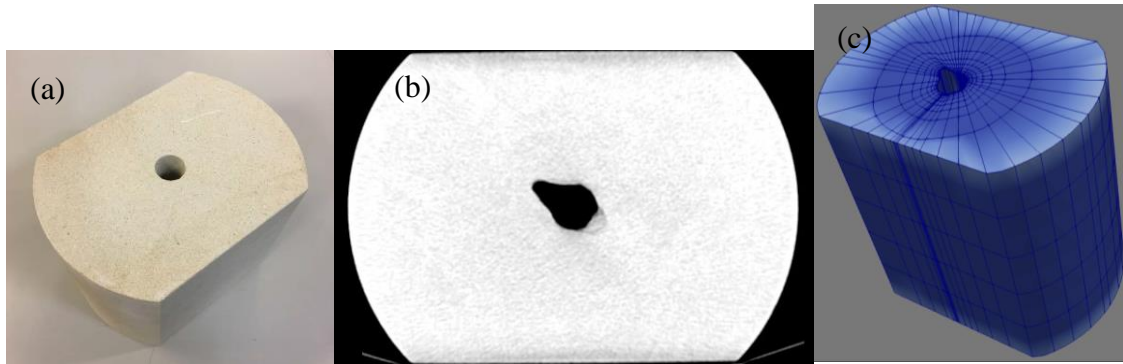
Figure 39. (a) An image of the core sample to be meshed. (b) a CT-scan at a cross-section of the post-test core sample. (c) The mesh of the core sample as produced by the mesh function showing boundaries and wellbore shape resembling those in the core image and CT-scan.

2.4 Development of a New Risk-Controlled Wellbore Stability Function

In many drilling scenarios that include deep wells and highly stressed environments, the mud weight required to completely prevent wellbore instability can be unpractically high. In such cases, what is known as risk-controlled wellbore stability criterion is introduced. This type of criterion allows for a certain level of shear failure and wellbore instability to take place. This means that the mud weight calculated using this criterion will only constrain wellbore instability to a certain manageable level, hence the name risk-controlled. Conventionally, the allowable level of wellbore instability in this type of models has always been based on the magnitude of the breakout angle (Furui et al. 2010, Kanfar et al. 2015, Ma et al. 2018), which is illustrated in **Figure 40**.
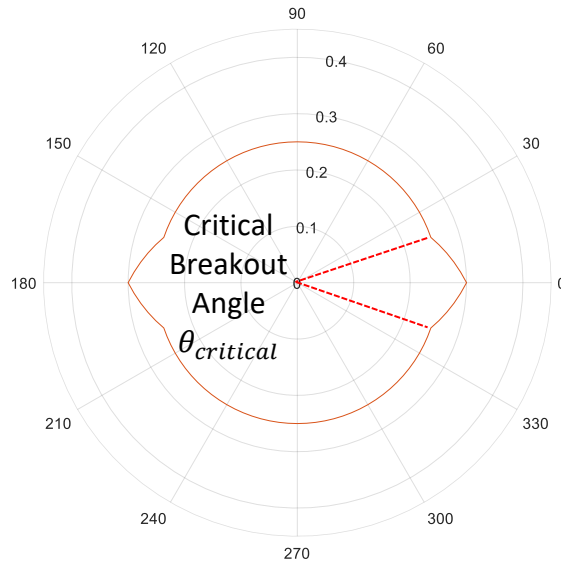
Figure 40. An illustration of the critical breakout angle used in risk-controlled wellbore stability models.

There are several methods proposed in the literature for determining the magnitude of the critical breakout angle. One in example in (Zoback 2007) where it is suggested that the wellbore will become unstable when the breakout angle exceeds 90°. A more sophisticated method based on drilling experience is used in (Morita and Nagano 2016), where the critical breakout angle to decide the threshold of wellbore instability is calculated based on the well section inclination angle γ as follows:

$$\theta_{critical} = 30^o + 60^o sin(\gamma), (\gamma = 0^o \ for \ horizontal \ wells) \tag{60}$$

However, the formulation of the integrated model in this work both calls and allow for a more sophisticated method of setting the wellbore instability risk-control parameter. Considering that the integrated model can predict shear failure zones that are non-uniform and non-symmetric across the sides of the wellbore, relying on the critical breakout angle as the risk-control parameter is no longer viable in most cases. This issue is more

prominent when examining the stability of a wellbore that is non-uniformly shaped to begin with. Another issue with considering the critical breakout as the constraint is that it only considers the circumferential extension of the instability zone while not paying similar consideration to the radial extension of the instability zones away from the wellbore wall. To improve on this and considering that the integrated model developed in this work is three dimensional with the ability to predict non-uniform failure, it is proposed to use shear failure zones mapping and cavings volume as the risk-control parameter. This prediction of cavings volume can be examined against a suitable hydraulics and mud rheology model to determine the critical cavings volume that can be tolerated without causing drilling troubles such as stuck pipe incidents, which is a more nuanced method of risk controlling wellbore instability.

The method employed here begins with extracting failed nodes at the end of a loading step in the integrated model and calculating the volume enclosed by these nodes to estimate the cavings volume. Since the distribution of failure zones predicted by the integrated model can be non-uniform and highly irregular, simplified symmetry rules cannot be used to calculate the area enclosed by failed nodes at each layer. Therefore, the implemented function will receive the information of the failed nodes and draw a polygon around each discrete failure region in each nodes layer. Then the area of each polygon is calculated by dividing it into discrete uniform geometric shapes. The definition of each polygon is based on the failed nodes information and on the size of the mesh. For example, if a group of failed nodes are surrounded by non-failed nodes, the failed nodes will be considered as a separate polygon, and hence, separate caving fragment. On the other hand,

72

if the failed zones are all interconnected to each other around the wellbore with no non-failed nodes intersecting them, a single polygon is defined around the wellbore to determine the failed region area. An example of both cases in shown in **Figure 41**.
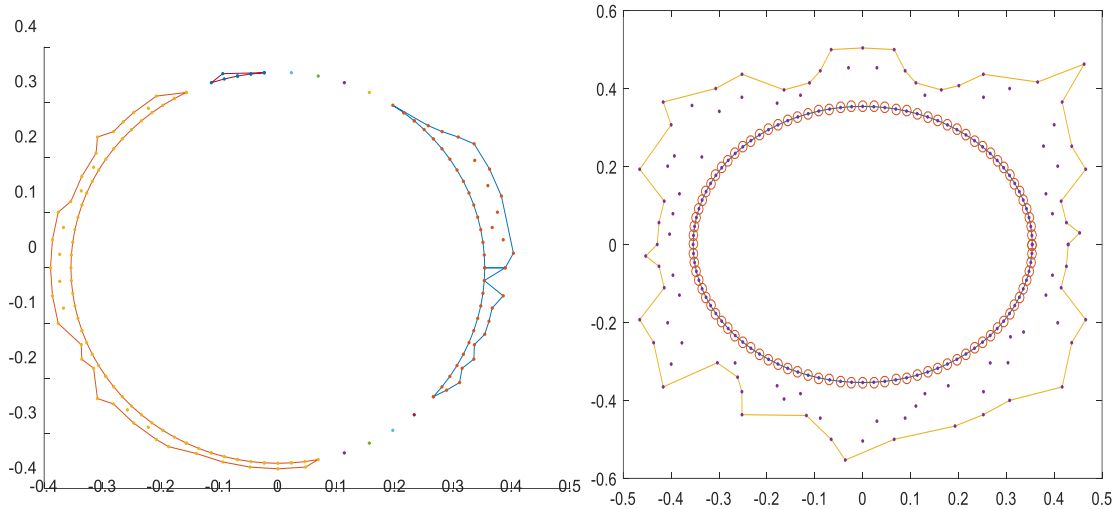
Figure 41. The outcome of the caving volume calculator function based on the extracted failed nodes around the wellbore. Left: separate polygons defining discrete failure zones or separate cavings fragments. Right: single polygon defining full wellbore failure/collapse.

After calculating the total area enclosed by the failed nodes across each layer, the volume is then calculated by multiplying the total area in a single nodes layer by half the thickness (or depth) of the corresponding layer from the mesh. Half the thickness of the layer is used here because the mesh layer consists of 20-node brick element, which is divided in two parts by three layers of nodes.

2.5 The Integration Process of the FEM and the Machine Learning Algorithms

Machine learning algorithms are primarily used in this work to improve the prediction of rock failure limits. As discussed previously, rock failure for drilling geomechanics applications is conventionally assessed using published failure criteria in conjunction with either an analytical or a numerical solution. However, recent technological developments in the industry in the area of data acquisition along with the ever-growing popularity of the use of machine learning present new potential solutions to many chronic problems. In terms of data acquisition, many logging while drilling tools have the ability to transmit density-based images in real-time every instance the drillstring is run in or out of hole. Since these tools are only usually run in reservoir sections, for non-reservoir sections, recent practices and advancements in shale shakers surveillance can provide valuable information about the changing state of the wellbore rock also in real-time while drilling (Du et al. 2020). By exploiting this data and employing machine learning algorithms through the proposed framework of physic-based data science, it is possible to significantly improve the way rock failure is predicted.

To be able to train the ML algorithms and produce reliable and useful predictions regarding rock failure, there needs to be a link between the ML algorithms and the FEM along with a clear definition of data that will be fed into the algorithm as a predictor dataset. The link between the FEM and ML along with the different formulations of the predictors datasets as required for different scenarios are the main topics that are discussed in this section.

2.5.1   Integration Process of FEM-ML Through Image Analysis

As mentioned in previous sections, the image analyzer is the link between the FEM and the machine learning algorithms (ML). Image analysis provides this link by classifying nodes failure state as observed from experiments or wells, which creates the necessary framework for a supervised training process of the machine learning algorithms. The training process is said to be supervised here because the outcome in terms of rock failure in the training dataset that is produced from offset wells or previous experiments is known through the observations made from images. Therefore, the machine learning algorithm will be trained using this dataset in a supervised manner to enable it to predict the outcome of rock failure in new wells or experiments whenever their predictors datasets are available.

The integration process begins by modelling the borehole in the FEM. Then, all the relevant stress, strain, displacement, and geometric shape indicators for all nodes from the FEM post processing are extracted and stored in the training dataset. The same step is repeated for all the available boreholes that were subjected to different conditions in terms of stress and borehole pressure and their nodal information is added to the same training dataset. Eventually, the final cumulative training dataset produced is used as the predictor dataset in the ML algorithm. As for the output variable, which is basically the node failure classification, it is produced by comparing the original borehole mesh to the final failed borehole mesh as interpreted from an image to filter out the failed nodes. The process for producing nodal failure classification from image interpretation is illustrated in **Figure 42**

where the nodes that are still visible in the borehole depicted to the right are the nodes that

will be classified as failed ones for the purpose of the supervised training process.
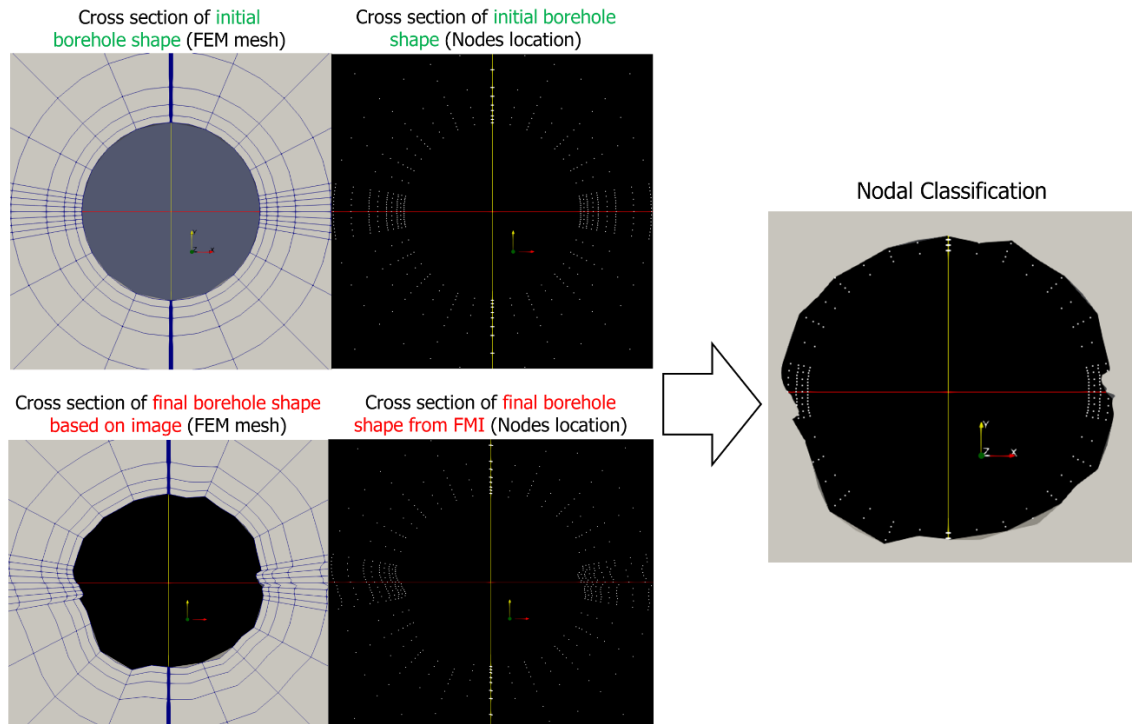
Figure 42. An illustration of the process undertaken to integrate the finite element model and the machine learning algorithm through nodes failure classification as interpreted from an image, where the nodes that are still visible as white dots in the wellbore to the right are classified as failed nodes.

The integration process is concluded by training the ML algorithm using the

dataset produced from the FEM and the image analysis along with the output variable

describing nodes failure classes. The trained model is eventually used to produce an

improved prediction of failure as depicted in **Figure 43**.
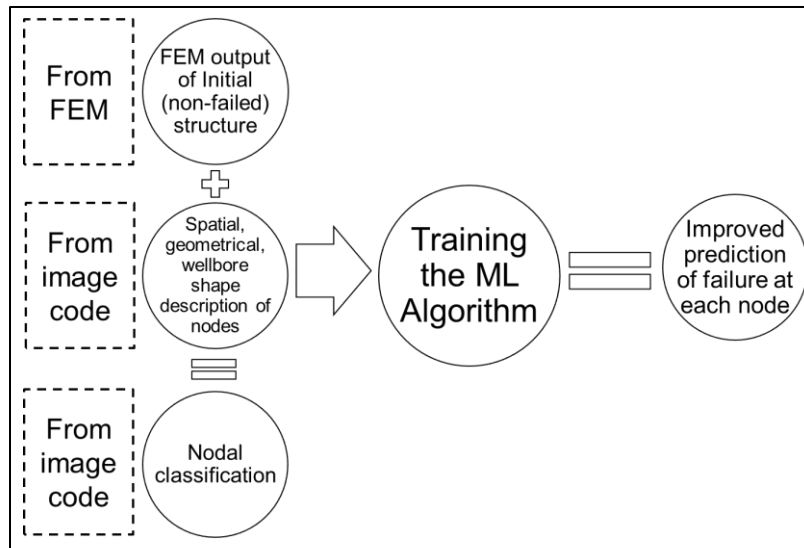
Figure 43. An illustration of the integration process between the FEM and ML.

2.5.2 Overview of the Used ML Algorithms and Cross-fold Validation Process

There are different machine learning algorithms used in this work. The selection of the suitable algorithm depends on its performance in terms of nodes failure classification accuracy. As detailed in the field applications chapter, a certain algorithm will yield higher accuracy for a certain application due to the nature of the training dataset produced for that application. Therefore, for each new training dataset produced from experimental or field data, several algorithms are trialed to determine the most suitable one. The algorithms that were routinely trialed on the datasets in this work are bootstrap aggregations of classification decision trees (also known as tree bagging), gradient boosting of classification decision trees, artificial neural networks, random forest, and support vector machines. Only the first three algorithms are overviewed in the next sections as they routinely produced the highest classification accuracy on all the datasets

produced for this work. Also, the process undertaken to train the algorithms and metrics used to evaluate their accuracy are overviewed in the next sections.

2.5.2.1 Classification Algorithms Accuracy Evaluation

The nature of all the datasets produced in this work, from both experimental work and field data, dictates that using a simple accuracy percentage as the metric for evaluating an algorithm performance is not sufficient. This is because the classification in output variable in all training datasets is heavily skewed towards the non-failed mesh nodes class. This is to be expected since the datasets are produced from several single well models that exhibit failure that is concentrated in the mesh nodes surrounding the wellbore, which means that the majority of nodes, which are placed away from the wellbore, are classified as non-failed nodes. The result of this skewed distribution of failure classes is that datasets can contain as little as 5% of nodes that are classified as failed nodes. Based on this, if an algorithm were to be predict that all nodes are non-failed nodes, its percentage accuracy can be as high as 95%. This is a high and an attractive accuracy for an algorithm that is objectively poor as it missed predicting all of the failed nodes, which are the main nodes of interest in this application. Therefore, other metrics for evaluating classification accuracy are required to enable the selection of the most suitable algorithm.

In data science, to provide a more representative measure of accuracy on skewed binary-class datasets, the confusion matrix and the parameters associated with it are used. Relying on the confusion matrix parameters allows for evaluating an algorithm ability to predict a certain class, such as the failed nodes class. The confusion matrix is illustrated in **Table 1**.

78

Table 1. Confusion matrix description.

		Predicted Classes	
		Non-failed Node (-1)	**Failed Node (1)**
Actual	**Non-failed Node (-1)**	True Negative (TN)	False Positive (FP)
Classes	**Failed Node (1)**	False Negative (FN)	True Positive (TP)

There are several parameters that can be based on the values of the confusion matrix to help better assess the performance of the algorithm. These parameters are:

$$Accuracy = \frac{TP + TN}{TP + FN + TN + FP} \tag{61}$$

$$Recall\ (True\ Positive\ Rate\ or\ TPR) = \frac{TP}{TP + FN} \tag{62}$$

$$Specificity\ (True\ Negative\ Rate\ or\ TNR) = \frac{TN}{TN + FP} \tag{63}$$

$$False\ Positive\ Rate\ (FPR) = \frac{FP}{FP + TN} \tag{64}$$

$$Precision\ (Positive\ Predictive\ Value, PPV) = \frac{TP}{TP + FP} \tag{65}$$

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{66}$$

One final parameter is produced by plotting the TPR from Equation (62) versus the FPR from Equation (64), which produces the Receiver Operating Characteristic (ROC) curve, and calculating the area under the curve. This parameter is denoted as the area under the ROC curve or AUROC and it evaluates the algorithm ability to provide a distinction between the two classes (James et al. 2017). In this work, the F1-score and the AUROC are the main metrics used for evaluating the classification accuracy of the trialed algorithms.

2.5.2.2 Multi-fold Outer and Inner Cross Validation

When training an algorithm based on a specific training dataset, there's a need to ensure that the trained model is generalizable. This means that the trained model should not be allowed to overfit on the training data, which can lead to poor classification performance when this trained model is used to generate prediction on new data. Two measures are taken to avoid overfitting and to ensure a generalizable model is generated. The first measure is to keep at least a single well or a single experiment data as the test dataset by not including them in the training dataset. Once the model is trained, its performance is then evaluated by how well it classifies the mesh nodes in the test dataset. The second measure is employing a multi-fold outer and inner cross validation process. In this process, the training dataset is split into a number of subsets or folds, which are known as the outer folds. Each outer fold is then split into a number of subsets of its own, which are known as the inner folds. The algorithm is then trained on all the inner folds except one, which is used to test the predictive performance of trained model. The process is repeated by alternating the inner fold used as the test. This inner validation is also repeated by alternating the outer fold used for creating the inner folds. In this work, the number of outer folds used are 5 and the number of inner folds ranges between 30 to 50 depending on the size of the dataset. An illustration of this process is shown in **Figure 44**.
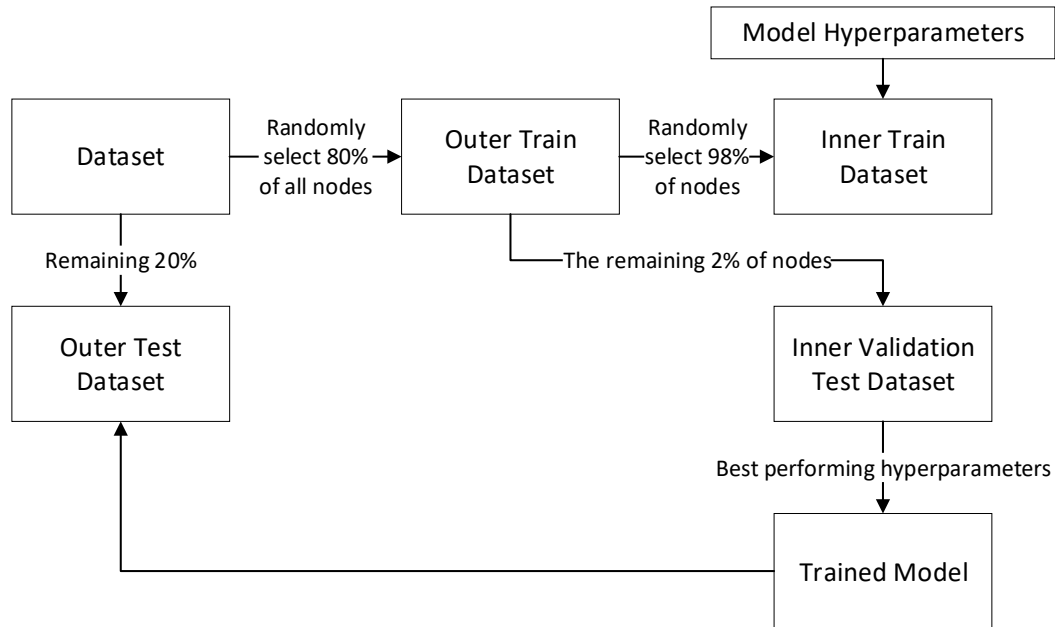
Figure 44. The multi-fold cross validation process employed in training the machine learning algorithms.

2.5.2.3 Bootstrap Aggregations of Classification Decision Trees (bagging)

This algorithm relies on decision trees as basis for classification. A single decision tress partitions the features space into m regions and makes the same classification decision for each mesh node that falls in the same region. By doing this, classification decision trees seek to minimize the classification error, which is:

$$E = 1 - max_k(\hat{p}_{mk}) \tag{67}$$

Where E is the classification error and $\hat{p}_{mk}$ is the probability that a mesh node from the training dataset is in the m region of the feature space from the k class. The features space in classification decision trees is split into regions based on one of two parameters, which are the Gini index and the Entropy. These are described as follows:

$$Gini\ index = \sum_{k=1}^{K} \hat{p}_{mk}(1 - \hat{p}_{mk}) \tag{68}$$

$$Entropy = \sum_{k=1}^{K} \hat{p}_{mk}\ log(\hat{p}_{mk}) \tag{69}$$

Where the Gini index evaluates the total variance across the two classes of mesh nodes failure, which is essentially a measure of the purity of each region in the feature space. For example, if there are a group of mesh nodes that fall within the same group or region due to their similar FEM output and geometrical data (i.e. features) and if these mesh nodes are all non-failed nodes, then this group is said to be pure and it has a $\hat{p}_{mk} = 1$ and a Gini index of zero. As for the Entropy, it has a similar interpretation where it will be zero for a pure group of mesh nodes. In this work, both of these are trialed when construction a decision tree-based model.

Once a tree is grown to make the required splits and defined regions that are as pure as possible, a classification decision is made for each mesh node. However, allowing the tree to grow and purify the regions as much as it can lead to overfitting and poor generalization of the model. To avoid this, the bootstrapping aggregation (bagging) algorithm is introduced. In trees bagging, trials are made by building several trees of the same depth and then having each tree vote on the failure class of each mesh node. This works because individual trees can vary from each other greatly. The number of trees built and the allowed depth of each tree are known as the hyperparameters of the algorithm. The training process is performed using the multi-fold cross-validation process with different hyperparameters and the best performing model is selected based on the accuracy

measure as shown in **Figure 45**. The implementation of this algorithm is done in Python

through the BaggingClassifier function from the sci-kit-learn public library.
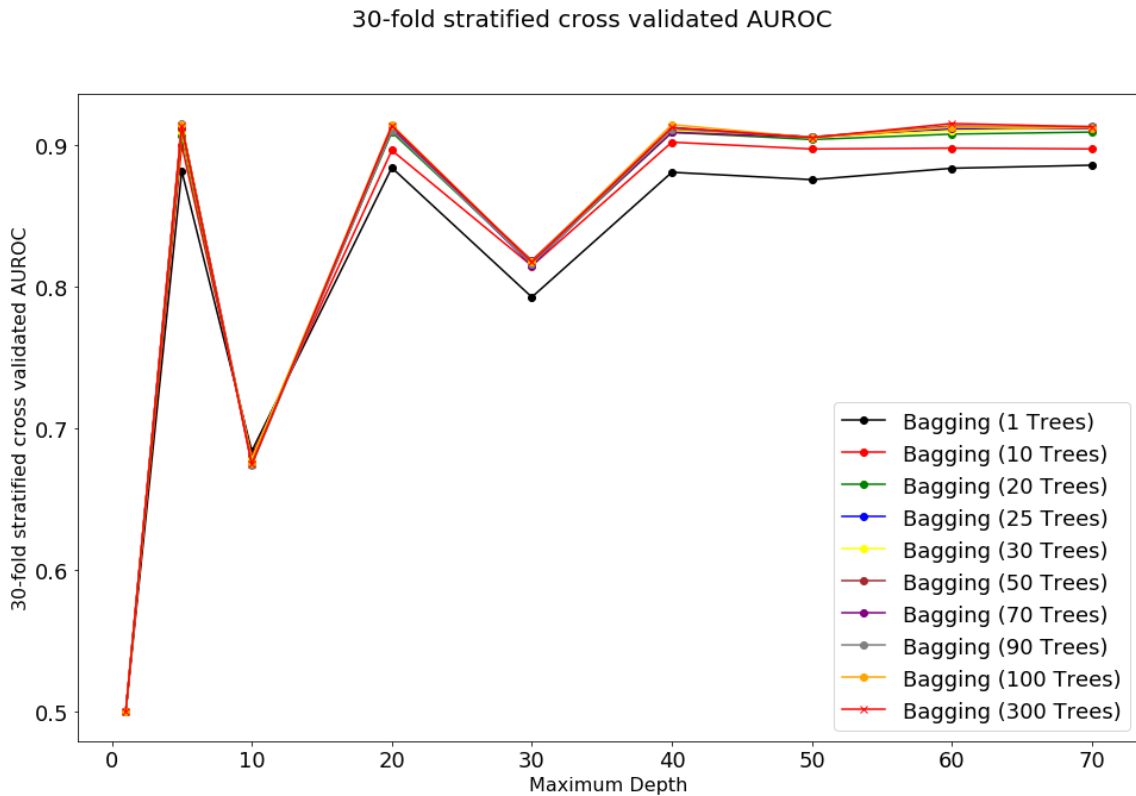
Figure 45. An example of the results of the multi-fold cross validation process for training a tree bagging model based on data from the field. Each data point represents a model that was trained with different hyperparameters.

2.5.2.4 Gradient boosting of classification decision trees

The boosted trees algorithm is considered because it offers two advantages over

bagging. The first one is that it allows for evaluating the importance of the features, which

can provide valuable insight on the factors that are affecting the mesh node failure the

most. The other advantage is that instead of building trees randomly, boosting builds tree

sequentially. This is done by first building a shallow classification tree and evaluating its error, then the next tree is built to improve on this error iteratively. The cross-validation process is performed using the same process used in trees bagging where the hyperparameters are trialed on the training dataset to find the best performing model as shown in **Figure 46**. The implementation of this algorithm is done in Python either through the AdaBoostClassifier function from the sci-kit-learn public library or through the XGBClassifier function from the xgboost public library.

		Max Depth of Trees					
		1	5	10	20	30	40	50
Number of Trees	1	0.5	0.88941454	0.98401464	0.99834578	0.99799514	0.998398	0.99854428
	5	0.5	0.90028199	0.99536496	0.99786867	0.99844467	0.99839639	0.99856813
	10	0.5	0.91436625	0.99838234	0.99841541	0.99827139	0.99859263	0.99849051
	20	0.5	0.94922655	0.99932039	0.99841414	0.9979138	0.99834719	0.99818843
	30	0.5	0.96745688	0.99954828	0.99834424	0.99824535	0.9980228	0.99839324
	40	0.5	0.9751295	0.99952378	0.99794159	0.99806981	0.99872009	0.99844333
	50	0.5	0.9804964	0.99934818	0.99814245	0.99849219	0.9980988	0.99836867

AUROC

Figure 46. An example of the results of the 50-fold cross validation process for training a tree boosting model (AdaBoostClassifier) based on data from experiments. Each colored cell point represents the AUROC for a model that was trained with different hyperparameters.

2.5.2.5 Artificial neural network

Artificial neural networks receive the features of the mesh nodes, with each feature parameter denoted as x_i for a total of D features, and then assigns a cell to each feature while adding extra cell as an intercept with x_0. These cells containing the features form what is known the input layer. The algorithm then assigns weights W_{ij} to each connection

84

between an input layer cell and a hidden layer cell, where the hidden layer contains M

cells with j as their index. The quantity in each hidden layer cell is determined as follows:

$$z_j = \sum_{i=0}^{D} (W_{ij}^{(1)} x_i) = W_{0j}^{(1)}(+1) + W_{1j}^{(1)} x_1 + \cdots + W_{ij}^{(1)} x_i + \cdots + W_{Dj}^{(1)} x_D \qquad (70)$$

Then, new weights denoted as $\alpha_j^{(2)}$ are assigned to each connection between a cell

in the final hidden layer and the output perceptron. These weights are determined based

on an activation function, which can be selected from a group of commonly used functions

for this purpose. An example of such function is the sigmoid:

$$sigmoid(\alpha) = s(\alpha) = \frac{1}{1+e^{-\alpha}} \qquad (71)$$

The hidden layers weights are calculated using the sigmoid function as follows:

$$\alpha_j^{(2)} = s(z_j) = s\left(\sum_{i=1}^{D} \left(W_{ij}^{(1)} x_i \right) + b_j^{(1)} \right) \qquad (72)$$

Therefore:

$$\alpha_j^{(2)} = \frac{1}{1 + exp(-W_{1j}^{(1)} x_1 - \cdots - W_{ij}^{(1)} x_i - \cdots - W_{Dj}^{(1)} x_D - b_j^{(1)})} \qquad (73)$$

Finally, the output perceptron for the classification decision for each mesh node is

determined as:

$$h_{W,b}(x) = \alpha_1^{(3)} = f\left(\sum_{j=1}^{M} (W_{1j}^{(2)} \alpha_j^{(2)}) + b_1^{(2)} \right) \qquad (74)$$

The structure of an illustrative neural network is shown in **Figure 47**.
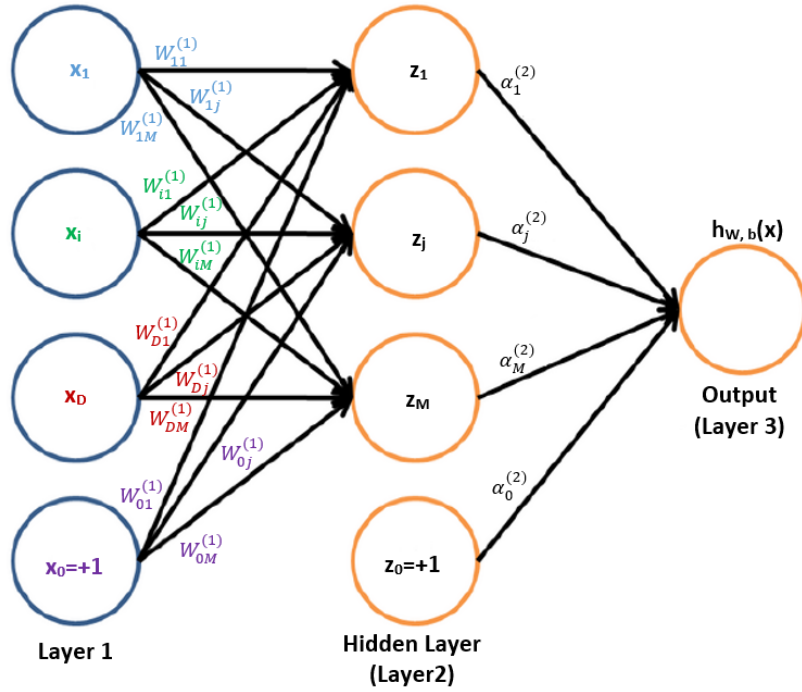
Figure 47. The structure of an example artificial neural network with symbols pertaining to the mathematical formulations discussed.

Now, the weights ($W_{ij}^{(1)}$ and $\alpha_j^{(l)}$) that will minimize the error in classifying the mesh nodes failure will need to be determined. This is done through a process of optimization using a gradient descent to find the minimum error in the cross-entropy function:

$$\epsilon(w) = -y_n \log[s(w^T x_n)] - (1 - y_n) \log[1 - s(w^T x_n)] \tag{75}$$

Where y_n is one instance of the output perceptron value, x_n is one instance of a feature, and w^T is the transpose of the connection weights matrix. The cross-entropy function is optimized using gradient descent as follows:

$$\frac{\partial \epsilon(w)}{\partial w_d} = (s(w^T x_n) - y_n) x_{nd} \tag{76}$$

Within the same process used for previous algorithms, a multi-fold cross validation process is performed with different hyperparameters by trialing different number of hidden layers, and different number of cells in each hidden layer along with other parameters. An example of the cross-validation process performed on lab experiments dataset is shown in **Figure 48** where four different hyperparameters are trialed. The implementation of this algorithm is done in Python through the TensorFlow public library.

#Nodes in hidden layer(s)	Dropout rate	Optimizer	F1 score		
			100th epoch	200th epoch	300th epoch
[8]	0	adam	0.927899687	0.95208463	0.96204107
[8]	0	rmsprop	0.936170213	0.953212726	0.958852868
[8]	0.25	adam	0.926799007	0.945611867	0.946395564
[8]	0.25	rmsprop	0.936517913	0.943231441	0.954403498
[8]	0.5	adam	0.933497537	0.964966195	0.974137931
[8]	0.5	rmsprop	0.942065491	0.947103275	0.948895899
[12]	0	adam	0.945794393	0.969734404	0.976514215
[12]	0	rmsprop	0.94959552	0.96457427	0.97149938
[12]	0.25	adam	0.933087784	0.964923077	0.969249692
[12]	0.25	rmsprop	0.955541641	0.972670807	0.975670618
[12]	0.5	adam	0.94081381	0.959158416	0.965346535
[12]	0.5	rmsprop	0.948092558	0.955927995	0.960845245
[8, 8]	0	adam	0.968381897	0.979490367	0.982630273
[8, 8]	0	rmsprop	0.955541641	0.962088254	0.969394129
[8, 8]	0.25	adam	0.951310861	0.965902046	0.967344424
[8, 8]	0.25	rmsprop	0.942602041	0.969811321	0.976426799
[8, 8]	0.5	adam	0.95890411	0.963467492	0.964956195
[8, 8]	0.5	rmsprop	0.943563729	0.937098845	0.948621554
[12, 8]	0	adam	0.96762142	0.97135741	0.97752809
[12, 8]	0	rmsprop	0.970825574	0.977061376	0.983219391
[12, 8]	0.25	adam	0.962088254	0.973074515	0.981998759
[12, 8]	0.25	rmsprop	0.969355847	0.970532915	0.978696742
[12, 8]	0.5	adam	0.952621605	0.956301457	0.959042218
[12, 8]	0.5	rmsprop	0.974968711	0.97882939	0.988220707
[16, 12]	0	adam	0.978193146	0.980876002	0.985776129
[16, 12]	0	rmsprop	0.969962453	0.980612883	0.9825
[16, 12]	0.25	adam	0.960501567	0.975761342	0.984558369
[16, 12]	0.25	rmsprop	0.978908189	0.986335404	0.98630137
[16, 12]	0.5	adam	0.971857411	0.981389578	0.986318408
[16, 12]	0.5	rmsprop	0.967418546	0.974358974	0.980124224

Figure 48. An example of the results of the 50-fold cross validation process for training an artificial neural network based on data from experiments. Each colored cell point represents a model that was trained with different hyperparameters.

2.5.3 Features List

In machine learning, and data science in general, the, nominally, independent variables, which form the predictor dataset are known as features. The features used in this work to train the ML algorithms are categorized into two main groups, which are the FEM-produced features and the image-analysis-produced features. **Table 2** lists those features along with their source. It should be noted that each separate mesh node, in all the modelled boreholes that are used for training, has a value assigned to it for each one of these features. This can mean that each mesh node can possibly have a completely unique set of features describing it. While this uniqueness can be useful in some formulations, it can also lead to suspect trained models in terms of failure prediction due to two issues, which are overfitting and lack of load dependency. Therefore, it is necessary to engineer these features for each training process depending on the nature of the dataset and the objective of failure prediction. These issues are discussed in the next section.

Table 2. Full list of the features available to train the machine learning algorithms

Feature	Source	Load Dependency
Principle stresses, $(\sigma_1, \sigma_2, \sigma_3)$	FEM post processing and constitutional modelling	Load dependent
Effective principle stresses, $(\sigma_1^e, \sigma_2^e, \sigma_3^e)$		
Conventional failure criterion, $(\tau - f)$		
Scalar plastic strain, ε_p		
Squared root of the second deviatoric stress invariant, $\sqrt{J_2}$		
Curve face orientation angle, CFO	Image interpretation analysis	Load independent
Normalized curve face orientation angle, CFO/β_{Node}		
Nodal radius, r_{Node}		
Derivative of borehole curve radius, r_w'		
Radius of curvature, RC		
Normalized radius of curvature, RC/β_{Node}		
Node circumferential angle, θ_{Node}		
Spatial variation and directional mechanical properties (anisotropy and heterogeneity indicators)	Lab tests	

2.5.4   Features Engineering and Selection

As discussed previously, when training an algorithm, it is necessary to ensure that the produced model is both generalizable and sufficiently accurate. However, as the FEM-ML integrated model can be used for different types of applications, in some cases, solely targeting a high generalizability and accuracy can be detrimental. There are certain applications of the integrated model that require dropping some features to amplify the influence of others. As expected, reducing the number of predictors can lead to less accuracy; but for the sake of producing a model that serves a specified application, this is a compromise that needs to be made in an engineered fashion. For example, in some applications where the integrated model objective is to predict the mud weight required to prevent wellbore instability, the FEM load-dependent features will either need to be the

only or the dominating features considered in the training process. Also, in certain cases where the data available for training are from offset wells that were drilled in a certain direction and inclination, while the failure prediction and mud weight to be determined are for a new well that will be drilled in a new direction, the features will need to filtered to ensure the prediction is accurate. Different types of these scenarios along with examples of the required features selection are discussed in the next sections.

2.5.4.1 Feature Selection for a Generalizable Failure Prediction (Load-dependent

Application)

To be able to produce a trained and integrated model that is aimed at predicting rock failure for heterogenous rock bodies or that is generalizable for different rock types, all displacement and strain related features will need to be excluded from the training process unless they are introduced in conjunction with spatial variations and directional mechanical properties (anisotropy and heterogeneity indicators). This is because rocks with different mechanical properties will fail at different threshold of strains and displacements, which means using these parameters as precursors of failure will not yield consistent and reliable results. However, strains and displacement related features can be used as features in the machine learning algorithm only in cases where spatial variations of mechanical properties such as the Youngs' modulus and the Poisson's ratio are also introduced alongside them as features. Therefore, excluding strains and displacement related features or including them along with variations of mechanical properties will ensure that the trained and integrated model is generalizable over different rock types. In the applications shown in this work, to be able to build a generalizable failure prediction

model, the training process is limited to a single conventional failure criterion $(\tau - f)$, principle stresses, and the scalar plastic strain. On the other hand, if the trained and integrated model is to be produced for a specific and homogenous rock body; theoretically, strains and displacement related features can be used as features.

2.5.4.2 Features Selection for Single Step Failure (Load-dependent Application)

In cases where the integrated model is to be used to calculate the mud weight required for a risk-controlled wellbore instability based on volume of cavings, the load-dependent features will need to be the only features considered in the training process. In these cases, it is assumed that the shape of the pre-failure wellbore is uniform (i.e. circular or elliptical) and the failure to be predicted is a single step failure as opposed to progressive failure. As can be seen from the list of available features, the load-dependent features are the ones produced by the FEM and its pre-processing calculations. Therefore, for this type of application, the features that can be included in the training process are all the stresses, the stress tensor invariants, scalar strains, and a conventional failure criterion. The reason that conventional failure criteria are included is that they are the only available feature that offers the consideration of strength parameters. For example, if the Mohr-Coulomb criterion is the one selected, this will allow for considering the influence of the strength parameters of this criterion, which are the cohesion and the internal friction angle, at the mesh nodes when training the algorithm and making new predictions using the integrated model.

The reason the geometric shape indicators features are excluded from the training process for this application is that they introduce a rigidity and lack of load dependency

when making predictions using the integrated model. Even though including them in the training process will most definitely produce a model with higher accuracy scores, it's detrimental to the application. To illustrate this issue, the results of wellbore failure prediction from the integrated model that was trained using both load-dependent and load-independent features are shown in **Figure 49** and **Figure 50**.
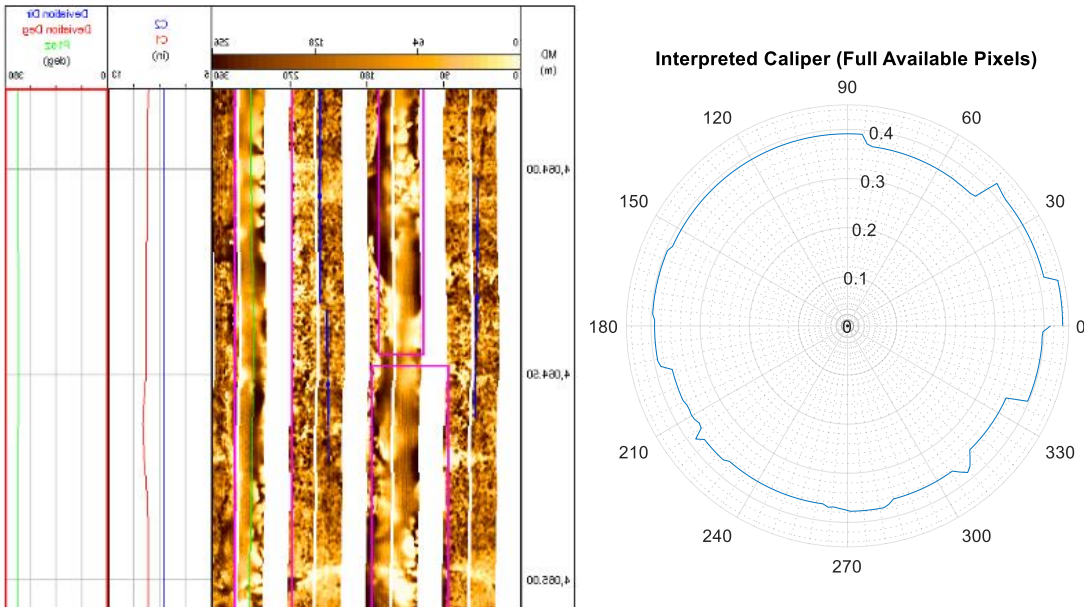
Figure 49. Failure as actually experienced by well-x at 9 ppg mud weight as shown in: Right: resistivity-based image and the multi-arm caliper reading. Left: the interpretation of the failed wellbore shape based on the resistivity image analysis.
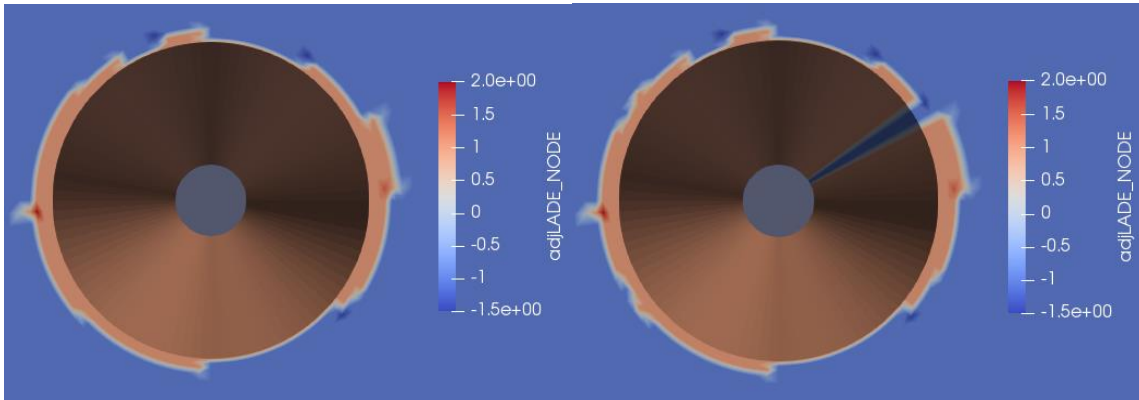
Figure 50. Failure prediction for Well-x using a model that was trained with both load dependent and independent features. Left: wellbore failure as predicted at a load step with 9 ppg mud weight and full in-situ stresses. Right: wellbore failure as predicted at a load step with 0 ppg mud weight and zero in-situ stresses.

Using a model that was trained with both load dependent and independent features to predict failure on well-x (training well) shows that it accurately predicts the actual shape of the failed wellbore as seen in the image log in **Figure 49**. However, the failure prediction shows little reliance on the load dependent features and more reliance on the load independent features as evidenced by the similar failure prediction when comparing the failure prediction at a load step with a mud weight of 9 ppg and full in-situ stresses to the one at another load step with zero mud weight and zero in-situ stresses as seen in **Figure 50**. This is a symptom of over reliance on the load independent features and it is clear that this model is not useful in determining the mud weight required to prevent wellbore failure.

To be able to produce a model that can be used for mud weight determination, the model will need to trained with load-dependent features. As mentioned before, reducing the number of features will reduce the model accuracy. This is evident here as the accuracy

dropped from 99% to 94% and the AUROC dropped from 99% to 92%. However, this is

still a sufficient accuracy as evident by the results shown in **Figure 51** where the case to

the left that has no mud weight or in-situ stresses applied. This case now correctly shows

that no wellbore failure is predicted. The case in the center in **Figure 51** shows that the

failure prediction sufficiently matches the failure that was actually experienced by the well

at the same actual mud weight used and the full in-situ stresses. The case to the right in

Figure 51, which evaluates wellbore failure at a higher mud weight of 9.8 ppg, shows that

the failure zone had been reduced. This is to be expected as higher mud weights reduce

wellbore instability. By comparing these results, it is clear that the trained model in now

fully load-dependent and can be used to calculate the mud weights required for wellbore

instability prevention.
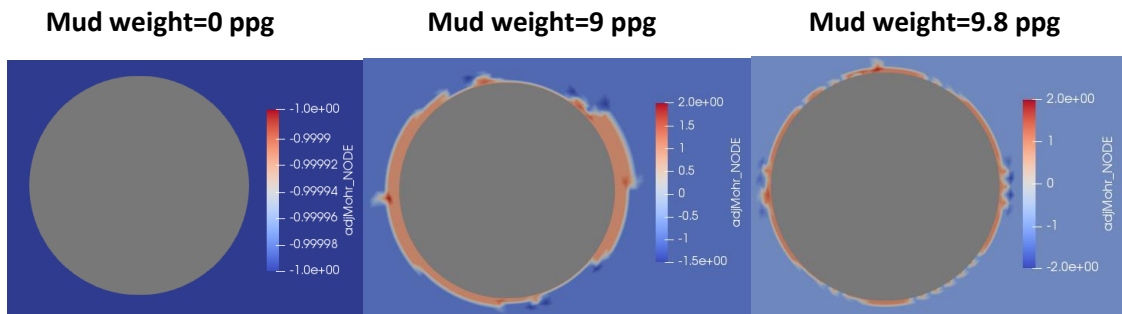
Figure 51. Failure prediction for Well-x using a model that was trained with only load dependent features. Left: no wellbore failure predicted at a load step with 0 ppg mud and 0 in-situ stresses. Center: wellbore failure as predicted at load step with 9 ppg mud weight and full in-situ stresses. Right: wellbore failure as predicted at load step with 9.8 ppg mud weight and full in-situ stresses.

2.5.4.3 Features Selection for Progressive Failure (Load-dependent Application)

This is for the cases where failure prediction needs to be made for the purpose of calculating the instability mud weight; but the starting wellbore shape is non-uniform. Such scenarios are applicable for the cases where predicting progressive failure steps is required. In these scenarios, the wellbore has already experienced some form of failure, which led to its non-uniform shape, and it is now desired to determine the mud weight required to prevent further failure. Since determining the mud weight is still the target, the trained model should still be load-dependent. However, predicting failure in a non-uniform wellbore is more challenging due to the lack of symmetry and training the model using the FEM-based features might not always be sufficient. In such cases, a trial and error process needs to be carried out where a limited number of load-independent features are introduced to the training process. The goal of the trial and error process is to reach a compromise between the load dependency of the model and the accuracy of the model. The final desired model in these scenarios is one that relies mostly on load-dependent features, considers minimal number of load-independent features, and has sufficient accuracy. An example of such case is shown in **Figure 52**. The model used for the failure prediction in this example was trained based on all the available load-dependent features and only two of the load-independent features, which are the derivative of borehole curve (r_w') and the normalized curve face orientation angle $(\frac{CFO}{\beta})$. The result of this training process is a model that is load-dependent and sufficiently accurate with an AUROC of 99% and an accuracy of 99%.
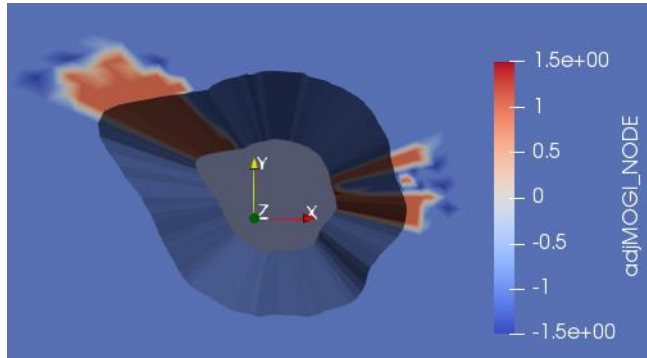
Figure 52. An example of failure prediction in a non-uniform borehole based on a model that was trained using all the available load-dependent features and a minimal number of load-independent features.

2.5.4.4 Features Selection for Non-uniform Borehole Prediction (Load-independent)

This scenario is useful for the applications where determining the mud weight is not of interest, but the shape of failure zone is. An example of these applications is cavings shape prediction. In these applications, it is desired to predict and match the shape of cavings that are observed on the shale shakers while drilling. This application is useful as a form of history matching for the integrated model to confirm its accuracy. In these scenarios, all the available features can be used to achieve the desired accuracy as there is no concerns regarding the load dependency of the model.

2.5.4.5 Features Selection for Predicting Failure in Deviated Wells (Load-dependent

   Application)

As discussed previously, features selection considerations are not only related to the objectives of the application for which the model is being trained, but they are also related to the nature of the dataset available for training. For example, in cases where the training data is available from vertical wells only and the failure prediction is desired for

deviated wells, the features will need to be filtered accordingly. In these cases, all the features that are direction-dependent will need to be removed from the training process. These features include the components of the stress and strain tensors and the directional displacements. This is because these parameters will behave differently at failure in a vertical well compared to a deviated well even if they are introduced in conjunction with spatial variations and directional mechanical properties (anisotropy and heterogeneity indicators). Therefore, using them to provide failure prediction in a well with a new direction and inclination will lead to inaccurate results. To illustrate this, stability mud weights polar charts are used. A stability mud weights polar charts represents the mud weight required at every possible azimuth at the circumferential angle of the chart and at every possible inclination at the radial axis of the chart. When the field is subjected to a strike-slip stress regime, where the maximum in-situ horizontal stress is greater than the vertical in-situ stress, the mud weights required for stability are always expected to increase as the well path approaches the direction of the minimum in-situ horizontal stress with increasing inclination. Following this trend, the maximum mud weight required for stability in this stress regime is expected to be at the direction of the minimum in-situ horizontal stress when the wellbore inclination is near 90°. An example of this expected trend in stability mud weights within a polar chart is shown in **Figure 53** where the mud weights are calculated based on the Lade failure criterion.
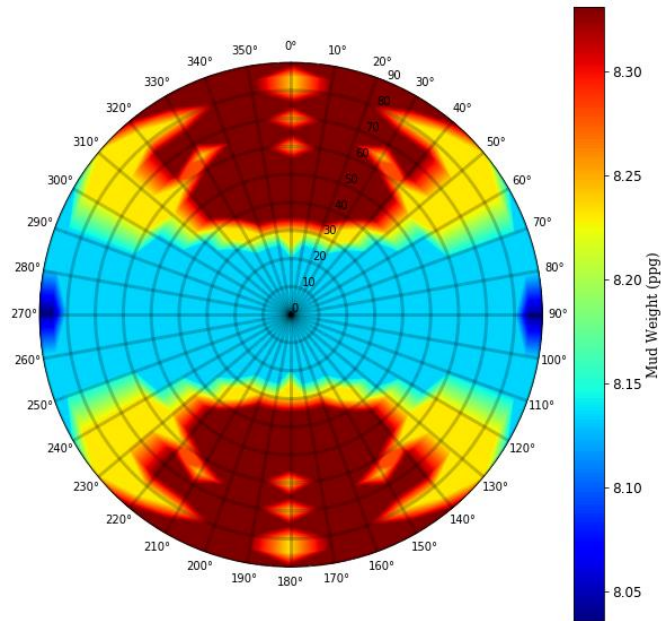
Figure 53. Polar chart for stability mud weight produced from the FEM standalone with the Lade failure criterion with no machine learning model. The allowed threshold of failure or the critical cavings volume for this chart is 1.5 ft³/12 ft interval (0.125 ft³/ft). The minimum horizontal stress direction is at the 0º circumferential angle and the stress regime is strike slip with (σ_H=20.8 ppg, σ_v=18 ppg, σ_h=12.1 ppg, =8.5, P_p=8.5 ppg, Depth=12.795 ft).

When producing a polar chart for the same field, but with a model that was trained with all the load-dependent features including the direction-dependent features, the charts in **Figure 54** are produced at two different critical cavings volumes. As can be seen, as the direction shifts toward the minimum in-situ stress and as the inclination increases, the stability mud weights show a drop in a consistent manner in both charts. Actually, the mud weight in a well in the direction of the minimum in-situ stress and an inclination of 90º is almost the minimum mud weight required for stability in the field in both charts. This is objectively the wrong result. In fact, the same polar chart shape and trends are consistently produced even with different in-situ stress values and allowed cavings volume. This shows

98

the effect of the direction-dependent features on skewing the stability mud weights for

deviated wells when the training data include vertical wells only.
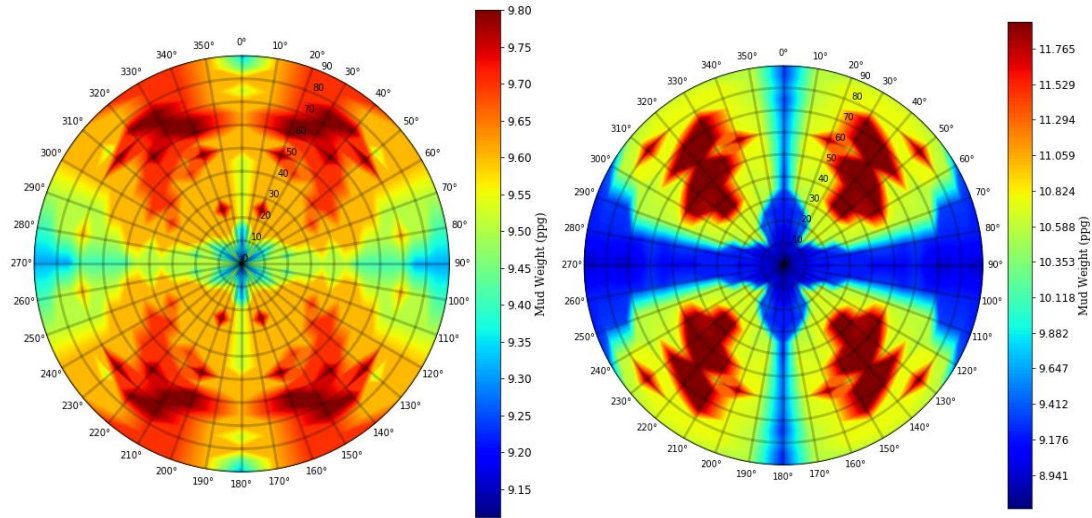
Figure 54. Polar charts for stability mud weight produced from the FEM and a ML that was trained with direction dependent features. The allowed threshold of failure or the critical cavings volume for the chart to the left is is 1.5 ft^3/12 ft interval (0.125 ft^3/ft) and for the chart to the right is 4 ft^3/12 ft interval (0.333 ft^3/ft). The minimum horizontal stress direction is at the 0° circumferential angle and the stress regime is strike slip with (σ_H=20.8 ppg, σ_v=18 ppg, σ_h=12.1 ppg, =8.5, P_p=8.5 ppg, Depth=12.795 ft).

Alternatively, when producing a polar chart for the same field, but with a model

that was trained with load-dependents features while excluding the direction-dependent

features, the chart in **Figure 55** is produced. As can be seen, now the trend of the stability

mud weights is behaving as expected by increasing as the direction shifts toward the

minimum in-situ stress and the inclination increases. Also, the mud weight in a well in the

direction of the minimum in-situ stress and an inclination of 90° is now the maximum mud

weight required for stability in the field, which is the expected trend.
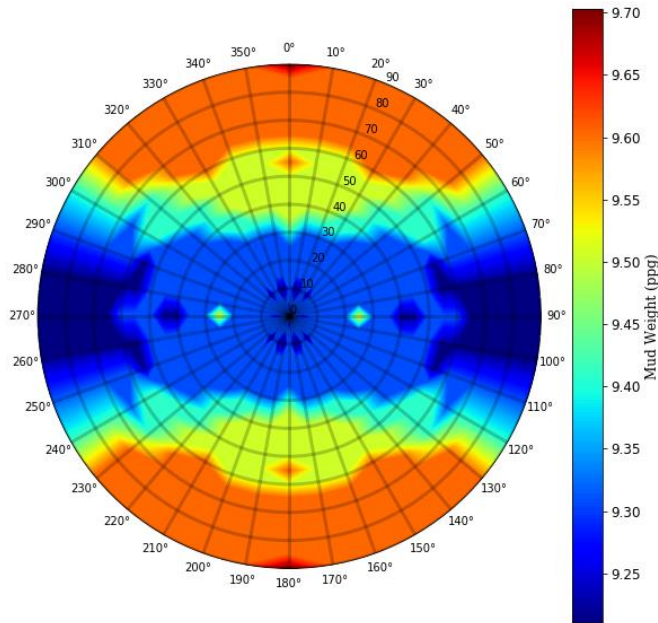
Figure 55. Polar chart for stability mud weight produced from the FEM and a ML that was trained with no direction dependent features. The allowed threshold of failure or the critical cavings volume for this chart is 1.5 ft³/12 ft interval (0.125 ft³/ft). The minimum horizontal stress direction is at the 0° circumferential angle and the stress regime is strike slip with (σ_H=20.8 ppg, σ_v=18 ppg, σ_h=12.1 ppg, =8.5, P_p=8.5 ppg, Depth=12.795 ft).

2.5.5 Feature Importance from Decision Tree Algorithms

As briefly mentioned at the beginning of the machine learning section, decision trees models offer a unique advantage by providing an evaluation of features importance. Knowing which features contributed the most to providing an accurate prediction of rock failure is useful as it helps gain more insight and understanding of the underlining mechanisms of failure. Decision trees evaluate feature importance through the concept of information gain. Through this concept, it is said that more information is gained from a certain feature when it reduces uncertainty. The uncertainty for each feature is evaluated mathematically either using the Gini index or the entropy as described in the bootstrap

100

aggregations of classification decision trees section. Both of these parameters are based on $\hat{p}_{mk}$, which is the probability that a mesh node from the training dataset is in the m region of the feature space from the k class. Therefore, the features that lead to more purity or less uncertainty ($\hat{p}_{mk}$ approaches 1, Gini index and entropy approach 0), are said to provide more information and will consequentially be assigned higher importance. Also, the feature importance is reflected in the feature position within the decision tree. More important features are placed higher within the trees as they provide more information and can be used to reduced uncertainty in the earlier splits of the tree.

An example of the evaluation of features importance is shown in **Figure 56**, where these values are from a model that was trained based on field offset wells data with the algorithm XGBoost. This trained model had an AUROC of 0.88. As can be seen in this figure, the highest contributor is the minimum principle stress. Also, the conventional failure criterion (Lade) and the maximum principle stress contributed significantly to failure prediction in this algorithm. The rest of the features are all load-independent features and, as can be seen, their contribution is small, especially when compared to the first three features, which are all load-dependent. Ensuring this type of reliance on load-dependent features ensures that the trained model is load-responsive as discussed previously. More detailed interpretations of these features importances are included in the next chapters.
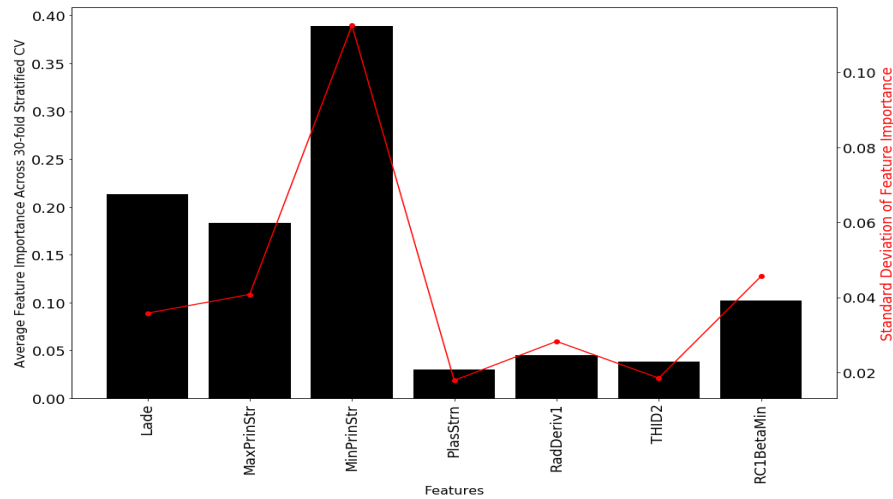
Figure 56. Features importance from the XGBoost algorithm that was trained based on field data.

2.6 Programming Languages and Tools

The integrated model consists of three major parts. These are: the initial meshing and the FEM, the image analysis and the remeshing functions, and the machine learning training algorithms. The programming languages and tools to develop and deploy the model are list in **Table 3**.

Table 3. Details of coding languages and tools used to develop and deploy the model

Process	Language	Software/Tool
Initial meshing code	Fortran	Compiled in Visual Studio
FEM		
Post-processor		
Experimental image interpretation	MATLAB	MATLAB R2018a
Field image interpretation		
Re-meshing based on image structure		
Processing of offset wells data and past experiments data for ML training		
Machine learning algorithms cross-validation, training, testing	Python	Anaconda3/Jupyter Notebook
3D Visualization	-	ParaView 5.6.0

CHAPTER III

INTEGRATED MODEL VALIDATION AND PROOF OF CONCEPT THROUGH

EXPERIMENTAL WORK

Since the modelling scheme proposed by the integrated model is different from the conventional methods for evaluating rock failure, it is necessary to provide a proof of concept and validation for the model results. While the integrated model is intended for field use, validation through experimental work is more valuable because the lab-based images are interpretable by the naked eye, unlike field-based logging images. Therefore, training the model with lab-based images will yield results that are more easily comparable. To achieve this, images from borehole failure lab experiments will be used to train the model. Then, a separate borehole failure test will be modelled using the trained integrated model and the results of failure prediction will be represented in 3D visualizations, which can be readily compared to the actual failure that was experienced by this separate borehole failure test.

The experimental work used in this part is based on partially published experiments by Morita and Nagano (2016) and Papamichos et al. (2019). These experiments are categorized into large scale borehole failure experiments and small scale true-triaxial borehole failure experiments. While the large scale experiments are not extensive enough to build a reliably trained integrated model, they do depict the process of progressive failure. Therefore, these are used only as a proof of concept and an illustration of the new modelling scheme. On the other hand, the small scale true-triaxial

experiments are extensive enough to train a reliable integrated model. These experiments also rely on high resolution CT-scans to evaluate the borehole failure in the post-test core samples. Therefore, it is deemed that these experiments are ideal for providing validation of the integrated model results.

3.1 Model Illustration and Proof of Concept Through Large Scale Borehole Failure Experiments

The partially published (Morita and Nagano 2016) experiments used in this part examine borehole failure in large rock specimens under two different horizontal stresses acting on the plane perpendicular to the borehole axis. In this section, details of the experimental setup are presented along with the results of supporting experiments for stress-strain curves and constitutive modelling. The final outcome of one of the experiments is shown through still frames from a video capturing the progressive failure steps of the borehole. Next, the still frames from the video along with the rock specimen data are used to train a suitable machine learning algorithm. Finally, an experiment is modelled in the integrated and trained FEM-ML model and its results are visualized in 3D, which is then compared to the actual failure experienced by the sample during the experiment.

3.1.1 Large Scale Experiments Details

The rock specimen used for this experiment is Yellow sandstone, which has a length of 10.5 inches, width of 10.5 inches, and thickness of 2.42 inches. The borehole has a diameter of 3 inches and the applied loads are 250 psi and 2,700 psi on the plane of the

borehole. The experimental setup allows for monitoring the progression of failure through a borehole slot as the stresses are applied in by the loading frame is shown in **Figure 57**.
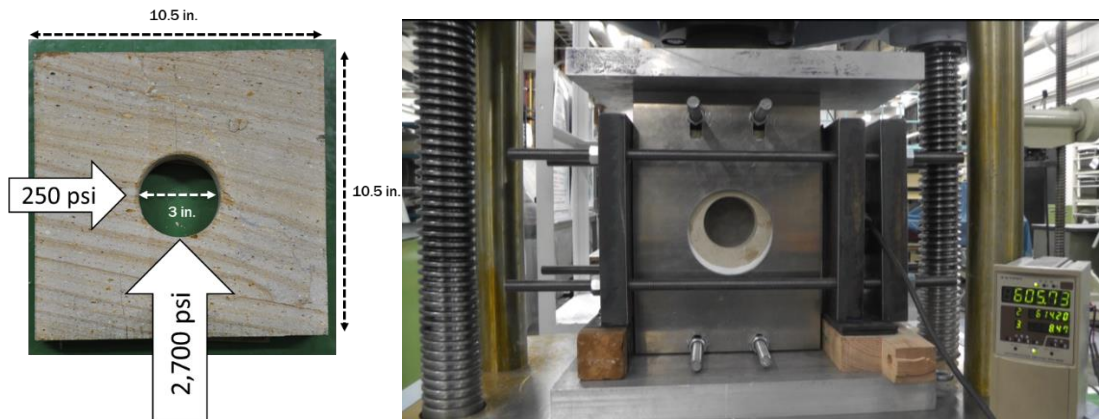
Figure 57. Left: the rock specimen dimensions and loads. Right: the experimental setup for the triaxial compression test of a large scale specimen with a borehole slot.

As the loads are being applied, a video recording captures the progression of failure zones around the borehole. At the end of the test, the specimen is dismounted and pictures are taken to document the failure zones across the borehole axis. The progression of the applied horizontal loads up to the failure point is shown in **Figure 58**. Also, a summary of the test results is shown in **Table 4** where the loads at the point of cracks initiation and at the point of breakage are recorded. Still frames are taken from the video to examine the separate steps of failure as they occur. The pictures taken after dismounting the specimen are used to corroborate the failure across the different layers of the borehole.
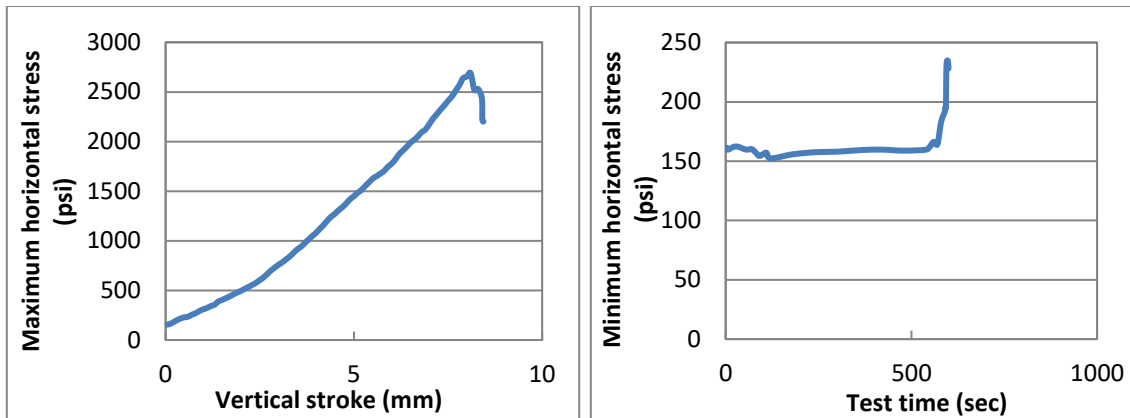
Figure 58. The progression of the applied loads on the rock specimen where the chart to the left shows the maximum horizontal stress and the chart to the right shows the minimum horizontal stress.

Table 4. A summary of the large scale experimental results.

Test progress	time (sec)	stress (psi)	Stress (psi) lateral	Stress ratio
Start	0	155.3	160.2	0.9696
Destruction (longitudinal crack)	577	2600	175.5	14.81
Destruction (breakout)	597	2689	244.9	10.98
Finish	599	2200	233.0	9.445

3.1.2 Stress-strain Curves and Constitutive Modelling

Standard triaxial tests are conducted on Yellow sandstone plugs to produce the stress strain curves shown in **Figure 59**. The tests are conducted at different confining pressures following the guidelines set by Morita and Nagano (2016) and Morita and Gray (1980) for constructing nonlinear constitutive stress-strain relations. These relations are used in the FEM post processor calculations to predict the stress state at each node based on the displacements that were determined by the FEM solver.
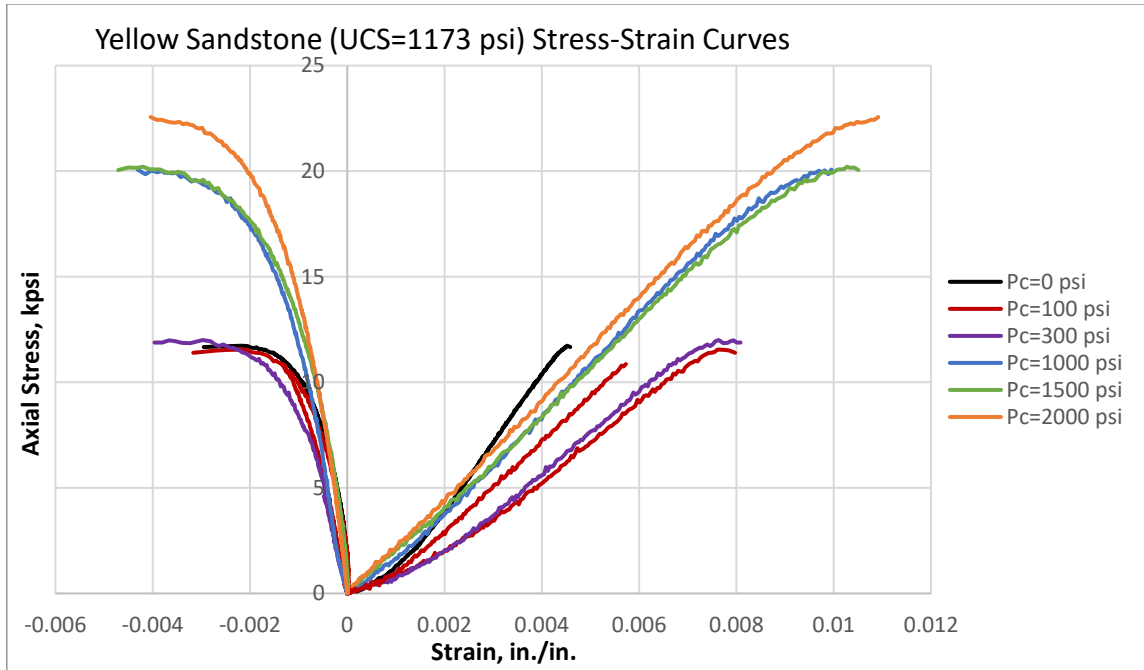
Figure 59. The stress-strain curves of the Yellow sandstone from triaxial tests at different confining pressures, which is used for constructing the nonlinear constitutive stress-strain relations.

3.1.3 Integrated Finite Element and Machine Learning Training and Results

Images taken as still frames from the video of the experiment are used to train the machine learning algorithm. The still frames of the captured progressive failure steps are shown in **Figure 60**. The image analysis for the purpose of mapping of borehole failure and eventually for training the machine learning algorithm is shown in **Figure 61**. It should be noted that the still frames are only capturing failure in the top layer or the surface of the specimen. For examining failure across the borehole axis and at non-surface layers of the specimen, post-test pictures are used.
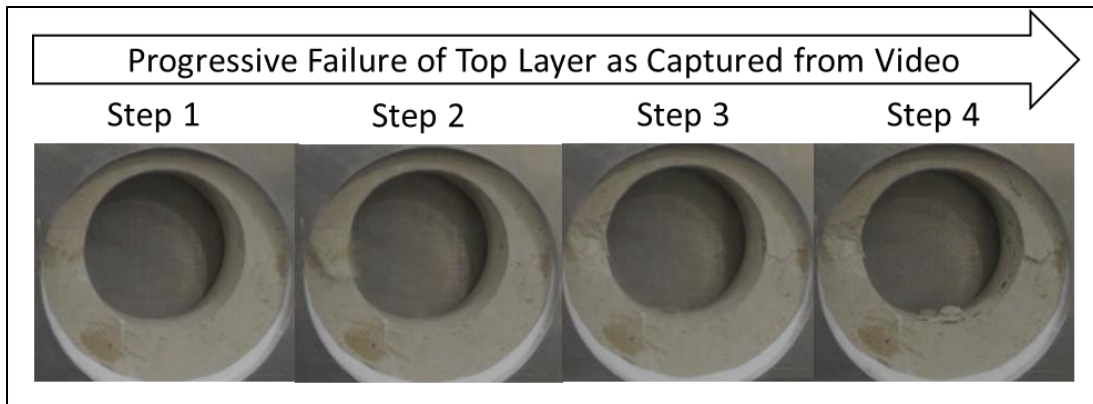
Figure 60. Still frames depicting the different progressive failure steps in the top layer of the large scale experiment.
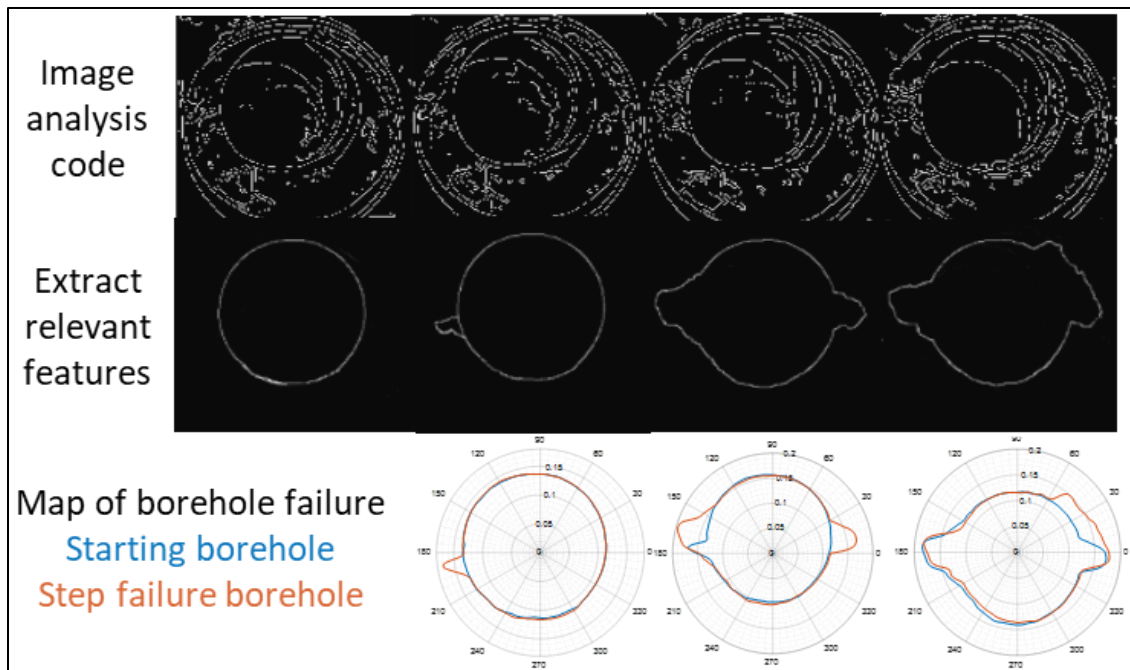
Figure 61. Image analysis of the still frames depicting the different progressive failure steps in the top layer of the large scale experiment.

Since failure in the center layers of the specimen is not visible in the video, it is captures as a single step failure from post-test images. The images describing the failure

at the center layers of the specimen are shown in **Figure 62** and their corresponding image

analysis are shown in **Figure 63**.

Figure 62. Borehole failure in center layers of the dismounted specimen in the large scale tests. Left: failure on the center of the left borehole wall. Center: failure on the center of the right borehole wall. Left: failure on the center of the left borehole wall.
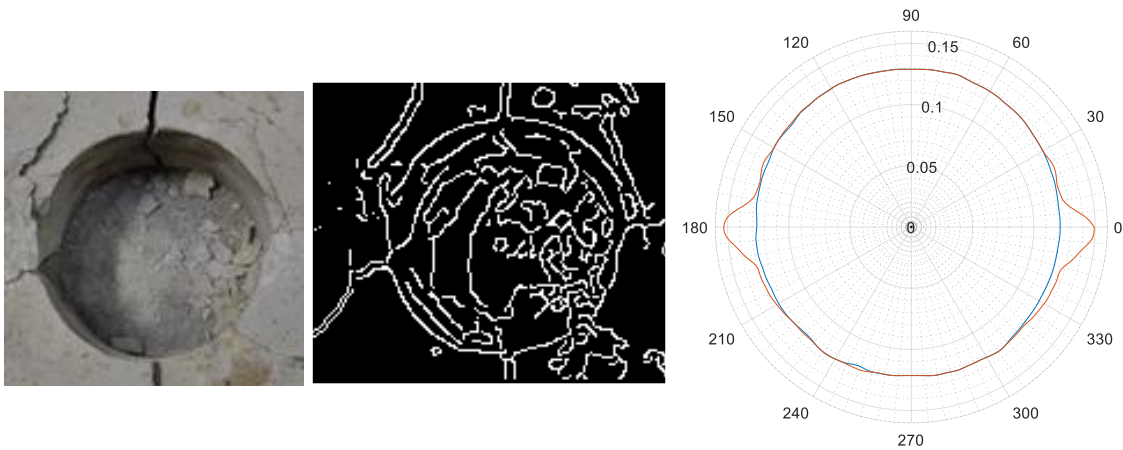
Figure 63. Image analysis and failure mapping of the failed zones around the borehole in the center layer as interpreted from post test pictures in the large scale experiments.

When modelling the specimen into the FEM, failure prediction can be produced

using a conventional failure criterion. The failure criterion selected here is Mohr-Coulomb

and the results of the FEM model and failure prediction is shown in **Figure 64**. It should

be noted that these results are based on the FEM and the Mohr-Coulomb criterion without the use of a machine learning algorithm. As can be seen from these results, the Mohr-Coulomb criterion predicts failure will take place at all mesh nodes at the end of the test.
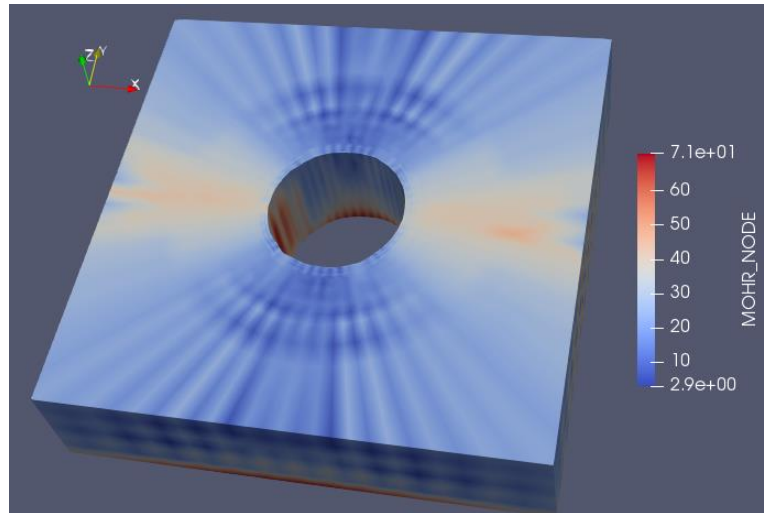
Figure 64. Prediction of failure in the large scale experiment using the FEM and the Mohr-Coulomb criterion (without the use of machine learning and image training).

After modeling the specimen into the FEM and using the output along with the images of the failed borehole to train the machine learning algorithm, new prediction of failure can now be produced. The training was carried out using the large scale experiments data only, which means that the dataset contains homogenous rock data. Therefore, the features used for training the model are the:

1. Mohr-Coulomb failure criterion

2. Principle stresses and plastic strain

3. Geometric shape indicators

110

To illustrate these results, the new failure prediction for different progressive failure steps and across different layers of the samples are shown. Starting with the progressive failure from step-2 to step-3 (progressive failure from step-1 to step-2 is not shown here because the failure zone is very small and was deemed of no value for the demonstration purpose targeted here), the improved failure prediction in both the top and center layers using the integrated FEM-ML model is shown in **Figure 65** and **Figure 66**. The failed areas in these figures are highlighted in red.
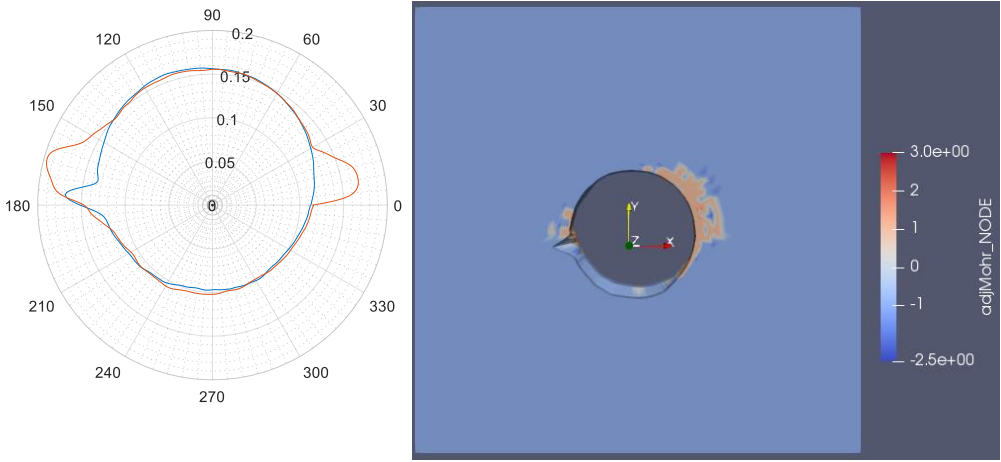
Figure 65. Left: map of failure zones around the borehole at the top layer of the large scale experiment for progressive failure step-2 to step-3. Right: prediction of failure at the top layer of the large scale experiment for progressive failure step-2 to step-3 using the trained integrated FEM-ML.
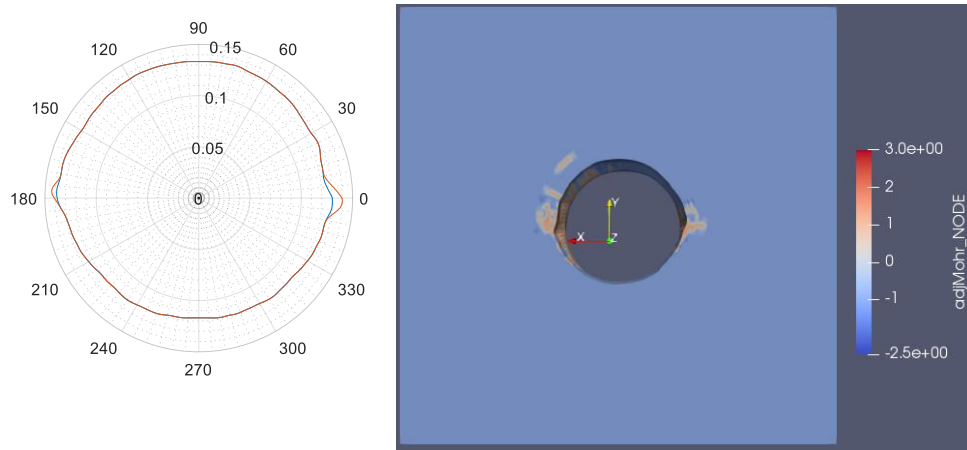
Figure 66. Left: map of failure zones around the borehole at the center layer of the large scale experiment for progressive failure step-2 to step-3. Right: prediction of failure at the center layer of the large scale experiment for progressive failure step-2 to step-3 using the trained integrated FEM-ML.

Now for the progressive failure from step-3 to step-4 (the final two steps), the improved failure prediction in both the top and center layers using the integrated FEM-ML model is shown in **Figure 67** and **Figure 68**.
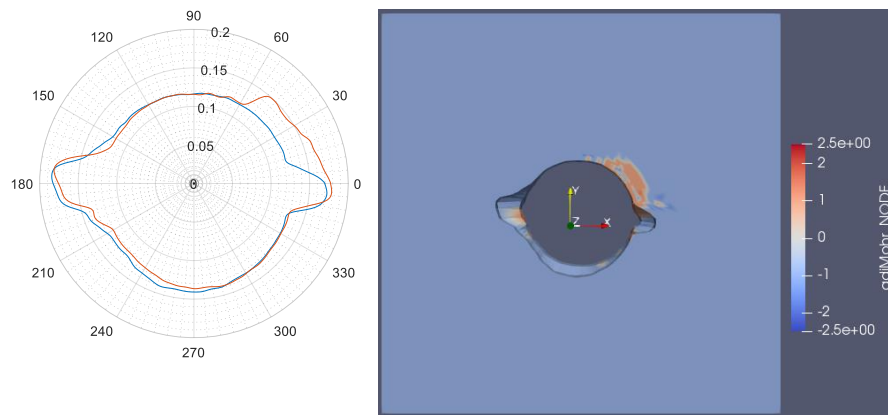
Figure 67. Left: map of failure zones around the borehole at the top layer of the large scale experiment for progressive failure step-3 to step-4. Right: prediction of failure at the top layer of the large scale experiment for progressive failure step-3 to step-4 using the trained integrated FEM-ML.
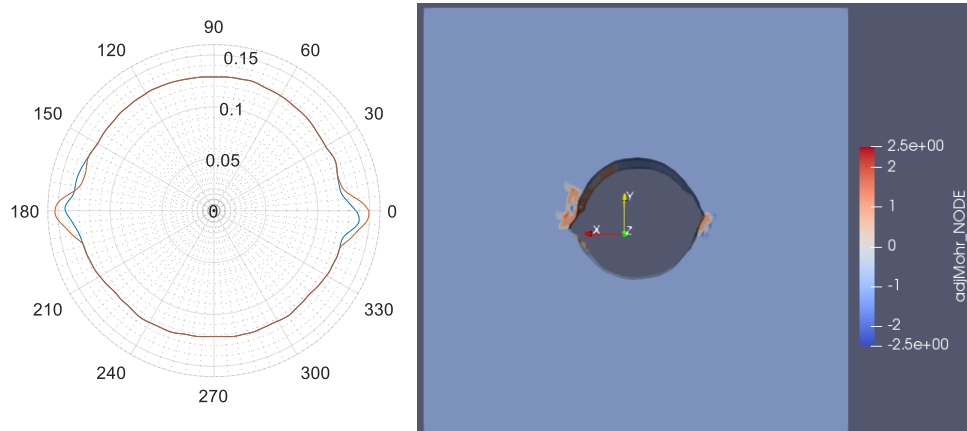
Figure 68. Left: map of failure zones around the borehole at the center layer of the large scale experiment for progressive failure step-3 to step-4. Right: prediction of failure at the center layer of the large scale experiment for progressive failure step-3 to step-4 using the trained integrated FEM-ML.

Finally, for the full failure from step-1 to step-4, the improved failure prediction in both the top and center layers using the integrated FEM-ML model is shown in **Figure 69** and **Figure 70**.
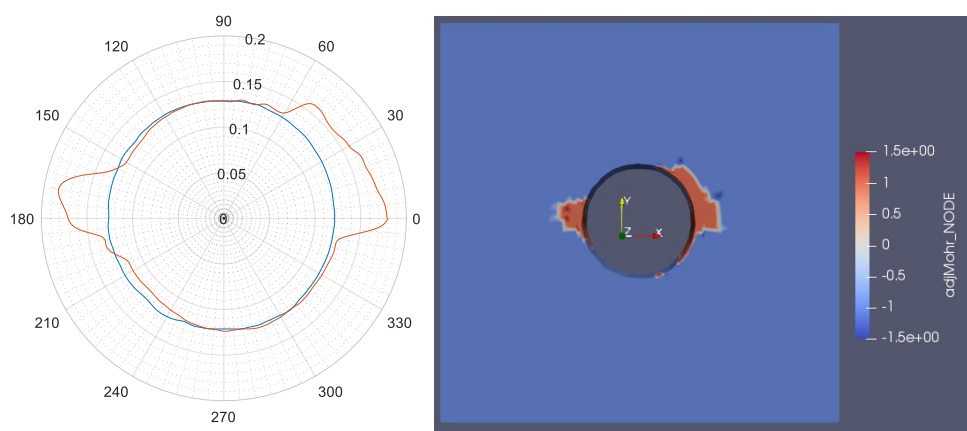
Figure 69. Left: map of failure zones around the borehole at the top layer of the large scale experiment for the full failure step-1 to step-4. Right: prediction of failure at the top layer of the large scale experiment for the full failure step-1 to step-4 using the trained integrated FEM-ML.
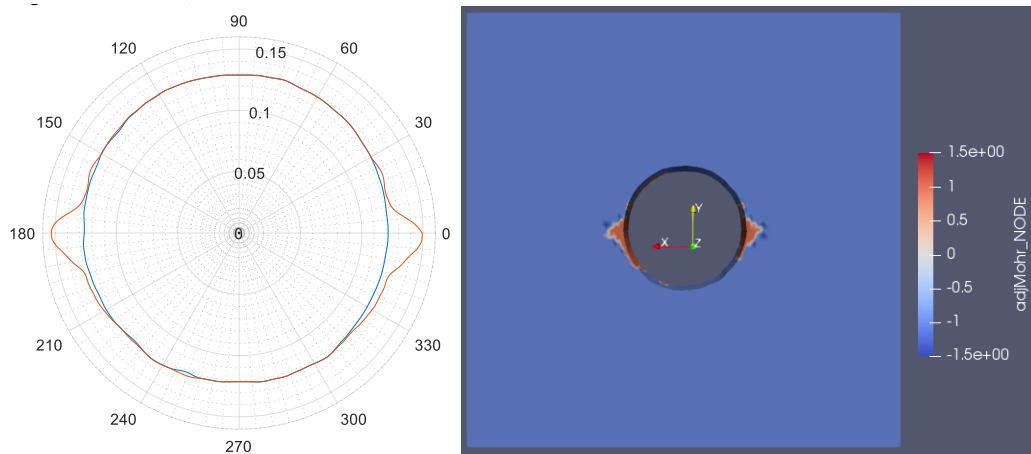
Figure 70. Left: map of failure zones around the borehole at the center layer of the large scale experiment for the full failure step-1 to step-4. Right: prediction of failure at the center layer of the large scale experiment for the full failure step-1 to step-4 using the trained integrated FEM-ML.

As can be seen from all the results for failure prediction using the integrated FEM-ML model, it agrees with the failure seen in the experiment pictures. The concertation of the failure zones as predicted by the integrated model is similar to the concentration of failure shown in the polar failure maps. These results show the potential that image training and physics-based machine learning have in substantially improving rock failure prediction Also, the 3D visualizations shown here illustrate the remeshing function ability to capture different non-uniform borehole shapes in different layers across the borehole axis for the purpose of modelling progressive failure.

As mentioned before, the large scale experiments are not extensive to train a reliable and generalizable machine learning algorithm. In fact, the model produced from the training process in the large scale experiment is, to a certain degree, overfitting. Despite this, these results are presented here to achieve three objectives. First, as a demonstrative exercise in the application of the integrated model. Second, to showcase the

model application in a progressive failure scenario. This is because progressive failure experiments are scarce and difficult to perform. Finally, to provide a proof of concept of the new modelling scheme proposed in this work. Moving forward, the same modelling steps shown in the large scale experiments are also followed for the other experimental and field data modelling efforts.

3.2 Integrated Model Validation Through True-Triaxial Experiments

True-triaxial experiments, that are partially published by Papamichos et al. (2019), are used for the purpose of validating the integrated FEM-ML model. These experiments are useful for this purpose because they are extensive enough in terms of the number of tests conducted, variations in tests conditions, and the quality of the failure monitoring image data. To illustrate the validation of the integrated model, the same process that was used for the large scale experiments in the previous section will also be used here, except instead of relying on still frames from videos depicting progressive failure, post-test CT-scans are used as the imaging data describing single step failure. Also, the data from these experiments are used to train three different models. The first model is a single step failure model for homogenous rock, which aims to predict the failure zones seen in CT-scans with a circular borehole as the starting borehole shape. This model's performance is evaluated comprehensively and compared to the performance of the conventional failure criterion. The second model is an illustration of the training process required for progressive failure prediction, which aims to predict the failure zones seen in CT-scans with a non-uniform borehole as the starting borehole shape. The final model is an illustration for the training

process with spatial variations and directional mechanical properties (anisotropy and heterogeneity indicators).

3.2.1 True-Triaxial Experiments Details

The experiments are conducted on Castelgate sandstone bored core samples. The shape and size of the rock samples used in these experiments are significantly different and smaller than those from the previous section. The shape and dimensions of the core samples are shown in **Figure 71**. The sample has a length and width of around 7.87 inches (200 mm), a slabbed outer diameter of around 5.50 inches (140 mm), and a borehole with a diameter is 0.87 inches (20 mm). The stresses $(\sigma_z, \sigma_r, \sigma_R)$ are applied in different combinations in the three directions as shown in **Figure 71** instead of a single confining pressure, hence the name true-triaxial. The stress σ_R is the maximum horizontal stress (perpendicular to the borehole axis), which is applied in direction of the slabbed face of the core.
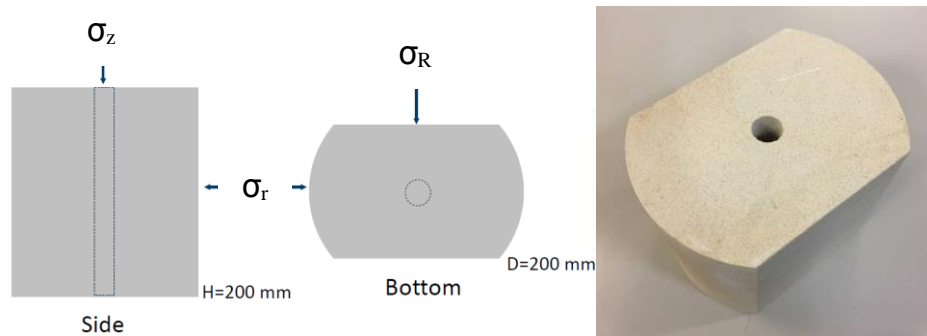
Figure 71. Dimensions and shape of the bored core sample used for the small scale true-triaxial tests.

The core samples are tested at different combinations of stresses using the true tri-axial test system shown in **Figure 72**. Six different tests are used for model training, testing, and validation. The details of the tests are shown in **Table 5**.
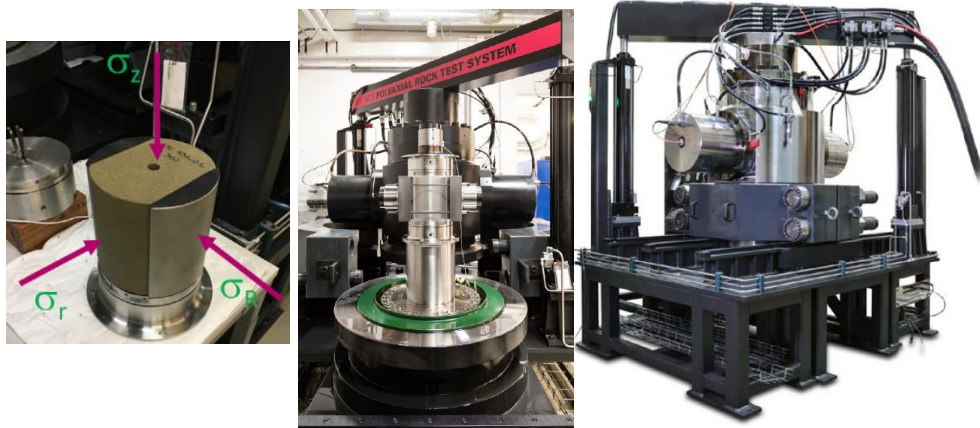
Figure 72. Left: the placement of the sample within the sleeve. Center: the interior of the cell with the sample in position. Right: the exterior of the system.

Table 5. Stress combinations and samples dimensions of the true-triaxial tests.

Experiment Number	$K_z=\sigma_z/\sigma_R$	$K_r=\sigma_r/\sigma_R$	Length, mm	Internal Diameter, mm	External Diameter, mm	Slabbed Length, mm
EXP-1	1	1	200.05	19.83	199.8	138.55
EXP-2	1	1/3	199.90	19.9	199.55	138.49
EXP-3	2/3	2/3	199.86	19.88	199.7	138.82
EXP-4	1	1/3	199.95	20	199.65	138.5
EXP-5	2/3	2/3	199.74	20.0	199.85	138.32
EXP-6	2	2/3	200.05	19.88	199.7	138.82

Once a test is completed, the sample with the failed borehole is removed and CT-scans are taken across 18 different cross sections of the sample. The distribution of the CT-scans along the borehole axis is as follows: the first scan is taken at 5 mm below the top surface and the remaining 17 scans are distributed equally along the remaining 195

mm length of the sample. An example of the CT-scans is shown in **Figure 73**, which

belongs to experiment-1.
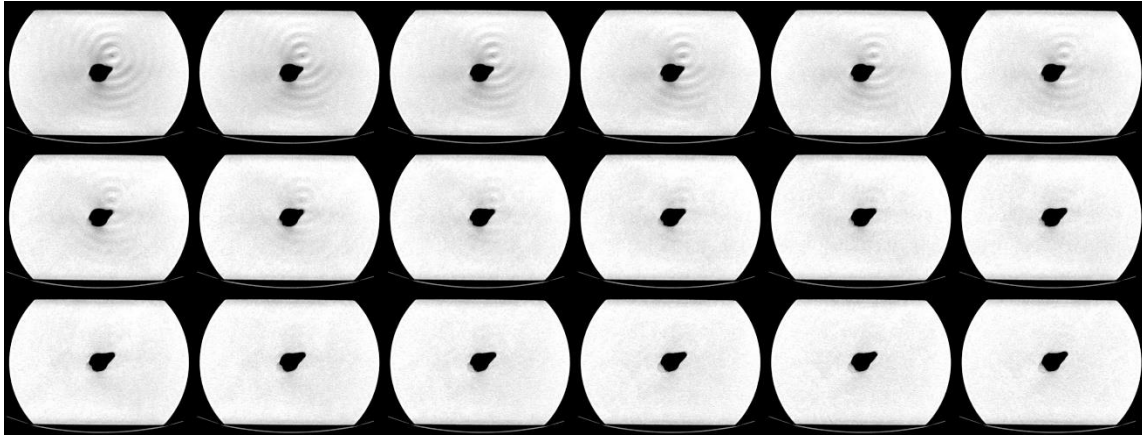
Figure 73. The post-test CT scans taken at 18 different cross sections along the axis of the borehole from EXP-1.

3.2.2 Stress-strain Curves and Constitutive Modelling

Standard triaxial tests are conducted on Castlegate sandstone plugs to produce the

stress strain curves shown in **Figure 74**. The tests are conducted at different confining

pressure following the guidelines set by Morita and Nagano (2016) and Morita and Gray

(1980) for constructing nonlinear constitutive stress-strain relations. These relations are

used in the FEM post processor calculations to predict the stress state at each mesh node

based on the displacements that were determined by the FEM solver.
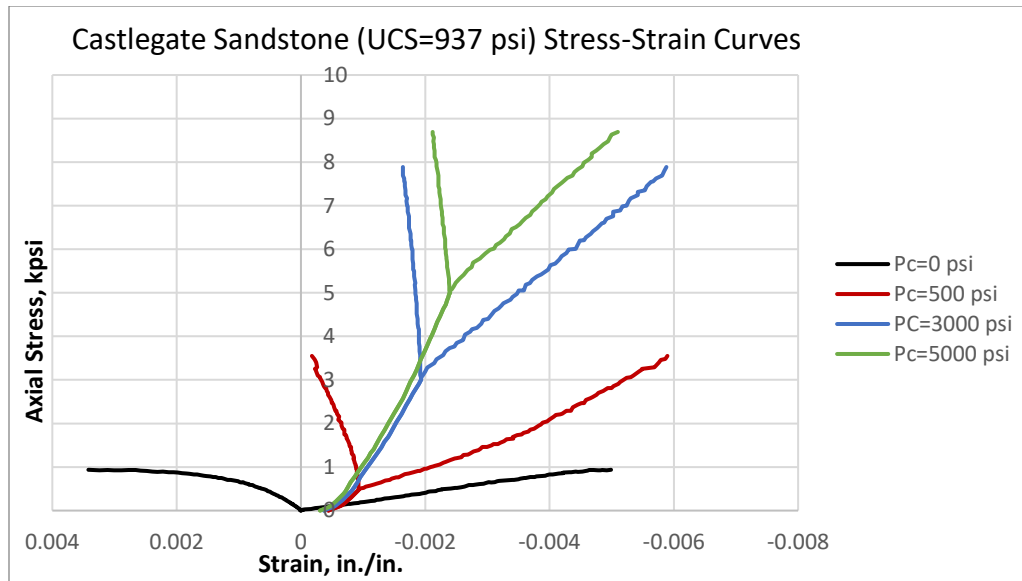
Figure 74. The stress-strain curves of the Castlegate sandstone from the true-triaxial tests at different confining pressures, which is used for constructing the nonlinear constitutive stress-strain relations.

3.2.3 Integrated Finite Element and Machine Learning Training and Results For Single

      Step Failure Prediction

     The first step is modelling all the available experiments in standalone FEM without the use of machine learning. Then, for the sake of comparison, three conventional failure criteria are used to predict failure in each experiment. The conventional failure criteria selected for this purpose are Mohr-Coulomb, Mogi, and Lade. The Mogi failure criterion is used as the main one for the purpose of comparison because it produced the best performance among the conventional failure criteria. The next step is to use the FEM output and the CT-scans from all the experiments to train the machine learning algorithm. Only one experiment, which is experiment-4, is not used in the training phase in order to reserve it for testing the predictive performance of the integrated model.

To interpret the CT-scans using the image analyzer function, only 5 representative CT-scans are selected. This is done because it was observed that CT-scans that were taken in close proximity to each other either showed no differences or only small differences in the borehole shape, which are not discernable by the meshing functions. Therefore, to conserve on computational cost, it was decided to create a mesh of five layers along the borehole axis reflecting the borehole shapes in only 5 representative CT-scans. The spatial distribution of the selected CT-scans within the core mesh is shown in **Figure 75**. An example of the mesh of the core sample from experiment-1 before the test is shown in **Figure 76** and the mesh after the test for experiment-1 along with the 5 representative CT-scans are shown in **Figure 77**.
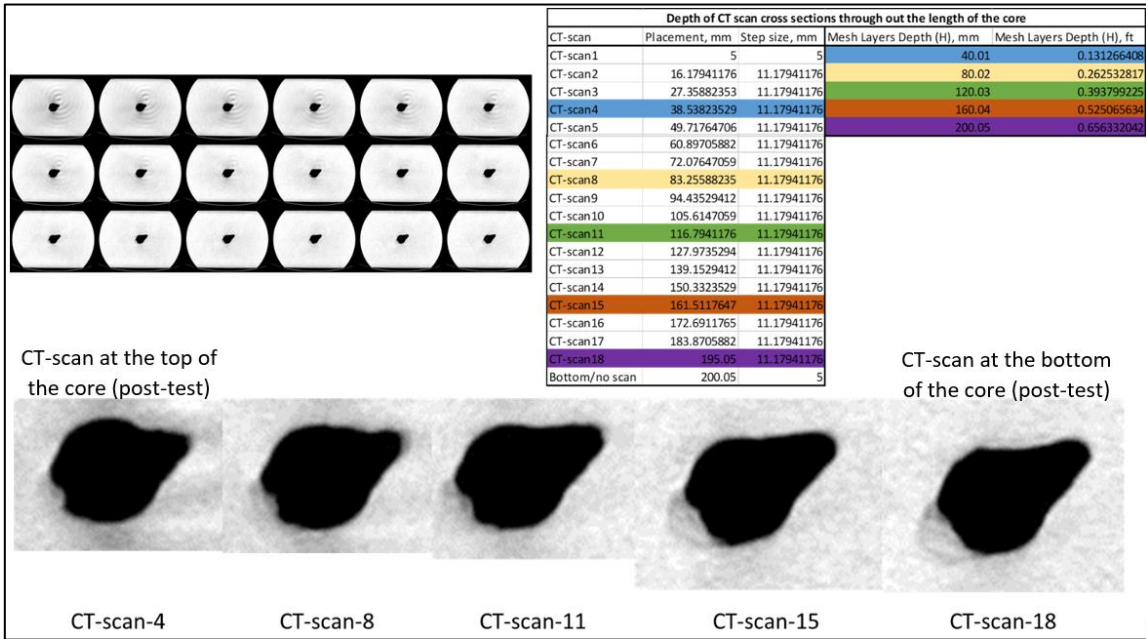
Figure 75. The spatial distribution of the selected CT-scans within the core mesh.

120
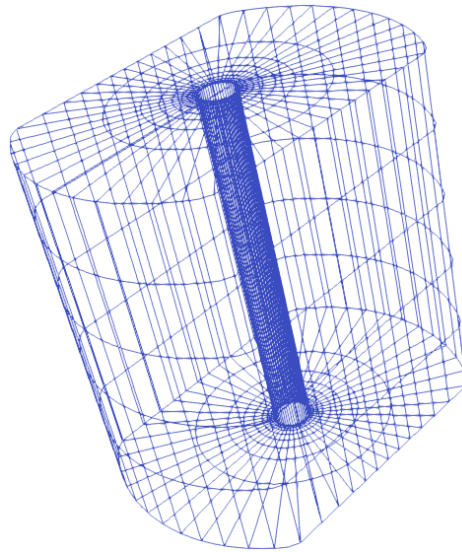
Figure 76. The pre-experiment mesh of the true-triaxial test core sample.
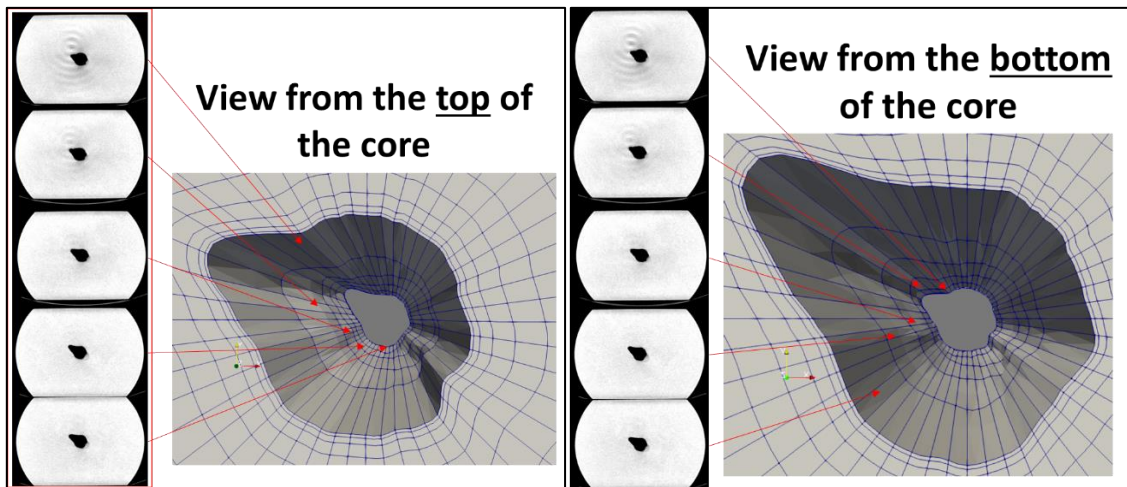
Figure 77. The post-experiment mesh from experiment-1 showing the different borehole shapes per layer based on 5 representative CT-scans taken across different cross-sections of the core sample along the borehole axis.

The interpretations of the 5 representative CT-scans test for experiment-1 are shown in **Figure 78**, which define the failure zones around the borehole.
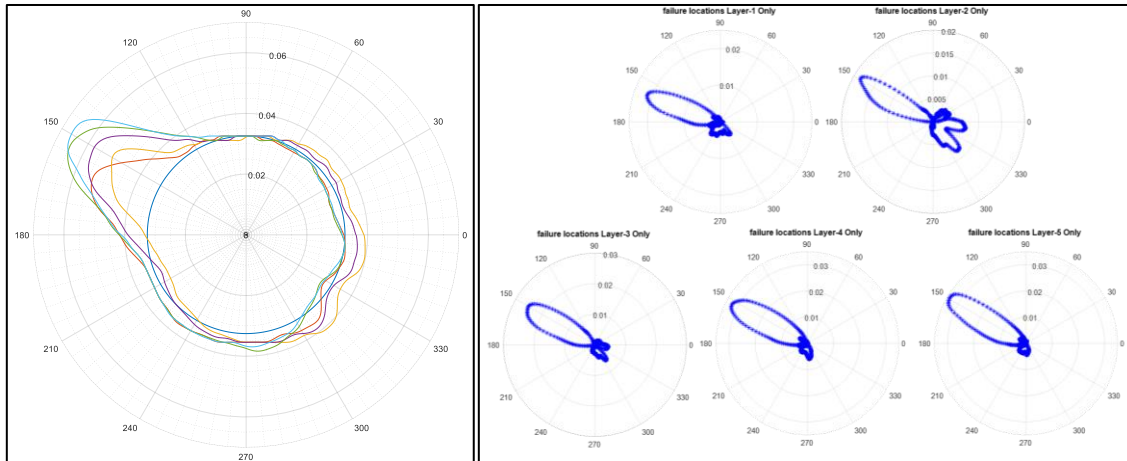
Figure 78. Left: The image analyzer results of interpreting the 5 representative CT-scans from experiment-1. Right: the definition of the location and extent of failure zones in the five cross-section layers along the borehole axis, which correspond to the CT-scans interpretations.

Once all the experiments which are intended for training the machine learning algorithm are processed, the size of the cumulative dataset added up to 59,280 data points. Each data point represents a mesh node and the features extracted for training the machine learning algorithm are all load-dependent features. This means that the training process was performed based on the mesh nodes information which are from the FEM only. Specifically, the features used for this training process are the:

1. Mogi failure criterion

2. Principle stresses

3. Plastic strain

The geometric shape indicators were excluded from the training process because the pre-test borehole shape is uniform (circle). Two more trained models for a generalizable failure criterion and for progressive failure prediction are discussed and
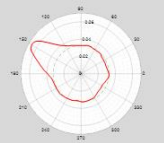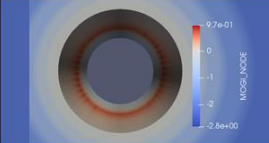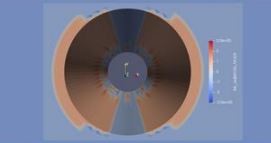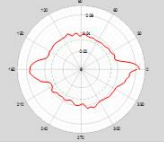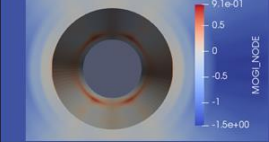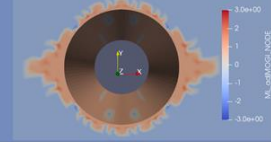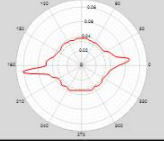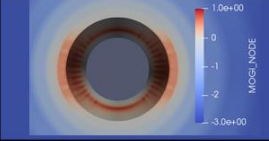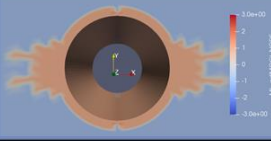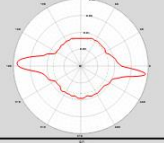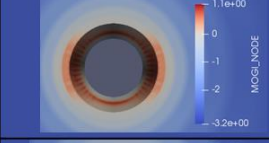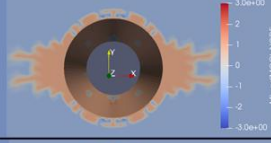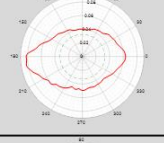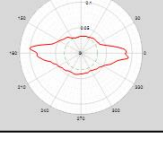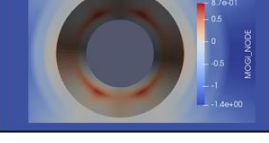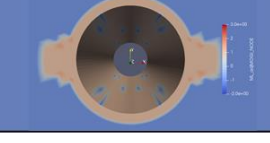
illustrated in the next section. The machine learning algorithms that were trialed in training are: bagging decision trees, neural networks, random forest, and gradient boosting decision trees with two separate algorithms (AdaBoost and XGBoost). All algorithms produced an accuracy of at least 80% and an AUROC of at least 0.75. The two best performing algorithms are bagging decision trees and neural networks both of which produced an accuracy of around 91% and an AUROC of around 0.88. Therefore, the displayed results for the integrated trained FEM-ML model are based on bagging decision trees. The result for failure prediction across a single layer for experiment-1 based on the standalone FEM with the Mogi failure criterion is shown to the left in **Figure 79**. The result for failure prediction across a single layer for experiment-1 based on the integrated FEM-ML is to the right shown in **Figure 79**. As can be seen, the conventional criterion predicted a minimal zone of failure which is uniformly distributed around the borehole (the area in red). On the other hand, the integrated FEM-ML predicts the failure zones in a manner that is more descriptive to what was experienced in the actual experiment.
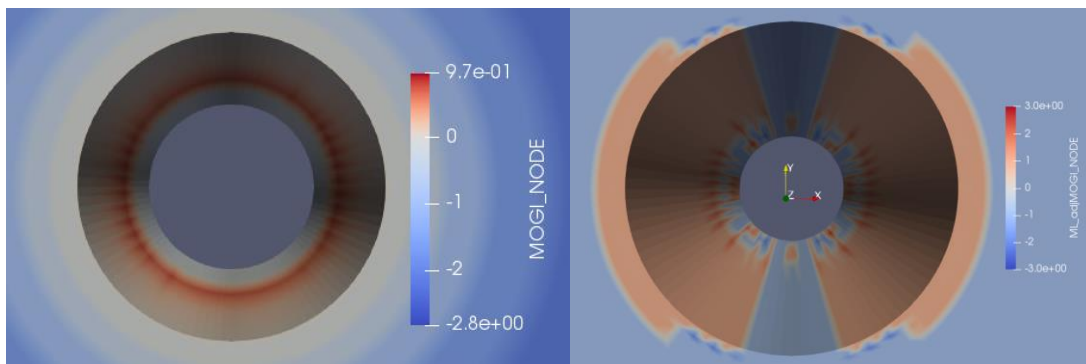
Figure 79. The results for failure prediction across a single layer for experiment-1. Left: result based on the standalone FEM with the Mogi failure criterion. Right: results based on the integrated FEM-ML model.

Similar interpretations can also be made from the results of the rest of the experiments as summarized in **Figure 80**. The summary shows, in the third column from the left, the maps of the failure zone for a single layer based on a single CT-scan for each experiment. In the fourth column from the left, the results for failure prediction across a single layer based on the standalone FEM with the Mogi failure criterion are shown. Finally, on the right most column, the results for failure prediction across a single layer based on the integrated FEM-ML are shown. For all experiments, the integrated FEM-ML model produced a more accurate prediction of failure zones as it corresponds more closely to the failure maps based on the CT-scans. It should be noted again that all experiments except experiment-4 were used for training the machine learning algorithm. Also, this summary visualizes the results for only a single layer for all experiments. Therefore, while this summarization method is useful for visually comparing the results, a more quantitative measure of the performance of all models is still needed in order to gain a comprehensive understanding of the improvements brought by the integrated FEM-ML model. The quantitative measure of the performance is the topic of the next section.

	Experiment Number	CT-scan Code Failure Map	Conventional Failure Criterion (Mogi)	FEM-ML Adjusted Failure Criterion
Used in training the machine learning models	Experiment 1			
	Experiment 2			
	Experiment 3			
	Experiment 5			
	Experiment 6			
Not used in training	Experiment 4			

Figure 80. Summary of the results of failure prediction across a single layer for all available true-triaxial experiments. The results compare failure prediction from the conventional failure criterion (Mogi) to the integrated FEM-ML model (single step failure).

3.2.4 Evaluation of the Performance of the Integrated Finite Element and Machine

Learning Model

After trialing different machine learning algorithms on the full true-triaxial experiments dataset, bagging of decision trees and neural networks produced the best

125

performance with an accuracy of around 91% and an AUROC of around 0.88. As mentioned before, the displayed results for failure prediction are based on the trained bagging of decision trees model with entropy as the decision tree splitting criterion. The hyperparameter tuning process and cross validation of the model was performed with a 30-fold stratified cross validation. The 30-fold cross validation means that the dataset was divided into 30 subsets and the algorithm was trained on 29 subsets and then tested on the last subset. This process is repeated with a different subset being used as the test subset each time. The performance of the algorithm is measured by averaging its accuracy scores and AUROC across the 30 folds. Stratification of the folds means that it is ensured that each subset includes the same ratio of failed to non-failed mesh nodes. The results of the 30-fold cross validation of bagging decision trees are shown in **Figure 81**. These results are used to pick the best performing hyperparameters for the model based on its accuracy, which in this case is evaluated using AUROC. Based on this, the best performing model has 300 bagged trees with a maximum tree depth of 50.
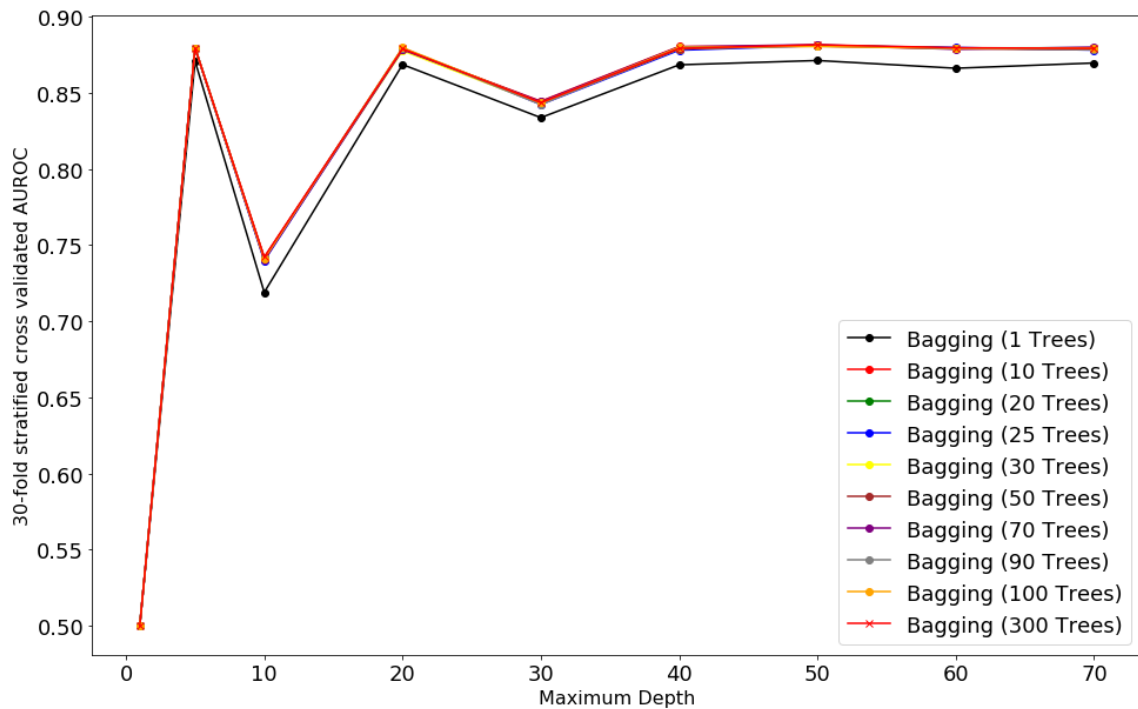
Figure 81. The results of the cross-validation process of the bagging decision trees ML algorithm based on the true-triaxial experiments dataset.

While the gradient boosted decision trees algorithms did not perform as well as bagging of decision trees and neural networks as they produced an accuracy of around 90% and an AUROC of around 0.88, they are still useful for the purpose of measuring features importance. Features importance is useful in creating a better understanding of factors that led to improving failure prediction. The features importance for this dataset is produced from gradient boosted decision trees algorithms, which is the XGBoost algorithm from the xgboost public library. The features importance from XGBoost is shown in **Figure 82**.
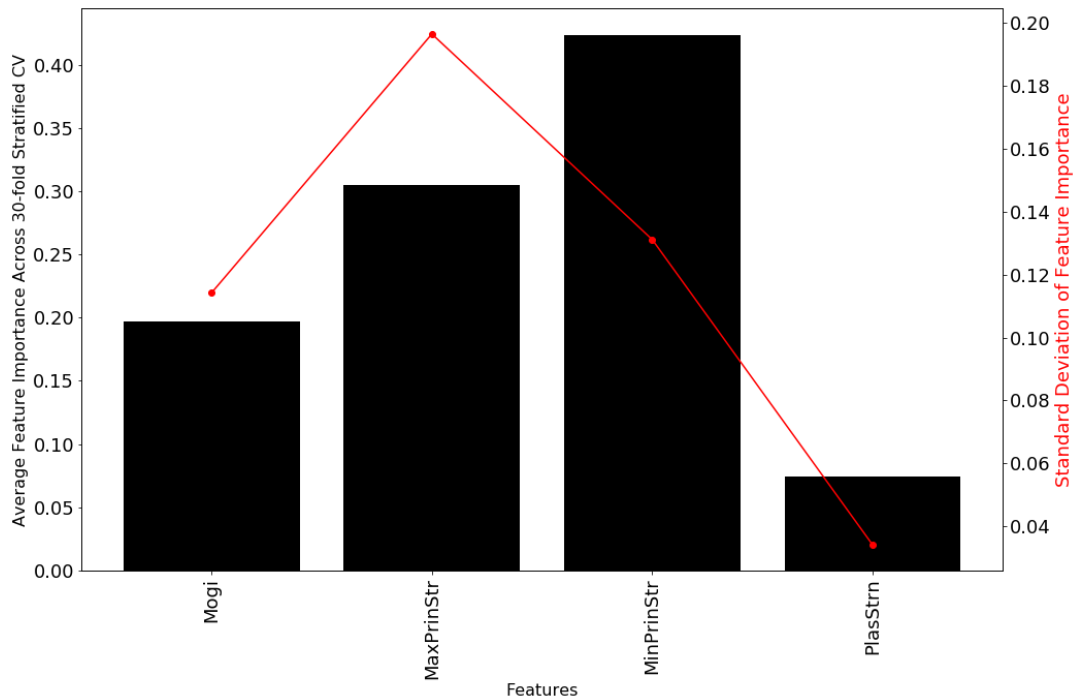
Figure 82. Features importance from the XGBoost algorithm that was trained based on true-triaxial lab experiments data.

As can be seen from both of the feature importance figures, the minimum principal stress has the highest importance. Another observation regarding features importance is that the importance of the conventional failure criterion (Mogi) is surprisingly low. This means that when the algorithm tries to use the conventional criterion to predict failure at the mesh nodes, they provided little contribution. This is surprising because the standard methods for deciding failure of rocks either analytically or numerically relies completely on the conventional criteria. Overall, the observations made based on features importance can be used to enhance the understanding of the failure mechanisms taking place. It should be noted that all of these features importance values are consistent as they were determined with a small standard deviation across the 30-folds as seen from red curves in **Figure 82**.

Now to be able to comprehensively compare the results of the standalone FEM with a conventional failure criterion against the results of the integrated FEM-ML, both models will need to be treated as mesh node failure classification algorithms. This is useful because it will enable the evaluation of all models using the confusion matrix parameters, which allows for a systematic comparison between them. Starting with the predictive performance in experiment-1, which is one of the experiments that were included in training the machine learning algorithm. The confusion matrix parameters, including accuracy, for experiment-1 are shown in **Figure 83**. It should be noted that experiment-1 is the worst performing training experiment.
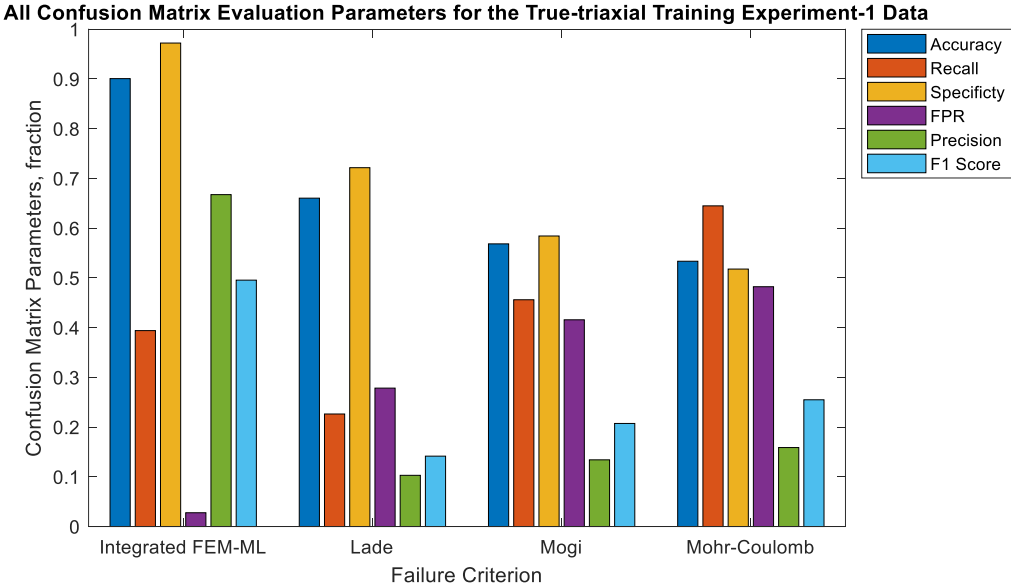
Figure 83. The confusion matrix parameters for experiment-1 (training experiment).

Based on all the different confusion matrix parameters, the integrated FEM-ML model performed considerably better than all the other conventional failure criteria in

predicting mesh nodes failure. To provide a physical meaning to these parameters, here are their definition in terms of mesh node classification being considered:

- Accuracy: Overall ratio of mesh nodes that were predicted correctly in terms of failure. It is desired to be as high as possible.

- Recall: A measure of the ability to predict failed mesh nodes correctly. It is desired to be as high as possible.

- Specificity: A measure of the ability to predict non-failed mesh nodes correctly. It is desired to be as high as possible.

- False positive rate (FPR): The ratio of non-failed mesh nodes that were predicted incorrectly. It is desired to be as low as possible.

- Precision: The ratio of the failed mesh nodes that were predicted correctly. It is desired to be as low as possible.

- F1-score: An overall measure of the predictive ability of the classification model. It is different than accuracy because it is not affected by the skewness of the dataset towards a certain class.

The predictive performance over the full dataset of the true-triaxial experiments is shown in **Figure 84**. This evaluates the performance over all the training experiments, which means that experiment-4 in not included here as it was not included in training the model. Looking at this chart it is clear that the integrated FEM-ML model performed considerably better than all the conventional failure criteria over the training dataset.
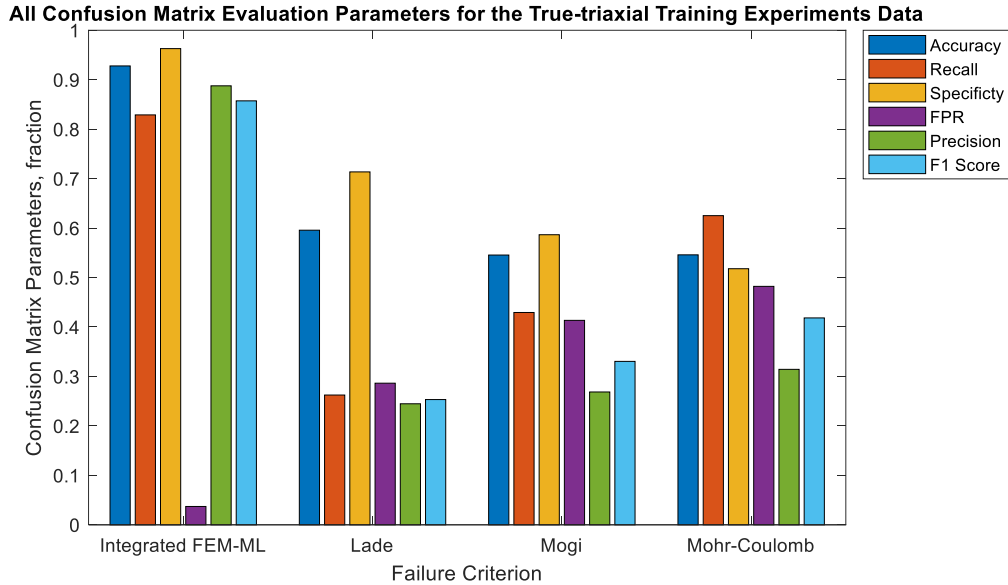
Figure 84. The overall confusion matrix parameters for all the training experiments (experiment-4 is not included).

However, the true measure of machine learning algorithms predictive performance is produced through data that was not seen by the algorithm during the training phase. For this reason, experiment-4 was excluded from training and the predictive performance over it is evaluated separately as shown in **Figure 85**. This chart now shows that the predictive performance of the integrated FEM-ML model has dropped marginally. Still, even with predictive performance drop in testing phase, the integrated model is still performing significantly better than the conventional failure criteria. This is true for all the different confusion matrix parameters. The only confusion matrix parameters in which conventional failure criteria performed comparably to the integrated model are specificity and recall. For specificity, the chart in **Figure 85** shows that the Lade failure criterion specificity is only 32% less than that of the integrated model. While this is still a

131

significantly lower score and poorer performance, it is comparable to the integrated model due to the nature of the dataset. The dataset is heavily skewed towards non-failed mesh nodes as they represent around 74% of the dataset. This means that the more optimistic (predicts less failure) models like Lade will always predict a relatively large number of non-failed mesh nodes, which can translate to a deceptively large specificity value. As for recall, the chart in **Figure 85** shows that the Mohr-Coulomb failure criterion recall is only 22% less than that of the integrated model. Again, while this is still a significantly lower score and poorer performance, it is comparable to the integrated model due to the Mohr-Coulomb failure criterion. The Mohr-Coulomb failure criterion tends to be more pessimistic (predicts more failure), which means it will always predict a relatively larger number of failed mesh nodes, which can translate to larger recall values.
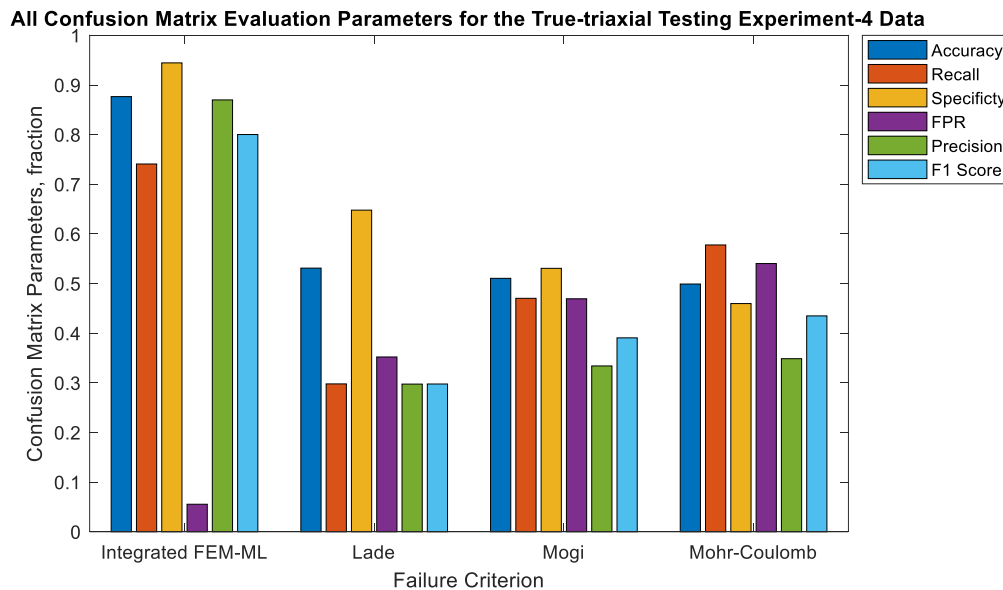
Figure 85. The confusion matrix parameters for experiment-4 (testing experiment).

Based on the discussion of the chart in **Figure 85**, it is clear that it is necessary to evaluate the overall performance of all considered models while paying special attention to the nature of the dataset and the models prediction tendencies. To achieve this, the F1-score is used as the main confusion matrix parameter to compare the predictive performance of the models. A summary of the F1-scores and precision values for the Mogi criterion and the integrated FEM-ML model is shown in **Table 6**. Precision is included because it is a measure of the model ability to predict failed nodes, which are the main point of concern from practical point of view. When examining the results summarized here, the most value is gained by comparing the performance of the model over the test experiment (experiment-4) as it has not been seen by the trained FEM-ML model. The performance of the test experiment shows that the integrated model has an F1-score that is 105% higher and a precision that is 156% higher than that of the conventional failure criterion. Based on the F1-score alone, this interpretation means the integrated FEM-ML has doubled the failure prediction accuracy that was provided by the conventional criteria.

Table 6. A summary of F1-scores and precision values for the Mogi Criterion and the integrated FEM-ML model.

		Conventional Failure Criterion (Mogi)		FEM-ML Adjusted Failure Criterion	
	Experiment Number	F1-score	Precision	F1-score	Precision
Used in training the machine learning models	Experiment-1	0.21	0.14	0.50	0.67
	Experiment-2	0.34	0.30	0.91	0.94
	Experiment-3	0.40	0.33	0.88	0.88
	Experiment-5	0.36	0.27	0.83	0.86
	Experiment-6	0.31	0.32	0.92	0.92
Not used in training	Experiment-4	0.39	0.34	0.80	0.87

3.2.5 Illustration of Other Training Procedures

The training procedure discussed in the previous section deals with predicting single step failure in a uniformly shaped (circular) borehole with a homogenous rock (Castlegate sandstone). Because of this, the model was trained based on load-dependent features only as the load-independent features are mostly constants for a circular borehole and are not useful as precursors for failure. Also, the assumption of homogeneity of the rock samples does not allow the prediction of asymmetric failure zones in a circular borehole, which can only be done with the introduction of spatial variations and directional mechanical properties (anisotropy and heterogeneity indicators). However, to illustrate the potential versatility of the modelling scheme used here and as discussed in the previous chapter, the selection of training features can be manipulated to fit other applications. This section discusses the training processes for two of these applications. The first one is

concerned with illustrating the prediction of progressive failure steps, as opposed to a single failure step. Progressive failure prediction entails dealing with an irregularly shaped boreholes, as opposed to the circular boreholes in single step failure prediction models discussed in the previous section. The second application is concerned with illustrating the training process for predicting asymmetric failure zones in a circular borehole with the introduction of spatial variations and directional mechanical properties (anisotropy and heterogeneity indicators).

3.2.5.1 An Illustration of Training for Progressive Failure Prediction

As mentioned at the end of the large scale experiments model discussion, progressive failure data are scarce in published literature and difficult to produce experimentally. They are difficult to produce experimentally because they would require applying certain loads to a specimen, extracting the specimen from the loading equipment to examine and image the failure zones, reinstalling the same specimen and applying higher loads, and finally extracting the specimen again to examine and image the progression of failure zones. Performing experimental work with these steps can be complicated and extracting and re-installing the specimen in between loading steps can significantly distort the results. Therefore, some improvisation is required to provide an illustration of progressive failure models. To achieve this, for each true-triaxial experiment, 2 CT-scans are selected with one designated as the initial pre-loading borehole shape; and the other as the post-loading borehole shape. Borehole failure for each true-triaxial experiment is assumed to be defined as the change in the borehole shape from

the first CT-scan to the second CT-scan. An illustration of this setup is shown in **Figure 86**.
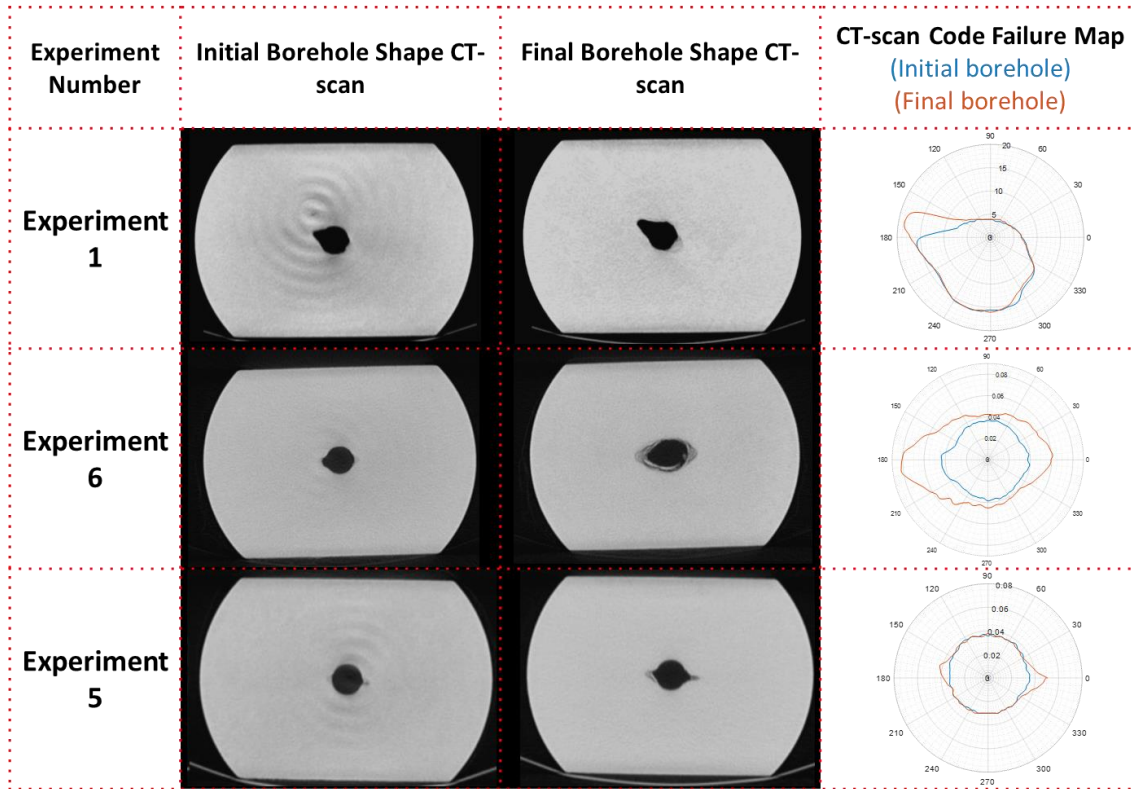
Figure 86. CT-scans and failure maps from true-triaxial experiments which are used to illustrate progressive failure model and irregular borehole failure prediction.

The consideration of an irregular borehole shape as the initial borehole shape in failure prediction dictates that the geometric shape indicators will now be introduced into the training features to ensure an accurate failure prediction. This is because both the standalone conventional failure criterion and the load-dependent training features cannot produce an accurate prediction of failure when the starting borehole shape is irregular or non-uniform. The consideration of the geometric shape indicators in the training process

ensures that the distribution of failure around the borehole irregularities is captured more accurately in the training process. Based on this, the training features used for the progressive failure application are:

1. Mogi failure criterion

2. Principle stresses

3. Plastic strain

4. Derivative of borehole curve, r_w'

5. Normalized radius of curvature, $\frac{RC}{\beta}$

6. Normalized curve face orientation angle, $\frac{CFO}{\beta}$

7. Nodal radius, r_{Node}
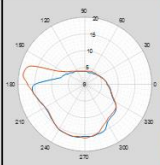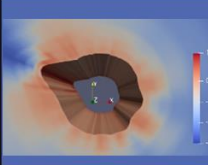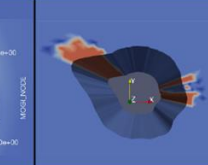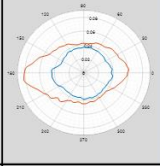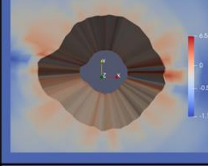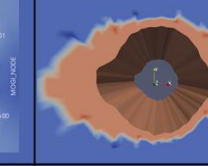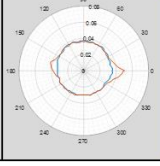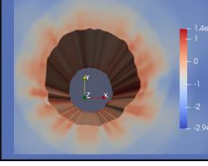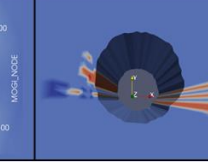
After trialing different machine learning algorithms on the full true-triaxial experiments dataset, gradient boosting of classification decision trees (AdaBoost) and neural networks produced the best performance with an accuracy of around 99% and an AUROC of around 0.99. The hyperparameter tuning process and cross validation of the generalizable model was performed with a 30-fold stratified cross validation as before. The results of the 30-fold cross validation of the AdaBoost algorithm are shown in **Figure 87**. These results are used to pick the best performing hyperparameters for the model based on its accuracy, which in this case is evaluated using AUROC. Based on this, the best performing model has 10 trees with a maximum tree depth of 30 as it produced the highest AUROC of 0.99.
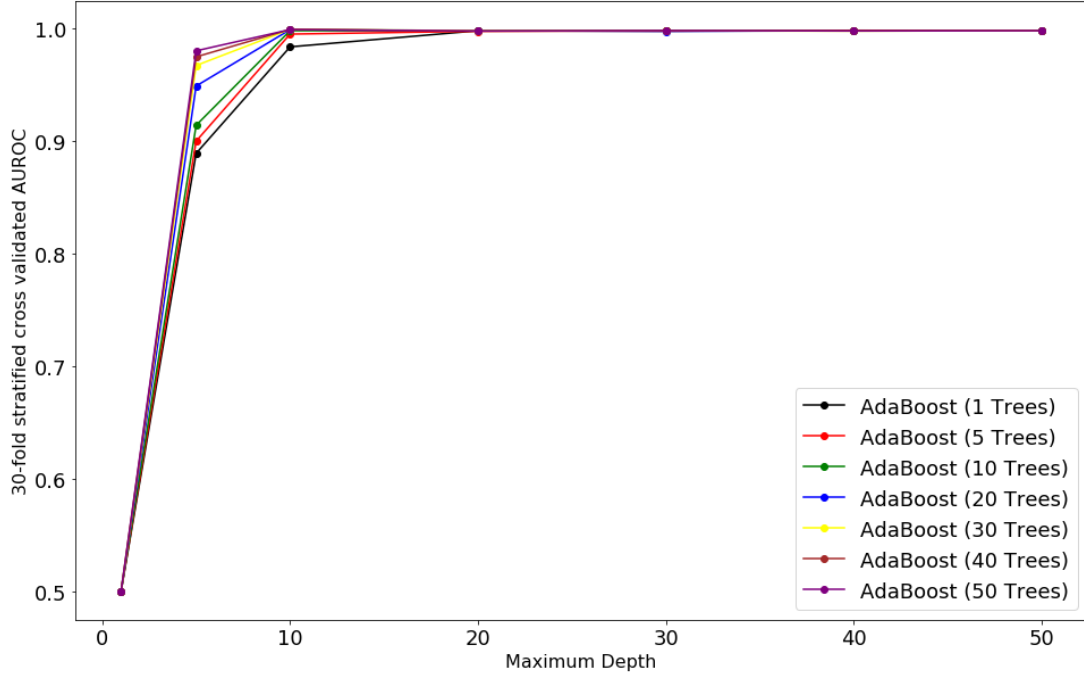
Figure 87. The results of the cross-validation process of AdaBoost ML algorithm based on the true-triaxial training experiments dataset for progressive failure prediction training.

The results of the integrated and trained FEM-ML model for progressive failure prediction are summarized in **Figure 88**. The summary shows, in the third column from the left, the maps of the failure zones progression from the initial CT-scan to the final CT-scan for each experiment. In the fourth column from the left, the results for progressive failure prediction based on the standalone FEM with the Mogi failure criterion are shown. In the fifth column from the left, the results for progressive failure prediction based on the integrated FEM-ML are shown. For all shown experiments, the integrated FEM-ML model produced a more accurate prediction of failure zones as it corresponds closely to the failure maps based on the CT-scans. It should be noted that when training the progressive failure model, experiment-1 and experiment-6 were used for training the

machine learning algorithm while experiment-5 was not included in training. Quantitatively, as shown in **Figure 88**, the training experiments yielded an AUROC of 0.99 while the test experiment yielded an AUROC of 0.95.

	Experiment Number	Image Code	Conventional Failure Criterion (Mogi)	FEM-ML Adjusted Failure Criterion	AUROC
Used in training the machine learning models	Experiment 1				0.99
	Experiment 6				0.99
Not used in training	Experiment 5				0.95

Figure 88. Summary of the results of failure prediction across a single layer for the true-triaxial experiments used for progressive failure modelling. The results compare failure prediction from the conventional failure criterion (Mogi) to the integrated FEM-ML model (progressive failure).

3.2.5.2 An Illustration of Training with Anisotropic Properties and Heterogeneity

  Indicators

   The assumption of homogeneity of the rock samples does not allow the prediction of asymmetric or irregular failure zones in a circular borehole, which can only be done with the introduction of spatial variations and directional mechanical properties

(anisotropy and heterogeneity indicators). It is useful to be able to predict the asymmetric failure zones in a manner that closely matches the failure zones distribution seen in the post-test CT-scans as this can improve the model accuracy considerably. This is especially useful in experiments that showed a localization of failure in certain zones around the borehole. An example of this is experiment-1 as shown in **Figure 89**. The localization of failure in top left quadrant of the borehole can be explained by one of two scenarios. One, the initial borehole shape might not have been perfectly circular due to the presence a notch in the top left quadrant of the borehole. However, since no pre-test CT-scan is available to confirm this scenario, it is discarded. Also, if the pre-test CT-scan was available and it confirms the presence of the notch, in that case the progressive failure model discussed in the previous section can account for the initial borehole irregularity to eventually predict the asymmetric failure zones in a manner that closely matches the failure zones distribution seen in the post-test CT-scans. The second scenario for explaining the localization of failure in experiment-1 is the influence of anisotropic properties or the presence of heterogeneities. Since this data is also not available, if it was to be manufactured or assumed, the anisotropic properties and the presence of heterogeneities can be introduced as a training feature for each mesh node. With these new features, the FEM-ML has the potential to predict failure zones that matches the post-test CT-scan as can be seen in the bottom of **Figure 89**.
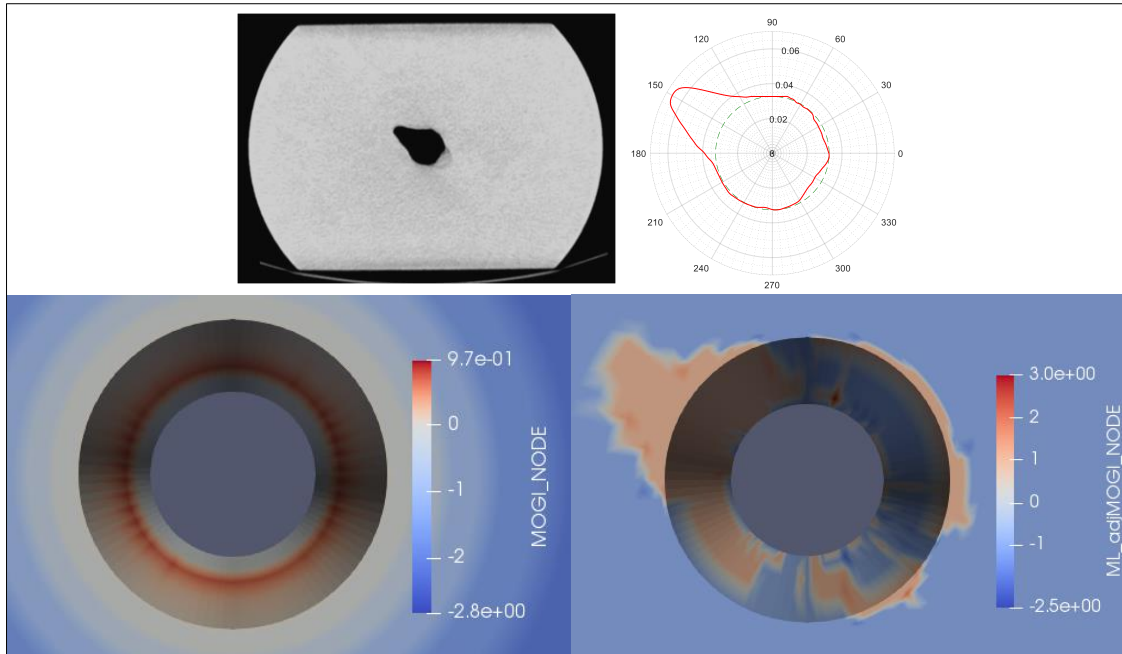
Figure 89. The potential results for failure prediction across a single layer for experiment-1. Top: post-test CT-scan and its corresponding failure map. Bottom left: result based on the standalone FEM with the Mogi failure criterion. Bottom right: potential results based on the integrated FEM-ML model with manufactured anisotropic properties and heterogeneity indicators.

Similar illustration can also be made from the results of the rest of the experiments as summarized in **Figure 90**. The summary shows, in the third column from the left, the maps of the failure zone for a single layer based on a single CT-scan for each experiment. In the fourth column from the left, the results for failure prediction across a single layer based on the standalone FEM with the Mogi failure criterion are shown. Finally, the right most column shows the potential results for failure prediction across a single layer based on the integrated FEM-ML that is trained with manufactured anisotropic properties and heterogeneity indicators.

	Experiment Number	CT-scan Code Failure Map	Conventional Failure Criterion (Mogi)	FEM-ML Adjusted Failure Criterion
Used in training the machine learning models	Experiment 1			
	Experiment 2			
	Experiment 3			
	Experiment 5			
	Experiment 6			
Not used in training	Experiment 4			

Figure 90. Summary of the potential results of failure prediction across a single layer for all available true-triaxial experiments. The results compare failure prediction from the conventional failure criterion (Mogi) to the integrated FEM-ML model if it is trained with manufactured anisotropic properties and heterogeneity indicators (single step failure).

CHAPTER IV

FIELD APPLICATIONS OF THE MODEL[*]

As the integrated model full structure is established through mathematical formulations and its performance is validated through lab experiments, all that remains is to provide an illustration of its field applications and the influence it can have on drilling operations. This is done in this chapter through the use of real-life offset or historical wells data from an offshore gas field in southeast Asia. The first step is using a portion of the available offset wells data as a training dataset for integrated model. This ensures that the model is trained based on a representative and sufficiently diverse dataset while also retaining a portion of the offset wells data to test the performance of the model. Once the integrated model is trained, its performance is evaluated to ensure its reliability. The training and evaluation processes are similar to the one shown in the experimental work-based model except for the type and source of data being used. Therefore, it is necessary to discuss the differences in sourcing and processing field-based data. The next step is showcasing the different scenarios where the use of the model is both applicable and impactful, especially when compared to the conventional geomechanics models being currently used in the industry. The first type of the integrated model field application scenario is real-time implementation. In this application, the area within a typical drilling

143

operations workflow where the model fits is discussed. Eventually, previously unattainable improvements that the integrated model brings to the table are illustrated. The same is also done for the other types of field application scenarios of the integrated model, which include examining the effect of varying intensity of wellbore enlargements as observed from shale shakers, examining the effect of wellbore size, and predicting progressive wellbore failure. Finally, while the modelling scheme proposed by the integrated model is a physics-based data science scheme, which means it does not require large datasets for training, it still requires a representative and sufficiently diverse dataset. Therefore, this chapter is capped with a discussion of potential applications for which no sufficient dataset is made available for this work.

4.1 Training the Integrated Model Based on Field Data

The process for training the integrated model based on field data is much similar to the process followed in the previous chapter. The main difference is that instead of using lab-based images such as CT-scans to provide the classification of mesh nodes for the supervised training, field-based image logs are used. Other differences include the sources of data required and the adjustment of the boundary conditions for the FEM to reflect the in-situ stress loading instead of lab equipment loading. Also, the field data collected has to be consistent in that it all belongs to a certain formation; and therefore, a certain type of rock. Based on this, training the integrated model follows these steps:

1. Collect drilling records from applicable offset wells to provide a description of wellbore rock failure that was experienced while drilling and the variations in mud weight used while drilling the interval or formation of interest

2. Collect the results of standard triaxial tests that were performed on cores from the formation of interest and use those for stress-strain constitutive modelling

3. Compile all the available image logs of the formation of interest. Since the training process is performed offline; i.e. not in real-time or while drilling, the logs can be from both LWD tools or wireline tools. Also, all types of image logs are of value and that includes resistivity-based images and ultrasonic-based images; however, ultrasonic-based images are preferred as they have no gaps, less noise, and offer more descriptive details of the state of the wellbore

4. Collect all the available multi-arm mechanical caliper logs data to be used for validating the image log interpretations

5. Collect in-situ stress measurements for each offset well ideally from mini-frac tests or alternatively from 1D mechanical earth models

6. Mesh each offset well data separately and run it through the FEM to extract all the load-dependent machine learning features at each mesh node

7. Perform image log interpretations and use the results to create mesh node failure classification

8. Compile all the data from both the FEM runs and the image log interpretations for all the available offset wells

9. Trial different machine learning algorithms on the compiled offset wells dataset by performing a multi-fold stratified cross validation and hyperparameter tuning

10. Select the best performing trained machine learning algorithm and import it to the FEM. Test its performance using a new offset well that was not used in the training process. If it shows satisfactory performance, designate it as the final trained and integrated FEM-ML. This is the model that will be used for evaluating rock failure and the drilling mud window for the formation of interest in all of the relevant field applications.

4.1.1 Sourcing and Processing Relevant Offset Wells Data

The sourcing and processing of relevant offset wells data as detailed in the previous discussion is illustrated here through real-life offset wells data from an offshore gas field in southeast Asia. This field targets a high strength sandstone gas reservoir. Since all the available data are from the reservoir section, the sandstone reservoir is the formation of interest; and training the integrated model is aimed to improve the drilling experience through this interval. The available data from this field consist of seven offset wells. Because the wellbore size used to drill the reservoir section is either 6" or 8.5", five wells will be used for training, while one 6" well and one 8.5" well will be used for testing the performance of the integrated model. It is worth to be noted that all the wells are vertically drilled and while the lack of diversity in the dataset in terms of drilling direction is challenging, especially for purely data driven machine learning, it is useful here as it will be used to illustrate the robustness of the physics-based data science approach being

followed here. This means that having a dataset that consists of vertical wells only is not a limiting factor when it comes to building a model that can predict rock failure at any drilling direction.

The first type of field data sourced from the offset wells is the wells profile. This includes the wellbore size, inclination, and azimuth. While it is true that all the available offset wells are vertical, it is not uncommon for vertically planned wells to experience some degrees of unintentional deviation while drilling. Therefore, it is essential that these deviations are extracted from the wellbore orientation surveys and reflected in offset wells data. The wellbore profile data for each offset well along with classification of training and testing offset wells are shown in **Table 7**.

Table 7. Details of offset wells profiles and training and testing classification

Well	Inclination, degrees	Azimuth, degrees	Wellbore Diameter, inches	Used for training?
Well-1	0º	0º	8.5	Yes
Well-2	7º	234º	6	Yes
Well-3	9.6º	275º	8.5	Yes
Well-4	5º	98º	8.5	Yes
Well-5	1º	87º	8.5	Yes
Well-6	1º	87º	8.5	No
Well-7	3º	17º	6	No

The next type of data to be sourced are the in-situ stress values and the definition of intervals where a representative image log is available. For the in-situ stress, the minimum horizontal stress values are gathered based on leak-off tests or mini-frac tests and the vertical overburden stress values are based on density log measurements. As for the maximum horizontal stress values and directions, these are constrained based on image

log analysis and stress polygon plots as shown in the example from well-1 in **Figure 91**. The brown lines on the stress polygon to the right in this figure represent contours of equal breakout width and the green lines are the constrained values of the maximum horizontal stress based on the breakout width observed from the image log. The breakout region is highlighted with pink boxes on the image log to the left in **Figure 91**.
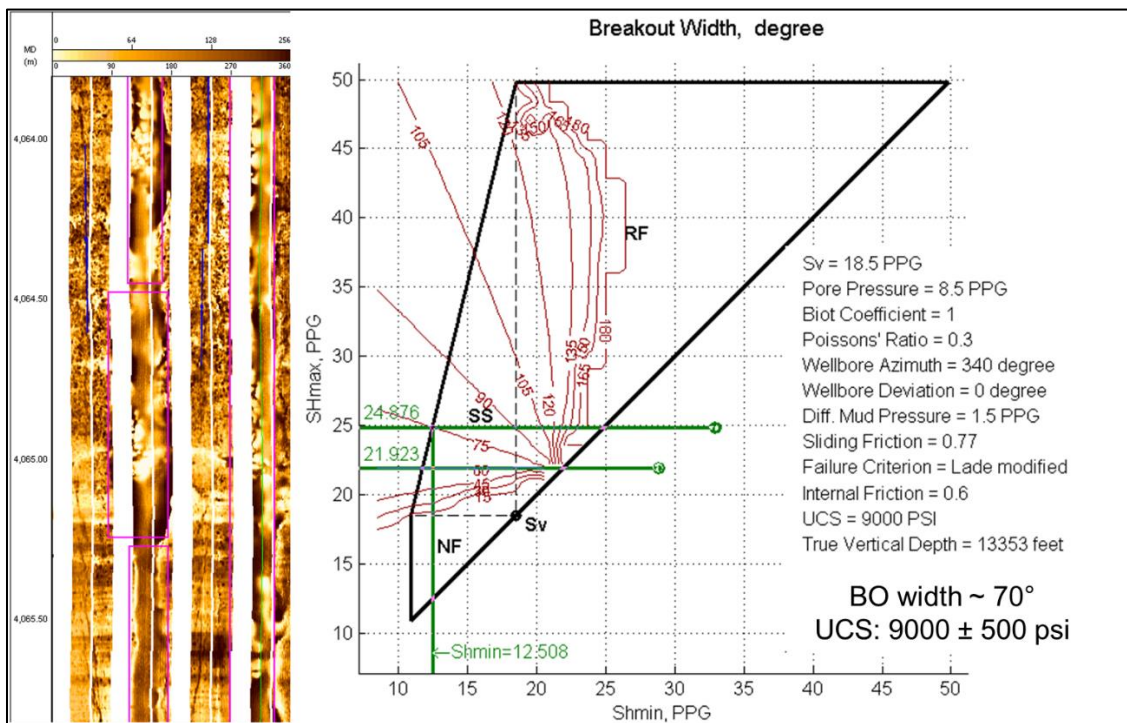
Figure 91. Using breakouts width to constrain σH magnitude in well-1. The brown lines on the stress polygon to the right represent contours of equal breakout width.

The interval of a representative image log is the depth that has an image log of acceptable quality in terms of image resolution, which exhibits wellbore failure that is representative of the failure throughout the wellbore section. All the image logs available are resistivity-based only and are taken by the formation micro-imager or micro-scanner

148

(FMI/FMS) tool. The in-situ stress values, the definition of depths of a representative image log, and the UCS values predicted from wireline logs for all the available offset wells data are shown in **Table 8**. The values of the in-situ stress and direction along with the wellbore profile information such as inclination and azimuth are used for preprocessing the loading for the FEM based on the equations discussed in the model development chapter (Equation (36) to Equation (41)). The predicted UCS values are used for interpolating the stress-strain curve for constitutional modelling.

Table 8. The in-situ stress values and their directions, the definition of intervals where a representative image log is available, and the predicted UCS for all the available offset wells

Well	Vertical Stress σ_V, ppg	Maximum Horizontal Stress σ_H and Its Direction, ppg and degrees	Minimum Horizontal Stress σ_h, ppg	Depth of a Representative FMI, ft	Predicted UCS, psia
Well-1	17.4	21.3, 42°N	12.65	13,337	9,000
Well-2	18	20.8, 56°N	12.1	12,795	9,000
Well-3	18.6	21.6, 57°N	12.2	12,909	8,000
Well-4	17.9	20.5, 32°N	12.38	12,790	9,000
Well-5	18	20.9, 59°N	12.2	12,994	9,000
Well-6	18	21.4, 54°N	12.2	12,558	5,000
Well-7	18.2	21.1, 53°N	12.4	13,337	10,000

The pore pressure values are also essential for calculating the effective stresses for the concept of poro-elasticity. These values are based on a modular formation dynamic tester (MDT) measurements from the formation of interest. The final type of information which is related to the loads applied to the wellbore is the mud weight. It is not uncommon for the mud weight to change throughout the interval. The mud weight is changed by the drilling engineer as a response to various events and occurrences that take place while

drilling, which include lost circulation, pack-offs, stuck-pipe incidents, and change in the drilling fluid formulation for chemical interaction issues. The different mud weights used to drill the formation of interest in the offset wells have to be reflected on the loading history in the FEM. This is necessary to capture the effect of stress hysteresis and more accurately model the stress state around the wellbore. The mud weights for the offset wells are best extracted from drilling summary plots. An example of a drilling summary plot for well-1 is shown in **Figure 92**, where the mud weight for the reservoir section (formation of interest) was initially 10.3 ppg; but it was eventually dropped twice to 10 ppg and 9.8 ppg respectively while drilling this section. The summary of pore pressures and mud weights for all the offset well is shown in **Table 9**. It is important to note that the mud weights used for the offset wells were determined to be sufficient for wellbore stability based on pre-drilling analysis by geomechanics and drilling engineers. This pre-drilling analysis originally performed by the field engineers is based on an analytical solution and 1D geomechanics earth model with the Lade failure criterion.
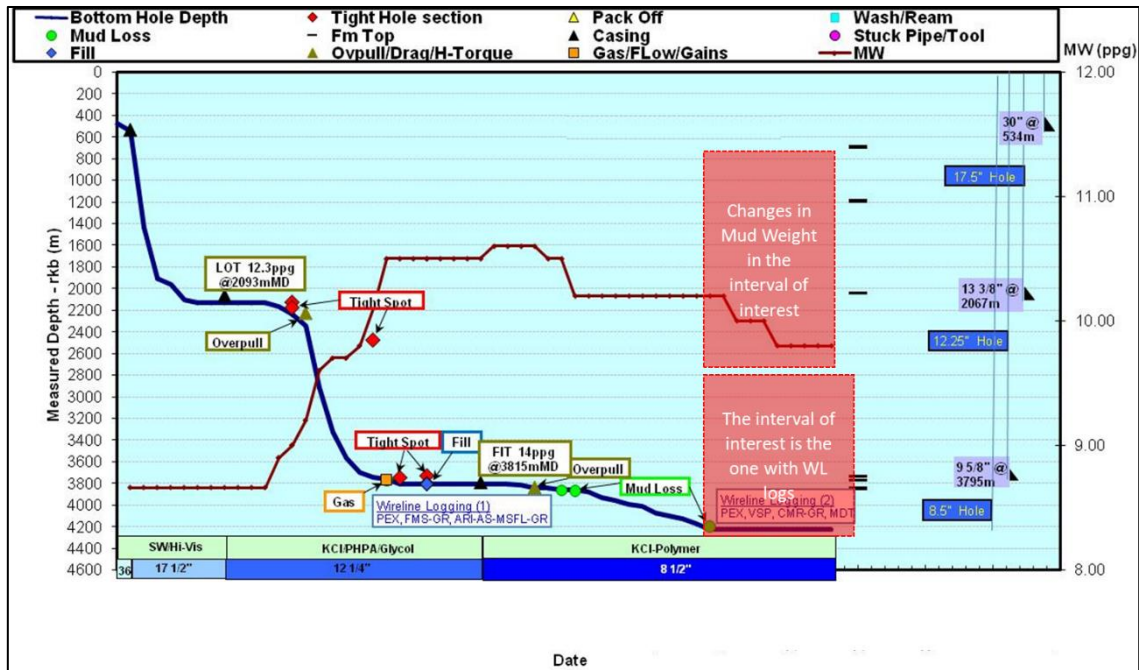
Figure 92. The drilling summary plot for well-1 used to extract variations in mud weight while drilling the formation of interest.

Table 9. Pore pressure values and the variation in the mud weight within the formation of interest for all the available offset wells

Well	Pore Pressure, ppg	Variations in mud weight, ppg
Well-1	8.5	10.3 → 10 → 9.8
Well-2	8.5	9
Well-3	8.6	9.6 → 9.8
Well-4	8.6	9.1 → 9.3
Well-5	8.5	9.25
Well-6	8.4	9.1 → 9.2
Well-7	8.5	9.2

The final step of offset wells data sourcing is summarizing drilling difficulties, specifically those which are related to rock failure. These can be collected from drilling morning reports, well completion report trouble summaries, and from multi-arm mechanical caliper log measurements. An example of a mechanical caliper log data for

well-1 is shown in **Figure 93**, where it illustrates the regions of enlargements and drilling induced fractures. The summary of the actual drilling difficulties encountered while drilling the formation of interest in each of the offset wells is shown in **Table 10**.
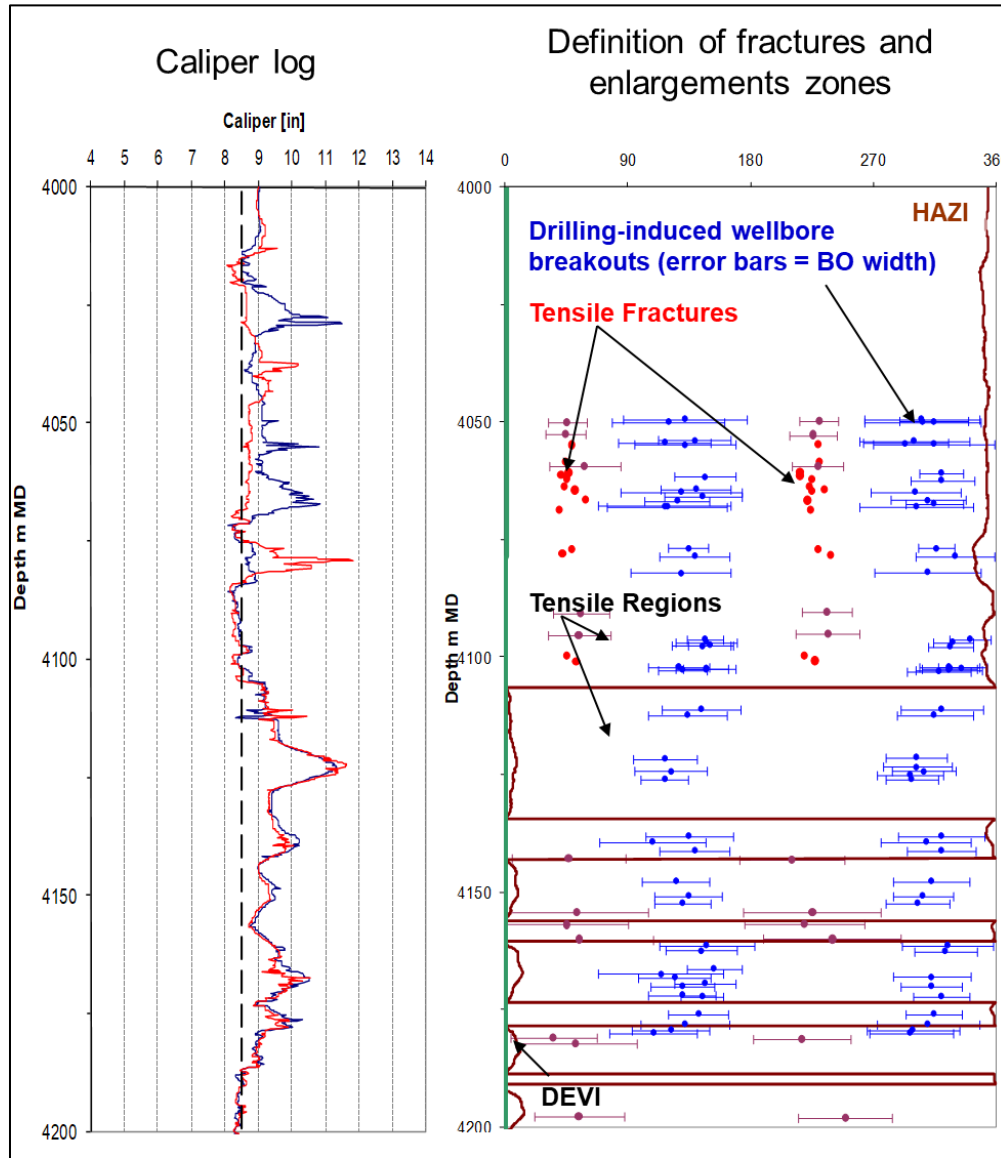
Figure 93. Well-1 wellbore failure description where: Left: Mechanical caliper log data. Right: Definition if enlargements and drilling induced fractures zones based on caliper data and FMI logs.

Table 10. The summary of the actual drilling difficulties encountered while drilling the formation of interest in each of the offset wells

Well	Summary of Drilling Difficulties Related to Rock Failure
Well-1	Wellbore instabilities. Major drilling problems were limited to tight hole, overpull and fill. Static losses reported in deeper section of the hole, which could be related to the occurrence of natural fractures.
Well-2	Wellbore instabilities. Multiple pack off, tight hole and fluid losses led to parted string and two sidetracks. Excessive cuttings and fill were seen as well.
Well-3	Tight hole and over pull. Casing string stuck and parted at 12,316 ft, so that 9 5/8'' casing was set at 12,320 ft, about 194 ft higher than hole depth. Hole packed off twice after drilling out the 9 5/8'' casing shoe at 12,513 ft and 12,536 separately. No losses were reported. Gas shows reported in deeper section related to natural fractures.
Well-4	Hole were sidetracked at 12,484 because the core barrel was lost in the hole. Fluid losses were reported during drilling the 8 ½'' hole section, which could be related to the occurrence of natural fractures. In the sidetrack, static losses reported in deeper sections of the hole related to natural fractures. Natural fractures on FMI at ~12,674 ft and 12,864 ft depths which could be related to the fluid losses occurred between 12,749 ft and 12,123 ft.
Well-5	Tight hole. Repeated stalling of drill string was observed while drilling to 10,417, hole nearly packed off, 13 ft fill. No fluid losses
Well-6	Overpull, excessive washing and reaming, fill and fluid losses. Serious fluid losses were experienced during drilling the 8 ½'' hole section, related to the of natural fractures.
Well-7	tight hole, over pull and fill. Excessive cavings were encountered at 9,842 ft. Hole was sidetracked from 5,426.5 ft because of a twisted off while drilling and the junk was left in hole. Fluid losses and gas were reported during drilling the 8 ½'' hole section, which could be related to natural fractures. Static losses reported in deeper related to natural fractures. Natural fractures are seen on FMI at ~12,927 ft where the fluid losses were reported.

The value of summarizing drilling difficulties from offset wells is in that it helps highlight the disparity between the actual mud weight that was used and the drilling difficulty caused by it against the mud weight that is calculated by the trained integrated model and the supposedly drilling difficulties it can help avoid.

Now that all the necessary data are collected, quantifying wellbore failure based on image logs is needed for the purpose of generating mesh node classification. As

mentioned previously, the interpretation of resistivity-based image logs, which are the only type available from the offset wells here, using the image analyzer function requires validation through mechanical caliper log data. This validation ensures that both the circumferential and radial extensions of wellbore enlargement zones as interpreted from the image log agree with those seen in the caliper log. An example of the validated FMI image log interpretation from well-1 is shown in **Figure 94**. The summary of wellbore enlargements extension as observed from FMI image logs for all the offset wells and within the formation of interest is shown in **Table 11**. The interpretation of the image log from well-1 shown in **Figure 94** is calibrated so that it reflects an enlargement angle of 70°.
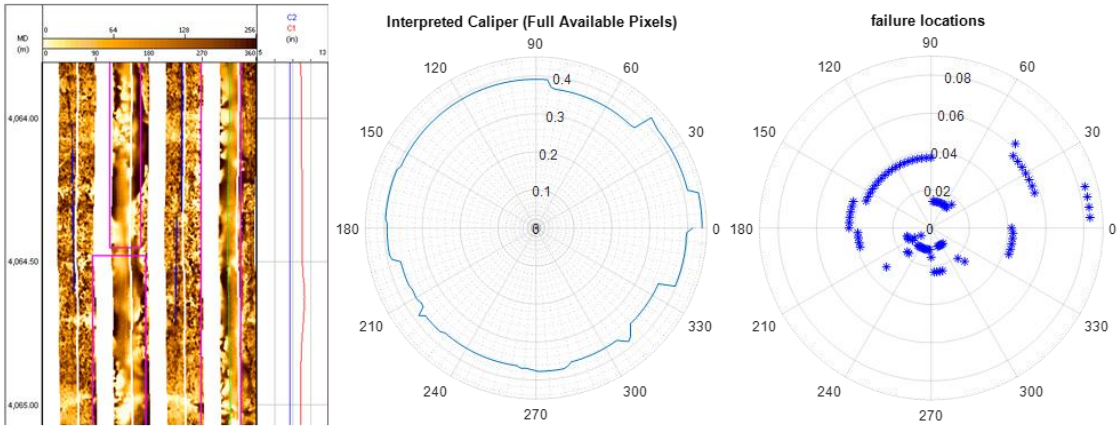
Figure 94. FMI image log interpretation for well-1 where: Left: the FMI image and caliper log with pink boxes highlighting areas of enlargement. Center: the results of the resistivity image log interpretation showing radial measurements in ft. in polar coordinates to reflect the actual shape of the wellbore cross-section. Right: polar chart defining the magnitude in ft. and location of enlargements around the wellbore.

154

Table 11. The summary of wellbore enlargements extension as observed from FMI image logs for all the offset wells and within the formation of interest

Well	Extension of Wellbore Enlargement (Breakout angle), degrees
Well-1	70°
Well-2	0° (no discernible enlargement zones)
Well-3	30°
Well-4	0° (no discernible enlargement zones)
Well-5	0° (no discernible enlargement zones)
Well-6	0° (no discernible enlargement zones)
Well-7	30°

It is worth noting that for the offset wells that showed 0° breakout angle in **Table 11**, it does not necessarily mean that there was no wellbore failure and the wellbore is in perfect condition. These wells still showed minimal and scattered enlargements around the wellbore, however, there were no well-defined zones of enlargements in the direction of the minimum horizontal stress, which are the only type of enlargements zones that are considered when referring to the breakout angle conventionally in the industry and literature. This means that the offset wells with 0° breakout angle can still show an irregular or non-uniform wellbore shape; albeit without discernible zones of concentrated enlargements in the direction of the minimum horizontal stress.

Now that all the data is compiled and processed for all the training offset wells as discussed thus far in this section, each well is then modelled in the FEM separately. Then, the output of the FEM, which describes the stresses, stress tensor invariants, and distance to the failure surface from a conventional failure criterion, is extracted from all training wells FEM results and compiled in a single dataset. This dataset consists of data points which represent a mesh nodes information for all training wells. The outcome of the FMI image log analysis is also compiled within the same dataset to represent mesh node failure

classification for each data point. Finally, this compiled dataset is used to train the machine learning algorithms and decide which algorithm is most suitable. Once the machine learning algorithm is selected, the trained machine learning algorithm is integrated with the FEM. Test wells—wells that were not included in training— are modelled with integrated and trained FEM-ML. If the performance on the test well is satisfactory, the integrated and trained FEM-ML can be used for predicting wellbore rock failure both in the pre-drilling phase and the while drilling phase.

4.1.2 Stress-strain Curves and Constitutive Modelling

Standard triaxial tests are conducted on the sandstone plugs from the gas reservoir (formation of interest in offset wells) to produce the stress strain curves. The curves shown in **Figure 95** are for a 5,000 psi UCS, while similar curves are produced for different UCS values. The tests are conducted at different confining pressure following the guidelines set by Morita and Nagano (2016) and Morita and Gray (1980) for constructing nonlinear constitutive stress-strain relations. These relations are used in the FEM post processor calculations to predict the stress state at each node based on the displacements that were determined by the FEM solver.
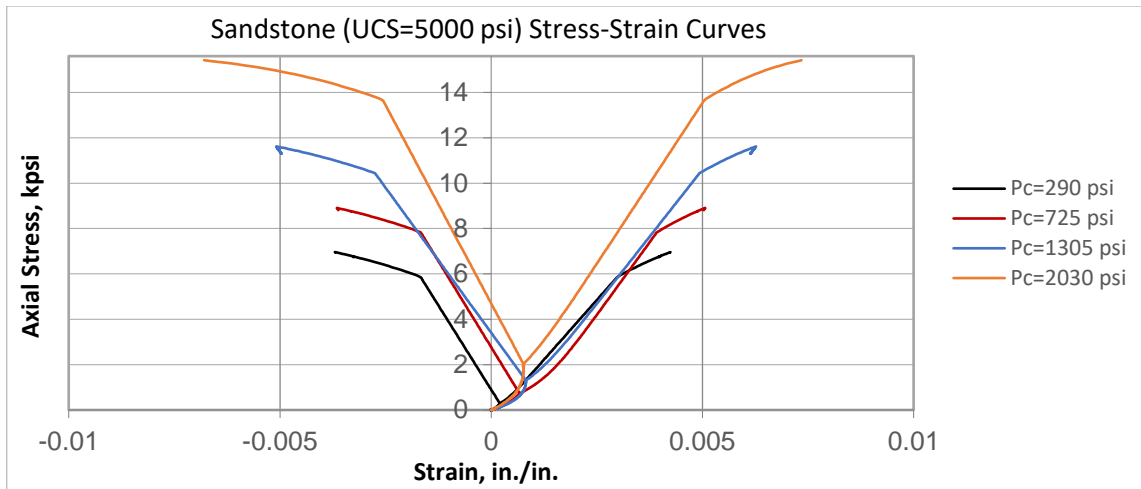
Figure 95. The stress-strain curves of the sandstone gas reservoir from the offset wells. Curves are from triaxial tests at different confining pressures, which is used for constructing the nonlinear constitutive stress-strain relations.

4.1.3 Evaluation of the Performance of the Integrated Finite Element and Machine

Learning Model

This section deals with two main field training application issues. The first issue is the application of training an integrated model that is generalizable for purpose of single step failure prediction. The second issue is training an integrated model that is able to predict progressive failure steps. The progressive failure prediction entails the consideration of an irregularly shaped wellbore as the starting step for failure prediction. This is as opposed to the single step failure prediction where the starting wellbore shape is circular. The main motivation for the juxtaposition of these two issues is to illustrate the effect of the wellbore shape irregularity. The generalizable integrated model for a single step failure prediction relies on a conventional failure criterion along with the principle stresses and plastic stain as features in the training process. This means that the geometric shape indicators play no role in the generalizable integrated model for a single step failure

157

prediction. However, for progressive failure prediction, using only a conventional failure criterion along with the principle stresses as features in the training process does not yield similar prediction accuracy. Therefore, as the starting wellbore shape becomes irregular, the geometric shape indicators will need to be introduced into the training process. As discussed in the second chapter, the geometric shape indicators are load-independent features. In cases where the trained model relies more heavily on these load-independent features, the trained model will be non-responsive to loads. This means it can only be used for defining the failure zones but not for determining the mud weight to prevent failure. To address this issue, when training a model for failure prediction in an irregularly shaped wellbore, different geometric shape indicators are introduced in the training process in a trial and error manner until a satisfactory compromise is reached between the desired accuracy and the trained model load responsiveness.

An illustration of the different stages of wellbore failure that are considered in this section is shown in **Figure 96**. This illustration is based on Well-1 from the offset wells data presented at the beginning of this chapter. For single step failure prediction models, the circular wellbore is considered as the starting wellbore shape and the FMI-based wellbore shape is the final shape that needs to be predicted. Progressive failure models consider an intermediate failure step, which practically is produced from an LWD images, as the starting wellbore shape and the FMI-based wellbore shape as the final wellbore shape.
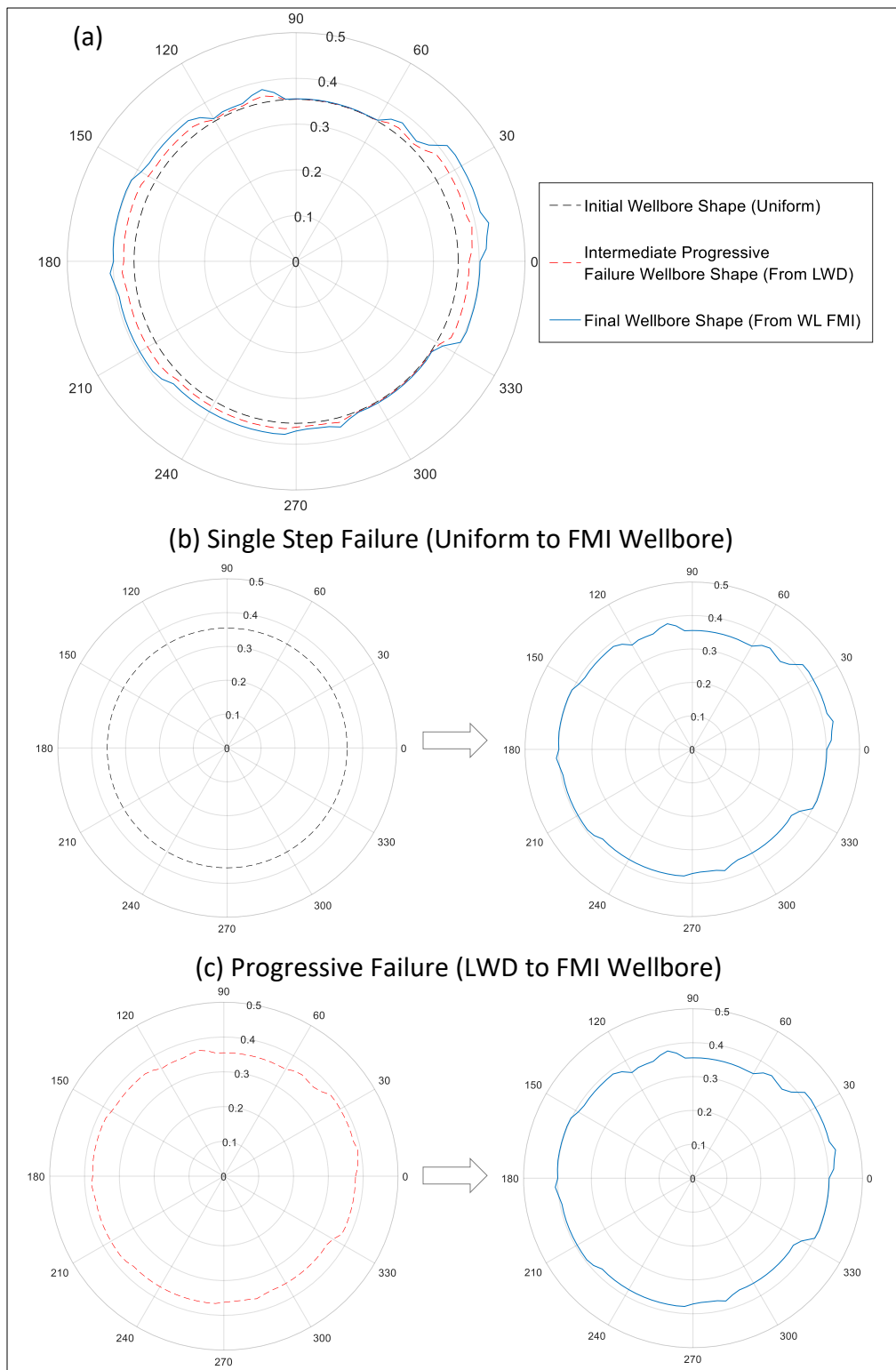
Figure 96. An illustration of the different stages of wellbore failure.

4.1.3.1 Generalizable Single Step Failure Prediction Model Training

There are two main issues that need to be considered for the model to be fully generalizable. The first issue is that the dataset will need to include information from different rock types. This is achieved here by combining the datasets from the large-scale experiments, the true-triaxial experiments, and from the field offset wells into a single dataset and training the model using it. The second issue, as discussed in the second chapter, there are certain features that will need to be excluded when the training dataset does not belong to a homogenous rock. These features include all displacement and strain parameters. The reason for this is heterogenous rock bodies fail at different strains and displacements according to the variations of their mechanical properties. Therefore, unless these variations are also included in the training process, the strain and displacement data cannot be used as training features. This second issue is resolved here by limiting the training features for the single step failure model to a conventional failure criterion, the principle stresses, and the plastic strain.

As mentioned at the beginning of this section, the main motivation for the juxtaposition of the generalizable and the progressive failure models is to illustrate the effect of the wellbore shape irregularity. The generalizable integrated model for a single step failure prediction does not rely the geometric shape indicators. In fact, for some geometric shape indicators, even if they were to be included in the training process, they will play no role in predicting failure. This is because these geometric shape indicators are constants for circular wellbores. The geometric shape indicator of the first derivative of the borehole curve radius (r_w') is perfect for illustrating this because of its mathematical

convivence. r'_w is always zero when the wellbore starting shape is circular, which is the case in single step failure prediction models.

The results of training based on the combined dataset of all available rock types for the purpose of a generalizable single step failure prediction are shown in **Figure 97**. The best performing model as shown in this plot is bagging of decision trees with a maximum AUROC of 0.87 and a maximum accuracy of 0.92.
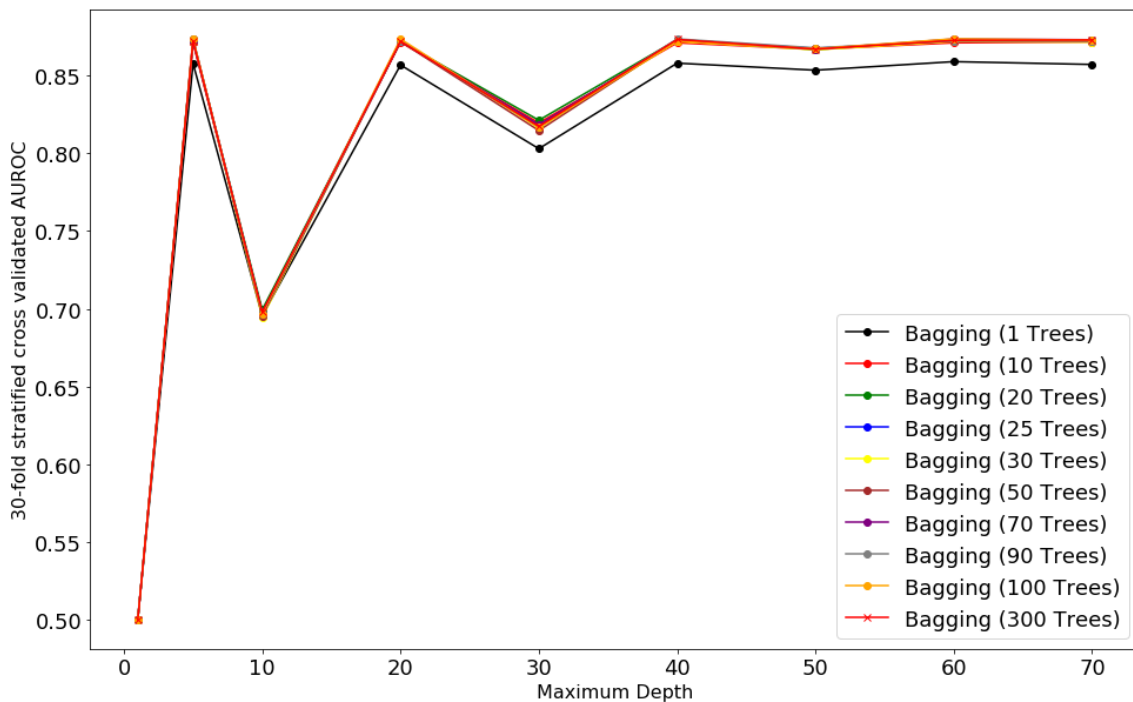
Figure 97. The results of the cross-validation process of the bagging decision trees ML algorithm based on the combined dataset from all available rock types for the purpose of a generalizable single step failure prediction training.

The contribution of each feature to failure prediction is produced as before using the XGBoost algorithm feature importance. As shown in **Figure 98**, even though the

geometric shape indicator of the first derivative of the borehole curve radius (r_w') is included in the training process, it has no effect on predicting failure.

Figure 98. Features importance from the XGBoost algorithm that was trained based on the combined dataset from all available rock types for the purpose of a generalizable single step failure prediction training.

As also can be seen from the feature importance in **Figure 98**, much like the case of the true-triaxial experiments, the minimum principal stress has the highest importance. Overall, the observations made based on feature importance can be used to enhance the understanding of the failure mechanisms taking place. For example, based on the features with the highest importance, the accuracy of a conventional criterion such as Lade's can be improved by adjusting its formulation.

Now to be able to comprehensively compare the results of the standalone FEM with a conventional failure criterion against the results of the integrated FEM-ML, all models will need to be treated as mesh node failure classification algorithms in similar fashion to what was done in the true-triaxial experiments analysis. This is useful because it will enable the evaluation of all models using the confusion matrix parameters, which allows for a systematic comparison between them. Starting with the predictive performance in well-1, which is one of the offset wells that were included in training the machine learning algorithm. The confusion matrix parameters, including accuracy, for well-1 are shown in **Figure 99**.
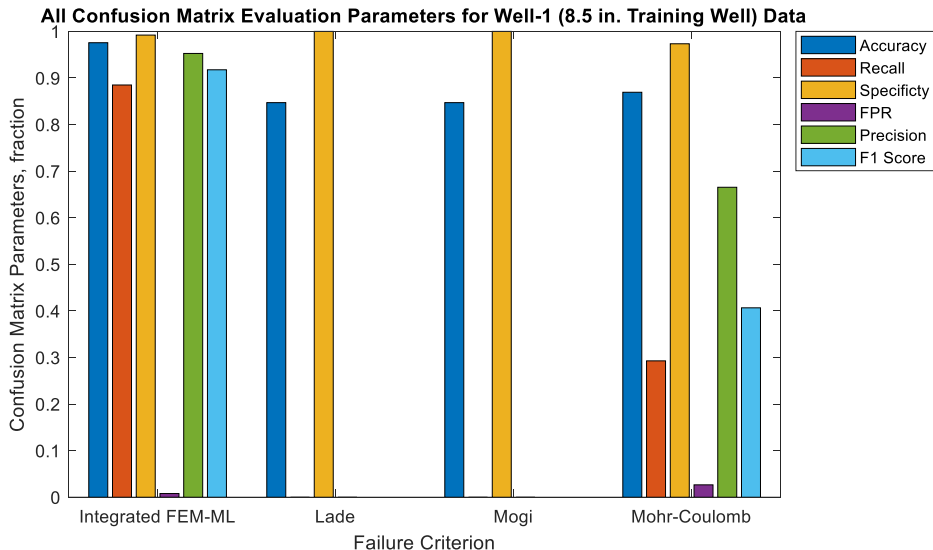
Figure 99. The confusion matrix parameters for well-1 (training well).

Based on all the different confusion matrix parameters, the integrated FEM-ML model performed considerably better than all the other conventional failure criteria in

predicting mesh nodes failure. The predictive performance over the full dataset of the training offset wells is shown in **Figure 100**. This evaluates the performance over all the training offset wells, which means that well-6 and well-7 are not included here. Looking at this chart, it is clear that the integrated FEM-ML model performed considerably better than all the conventional failure criteria over the training dataset.
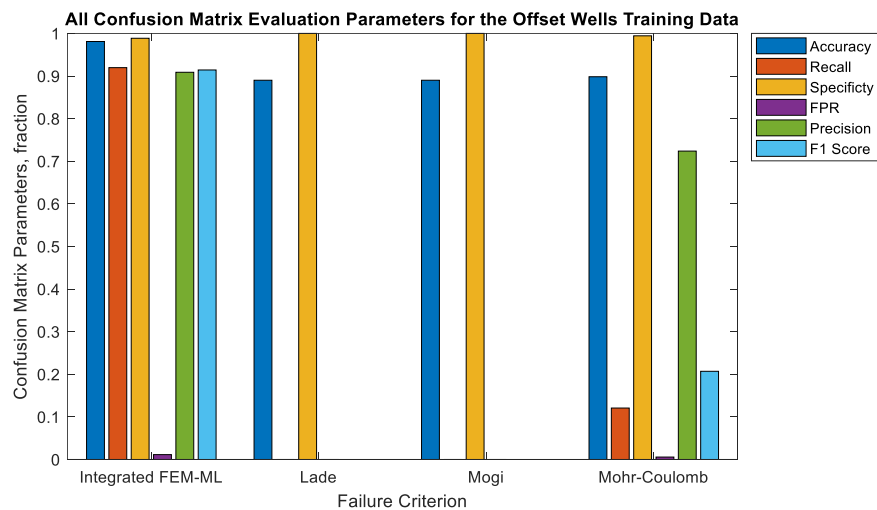
Figure 100. The overall confusion matrix parameters for all the training offset wells (well-6 and well-7 are not included).

However, the true measure of machine learning algorithms predictive performance is produced through data that was not seen by the algorithm during the training phase. For this reason, well-6 and well-7 were excluded from training and the predictive performance over them is evaluated separately as shown in **Figure 101** for well-6 and in **Figure 102** for well-7. It should be noted that well-6 has an 8.5 inches wellbore and well-7 has a 6 inches wellbore. These charts now show that the predictive performance of the integrated FEM-ML model has dropped noticeably. Still, even with predictive performance drop in

164

the testing phase, the integrated model is still performing significantly better than the conventional failure criteria. However, the comparison in performance between the conventional failure criteria and the integrated FEM-ML model is more nuanced in this dataset. There are three confusion matrix parameters in which conventional failure criteria performed comparably to the integrated model. These are the accuracy, specificity, and precision for Mohr-Coulomb only. For specificity, the charts in **Figure 101** and **Figure 102** show that all the conventional failure criteria specificity are marginally higher than that of the integrated model. The comparable specificity values are due to the nature of the dataset. The training dataset is heavily skewed towards non-failed mesh nodes as they represent around 87% of the dataset. As for the test wells, well-6 dataset is 93.22% non-failed nodes and well-7 dataset is 93.62% non-failed nodes. The uneven distribution of classes is to be expected since only the mesh nodes around the wellbore will fail and mesh nodes away from the wellbore all the way to the boundary will almost always be non-failed nodes in usual circumstances. This skewness of the dataset means a model that can merely predict only non-failed nodes correctly will always yield a high specificity. The same argument applies to the comparable accuracy values. As for the precision values from the Mohr-Coulomb criterion as seen in the charts in **Figure 101** and **Figure 102**, they are higher than those of integrated FEM-ML. In fact, in well-6, Mohr-Coulomb precision is 34% higher than that of the integrated FEM-ML. As mentioned before and as shown in the image analysis summery previously, well-6 image logs showed minimal and scattered enlargements around the wellbore, which meant that only 6.78% of its nodes are

failed nodes. The small number of mesh nodes to be classified as failed means that classifying a small portion of them incorrectly will lead to a significant drop in precision.
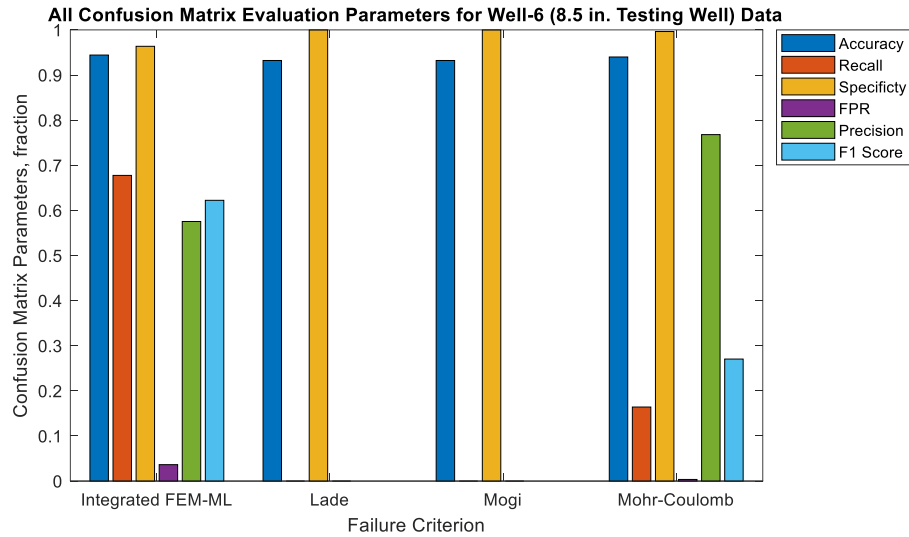
Figure 101. The confusion matrix parameters for well-6 (testing well).
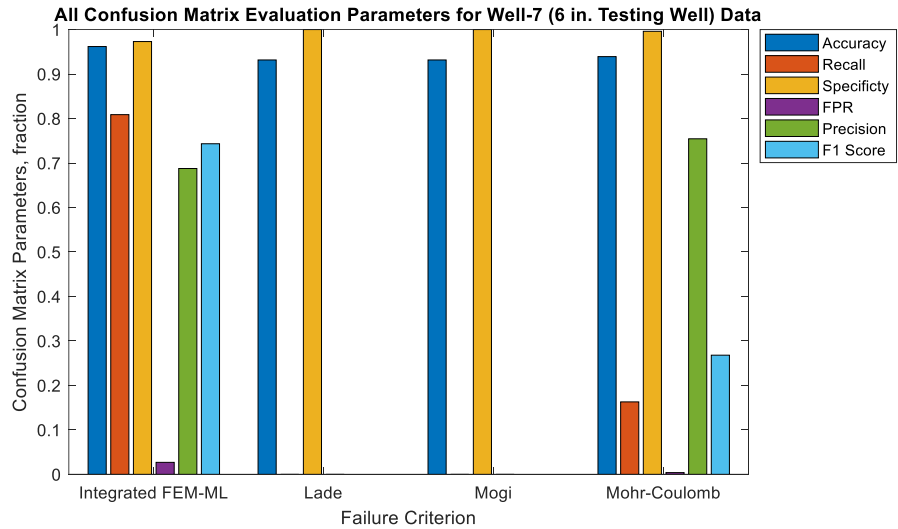
Figure 102. The confusion matrix parameters for well-7 (testing well).

Based on the discussion of the charts in **Figure 101** and **Figure 102**, it is clear that it is necessary to evaluate the overall performance of all the considered models while paying special attention to the nature of the dataset and the models prediction tendencies. To achieve this, the F1-score is used as the main confusion matrix parameter to compare the predictive performance of the models. The performance of the test wells shows that the integrated model has an F1-score that is 130% higher in well-6 and 178% higher in well-7 than that of the Mohr-Coulomb criterion. Based on the F1-score alone, this interpretation means the integrated FEM-ML has more than doubled the failure prediction accuracy that was provided by the conventional criteria. The low F1-score of the Mohr-Coulomb criterion contrasts its high precision due to the conservative nature of this criterion. As discussed in the true-triaxial experiments section, the Mohr-Coulomb criterion tends to be more conservative than other criteria in predicting failure because it does not consider the strengthening effect of the intermediate principle stress (Al-Ajmi and Zimmerman 2005). This explains the high precision and low F1-score of the Mohr-Coulomb criterion as it tends to predict more failure than might actually be occurring. Within the context of this application. This means the Mohr-Coulomb criterion predicts a high percentage of the failed nodes leading to a high precision, while incorrectly predicting a portion of non-failed nodes as failed nodes leading to a lower F1-score.

Other than the confusion matrix parameters, the improved performance provided by the integrated FEM-ML model can more practically be evaluated based on the mud weights it recommends, which can be compared against the mud weight recommended by

167

the conventional criterion in the offset wells along with the drilling difficulties that resulted from the conventional criterion determined mud weight. Such comparisons are carried out in the next sections.

4.1.3.2 Progressive Failure Prediction Model Training

For progressive failure models, the starting wellbore shape is an intermediate failure step, which shows an irregularly shaped wellbore. In practical setting, this intermediate failure step is ideally obtained from an LWD density-based image. However, since the dataset of offset wells introduced at the beginning of this chapter includes only wireline resistivity-based FMI images, some improvisation is needed to re-imagine the intermediate failure step that would have been provided by the LWD tool. This is done by interpreting the FMI-based wellbore shape and reducing the radial extension of failure zones by half. This works to create a manufactured intermediate failure step, which can be used to illustrate the progressive failure prediction model. The wellbore shapes from both the FMI-based image and the manufactured LWD image for four different offset wells from the dataset are shown in **Figure 103**.
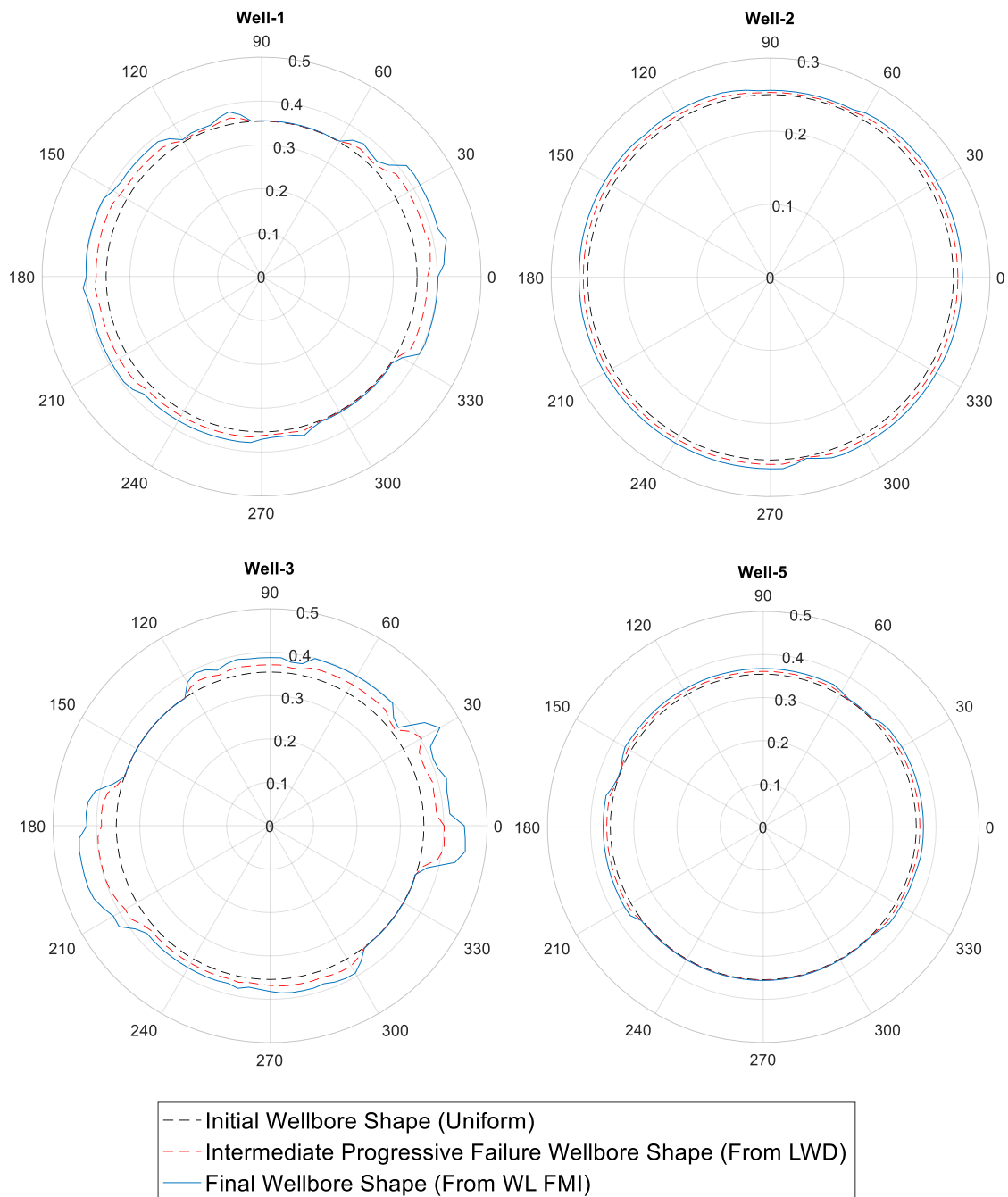
Figure 103. The wellbore shapes from both the FMI-based image and the manufactured LWD image for four different offset wells from the dataset.

Since the starting wellbore shape for progressive failure models is irregular, the geometric shape indicators will need to be introduced to the training process. As discussed previously, the geometric shape indicators are load-independent features. In cases where the trained model relies more heavily on these load-independent features, the trained model will be non-responsive to loads. To address this issue, when training a model for failure prediction in an irregularly shaped wellbore, different geometric shape indicators are introduced in the training process in a trial and error manner until a satisfactory compromise is reached between the desired accuracy and the trained model load responsiveness. The load responsiveness of each trained model is evaluated by examining the failure prediction results at different mud weights as shown in the second chapter. The summary of the trial and error process based on the combined dataset for progressive failure is shown in **Table 12**.

Table 12. Summary of the trial and error training process based on the combined dataset for progressive failure prediction.

ML Training Algorithm	Irregularity Related Training Features	AUROC	Is the Model Load-Responsive?
AdaBoost	r_w'	0.8548	Yes
AdaBoost	$r_w', RC/\beta_{Node}, CFO/\beta_{Node}, \theta_{Node}$	0.8905	Yes
XGBoost	r_w', CFO	0.8726	Yes
XGBoost	r_w', RC, r_{Node}	0.8909	No
XGBoost	$r_w', RC/\beta_{Node}, \theta_{Node}$	0.8824	Yes
Bagging	r_w'	0.7634	Yes
Bagging	r_w', RC, r_{Node}	0.9916	No
Bagging	$r_w', RC/\beta_{Node}, \theta_{Node}$	0.9041	Yes

Examinations of the contributions of each of the trialed irregularity related training features combinations are shown in **Figure 104**, **Figure 105**, **Figure 106**, and **Figure 107**.

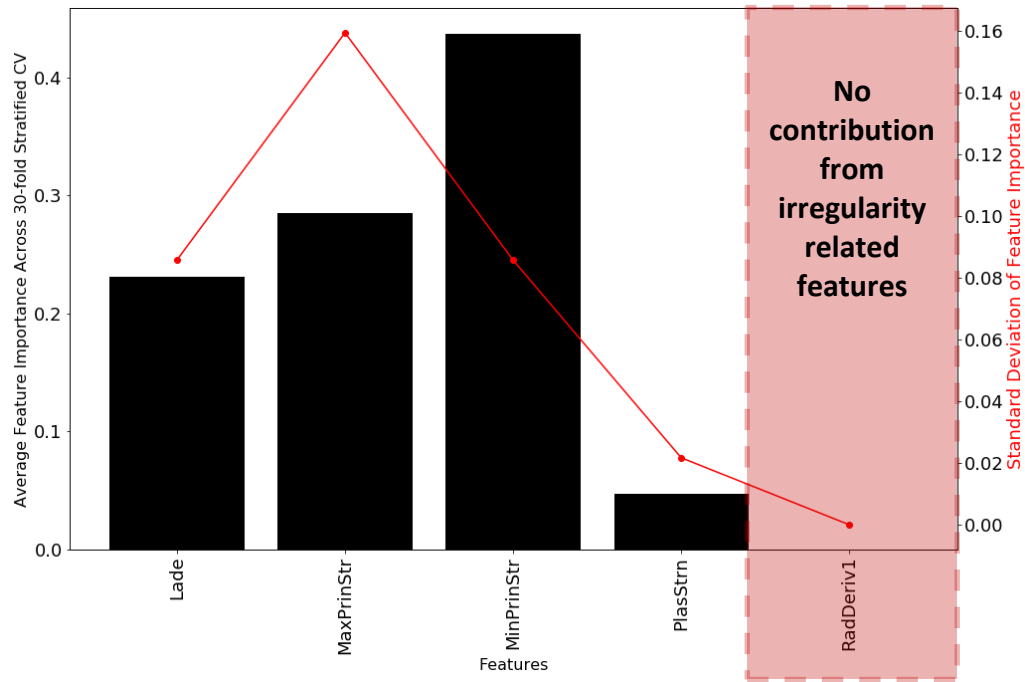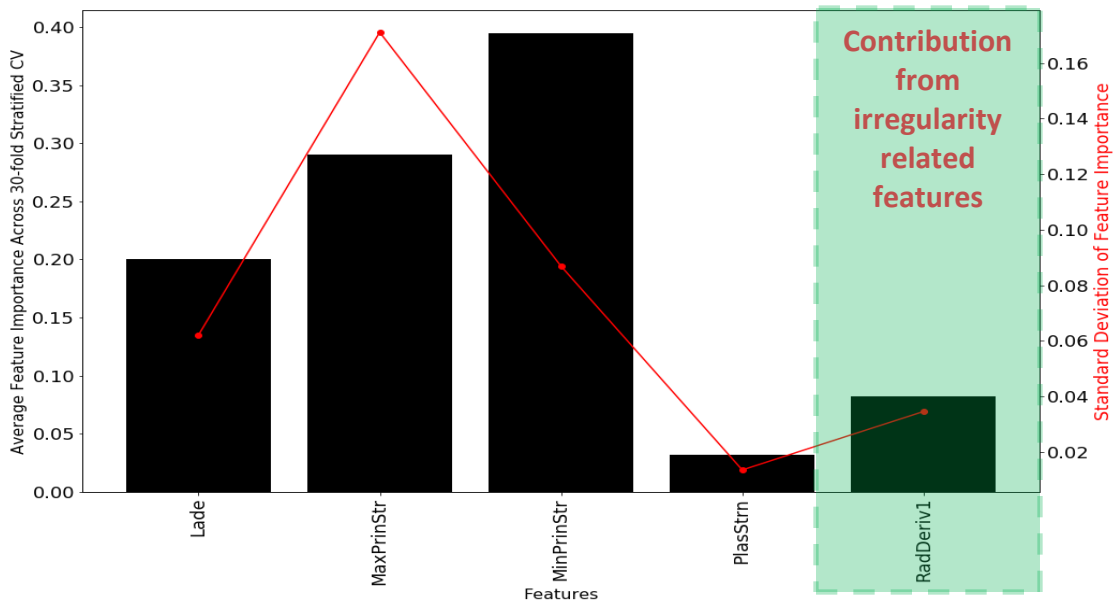**Figure 104. Features importance from the AdaBoost algorithm that was trained based on the combined dataset from all available rock types for the purpose of a generalizable progressive failure prediction training with the irregularity related feature (r_w')**
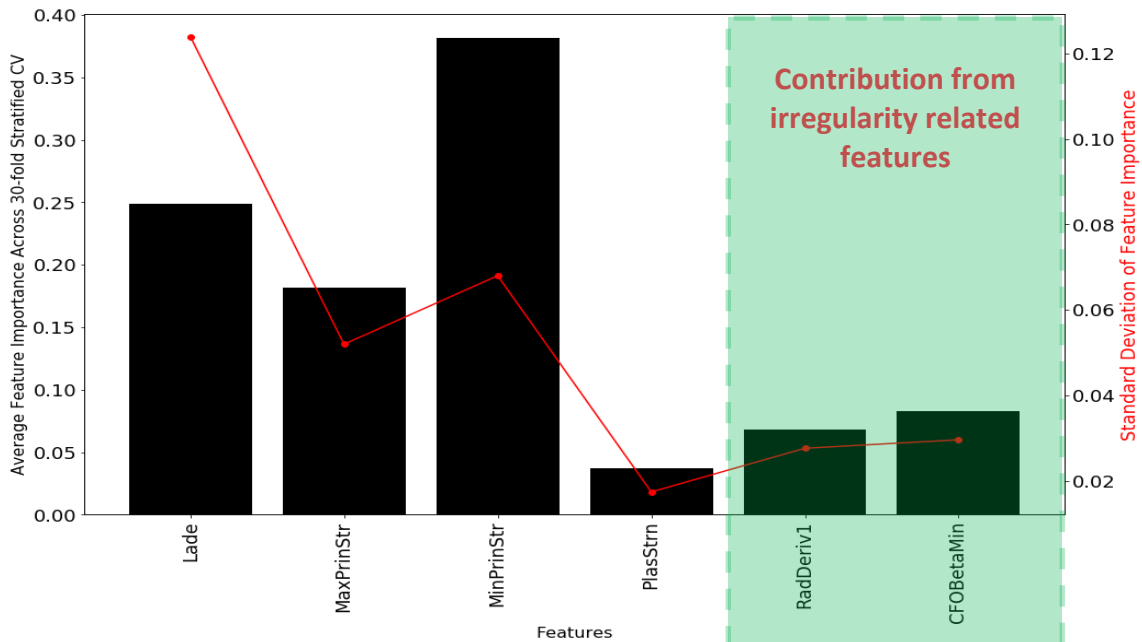
Figure 105. Features importance from the XGBoost algorithm that was trained based on the combined dataset from all available rock types for the purpose of a generalizable progressive failure prediction training with the following irregularity related features: (r_w' and CFO/ β_{Node}).
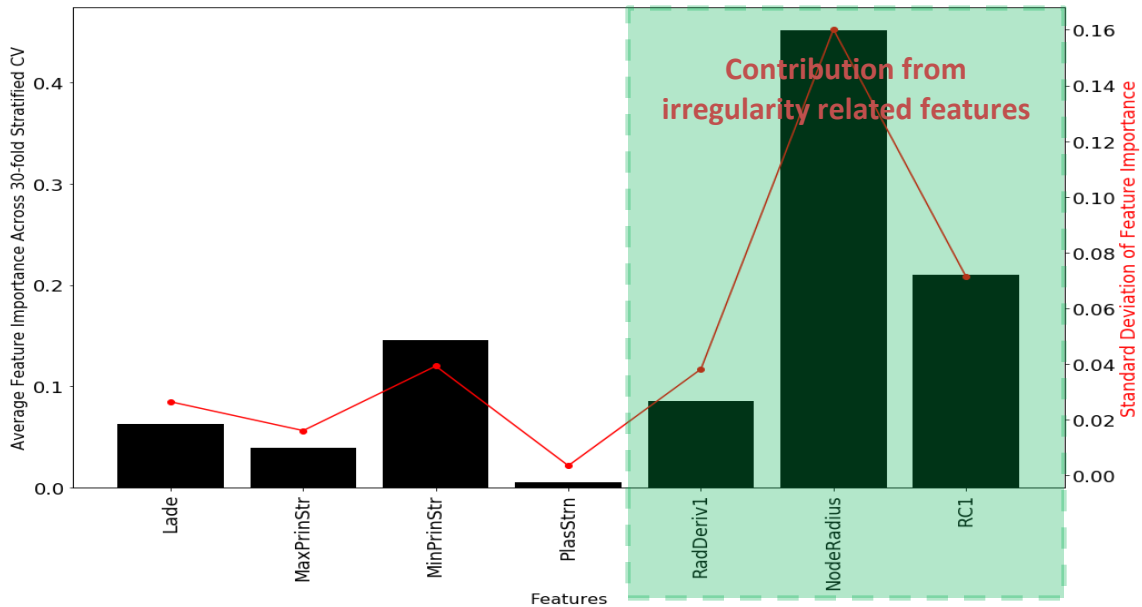
Figure 106. Features importance from the XGBoost algorithm that was trained based on the combined dataset from all available rock types for the purpose of a generalizable progressive failure prediction training with the following irregularity related features: (r_w', r_{Node}, RC).
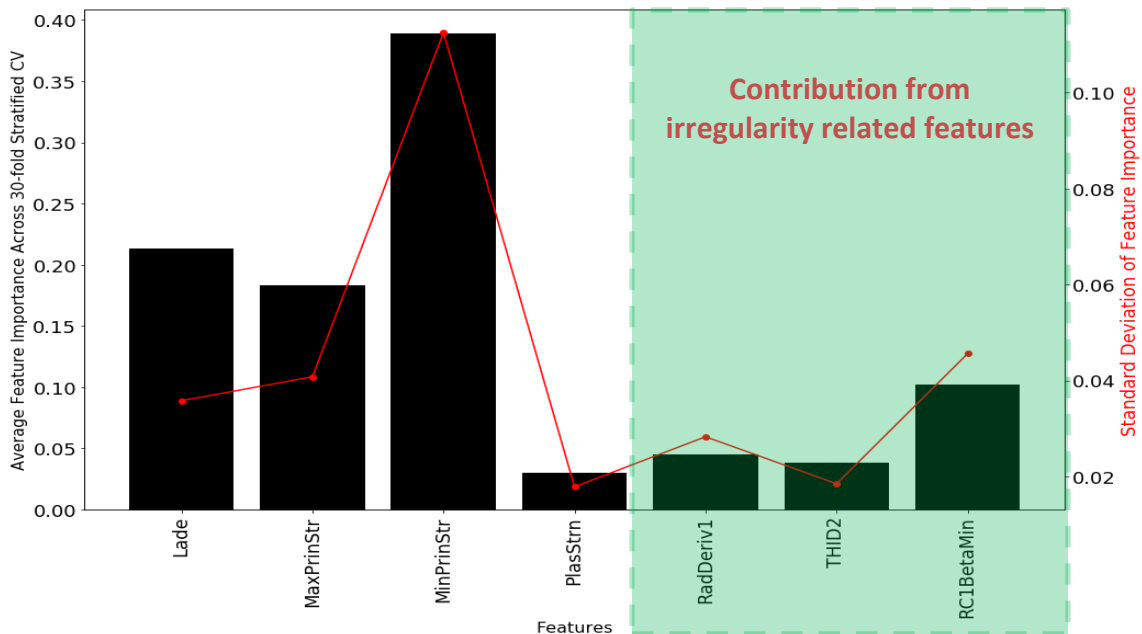
Figure 107. Features importance from the XGBoost algorithm that was trained based on the combined dataset from all available rock types for the purpose of a generalizable progressive failure prediction training with the following irregularity related features: (r_w', θ_{Node}, RC/ $\boldsymbol{\beta}_{Node}$).

Examining the contributions of each of the trialed irregularity related training features combinations as shown in **Figure 104**, **Figure 105**, **Figure 106**, and **Figure 107** helps to understand each model load responsiveness. For example, when training the model with a combination of features that include (r'_w, RC, r_{Node}) as shown in **Figure 106**, it can be seen that the failure prediction is being overwhelmed by the r_{Node}. Overlying on the r_{Node} feature reduces the reliance on the load-dependent features (Lade criterion, principle stresses, and plastic strain) quite severely to the point where failure prediction as a whole becomes load-independent, which leads to a load-nonresponsive model. Now, while a model that was trained with this combination of features yielded an AUROC of more than 0.99 as shown in **Table 12**, it is not useful for the purpose of enhancing the estimation of mud weights required to prevent failure. On the other hand, when training the model with a combination of features that include $(r'_w, RC/\beta_{Node}, \theta_{Node})$ as shown in **Figure 107**, it can be seen that the failure prediction is still relying mostly on the load-dependent features while the irregularity related features are only providing an adjustment. A model that was trained with this combination of features yielded an AUROC of more than 0. 92 using the Bagging algorithm as shown in **Table 12**. Therefore, this algorithm is said to provide the best compromise between accuracy and loading responsiveness. The accuracy measure results from the cross-validation training of this model (bagging decision trees with $r'_w, RC/\beta_{Node}, \theta_{Node}$) are shown in **Figure 108**.
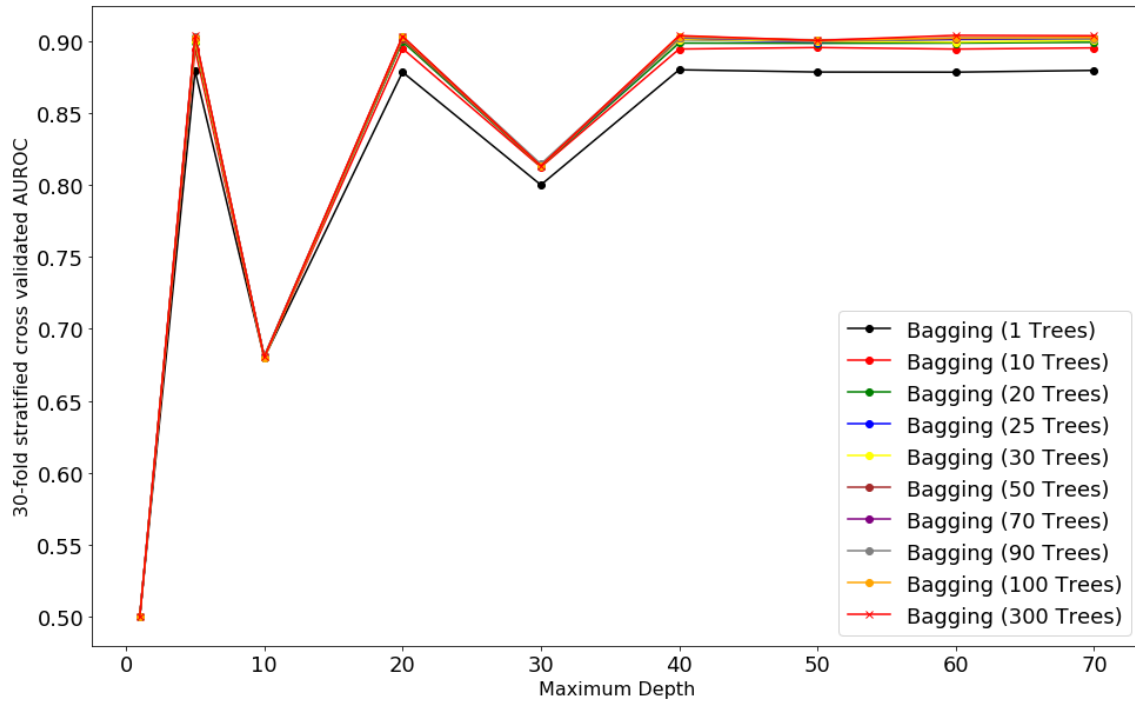
Figure 108. The results of the cross-validation process of the bagging decision trees ML algorithm based on the combined dataset from all available rock types for the purpose of a progressive failure prediction training. The features used for training are (Lade criterion, principle stresses, plastic strain, r_w', θ_{Node}, and RC/β_{Node}).

Figure 109 provides 3D visualizations of progressive failure prediction using both a standalone failure criterion and the selected trained and integrated model. As shown at the top of **Figure 109**, the goal is to predict the progressive failure step with the manufactured LWD wellbore shape as the starting wellbore and the WL FMI-based image as the final wellbore shape. The failure prediction using the standalone conventional failure criterion (Lade) is shown at the bottom left side in **Figure 109** as a top 3D view of the wellbore. This shows that the Lade criterion predicts no further failure to take place as signified by the absence of red zones around the wellbore. The failure prediction using the

174

integrated model is shown at the bottom right side in **Figure 109**, where it corresponds accurately with the actual failure that is observed in the illustration at the top of the figure.
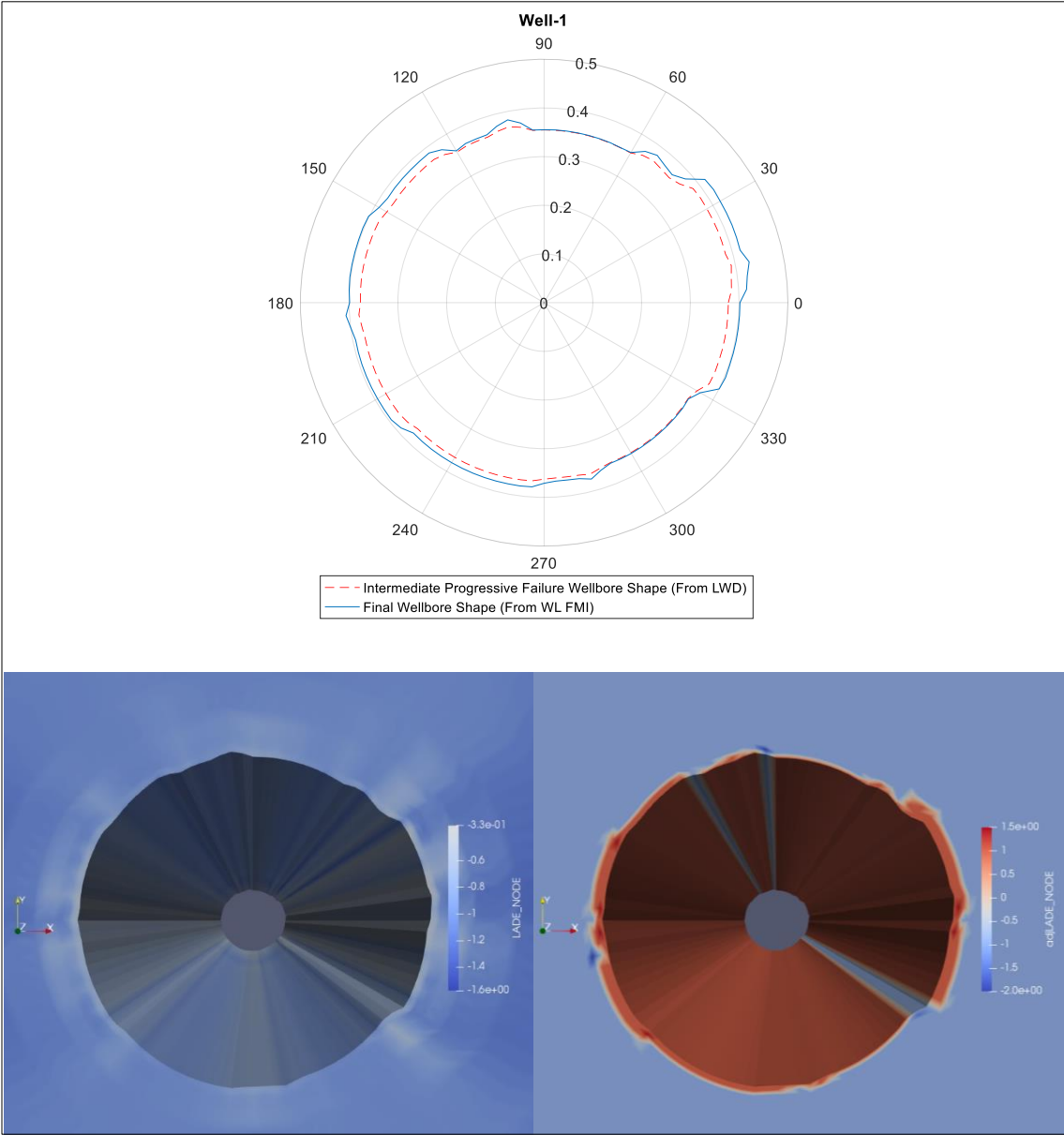
Figure 109. Top: progressive failure prediction desired (from LWD to WL FMI). Bottom left: Well-1 progressive failure prediction using the conventional Lade failure criterion. Bottom right: Well-1 progressive failure prediction using the integrated FEM-ML.

Following the same procedure used for analyzing the performance of the single step failure prediction model, the results of the standalone FEM with a conventional failure criterion is compared against the results of the integrated FEM-ML for progressive failure. Starting with the predictive performance in well-1, which is one of the offset wells that were included in training the machine learning algorithm. The confusion matrix parameters, including accuracy, for well-1 are shown in **Figure 110**.
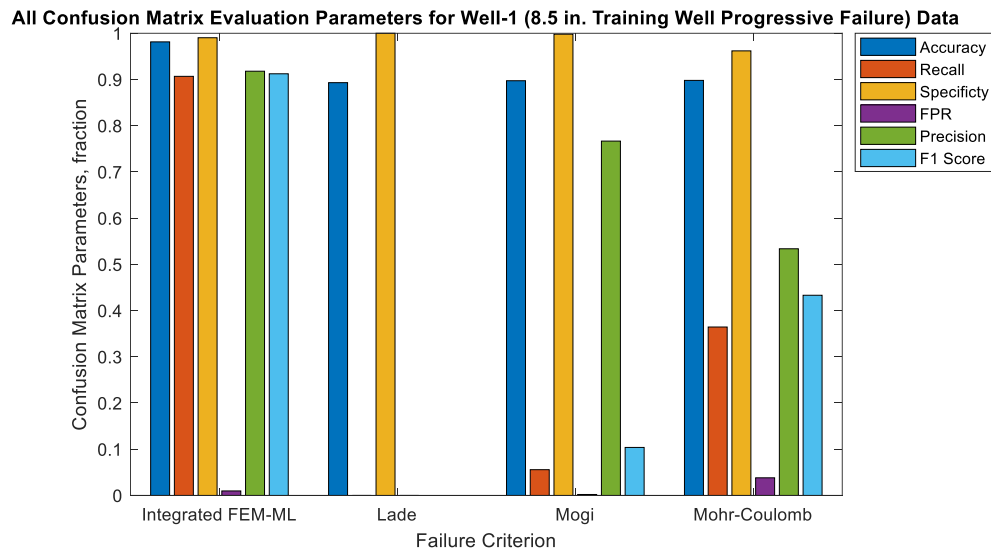
Figure 110. The confusion matrix parameters for well-1 (training well-progressive failure).

Based on all the different confusion matrix parameters, the integrated FEM-ML model for progressive failure performed considerably better than all the other conventional failure criteria. The predictive performance over the full dataset of the training offset wells is shown in **Figure 111**. This evaluates the progressive failure prediction performance over all the training offset wells, which means that well-6 and well-7 are not included here.

176

Looking at this chart, it is clear that the integrated FEM-ML model performed considerably better than all the conventional failure criteria over the training dataset.
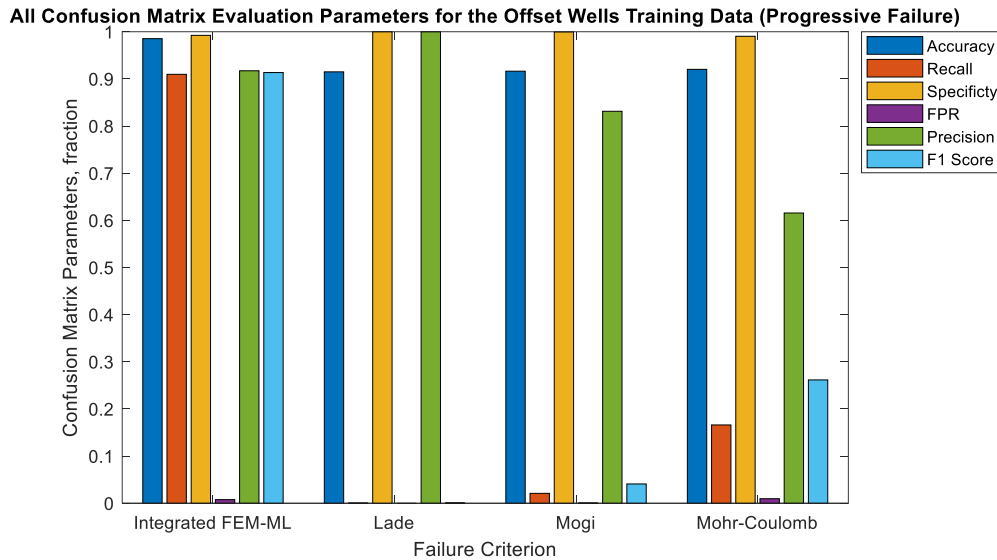
Figure 111. The overall confusion matrix parameters for all the training offset wells for progressive failure (well-6 and well-7 are not included).

Results for wells excluded from training are shown in **Figure 112** for well-6 and in **Figure 113** for well-7. These charts now show that the progressive failure predictive performance of the integrated FEM-ML model has dropped noticeably. Still, even with predictive performance drop in testing phase, the integrated model is still performing significantly better than the conventional failure criteria. Upon examining the results for progressive failure in well-6 and well-7, the discussion of the results from the single step failure model, previously shown in **Figure 101** and **Figure 102**, also applies here.

177
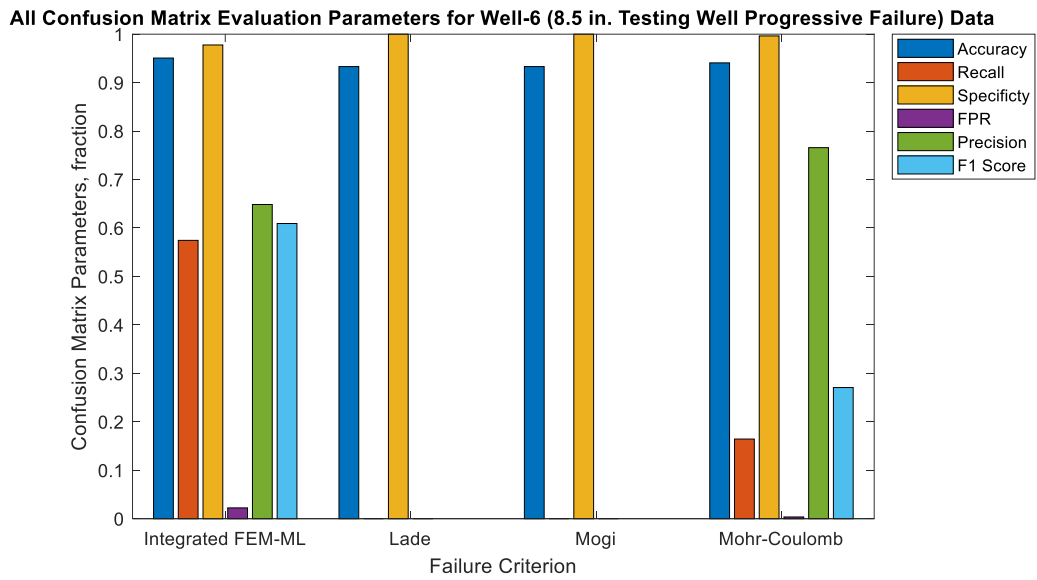
Figure 112. The confusion matrix parameters for well-6 (testing well- progressive failure).
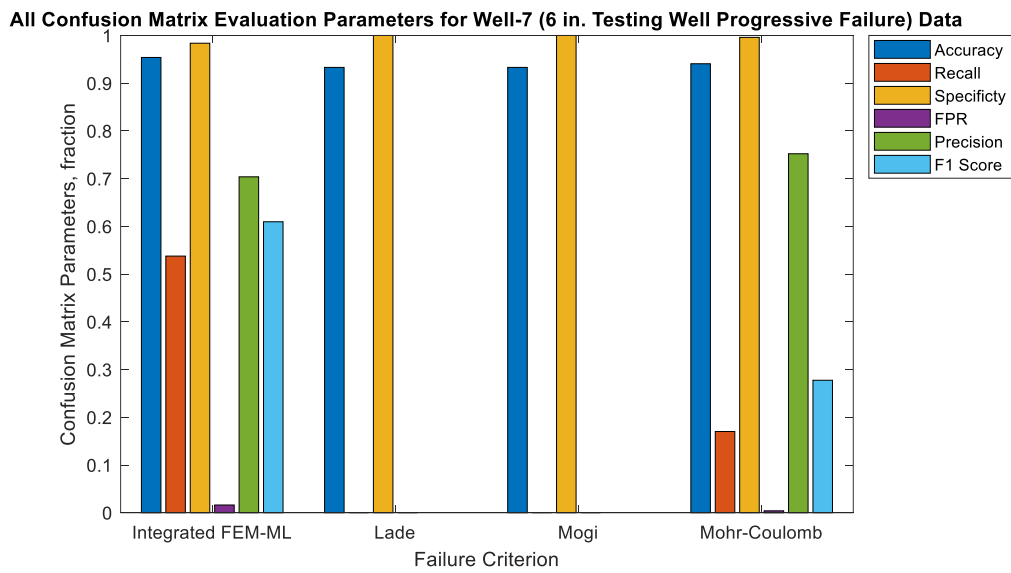
Figure 113. The confusion matrix parameters for well-7 (testing well- progressive failure).

For the overall performance of the progressive failure model, the test wells results show that the integrated model has an F1-score that is 125% higher in well-6 and 120% higher in well-7 than that of the Mohr-Coulomb criterion. Based on the F1-score alone, this interpretation means the integrated FEM-ML has more than doubled the progressive failure prediction accuracy that was provided by the conventional criteria.

4.2 Real-time Implementation of the Integrated Model

This particular field application illustrates the model ability to respond to various drilling events as they occur to redefine the drilling window limits. The model is able to achieve this type of while-drilling or real-time response using the image analyzer and the wellbore re-meshing functions discussed previously in the model development chapter. To provide a display of this field application, different scenarios that represent typical drilling events are introduced into the model. The changes of the drilling window with the introduction of each scenario is then produced and examined. The well data used in this illustration is from offset well-2 from the previous section. The full analysis for this part was previously published by AlBahrani and Morita (2020).

4.2.1 Real-time Application Scenarios

To provide an understanding of how the drilling window of a certain interval changes, four different case scenario that are based on field data are used here. These scenarios are based on a vertical well drilled to a depth of around 13,000 ft targeting a sandstone gas reservoir (well-2). Initial loading input data of the well are shown in **Table 13**.

Table 13. The initial loading parameters for all the case scenarios considered for the real-time application of the model

Loading Parameter	Load Magnitude, ppg
Pore pressure	8.5
Overburden (vertical) stress	18
Maximum horizontal stress	20.8
Minimum horizontal stress	12.1
Mud weight used	9

While drilling the 6" hole section in the sandstone gas reservoir, the well experienced several severe events that suggest wellbore instabilities. These include multiple pack off, tight hole, and fluid losses, which led to parted string and two sidetracks. Excessive cuttings and fill were seen as well. Both scattered wellbore enlargements and drilling induced fractures were observed in the FMI image log of this well as shown in **Figure 114**.
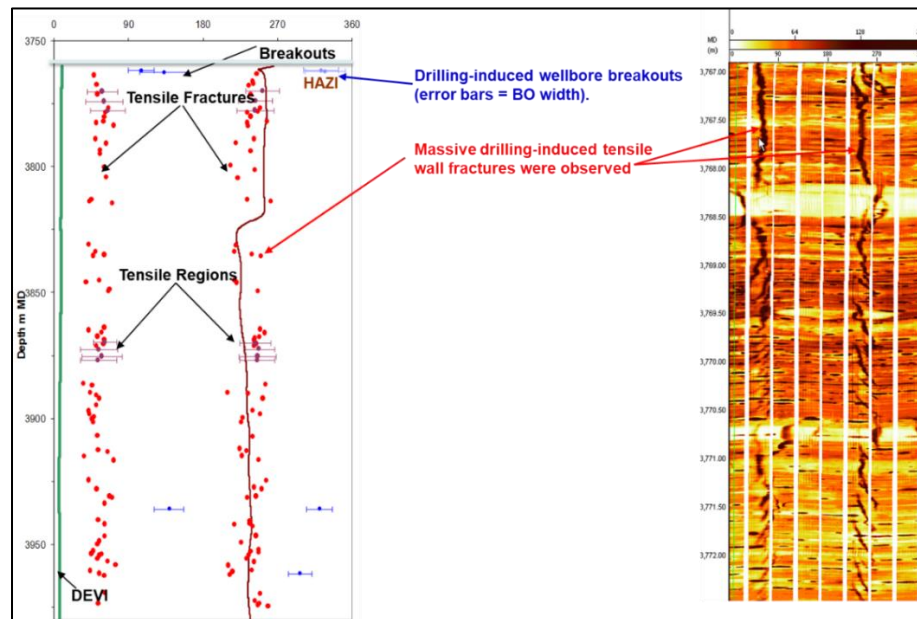
Figure 114. The FMI image log of well-2, which is used for the illustration of the real-time application scenarios.

Since no LWD images are available, there is no knowledge of which mode of failure (enlargement or fractures) had occurred first in this well. Hence, four different case scenarios are assumed with alternating order of occurrence of each mode of wellbore failure to illustrate the real-time application. The case scenarios are described in **Table 14**.

Table 14. Description of the case scenarios used to illustrate the real-time application of the model

Case Scenario Number	Wellbore Description
1	Perfectly Circular Wellbore with No Fractures (Base Case)
2	FMI Image Log Based Wellbore with No Fractures
3	Perfectly Circular Wellbore with Fractures
4	FMI Image Log Based Wellbore with Fractures

Each one of these case scenarios is run into the FEM using the image analyzer and re-meshing functions to produce representations of the changes to the drilling window of the interval of interest. For failure criteria, only results from conventional failure criterion are displayed in the illustration of the real-time applications of the model to keep a focus on analyzing the standalone effect re-meshing while drilling. For the specific conventional failure criterion, it was determined that the Lade failure criterion is the one that most closely agrees with the reservoir rock failure data such as leakoff test (LOT) and triaxial core experiments. In this section, the results for both the Lade failure criterion with K=0.9 and the Mogi failure criterion will be presented to provide a perspective for comparison considering that Mogi is more commonly used in the industry.

The mesh and re-meshes of all the case scenarios are shown in **Figure 115**. For cases 1 and 2, a uniform mesh is used to capture failure evenly around the wellbore. As

for the fractured cases of 3 and 4, the mesh is finer in the maximum horizontal stress

direction (y-axis direction) to enable capturing the geometry of drilling induced fractures.

Also, in cases 3 and 4, the mesh is finer in the minimum horizontal stress direction (x-axis

direction) to allow for monitoring instability shear failure while still reducing computation

cost by using a coarser mesh elsewhere. It should also be noted that enlargements, as per

the observations made from the image log in **Figure 114**, can occur anywhere on the

wellbore and not only at 90° from induced fractures.
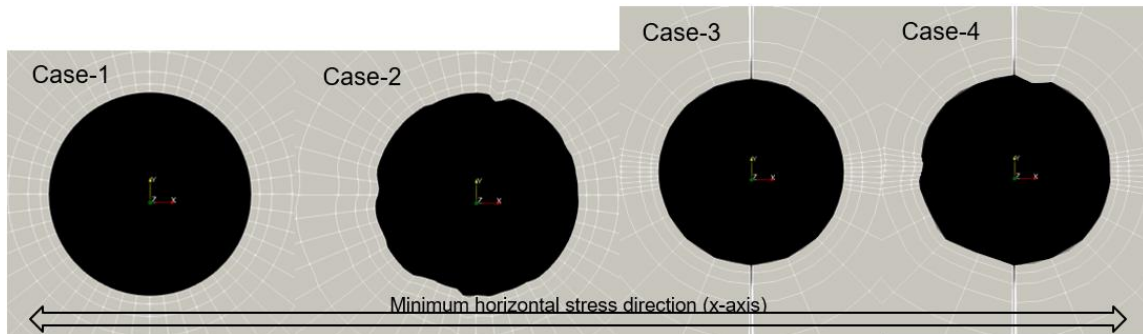
Figure 115. A 2D top view of the meshes of case scenarios 1 to 4 from the left to the right respectively, which are used to illustrate the real-time application of the model.

There are two points of consideration that need to be addressed before analyzing

the results of the drilling window variations between the case scenarios. First, the way

fractures are modelled into the wellbore here can consider the presence of drilling induced

fractures at the maximum horizontal stress direction, or natural fractures at random

directions. The nodal coordinates are simply re-configured to reflect the presence of the

fracture while still honoring the discontinuity of the fracture. This means that fractures are

modelled as a completely static discontinuity, meaning, the model assumes that the

fracture in these particular scenarios have an infinite fracture toughness. This is similar, but not identical, to what was done in some published works where a static fracture was modelled to assess its effect on wellbore strengthening (Y. Feng et al. 2015, Salehi and Nygaard 2010). Therefore, the model here allows for no further fracture growth or evolution as the dimensions of the fracture will be measured or assumed based on image logs, field experience, and published data. This re-meshing for the fractures, as shown in **Figure 116**, exposes the faces of the fracture to the same distributed face load acting on the wellbore wall as a result of hydrostatic pressure of the mud.
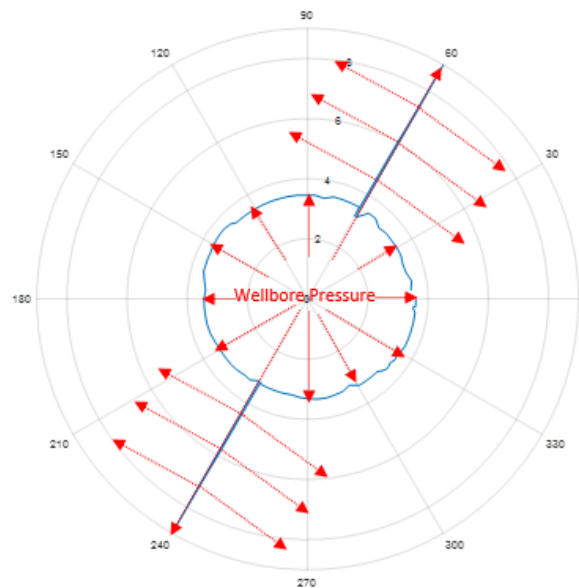
Figure 116. An illustration of how the wellbore pressure (mud weight) is acting on both the wellbore walls and on the faces of the fractures in the form of distributed face load in the FEM formulation.

The second point of consideration is the determination of the drilling window limits in the fractured cases. For the maximum mud weight, no new value is estimated for

the fractured cases as this limit is assumed to have already been violated and the only maximum limit that remains is the fracture extension or propagation pressure. But since the fractures modelled here are static, this is beyond the scope of the model. Also, in case of encountering or causing a fracture while drilling, the common response is to lower the mud weight to avoid exacerbating the loss of circulation. This means that the new allowable maximum mud weight is no longer of interest from a drilling perspective. As for the minimum mud weight, the determination of this limit ignores the failure that occurs on the inside of the fracture walls and near the fracture mouth. The minimum mud weight limit in these cases is still determined at the onset of wellbore instability failure, which in these cases is localized in the direction of the minimum horizontal stress (x-axis) because the well is placed vertically in a strike-slip regime. If failures on the inside of the fracture walls and near the fracture mouth were to be considered as the onset of failure that determines the mud weight, the resulting drilling window will become extremely narrow to nonexistent. It is also worth noting that for all case scenarios, the determination of the maximum and minimum mud weights is performed conservatively at the onset of failure (no- risk control). This means that the mud weight is set at the first step that shows failure criterion value above zero.

4.2.2 The Model Results and Their Influence on Field Response

The following are the results of the four case scenarios. The 3D images show the values of the failure criterion calculated at the nodal points of the 3D block of the wellbore at the onset of wellbore instability. Positive values indicate failed location, which are observed in the minimum horizontal stress direction, which is the x-axis in this model.
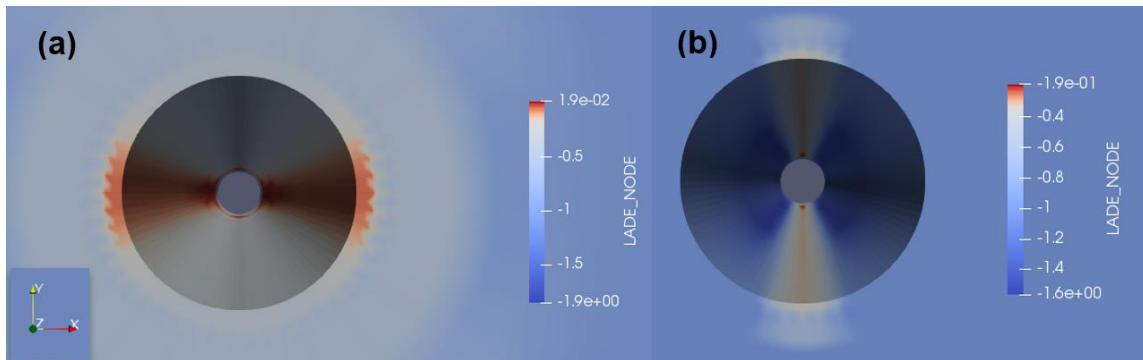
Figure 117. 3D top view of case-1 (perfectly circular with no fractures) wellbore at (a) Failure criteria at the onset of wellbore instability. (b) Failure criteria at max. mud weight.
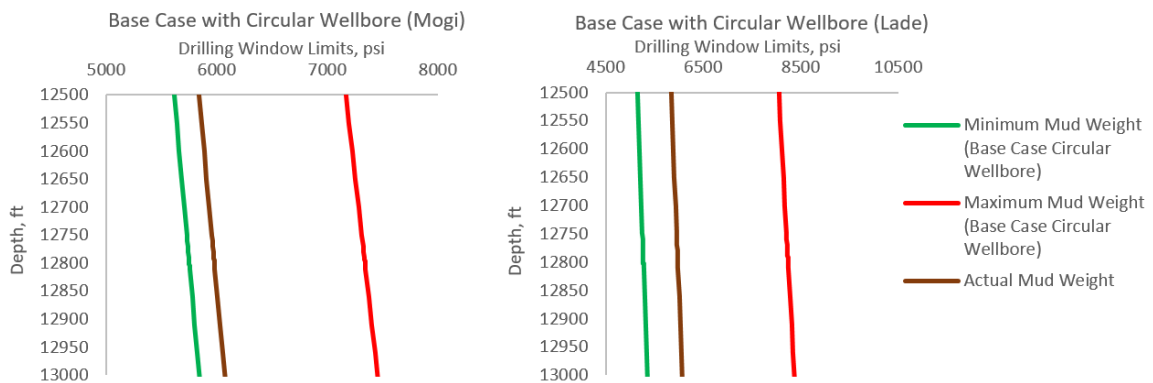
Figure 118. The drilling window of case-1 (perfectly circular with no fractures) based on both the Mogi (left) and the Lade (right) criteria with actual mud weight used.

When discussing the results, only the Lade failure criterion-based drilling window will be referred to here. As mentioned previously, this is because it was determined that the Lade failure criteria is the one that most closely agrees with the reservoir rock failure data such as leakoff test and triaxial core experiments. The Mogi failure criterion is presented to provide a perspective for comparison considering that Mogi is more commonly used in the industry than Lade.

When comparing the drilling window of Case-1 (perfectly circular wellbore) in **Figure 118** to Case-2 (FMI based wellbore) in **Figure 120**, it can be seen that modelling

the irregular shape of the wellbore led to a significantly narrower window. In Case-2, higher mud weight is now required to prevent wellbore instability, while fracturing is expected to occur at a lower value than determined in Case-1. This result is logical since a departure from a circular shape, an inherently stable shape, to a non-uniform and irregular shape will make the structure weaker. The irregularities around the wellbore introduce weak zones that allow for failure to occur at a lower threshold. The effect of these irregularities can be seen when comparing **Figure 117** of Case-1 to **Figure 119** of Case-2 at both the minimum mud weight and at the maximum mud weight. This shrinkage of the drilling window provides a path to explain the mechanism leading to the concurrence of both instabilities and drilling induced fractures within the same interval. With this narrow drilling window, even with a constant, and presumably safe, mud weight, minor fluctuations in the bottomhole pressure can lead to violating both limits and inducing enlargements and fractures at the same time.
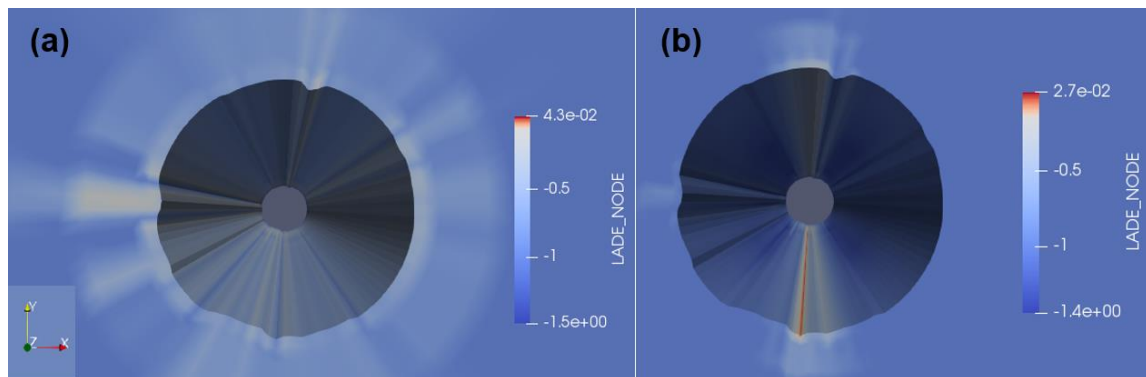
Figure 119. 3D top view of case-2 (FMI image-based wellbore with no fractures) wellbore at (a) Failure criteria at the onset of wellbore instability. (b) Failure criteria at max. mud weight.
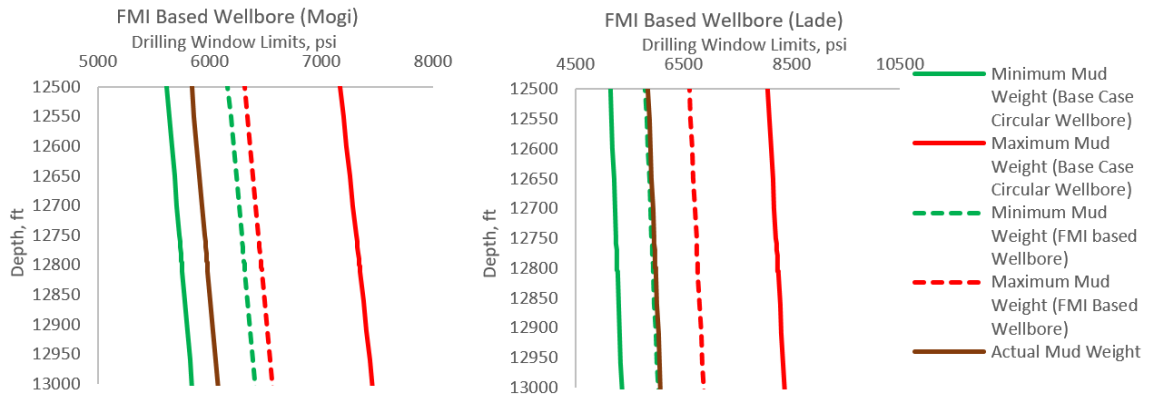
Figure 120. The drilling window of case-2 (FMI image-based wellbore with no fractures). Solid lines correspond to case-1 (perfectly circular wellbore).
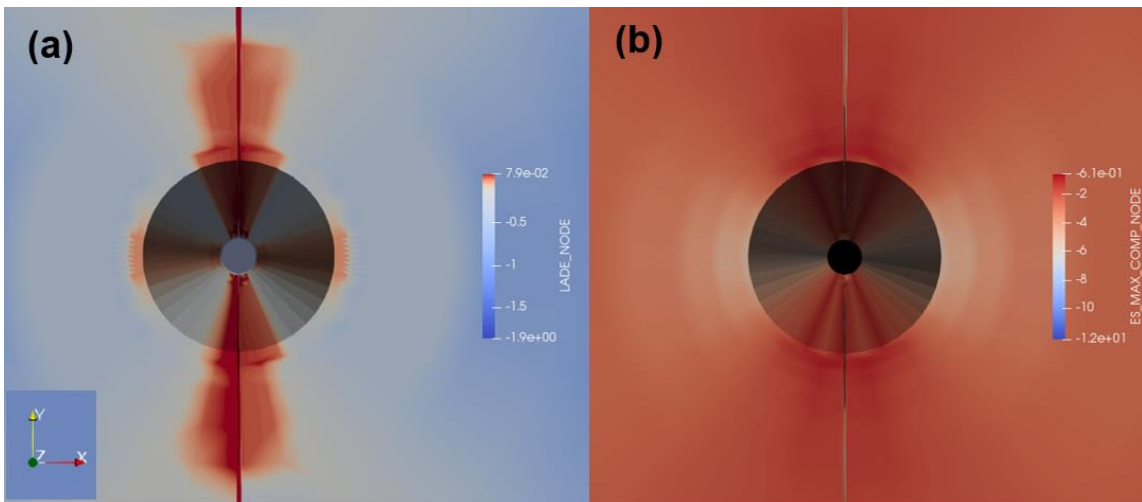
Figure 121. 3D top view of case-3 (perfectly circular with fractures) wellbore. (a) Failure criteria at the onset of wellbore instability. (b) Tangential stress at actual mud weight.
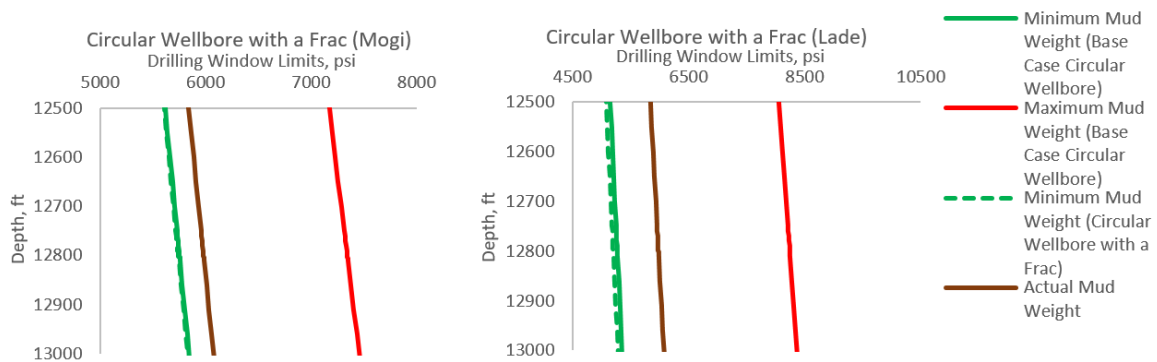
Figure 122. The drilling window of case-3 (perfectly circular with fractures). Solid lines correspond to case-1 (perfectly circular wellbore).

As for Case-3 (perfectly circular wellbore with drilling induced fractures) in **Figure 122**, the introduction of a fracture into the model led to a slightly lower mud weight for instability. The introduction of the fracture into the structure works to introduce extra compression (signified by the negative stress values or blue zones as shown in **Figure 121 (b)**) in the tangential stress around the wellbore. This increase in stress concentration means that the zone in the minimum stress direction is more compressed than before, which means that it is now more likely to experience shear failure than before. To analyze the effect of the introduction of a fracture, **Figure 125** compares the tangential stress of a circular wellbore with and without a fracture. As shown here and as mentioned before, the introduction of the fracture does increase the state of compression around the wellbore significantly. This increased compression and likelihood of shear failure seems to be contradictory when examined against the slightly lower stability mud weight determined in **Figure 122**, however, this contradiction can be cleared when considering the double action of the wellbore pressure in this scenario. As depicted in **Figure 116**, in a fractured wellbore, the wellbore pressure acts both on the face of the wellbore walls and on the face

188

of the fractures. While the action of the wellbore pressure on the wellbore wall works to provide a stabilizing effect to counter shear failure, on the other side, it also works to elevate compression around the wellbore by loading the fracture face. This means that in a fractured wellbore scenario, there is a sweet spot of wellbore pressure that provides a balance between the support provided to the wellbore wall and the compressional load applied to the fracture face. Such balance between these loads can only be evaluated with a numerical method such as the FEM. Operationally, this result means that when experiencing loss of circulation, or any other signs of a fractured wellbore, lowering the mud weight to avoid exacerbating the loss of circulation should be faced with caution of violating instability minimum mud weight limit. Without a case specific model for a fractured wellbore, there is no reliable way to know if the minimum mud weight for stability has remained the same or shifted up or down.
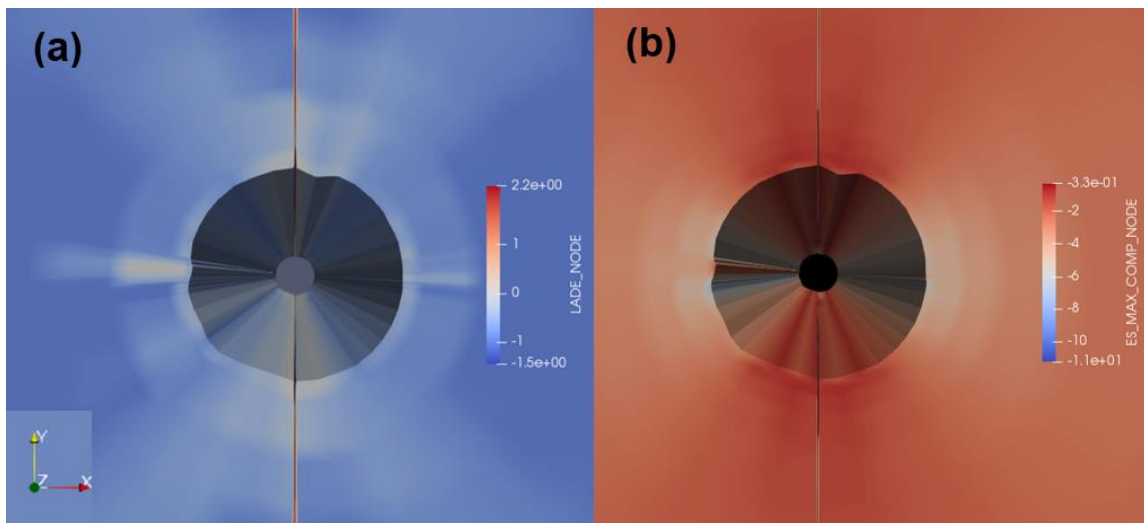
Figure 123. 3D top view of case-4 (FMI image-based wellbore with fractures) wellbore. (a) Failure criteria at the onset of wellbore instability. (b) Tangential stress at actual mud weight.
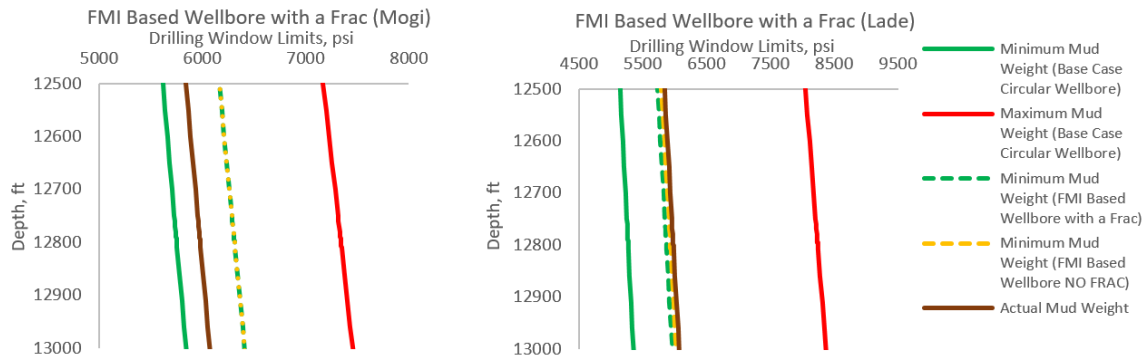
Figure 124. The drilling window of case-4 (FMI image-based wellbore with fractures). Solid lines correspond to case-1 (perfectly circular wellbore).
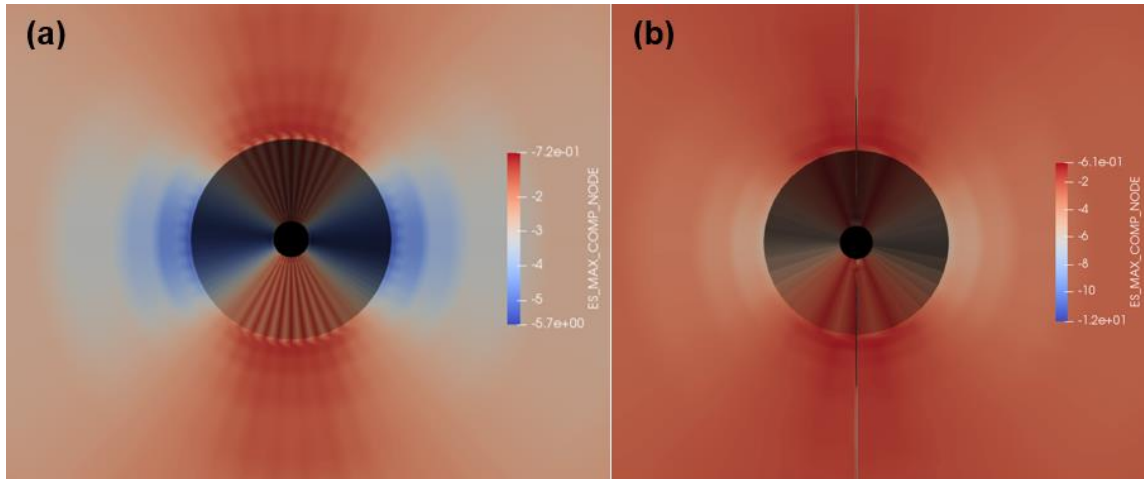
Figure 125. A 3D top view comparison of (a) the tangential stress of a perfectly circular wellbore with no fractures (case-1) evaluated at the actual mud weight value of 9 ppg and (b) the tangential stress of a perfectly circular wellbore with fractures (case-3).

Comparing Case-4 (FMI based wellbore with drilling induced fractures) in **Figure 124** to Case-1, the plot shows that the minimum mud weight in Case-4 is now at a higher value, which again creates a narrower window. Just as in the comparison between Case-1 and Case-3, this shift in the minimum mud weight can be attributed to weaker irregular structure of the wellbore. When comparing the results of Case-4 to Case-2 to examine the

standalone effect of introducing a fracture into an FMI based wellbore, it can be seen that the minimum mud weight in Case-4 is at a slightly lower value than in Case-2, which creates a slightly larger drilling window. Again, just as in the comparison between Case-1 and Case-3, this shift in the minimum mud weight from Case-4 to Case-2 can be attributed to the double action of the wellbore pressure despite the increase in compression around the wellbore as shown in **Figure 123 (b)**.

4.2.3 Summary of Advantages of the Real-time Application of the Model

Overall, these results show just how dynamic the size of the drilling window is. There are various events, beyond the selection of the mud weight, which can change the structural integrity of the wellbore. These events can include, but definitely is not limited to, surge and swab pressures, drilling through rubblized zones, wellbore failure at bedding planes, key-seating caused by the drill string resting on one side of the wellbore, and drillstring vibrations. Should any signs corresponding to these events be detected in real-time while drilling, drilling engineers can use whatever available data, such as LWD wellbore images and shale shakers surveillance, to estimate the new geometry of the wellbore and determine the new drilling window limits by which they can continue to operate with more confidence. Undetected changes to the drilling window can lead to major non-productive time (NPT) incidents such as the case in the offset well used for illustration here. As mentioned previously, this offset well (well-2) experienced several severe events that suggest wellbore instabilities. It experienced multiple pack off, tight hole and fluid losses, which led to parted string and two sidetracks. Excessive cuttings and fill were seen as well. These drilling difficulties were experienced all the while operating

with a mud weight that falls within the pre-drilling determined drilling window. Based on the modelling results shown here, considering and monitoring the dynamic nature of the drilling window has the potential to be a major step towards eliminating these NPT incidents.

The results of the FEM with the proposed re-meshing scheme and image log analysis show that the drilling window is a dynamic parameter that can change with the progression of the drilling process. While the pre-drilling determined mud weights limits are a good starting point, it should always be considered that these can change with time. Therefore, signs indicating changes to wellbore geometry should be used to estimate the changes to the mud weights limits. Also, the modelling results offer an explanation to the phenomenon of the concurrence of drilling induced fractures and wellbore enlargements within the same interval while drilling with a constant mud weight. Beyond the operational practices-caused fluctuations of bottomhole pressure, the model results show how considering the changing wellbore geometry can serve to narrow the drilling window, which in turn will effectively increase the chances of these two opposite modes for failure to take place at the same time.

There are several field operational implications to the changes in the drilling window observed here. In cases where excessive cavings and abnormal drill-bit-produced rock cuttings are observed in the shale shakers, there should be more concern of inducing further instability and loss of circulation or drilling induced fractures. This is because cavings are an indication of a new wellbore geometry, and as seen previously, this can

mean that drilling window is now narrower. Therefore, adjustment to mud weight should be performed accordingly.

4.3 Examining the Effect of Wellbore Enlargement Intensity

In this field application of the integrated model, it is assumed that no LWD based image logs are available while drilling the interval of interest. However, shale shakers surveillance is proposed here as the only alternative to monitoring the changes in the wellbore geometry. Such scenarios are applicable when drilling overburden formations (non-reservoir sections), where no logging tools are run. It should be noted that the FEM-ML model for this application is trained as before using the following features:

1. Lade failure criterion
2. Principal stresses
3. Plastic strain

4.3.1 Scenarios for Varying Intensity of Wellbore Enlargement Based on Shale Shaker Observations

In a similar fashion to the previous filed application, the integrated model application is illustrated through modelling different case scenarios of varying intensity of wellbore enlargements or breakouts. The description of the wellbore shape in these case scenarios is show in **Figure 126**. The depicted case scenarios assume that for each case, the shape of the enlarged wellbore is inferred based on observations of cavings on the shale shakers. The area of enlargement is assumed to always be in the direction of the minimum horizontal stress, which is the x-axis in these cases. The initial input data for all

the cases are based on the same offset well that was used to illustrate the real-time application of the model (well-2). The wellbore uniform diameter is 6 inches in the base case. Starting with enlargement intensity case-1, both the radial and circumferential extension of the enlargement zones are increasing from one case to the next one
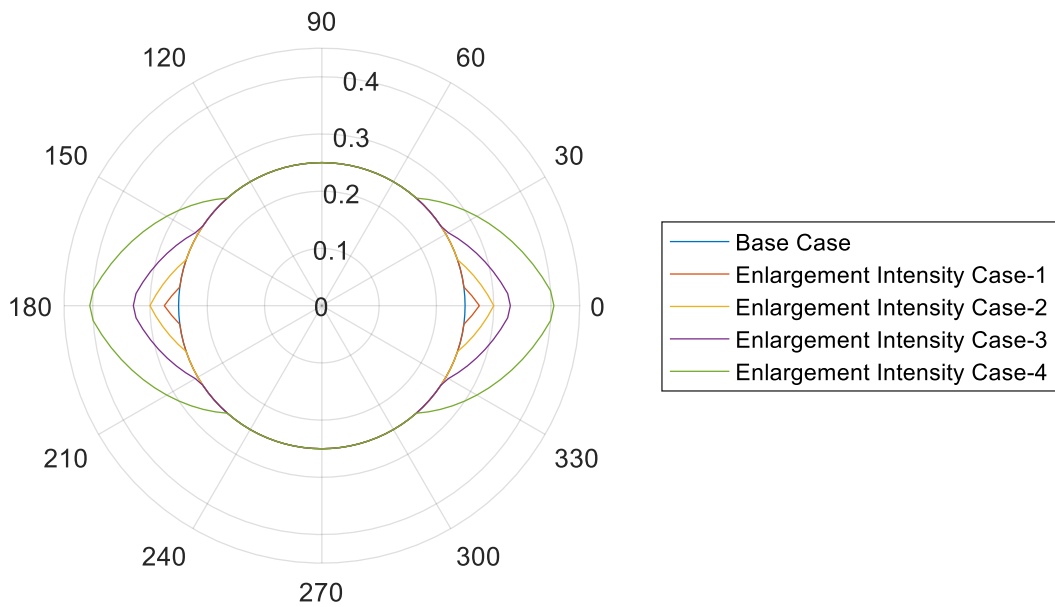
Figure 126. Wellbore case scenarios with varying intensity of enlargements. Wellbore radial measurements are in ft.

To illustrate the application of the integrated model, risk-controlled stability mud weight polar charts are produced. The risk control means the stability mud weights are calculated at a limit that allows a maximum value of cavings volume. A sensitivity analysis is performed by producing the polar charts with different limits of cavings volume risk control. The main objective of the sensitivity analysis is enabling the use of a suitable

hydraulics model that can help decide the safe maximum value of cavings, which will not cause stuck-pipe incidents, tight spots, or hole cleaning issues.

The risk-controlled stability mud weights are calculated using the integrated model that was trained in a similar fashion to the training process in the first section of this chapter. The only difference is that the machine learning algorithm in this application is trained using features that are load-dependent, but not direction dependent. This is essential because the training process was carried out using all vertical wells, while the polar charts require the determination of mud weights at all possible directions and inclinations. Training the algorithm using only scalar and load-dependent features ensures that having only vertical offset wells is not a limiting factor as discussed in the machine learning features selection and engineering section of the model development chapter.

4.3.2 The Integrated Model Results Through Dynamic and Risk-controlled Polar Charts

For the sake of consistency and to ensure the results re-producibility, it should be noted that the stability mud weights are calculated by considering the effect of stress hysteresis. This is done by first subjecting the wellbore to an initial overbalance mud weight load that will ensure shear failure is minimized. Then, the mud weight load is dropped in steps and the cavings volume is calculated at each step. Employing this method of calculating the stability mud weight means that the initial overbalance mud weight will affect the caving volumes produced at the subsequent steps of mud weight reduction. Therefore, the initial overbalance mud weight used to minimize shear failure is stated for each case scenario.

The polar charts for each case scenario are produced at different maximum allowed cavings volumes. When reading the charts, it should be noted that the color gradient legend is different from one chart to the others and a specific color can mean different mud weights in different charts. Also, the horizontal or x-axis direction of all the polar charts represents the direction of the maximum horizontal stress. This places the maximum horizontal stress in the 90º to 270º direction and the minimum horizontal stress in the 0º to 180º direction. An example mud weight polar chart with general reading instructions that are applicable to all the other polar charts in this section is shown in **Figure 127**.
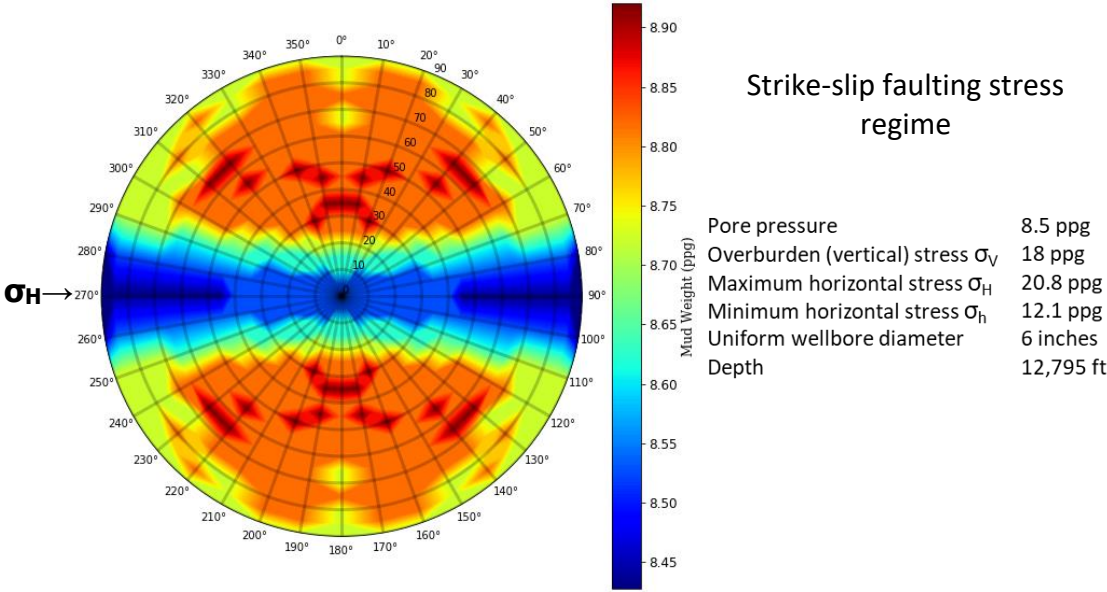
Figure 127. An example of a stability mud weight polar chart with general reading instructions that are applicable to all the polar charts displayed in the 'Examining the Effect of Wellbore Enlargement Intensity' section.

4.3.2.1 Base Case

There are two goals for analyzing the results of the base case, which is the case with a uniform wellbore with no enlargements. The first goal is to provide a reference point or a benchmark for the cases with varying enlargement intensity. The second goal is to highlight the differences in the stability mud weight determined by the conventional failure criterion (Lade) and the one determined by the trained and integrated FEM-ML. The Lade failure criterion was used by the offset wells engineers to determine the stability mud weight for drilling all the offset wells used for training and testing the model. As mentioned in the previous discussion of the offset wells, the pre-drilling-Lade determined stability mud weights were over optimistic as they were not enough to arrest the wellbore rock failure as evidenced by the drilling difficulties encountered. This over-optimism is fixed by the integrated FEM-ML as shown in **Figure 128**, **Figure 129**, and **Figure 130**.
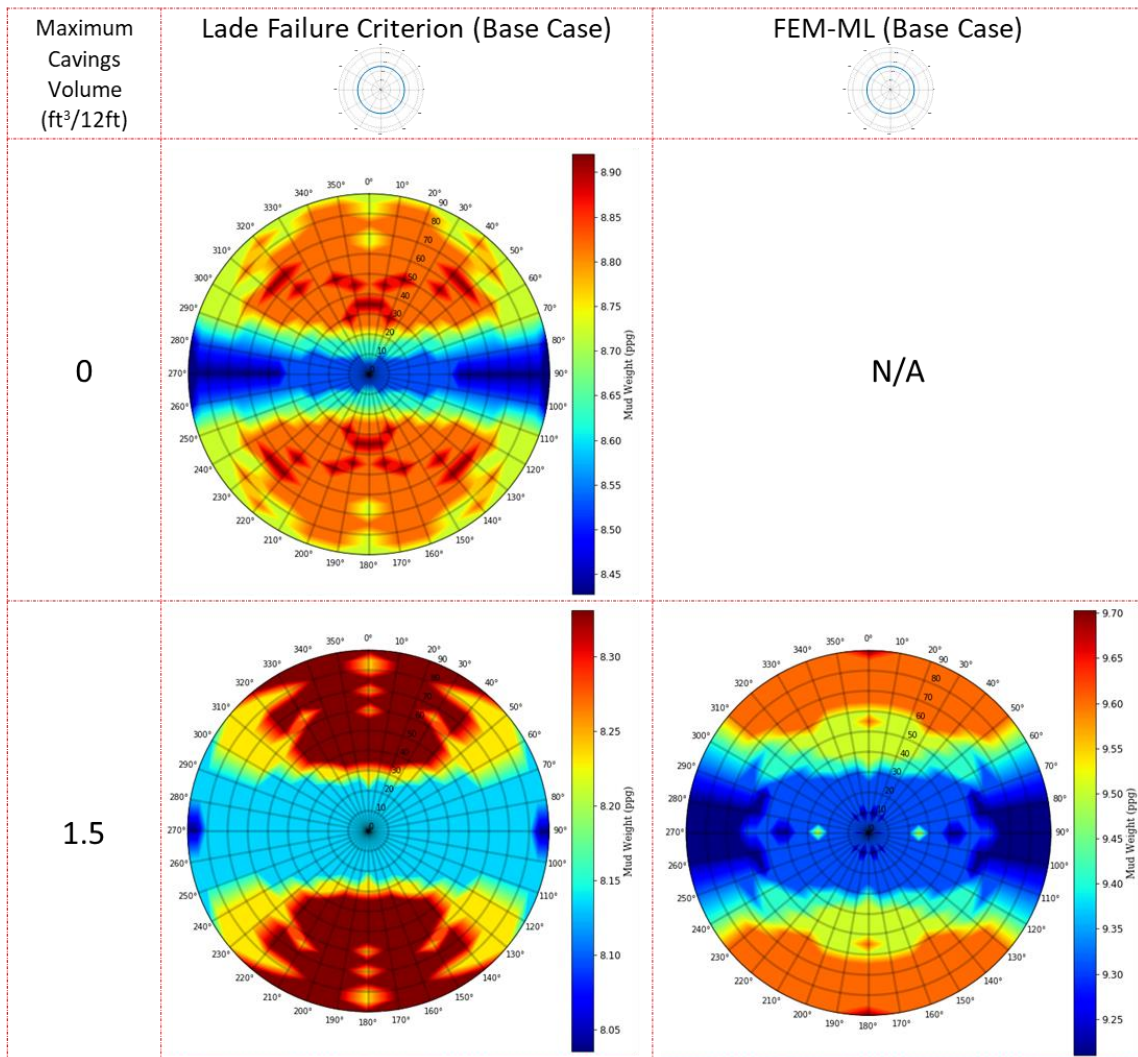
Figure 128. Stability mud weights polar charts for the base case scenario as determined by both the Lade failure criterion and the integrated and trained FEM-ML. The charts are risk-controlled at under two maximum allowed caving volumes, which are 0 ft³/12ft and 1.5 ft³/12ft.

When examining the stability mud weight polar charts in **Figure 128**, it can be seen that, in general, the integrated and trained FEM-ML model determines the stability mud weight to be around 1 ppg higher than the one determined by the Lade failure criterion. The overbalance initial mud weight that minimizes shear failure all the way

down to no cavings in the Lade criterion is 9.5 ppg. On the other hand, the FEM-ML results show that it is not possible to completely prevent cavings. The FEM-ML shows that the caving volume can be minimized to under 1.5 ft^3/12ft with an initial overbalance mud weight of 9.8 ppg. The cavings volumes are at determined per 12 ft intervals as this is the length of interval that was modelled in the FEM initially. This can be changed to fit the requirements of the hydraulics model that will be used to set the critical maximum cavings model.

The polar charts are produced at different allowed maximum caving volumes. It can be seen in **Figure 128**, **Figure 129**, and **Figure 130** that as the allowed maximum caving volumes is increased, the stability mud weight requirements drop for both the Lade criterion and the FEM-ML. For a closer examination of these results, a horizontal well (90° inclination on the radial axis of the polar chart) in the direction of the minimum horizontal stress (0° azimuth on the circumferential axis of the polar chart) is considered. According to the Lade criterion, this well requires mud weights of 8.46, 8.08, 7.89, 7.89, 7.89, 7.79, and 7.79 ppg for arresting the cavings volume to under 0, 1.5, 2, 3, 4, 5, 6 ft^3/12ft respectively. For the same well, according the integrated FEM-ML, this well requires mud weights of 9.70, 9.60, 9.40, 9.31, 9.31, and 8.91 ppg for arresting the cavings volume to under 1.5, 2, 3, 4, 5, 6 ft^3/12ft respectively. This comparison shows that the conventional failure criterion is underestimating the mud weight required for stability by more than 1.5 ppg. Also, this contrast and over-optimism of the conventional criterion can lead to excessive cavings that can exceed 6 ft^3/12ft.
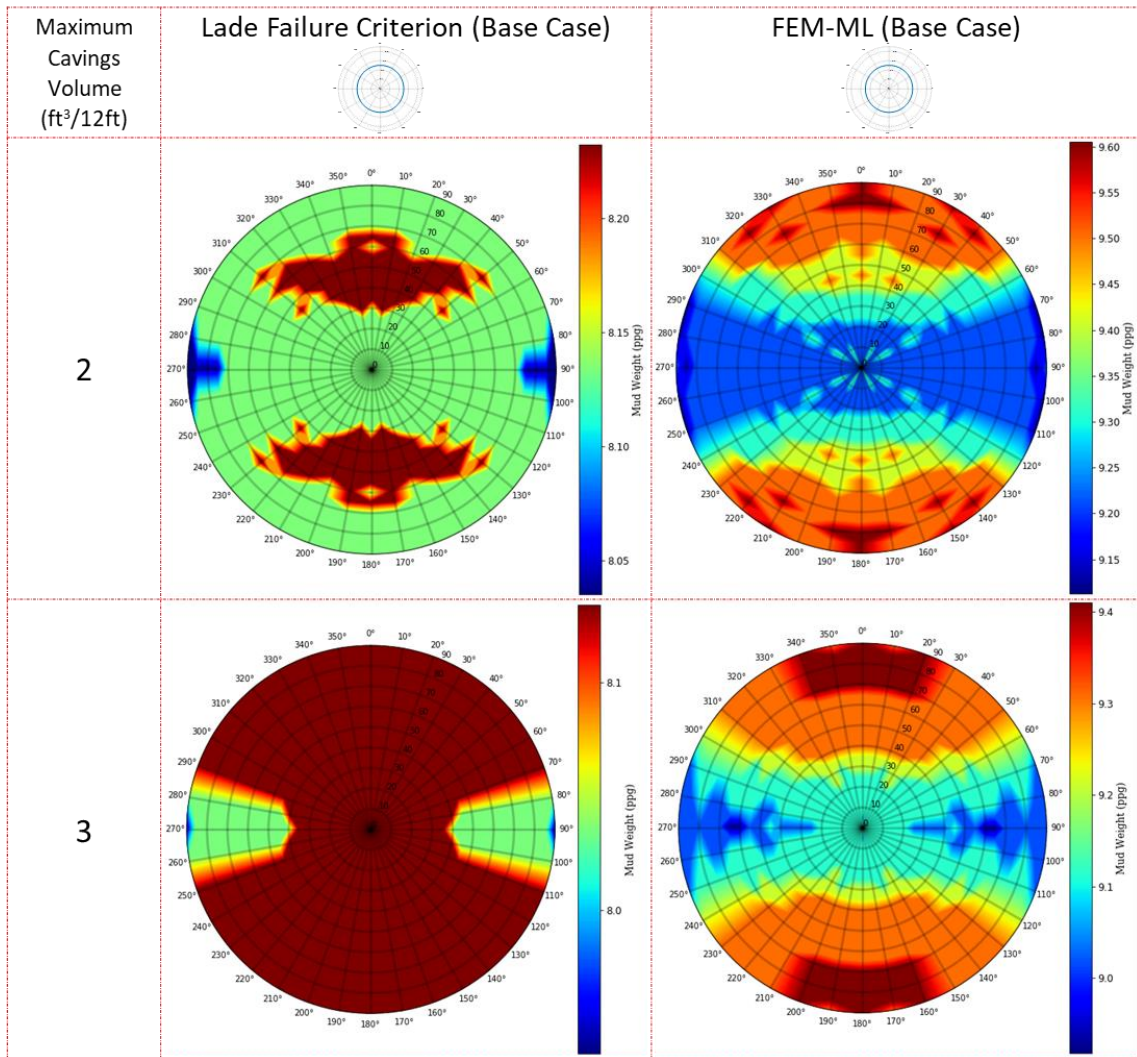
Figure 129. Stability mud weights polar charts for the base case scenario as determined by both the Lade failure criterion and the integrated and trained FEM-ML. The charts are risk-controlled at under two maximum allowed caving volumes, which are 2 ft³/12ft and 3 ft³/12ft.

Figure 130. Stability mud weights polar charts for the base case scenario as determined by both the Lade failure criterion and the integrated and trained FEM-ML. The charts are risk-controlled at under three maximum allowed caving volumes, which are 4 ft³/12ft, 5 ft³/12ft, and 6 ft³/12ft.

4.3.2.2 First Enlargement Intensity Case

The first case scenario of wellbore enlargements assumes small enlargements in the direction of the minimum horizontal stress as shown in **Figure 131**. This scenario can be identified while drilling through examining the shape and dimension of cavings in the shale shakers.
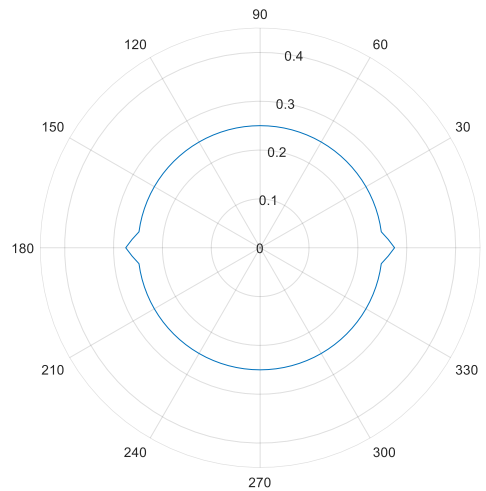
Figure 131. The shape of the wellbore in the first enlargement case scenario. Wellbore radial measurements are in ft.

Starting from this case scenario and on, the comparison is made between the stability mud weight polar charts for the base case (the wellbore with no enlargements) and the enlarged case scenario. For both cases, only the integrated FEM-ML is used to determine the stability mud weights and create the polar charts.

In this first case scenario of wellbore enlargement, the FEM-ML results show that it is not possible to completely prevent further cavings. The FEM-ML shows that the cavings volume can be minimized to under 1.5 ft^3/12ft with an initial overbalance mud

202

weight of 9.8 ppg. When comparing the charts at all risk controls as shown in **Figure 132,**

Figure 133, and **Figure 134** it can be seen that, in general, the first case will require the

stability mud weight to be raised by around 0.2 ppg from that of the base case.
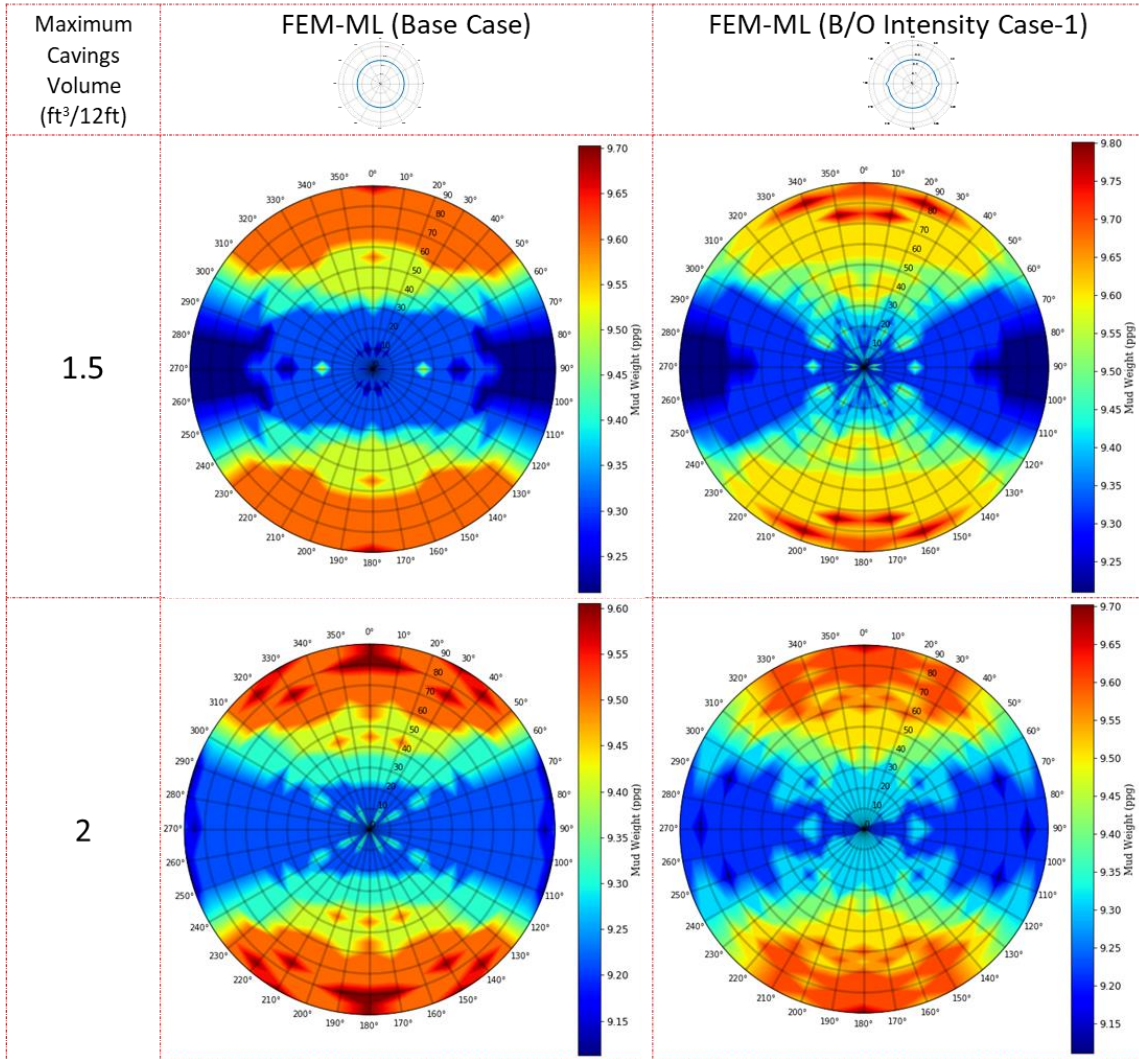
Figure 132. Stability mud weights polar charts for the first case scenario of wellbore enlargement as determined by the integrated and trained FEM-ML compared against those of the base case. The charts are risk-controlled at under two maximum allowed caving volumes, which are 1.5 ft³/12ft and 2 ft³/12ft.

For a closer examination of the results in **Figure 132, Figure 133**, and **Figure 134**, a horizontal well (90° inclination on the radial axis of the polar chart) in the direction of the minimum horizontal stress (0° azimuth on the circumferential axis of the polar chart) is considered. In the base case and according to the integrated FEM-ML, this well requires mud weights of 9.70, 9.60, 9.41, 9.31, 9.31, and 8.91 ppg for arresting the cavings volume to under 1.5, 2, 3, 4, 5, 6 ft^3/12ft respectively. For the same well but for the first case scenario of wellbore enlargement and according the integrated FEM-ML, this well requires mud weights of 9.70, 9.70, 9.62, 9.41, 9.31, and 9.12 ppg for arresting the cavings volume to under 1.5, 2, 3, 4, 5, 6 ft^3/12ft respectively. Alternatively, when considering a vertical well instead, in the base case and according to the integrated FEM-ML, this well requires mud weights of 9.21, 9.21, 9.12, 8.92, 8.92, and 8.92 ppg for arresting the cavings volume to under 1.5, 2, 3, 4, 5, 6 ft^3/12ft respectively. For the vertical well but for the first case scenario of wellbore enlargement and according the integrated FEM-ML, it requires mud weights of 9.31, 9.31, 9.21, 9.12, 8.92, and 8.92 ppg for arresting the cavings volume to under 1.5, 2, 3, 4, 5, 6 ft^3/12ft respectively. These comparisons mean that when cavings, which are of dimensions that are similar to this case scenario, are observed on the shale shakers, the mud weight should be only raised by 0.1 to 0.2 ppg to prevent further wellbore enlargements. The small changes in stability mud weight in this case scenario are expected due to the small wellbore enlargement.
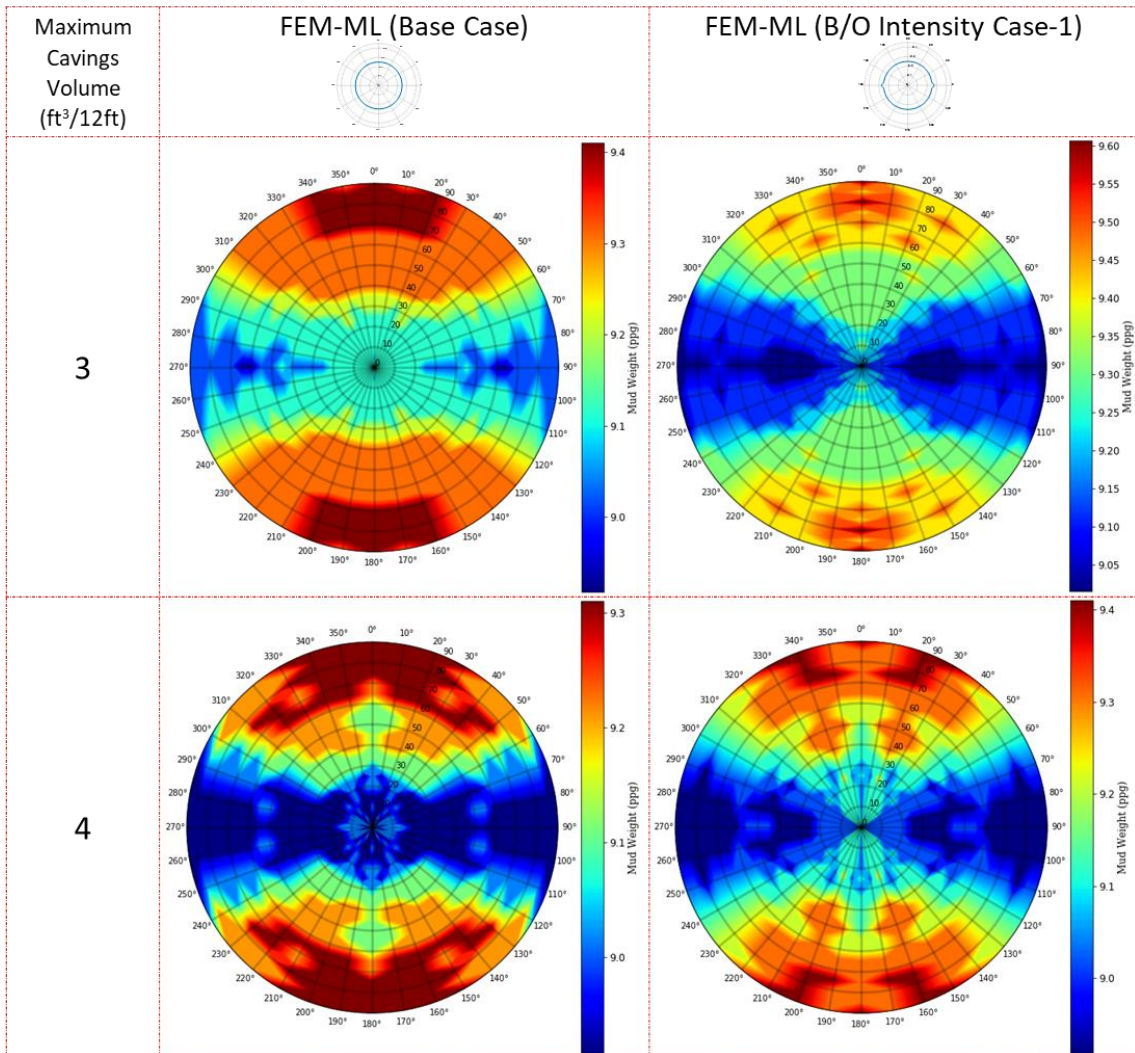
Figure 133. Stability mud weights polar charts for the first case scenario of wellbore enlargement as determined by the integrated and trained FEM-ML compared against those of the base case. The charts are risk-controlled at under two maximum allowed caving volumes, which are 3 ft³/12ft and 4 ft³/12ft.
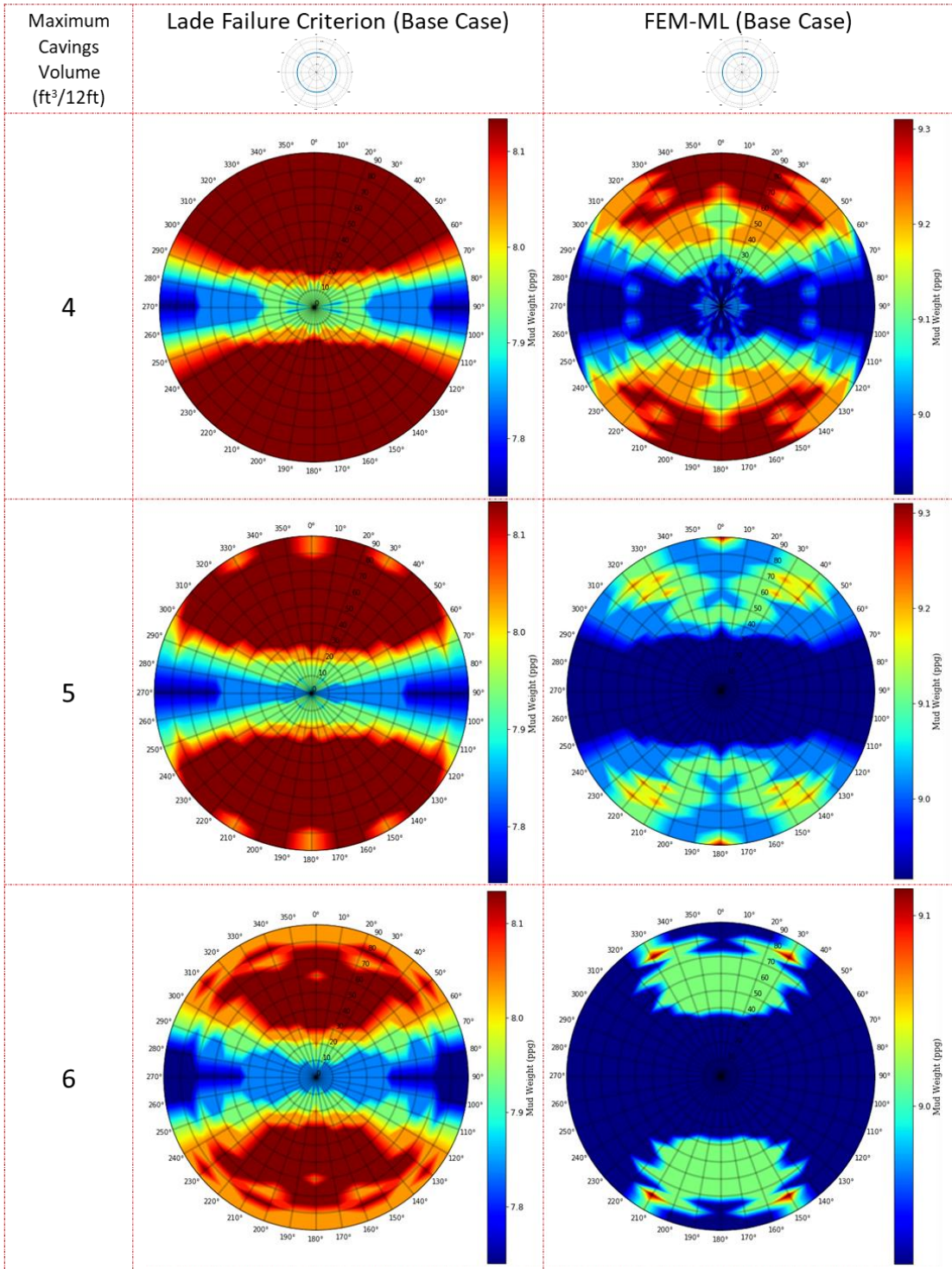
Figure 134. Stability mud weights polar charts for the first case scenario of wellbore enlargement as determined by the integrated and trained FEM-ML compared against those of the base case. The charts are risk-controlled at under two maximum allowed caving volumes, which are 5 ft³/12ft and 6 ft³/12ft.

4.3.2.3 Second Enlargement Intensity Case

The second case scenario of wellbore enlargements assumes enlargements in the direction of the minimum horizontal stress as shown in **Figure 135**. This scenario can be

identified while drilling through examining the shape and dimension of cavings in the shale shakers.
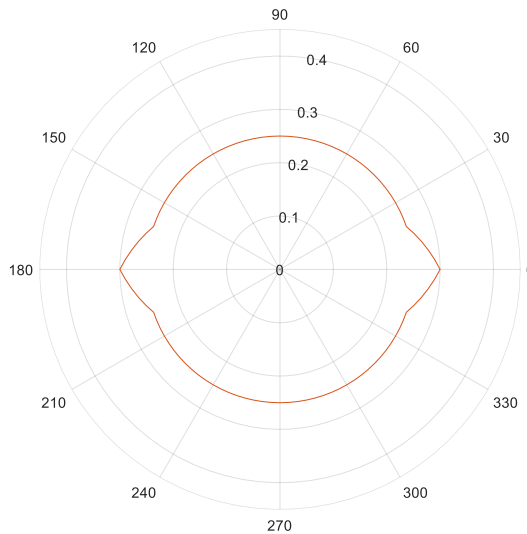
Figure 135. The shape of the wellbore in the second enlargement case scenario. Wellbore radial measurements are in ft.

In this second case scenario of wellbore enlargement, the FEM-ML results show that it is not possible to completely prevent further cavings. The FEM-ML shows that the cavings volume can be minimized to under 1.5 ft^3/12ft with an initial overbalance mud weight of 9.8 ppg. When comparing the charts at all risk controls as shown in **Figure 136, Figure 137**, and **Figure 138** it can be seen that, in general, the second case will require the stability mud weight to be raised by around 0.2 to 0.4 ppg from that of the base case.
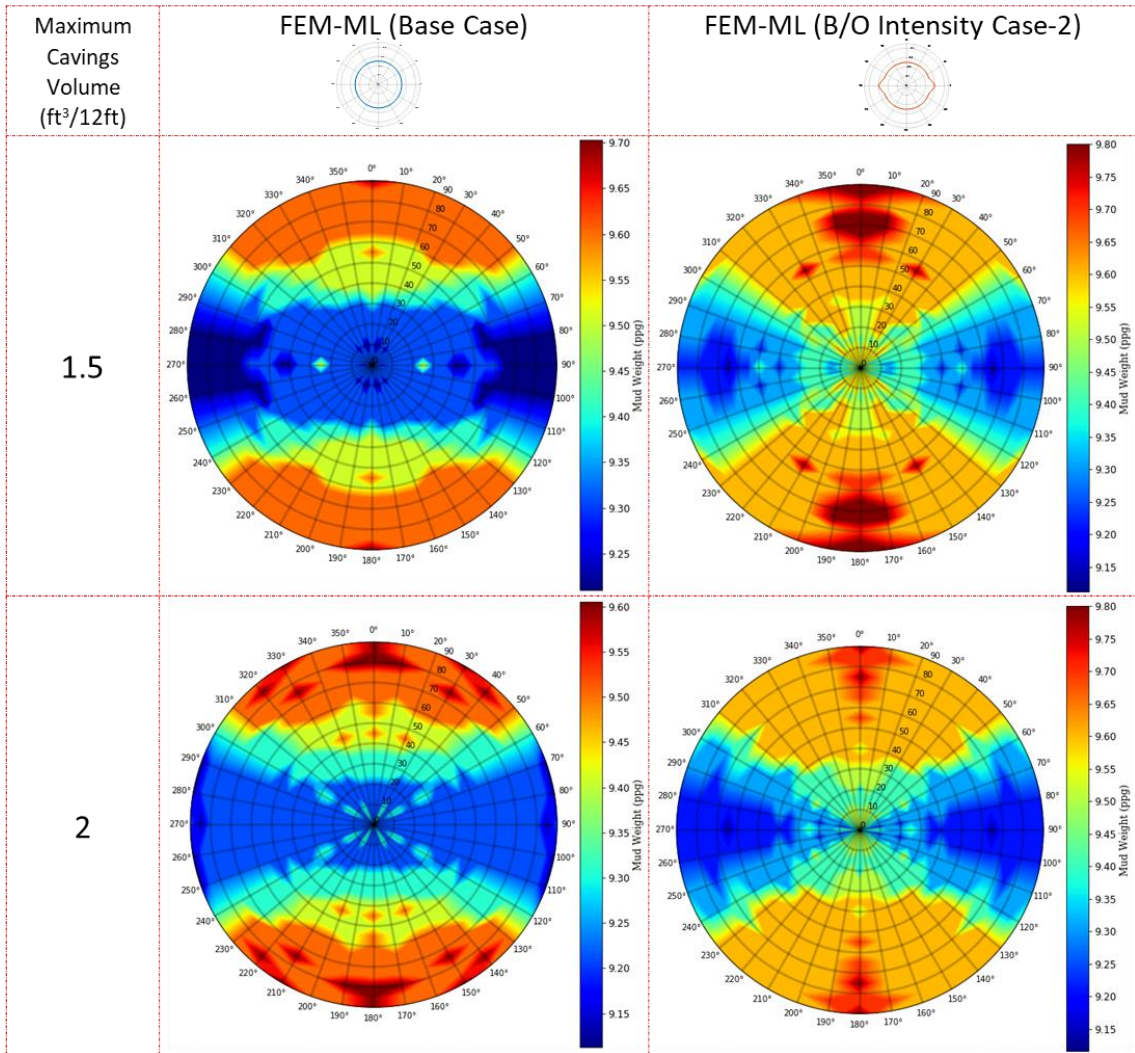
Figure 136. Stability mud weights polar charts for the second case scenario of wellbore enlargement as determined by the integrated and trained FEM-ML compared against those of the base case. The charts are risk-controlled at under two maximum allowed caving volumes, which are 1.5 ft³/12ft and 2 ft³/12ft.

For a closer examination of the results in **Figure 136, Figure 137**, and **Figure 138**, a horizontal well (90° inclination on the radial axis of the polar chart) in the direction of the minimum horizontal stress (0° azimuth on the circumferential axis of the polar chart) is considered. In the base case and according to the integrated FEM-ML, this well requires

mud weights of 9.70, 9.60, 9.41, 9.31, 9.31, and 8.91 ppg for arresting the cavings volume to under 1.5, 2, 3, 4, 5, 6 ft^3/12ft respectively. For the same well but for the second case scenario of wellbore enlargement and according the integrated FEM-ML, this well requires mud weights of 9.80, 9.80, 9.61, 9.51, 9.31, and 9.31 ppg for arresting the cavings volume to under 1.5, 2, 3, 4, 5, 6 ft^3/12ft respectively. Alternatively, when considering a vertical well instead, in the base case and according to the integrated FEM-ML, this well requires mud weights of 9.21, 9.21, 9.12, 8.92, 8.92, and 8.92 ppg for arresting the cavings volume to under 1.5, 2, 3, 4, 5, 6 ft^3/12ft respectively. For the vertical well but for the second case scenario of wellbore enlargement and according the integrated FEM-ML, it requires mud weights of 9.61, 9.61, 9.31, 9.21, 9.12, and 9.12 ppg for arresting the cavings volume to under 1.5, 2, 3, 4, 5, 6 ft^3/12ft respectively. These comparisons mean that when cavings, which are of dimensions that are similar to this case scenario, are observed on the shale shakers, the mud weight should be only raised by 0.2 to 0.4 ppg to prevent further wellbore enlargements.
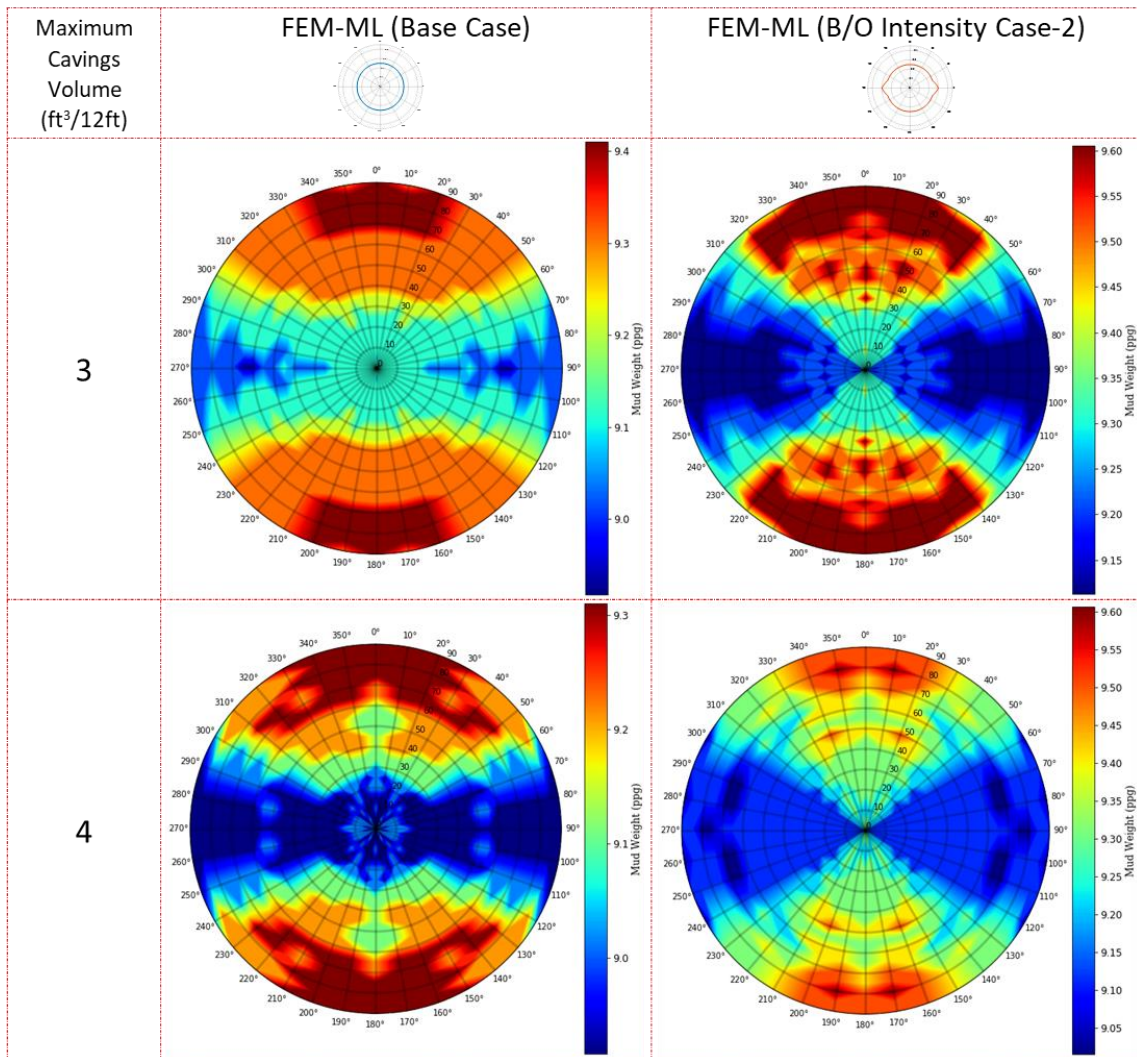
Figure 137. Stability mud weights polar charts for the second case scenario of wellbore enlargement as determined by the integrated and trained FEM-ML compared against those of the base case. The charts are risk-controlled at under two maximum allowed caving volumes, which are 3 ft³/12ft and 4 ft³/12ft.

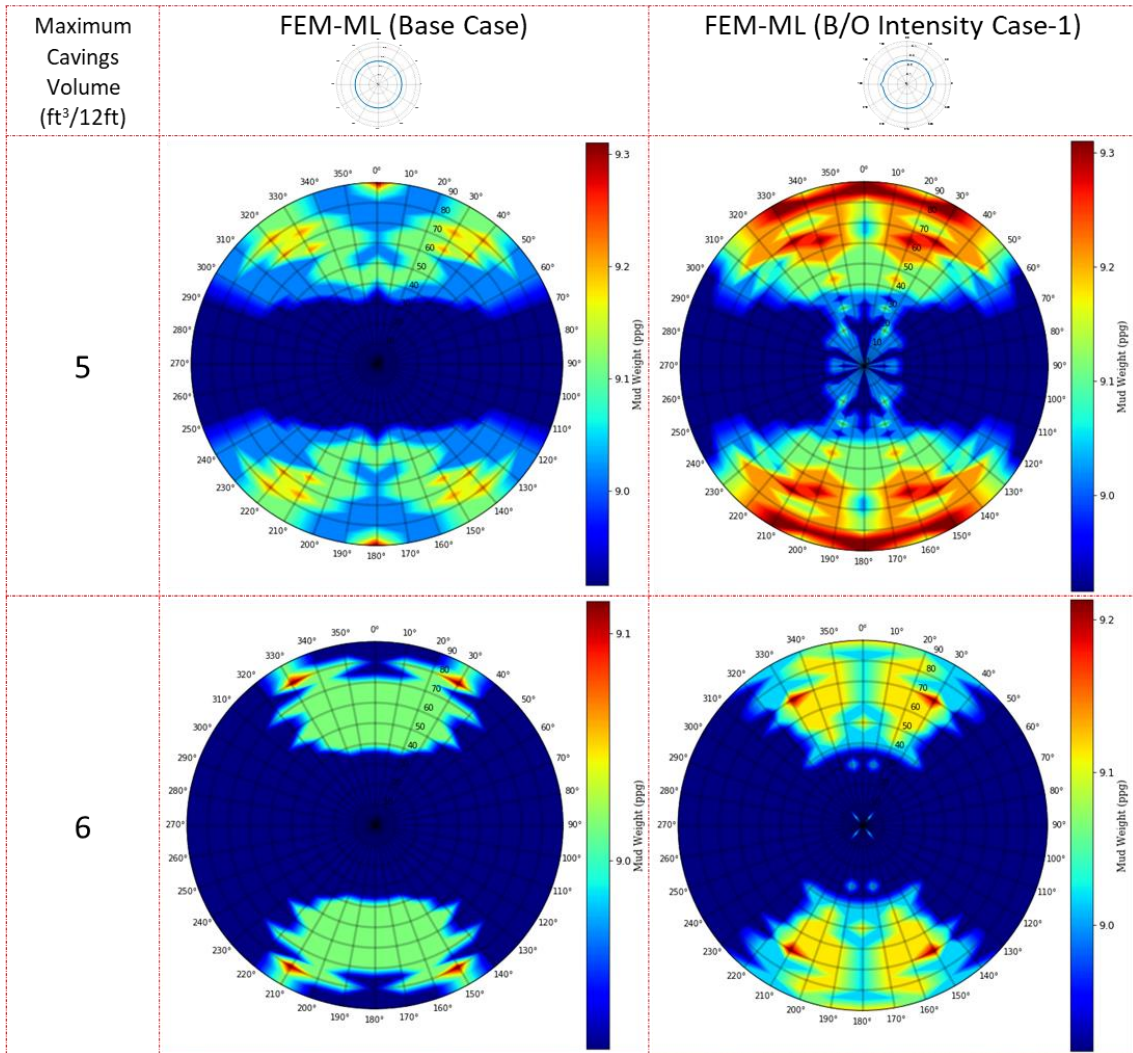| Maximum Cavings Volume (ft³/12ft)	FEM-ML (Base Case)	FEM-ML (B/O Intensity Case-2)
5		
6		

Figure 138. Stability mud weights polar charts for the second case scenario of wellbore enlargement as determined by the integrated and trained FEM-ML compared against those of the base case. The charts are risk-controlled at under two maximum allowed caving volumes, which are 5 ft³/12ft and 6 ft³/12ft.

4.3.2.4 Third Enlargement Intensity Case

The third case scenario of wellbore enlargements assumes enlargements in the direction of the minimum horizontal stress as shown in **Figure 139**. This scenario can be

identified while drilling through examining the shape and dimension of cavings in the shale shakers.
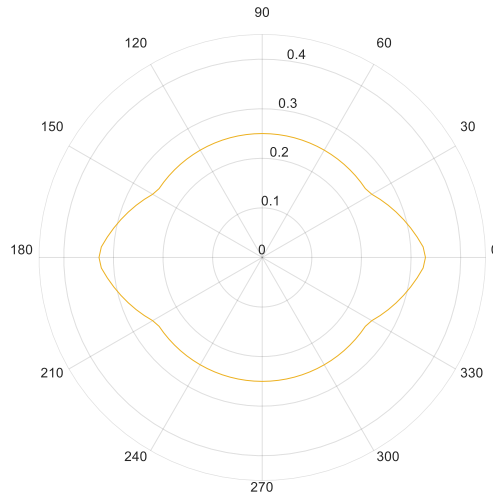
Figure 139. The shape of the wellbore in the third enlargement case scenario. Wellbore radial measurements are in ft.

In this third case scenario of wellbore enlargement, the FEM-ML results show that it is not possible to completely prevent further cavings. The FEM-ML shows that the cavings volume can be minimized to under 1.5 ft^3/12ft with an initial overbalance mud weight of 12 ppg. When comparing the charts at all risk controls as shown in **Figure 140, Figure 141**, and **Figure 142** it can be seen that, in general, the third case will require the stability mud weight to be raised by around 2 ppg from that of the base case.
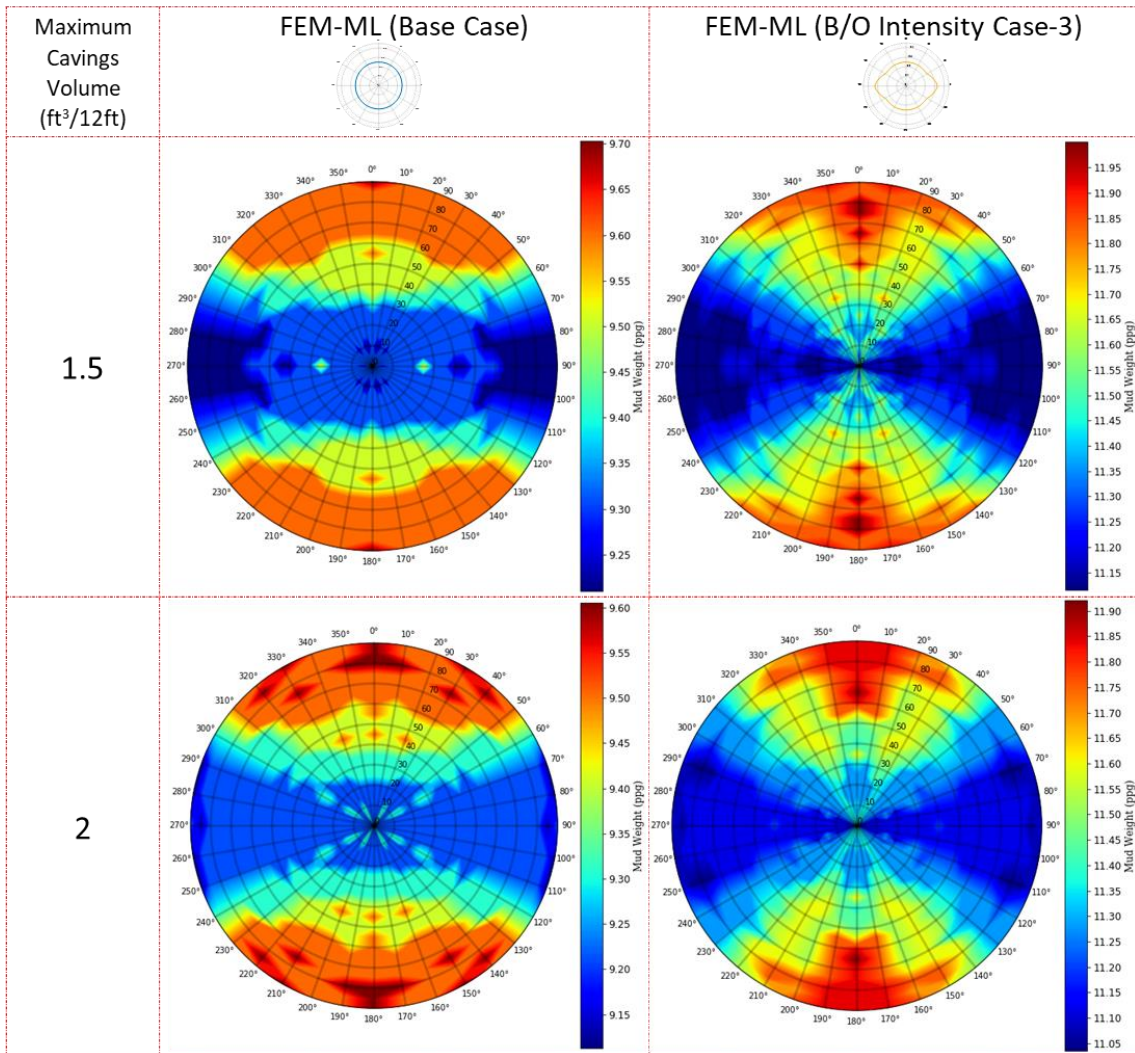
Figure 140. Stability mud weights polar charts for the third case scenario of wellbore enlargement as determined by the integrated and trained FEM-ML compared against those of the base case. The charts are risk-controlled at under two maximum allowed caving volumes, which are 1.5 ft³/12ft and 2 ft³/12ft.

For a closer examination of the results in **Figure 140, Figure 141**, and **Figure 142**, a horizontal well (90° inclination on the radial axis of the polar chart) in the direction of the minimum horizontal stress (0° azimuth on the circumferential axis of the polar chart) is considered. In the base case and according to the integrated FEM-ML, this well requires

mud weights of 9.70, 9.60, 9.41, 9.31, 9.31, and 8.91 ppg for arresting the cavings volume to under 1.5, 2, 3, 4, 5, 6 ft^3/12ft respectively. For the same well but for the third case scenario of wellbore enlargement and according the integrated FEM-ML, this well requires mud weights of 11.84, 11.84, 11.76, 11.60, 11.44, and 11.36 ppg for arresting the cavings volume to under 1.5, 2, 3, 4, 5, 6 ft^3/12ft respectively. Alternatively, when considering a vertical well instead, in the base case and according to the integrated FEM-ML, this well requires mud weights of 9.21, 9.21, 9.12, 8.92, 8.92, and 8.92 ppg for arresting the cavings volume to under 1.5, 2, 3, 4, 5, 6 ft^3/12ft respectively. For the vertical well but for the third case scenario of wellbore enlargement and according the integrated FEM-ML, it requires mud weights of 11.28, 11.20, 11.12, 11.12, 11.12, and 11.04 ppg for arresting the cavings volume to under 1.5, 2, 3, 4, 5, 6 ft^3/12ft respectively. These comparisons mean that when cavings, which are of dimensions that are similar to this case scenario, are observed on the shale shakers, the mud weight should be only raised by around 2 ppg to prevent further wellbore enlargements.

Figure 141. Stability mud weights polar charts for the third case scenario of wellbore enlargement as determined by the integrated and trained FEM-ML compared against those of the base case. The charts are risk-controlled at under two maximum allowed caving volumes, which are 3 ft³/12ft and 4 ft³/12ft.
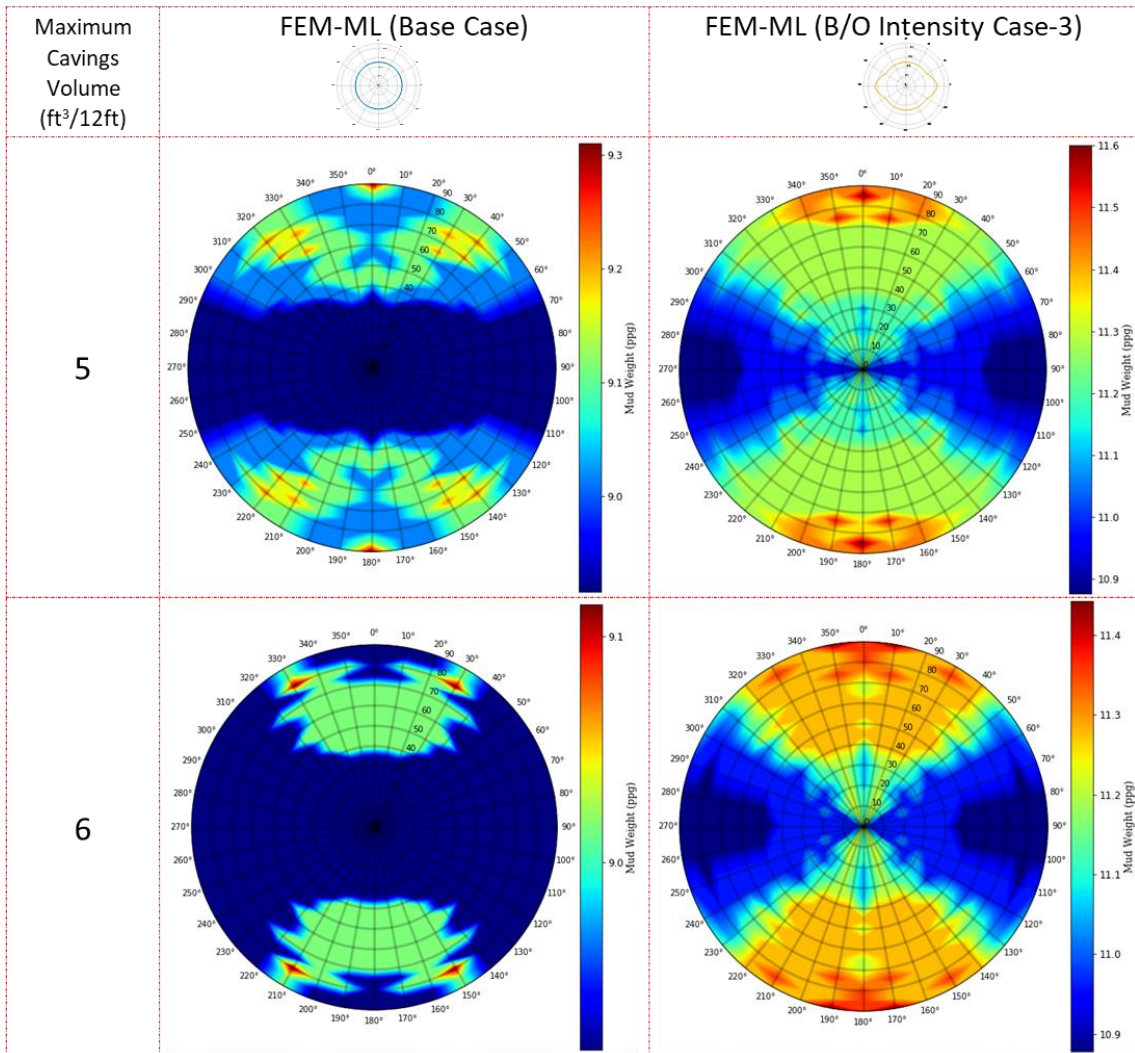
Figure 142. Stability mud weights polar charts for the third case scenario of wellbore enlargement as determined by the integrated and trained FEM-ML compared against those of the base case. The charts are risk-controlled at under two maximum allowed caving volumes, which are 5 ft³/12ft and 6 ft³/12ft.
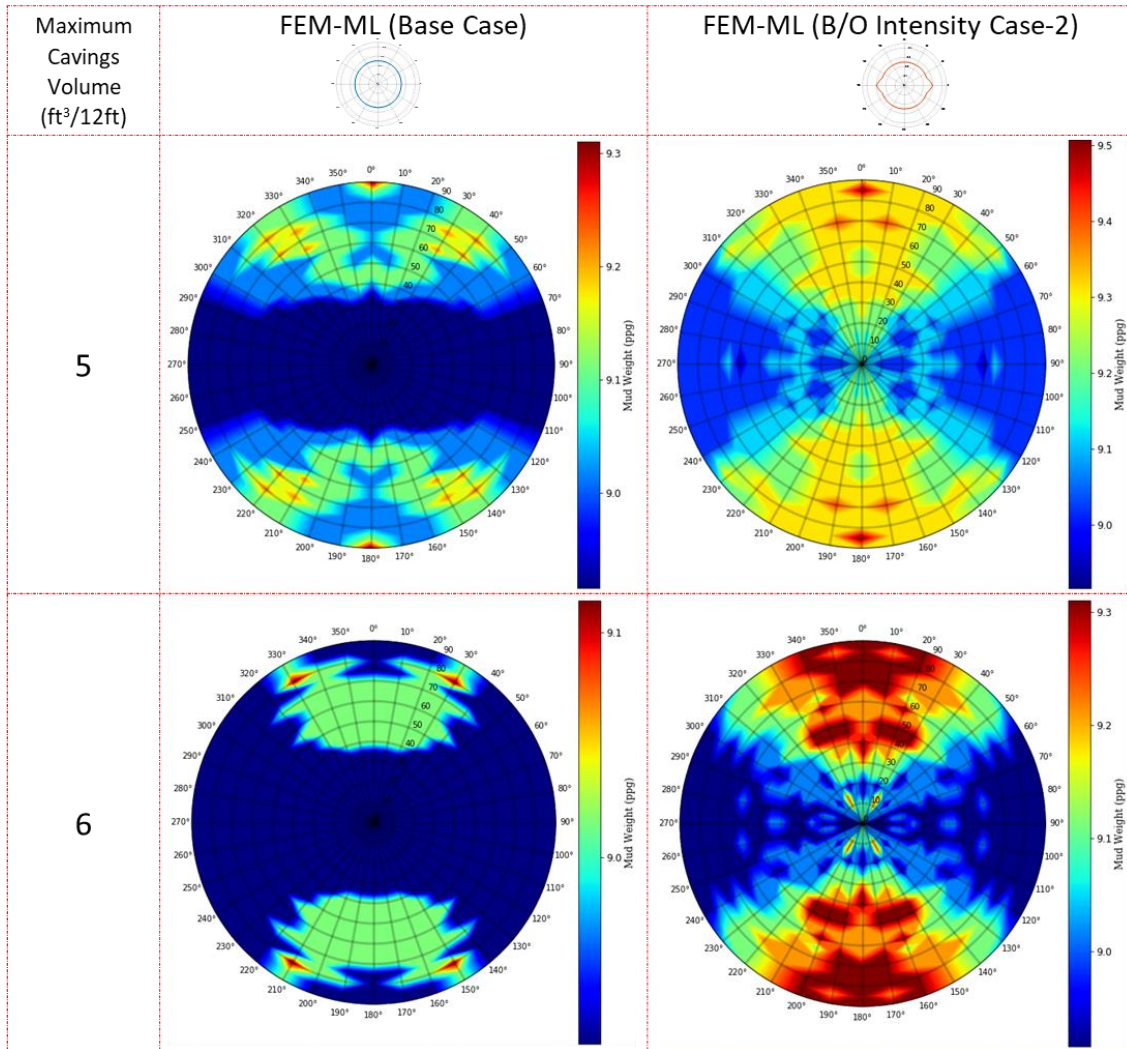
4.3.2.5 Fourth Enlargement Intensity Case

The fourth case scenario of wellbore enlargements assumes very extensive enlargements in the direction of the minimum horizontal stress as shown in **Figure 143**. This scenario might not be realistically possible in terms of both occurrence and detection

from shale shakers surveillance. This is because the large cavings will most likely break apart gradually, which means intact cavings will not be seen on the shale shakers to confirm the occurrence of this case scenario. However, it is included in this section to complete the sensitivity analysis with exaggerated enlargements.
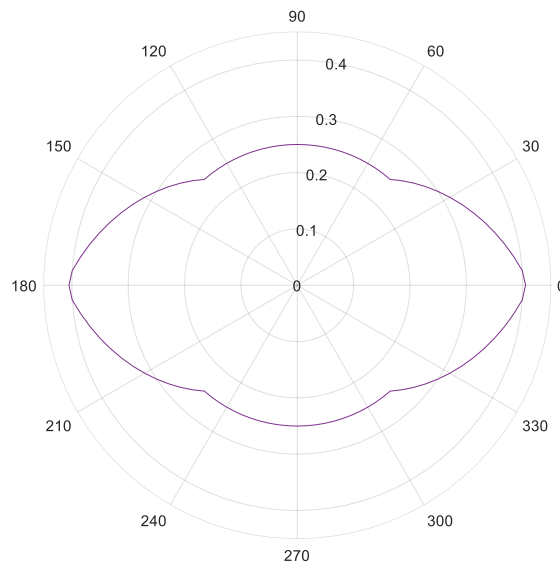
Figure 143. The shape of the wellbore in the fourth enlargement case scenario. Wellbore radial measurements are in ft.

In this fourth case scenario of wellbore enlargement, the FEM-ML results show that it is not possible to completely prevent further cavings. The FEM-ML shows that the cavings volume can be minimized to under 1.5 ft^3/12ft with an initial overbalance mud weight of 13.2 ppg. When comparing the charts at all risk controls as shown in **Figure 144**, **Figure 145**, and **Figure 146** it can be seen that, in general, the fourth case will require the stability mud weight to be raised by around 3.5 ppg from that of the base case.
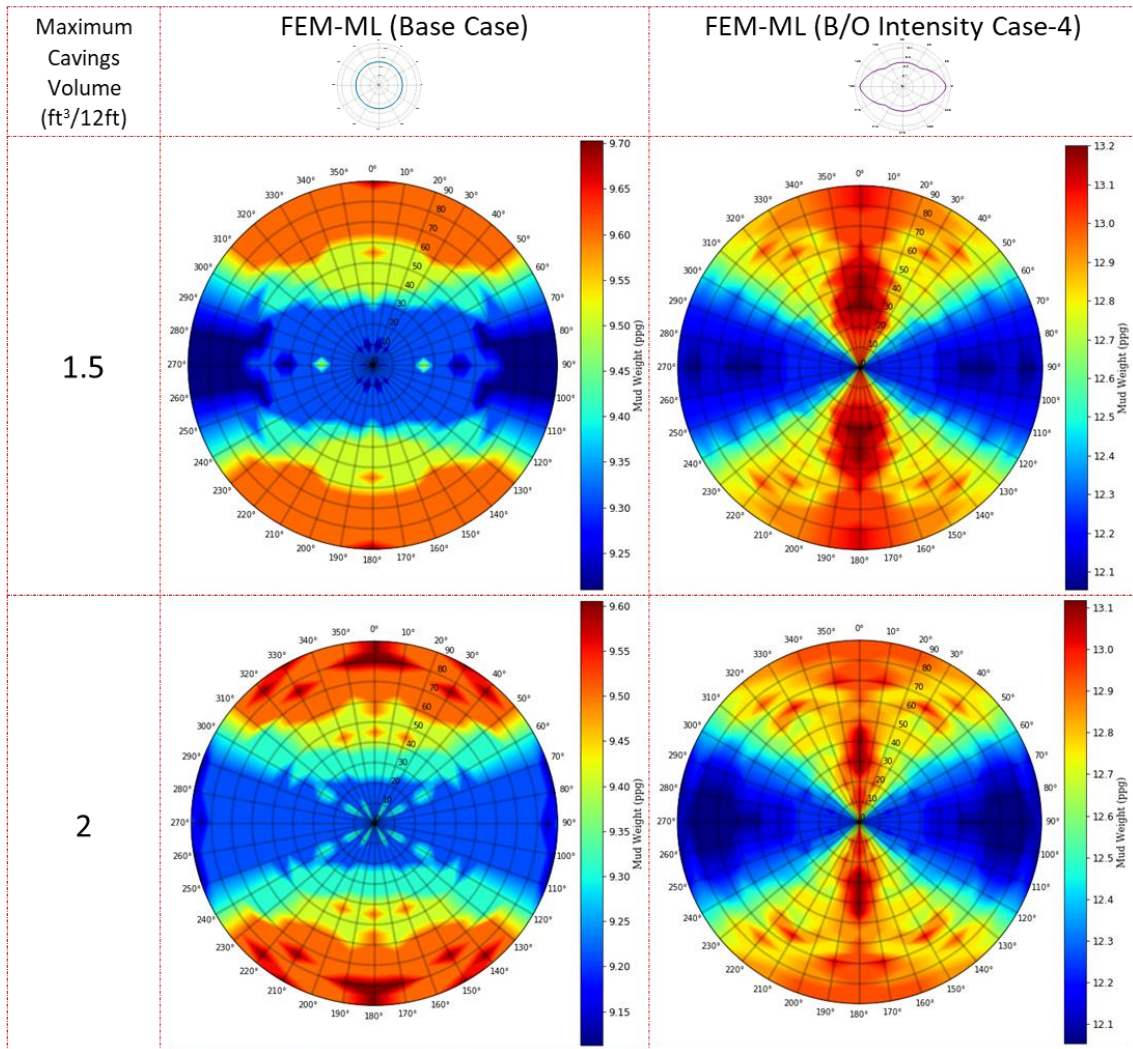
Figure 144. Stability mud weights polar charts for the fourth case scenario of wellbore enlargement as determined by the integrated and trained FEM-ML compared against those of the base case. The charts are risk-controlled at under two maximum allowed caving volumes, which are 1.5 ft³/12ft and 2 ft³/12ft.

For a closer examination of the results in **Figure 144**, **Figure 145**, and **Figure 146**, a horizontal well (90° inclination on the radial axis of the polar chart) in the direction of the minimum horizontal stress (0° azimuth on the circumferential axis of the polar chart) is considered. In the base case and according to the integrated FEM-ML, this well requires

mud weights of 9.70, 9.60, 9.41, 9.31, 9.31, and 8.91 ppg for arresting the cavings volume to under 1.5, 2, 3, 4, 5, 6 ft^3/12ft respectively. For the same well but for the fourth case scenario of wellbore enlargement and according the integrated FEM-ML, this well requires mud weights of 13.11, 12.94, 12.85, 12.67, 12.89, and 12.59 ppg for arresting the cavings volume to under 1.5, 2, 3, 4, 5, 6 ft^3/12ft respectively. Alternatively, when considering a vertical well instead, in the base case and according to the integrated FEM-ML, this well requires mud weights of 9.21, 9.21, 9.12, 8.92, 8.92, and 8.92 ppg for arresting the cavings volume to under 1.5, 2, 3, 4, 5, 6 ft^3/12ft respectively. For the vertical well but for the fourth case scenario of wellbore enlargement and according the integrated FEM-ML, it requires mud weights of 13.03, 12.85, 12.76, 12.32, 12.32, and 12.23 ppg for arresting the cavings volume to under 1.5, 2, 3, 4, 5, 6 ft^3/12ft respectively. These comparisons mean that when enlargements, which are of dimensions that are similar to this case scenario, are identified, the mud weight should be only raised by around 3.5 ppg to prevent further wellbore enlargements.
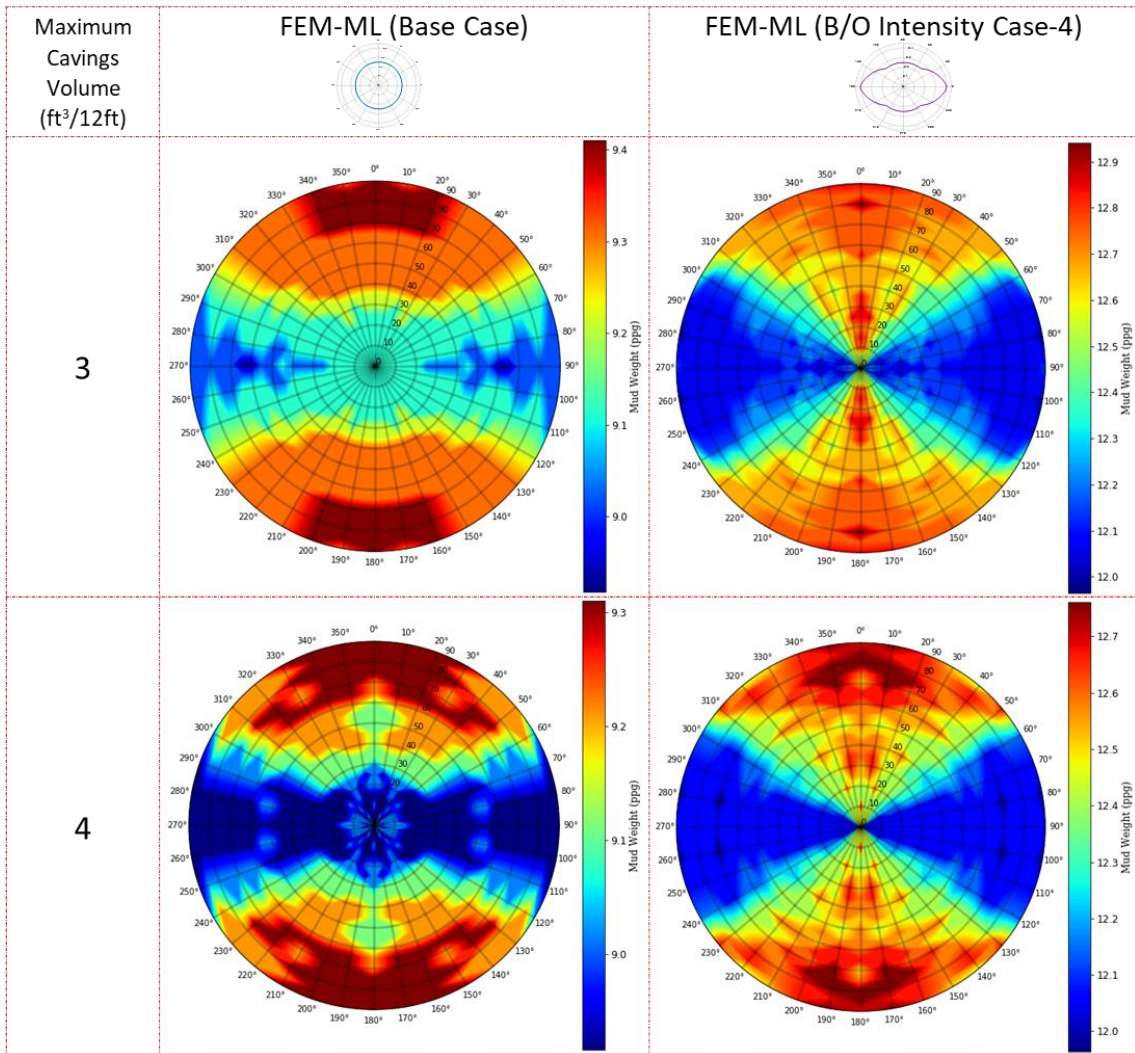
Figure 145. Stability mud weights polar charts for the fourth case scenario of wellbore enlargement as determined by the integrated and trained FEM-ML compared against those of the base case. The charts are risk-controlled at under two maximum allowed caving volumes, which are 3 ft³/12ft and 4 ft³/12ft.
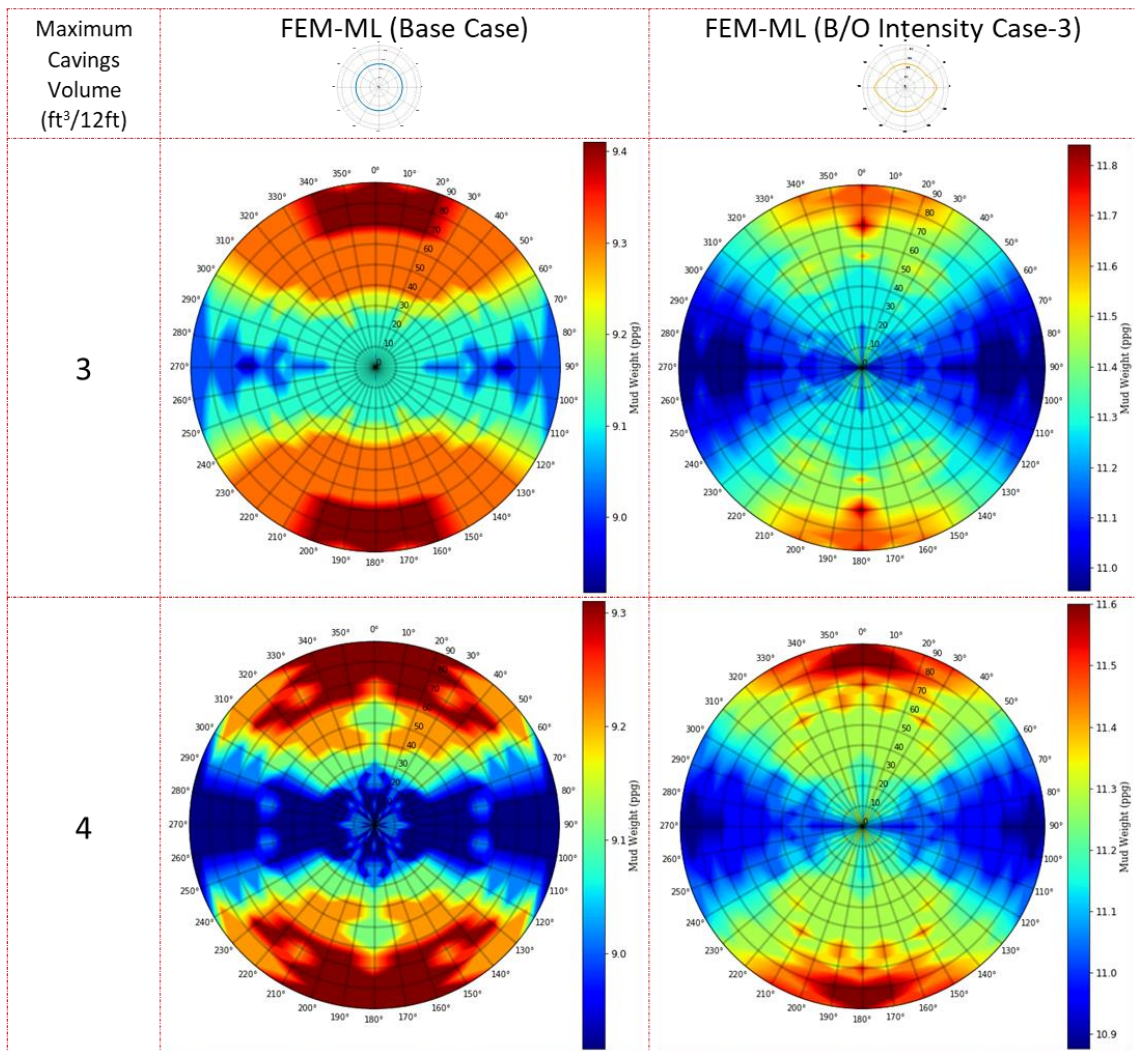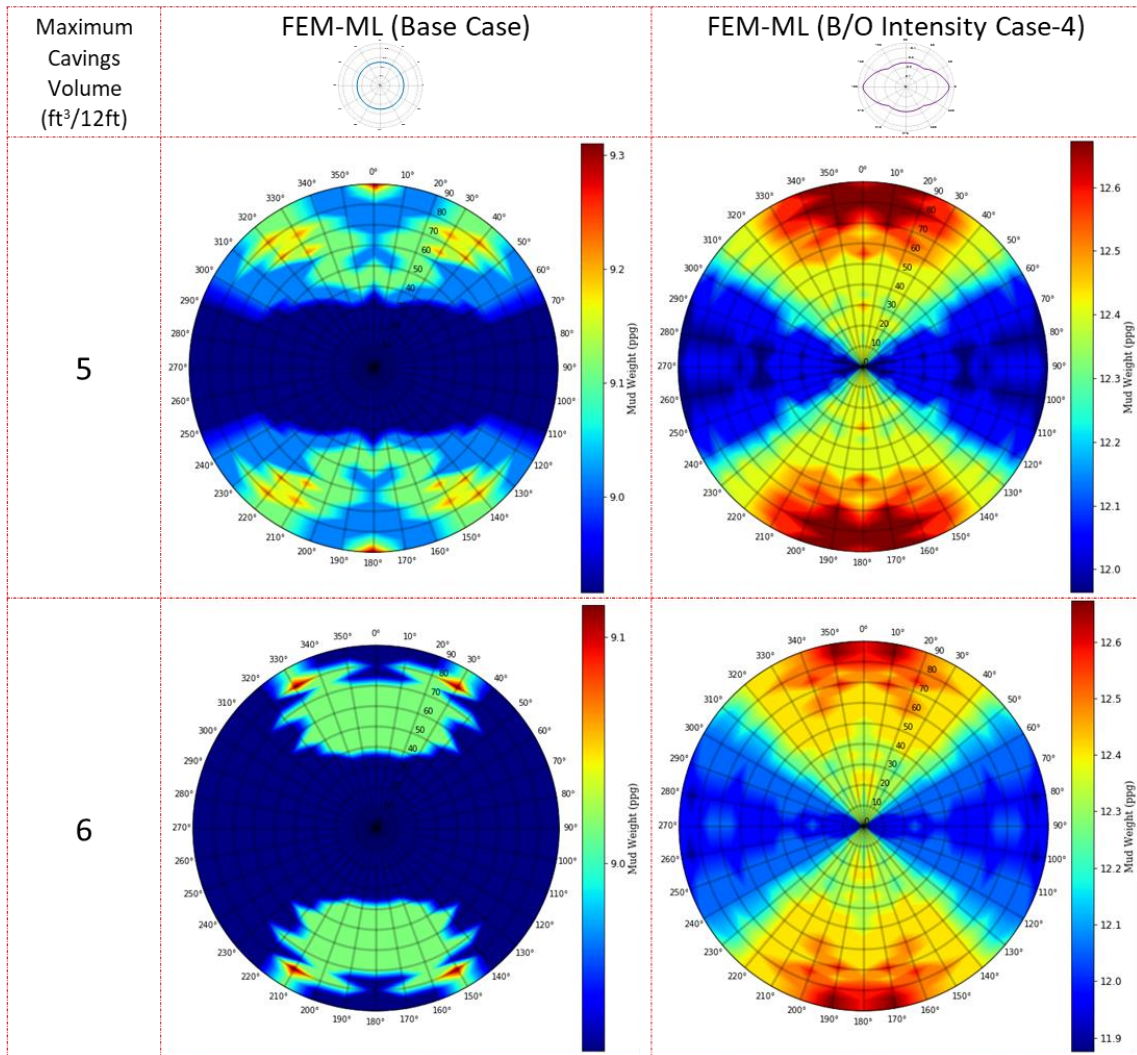
Figure 146. Stability mud weights polar charts for the fourth case scenario of wellbore enlargement as determined by the integrated and trained FEM-ML compared against those of the base case. The charts are risk-controlled at under two maximum allowed caving volumes, which are 5 ft³/12ft and 6 ft³/12ft.

4.3.2.6 Summary of the Results for All Cases for a Horizontal Well and a Vertical Well

A summary of the results from all the previous cases and based on the close examination of a horizontal well in the direction of the minimum horizontal stress and for a vertical well are in **Table 15**, **Table 16**, **Table 17**, and **Table 18**. These results detail the

221

stability mud weight requirements for all wellbore enlargement case scenarios for the horizontal well in **Table 15** and for the vertical well in **Table 17**. Both of which are at different risk-control of caving volume. The results also detail the required raise of stability mud weight from that of the base case (uniform wellbore) to all the wellbore enlargement case scenarios. The mud weight raise required for the horizontal well are in **Table 16** and the ones for the vertical well are in **Table 18**.

Table 15. Summary of the stability mud weights required for a horizontal wellbore in the direction of the minimum horizontal stress. The stability mud weights are from the polar chart and are based on the integrated FEM-ML.

Maximum Allowed Cavings Volume, ft^3/12 ft	Stability Mud Weights for A Horizontal Well in the Minimum Horizontal Stress Direction (FEM-ML), ppg				
	Base Case	Case 1	Case 2	Case 3	Case 4
1.5	9.7022	9.7022	9.8002	11.8403	13.1123
2	9.6042	9.7022	9.8002	11.8403	12.9363
3	9.4082	9.6042	9.6042	11.7603	12.8483
4	9.3102	9.4082	9.5062	11.6003	12.6723
5	9.3102	9.3102	9.3102	11.4402	12.5843
6	8.9182	9.1142	9.3102	11.3602	12.5843

Table 16. Summary of the stability mud weights raises from those of the base case for a horizontal wellbore in the direction of the minimum horizontal stress. The stability mud weights are from the polar chart and are based on the integrated FEM-ML.

Maximum Allowed Cavings Volume, ft^3/12 ft	Required Stability Mud Weight Raise for A Horizontal Well in the Minimum Horizontal Stress Direction (FEM-ML), ppg				
	Base Case	Case 1	Case 2	Case 3	Case 4
1.5	0	0	0.098	2.1381	3.4101
2	0	0.098	0.196	2.2361	3.3321
3	0	0.196	0.196	2.3521	3.4401
4	0	0.098	0.196	2.2901	3.3621
5	0	0	0	2.13	3.2741
6	0	0.196	0.392	2.442	3.6661

Table 17. Summary of the stability mud weights required for a vertical wellbore. The stability mud weights are from the polar chart and are based on the integrated FEM-ML.

Maximum Allowed Cavings Volume, ft^3/12 ft	Stability Mud Weights for A Vertical Well (FEM-ML), ppg				
	Base Case	Case 1	Case 2	Case 3	Case 4
1.5	9.2122	9.3102	9.6042	11.2802	13.0243
2	9.2122	9.3102	9.6042	11.2002	12.8483
3	9.1142	9.2122	9.3102	11.1202	12.7603
4	8.9182	9.1142	9.2122	11.1202	12.3203
5	8.9182	8.9182	9.1142	11.1202	12.3203
6	8.9182	8.9182	9.1142	11.0402	12.2323

Table 18. Summary of the stability mud weights raises from those of the base case for a vertical wellbore. The stability mud weights are from the polar chart and are based on the integrated FEM-ML.

Maximum Allowed Cavings Volume, ft^3/12 ft	Required Stability Mud Weight Raise for A Vertical Well (FEM-ML), ppg				
	Base Case	Case 1	Case 2	Case 3	Case 4
1.5	0	0.098	0.392	2.068	3.8121
2	0	0.098	0.392	1.988	3.6361
3	0	0.098	0.196	2.006	3.6461
4	0	0.196	0.294	2.202	3.4021
5	0	0	0.196	2.202	3.4021
6	0	0	0.196	2.122	3.3141

4.4 Examining the Wellbore Size Effect

Deciding the wellbore size is a crucial step in drilling engineering design as different wellbore sizes within the same formation and at the same depth can experience rock failure at different loading limits. Therefore, it is common to develop empirical correlations that evaluate the wellbore size effect on stability and induced fractures. Morita (2004) combines the use of laboratory and field data to derive an empirical correlation for the wellbore size effect at different stress regimes and for wells drilled in different directions. The empirical correlation relies on the critical plastic strain as the failure criterion, which requires the determination of empirical fitting constants and a breakout angle. It also relies on the assumption that wellbore shape is circular and breakouts take place symmetrically. Similar analysis can be performed using the integrated model, which is trained based on different borehole sizes. It is crucial to note that the actual wellbore size is not a distinct feature in the training process in the machine learning algorithm, which is unlike the methods used in empirical correlations. It is also not a parameter in

conventional failure criteria. For both methods, failure is evaluated at every mesh node based on its stress state, which in turn is dependent on structural considerations such as the wellbore size. Not including the wellbore size as a distinct feature in the training process or as a parameter in the conventional failure criteria ensures that the final model is generalizable to all wellbore sizes. Another advantage of the integrated model is that it does not rely on the assumption that breakout take place symmetrically. Instead, wellbore deformation is evaluated as total caving volumes, which can be distributed irregularly around the wellbore. The integrated FEM-ML model ability to evaluate the wellbore size effect is examined in this section by using Well-2 from the offset wells data presented at the beginning of this chapter. The integrated model results are examined against those from a conventional failure criterion (Lade). The basic Well-2 information used to perform the analysis are shown in **Table 19**.

Table 19. The Well-2 information used as basis for the wellbore size effect analysis

Parameter	Value
Pore pressure and mud weight, ppg	(8.5) (9.2)
Overburden (vertical) stress, ppg	18
Maximum horizontal stress, ppg	20.8
Minimum horizontal stress, ppg	12.1
Depth and wellbore size, ft and in.	(12,795) (6)
Actual cavings volume, ft^3/12ft	3

Variations in risk-controlled stability mud weights for a vertical well, which are estimated at different wellbore sizes using both the integrated FEM-ML and the Lade failure criterion are shown is **Figure 147** and **Figure 148** respectively. Both results are examined against Well-2 actual caving volume.

Wellbore Deformations Represented as Caving Volumes at Different Wellbore Sizes for Vertical Wells with FEM-ML
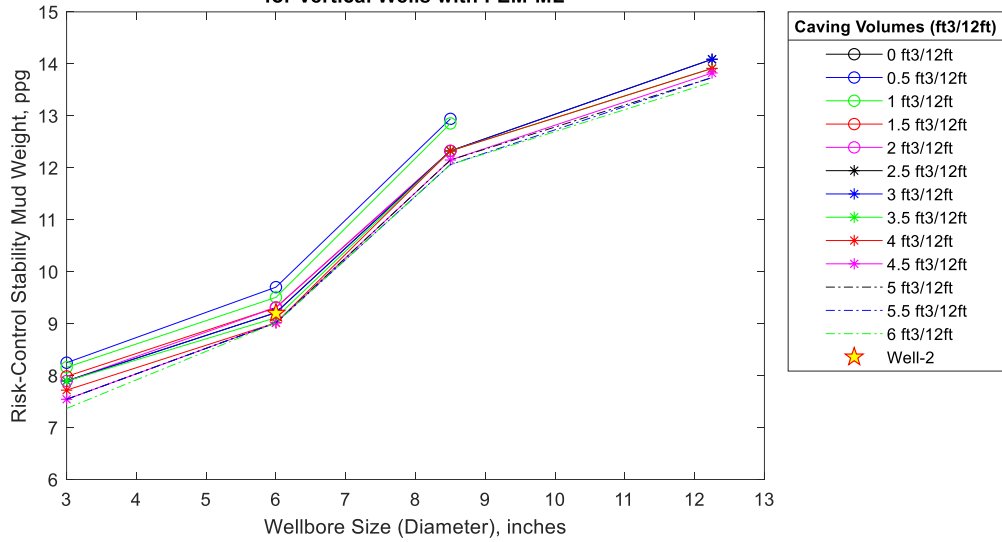
Figure 147. Variations in risk-controlled stability mud weights for a vertical well, which are estimated at different wellbore sizes using the integrated FEM-ML and examined against Well-2 actual caving volume.

Wellbore Deformations Represented as Caving Volumes at Different Wellbore Sizes for Vertical Wells with Lade
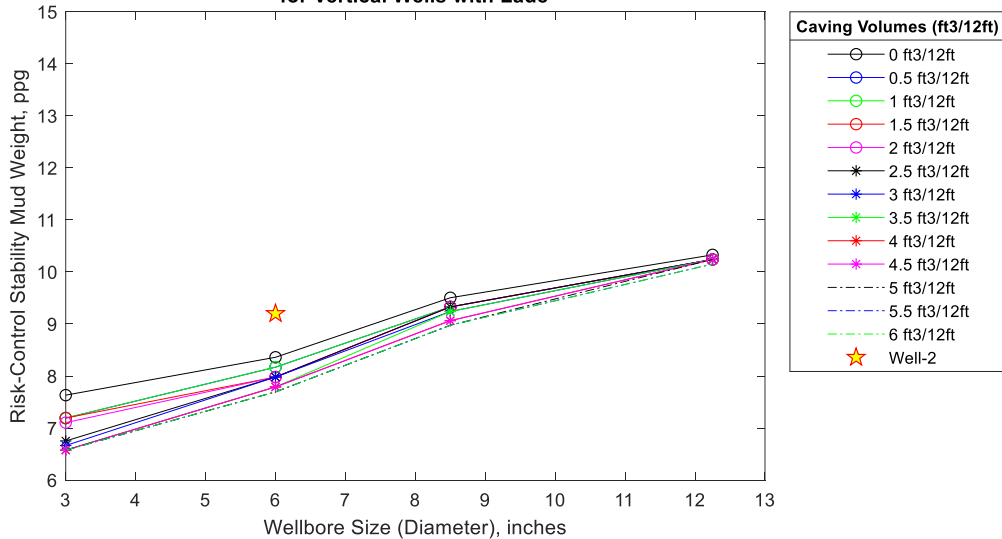
Figure 148. Variations in risk-controlled stability mud weights for a vertical well, which are estimated at different wellbore sizes using the Lade failure criterion and examined against Well-2 actual caving volume.

The wellbore sizes examined here include common wellbore sizes, which are 6 in., 8.5 in., and 12.25 in. and a smaller wellbore size of 3 in. It should be noted that the wellbore is in a strike-slip stress regime, which means the vertical stress is the intermediated stress. The well is placed at a depth of 12,795 ft. When examining the trends in both **Figure 147** and **Figure 148**, it can be seen that as the wellbore size increases, the amount of deformation and risk-controlled stability mud weight also increase. This is the expected trend and ensuring that the integrated FEM-ML model follows it is the first validation step in this analysis. When comparing the magnitude of risk-controlled mud weights in **Figure 147** and **Figure 148**, it is clear that the integrated model generally predicts higher mud weights. This is validated by the actual cavings volume from Well-2. As shown in **Figure 147**. The integrated model results show that the 6 in. wellbore of Well-2 produces 2.5 to 3 ft^3/12ft cavings when drilled with a 9.2 ppg mud weight. It also shows that it is not possible to reduce the deformation to less than 0.5 ft^3/12ft for the sizes of 3 in., 6 in., and 8.5 in. As for the 12.25 in. wellbore size, the integrated model shows that it is not possible to reduce the deformation to less than 2.5 ft^3/12ft. On the other hand, the Lade determined deformations are more optimistic as it estimates no deformations to take place for Well-2 at the same mud weight. It also shows that it is possible to completely prevent deformations for all the examined wellbore sizes.

To examine the effect of wellbore orientation on the size effect analysis, variations in risk-controlled stability mud weights for a horizontal well in the minimum stress direction, which are estimated at different wellbore sizes using both the integrated FEM-ML and the Lade failure criterion, are shown is **Figure 149** and **Figure 150** respectively.
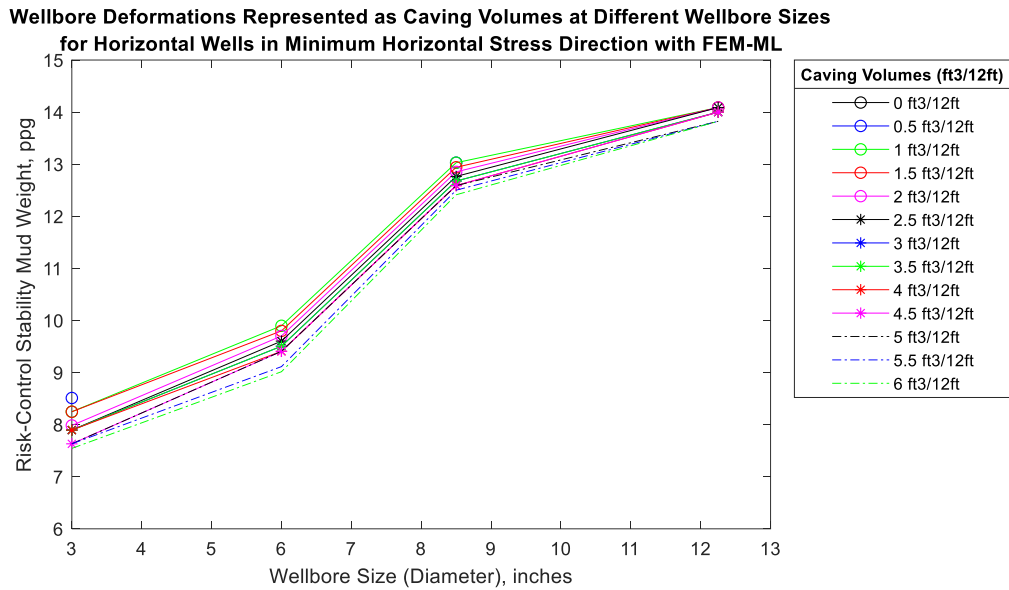
Figure 149. Variations in risk-controlled stability mud weights for a horizontal well in the minimum stress direction, which are estimated at different wellbore sizes using the integrated FEM-ML.
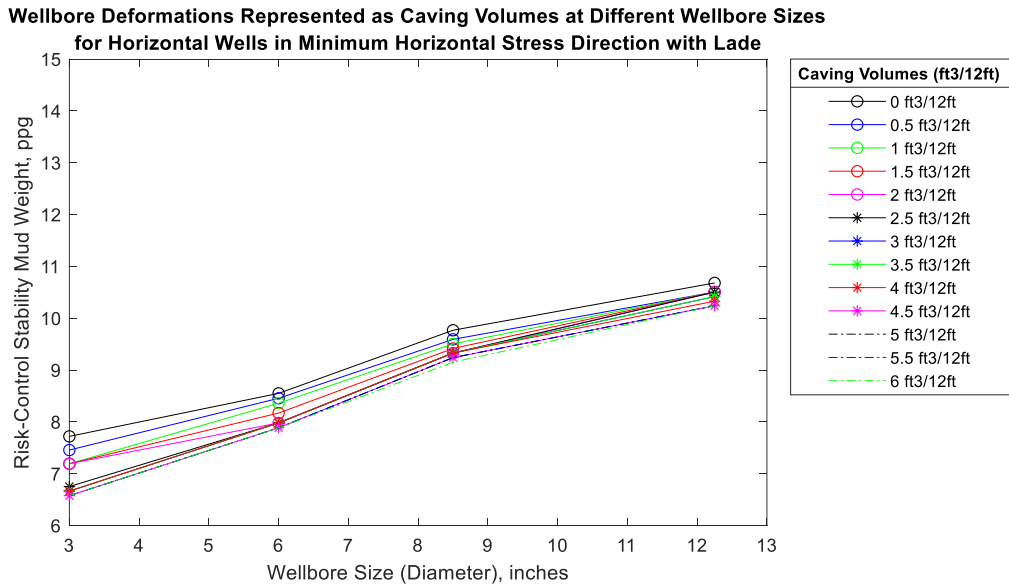
Figure 150. Variations in risk-controlled stability mud weights for a horizontal well in the minimum stress direction, which are estimated at different wellbore sizes using the Lade failure criterion.

The plots for the horizontal well cases are produced using the same the integrated FEM-ML used for producing the vertical well plots. This is possible because as discussed in the second chapter in section 2.5.4.5 and as shown in the polar charts displayed in previous sections, when the model is trained with features that are direction-independent, it is able to capture the variations in mud weights required for deviated wells. This is possible even if the training data contains only vertical wells.

When examining the trends in both **Figure 149** and **Figure 150**, it can be seen again that as the wellbore size increases, the amount of deformation and risk-controlled stability mud weight also increase. When comparing the mud weights for the vertical wells in **Figure 147** and **Figure 148** to the horizontal wells in **Figure 149** and **Figure 150**, both models estimate higher mud weights and more deformations in horizontal wells. This is the expected trend for a wellbore in a strike-slip stress regime. When comparing the magnitude of risk-controlled mud weights in **Figure 149** and **Figure 150**, it is again clear that the integrated model generally predicts higher mud weights. It also shows that it is not possible to reduce the deformation to less than 0.5 ft^3/12ft for all the examined wellbore sizes. The Lade criterion determined deformations are again more optimistic as it shows that it is possible to completely prevent deformations for all the examined wellbore sizes.

4.5 Other Potential Applications of the Integrated Model

To conclude the field applications of the model in this chapter, applications for which no sufficient dataset is made available for this work are discussed. This is because

while the modelling scheme proposed by the integrated model is a physics-based data science scheme, which means it does not require large datasets for training, it still requires a representative, sufficiently diverse, and descriptive dataset. Such applications include cavings shape prediction and predicting the sanding flow rate.

4.5.1 Cavings Shape Prediction and Matching

Since the integrated model allows for accurate prediction and definition of failure zones around the wellbore no matter how irregularly shaped they are, this can be exploited to anticipate the shapes of cavings that will be seen on the shale shakers. Considering that the different shapes of cavings are correlated to different modes of rock failure as detailed by Skea et al. (2018), anticipating these shapes through the integrated model and matching them to shale shakers observations can help improve the model accuracy. Much like history matching in reservoir modelling, matching 3D visualizations of cavings with those observed in shale shakers can be used to identify inaccuracies in the model to adjust it while drilling. The integrated model when trained with both load-dependent and load-independent features can define the caving shapes to a high degree of accuracy. Shale shakers surveillance reports and cavings images from offset wells can be used for the purpose of training the integrated model on cavings shapes prediction. Then, in a while-drilling scenario, a video feed from a camera pointed at the shale shakers can be utilized to first identify cavings and characterize their structure through a separately trained convolutional neural network and principle-component analysis (CNN-PCA) in a similar fashion to the work performed by Du et al. (2020). The cavings identified from this feed are then to be used to either confirm the cavings shape which was predicted by the

230

integrated model, or to find the reason for the discrepancy and adjust the model to match the cavings shape while drilling.

4.5.2 Modelling for Sand Flow Rate

The integrated model, when trained properly, can work to map shear failure zones around a borehole or a perforation. This type of analysis is applicable in predicting sand production as the flown sand particles can be a result of shear failure. Therefore, by training the integrated model using sand flow rates or preferably lab-based images of core samples experiencing sanding, the integrated model have the potential to, along with a proper bridging criterion, predict sand flow rate with a high degree of accuracy.

CHAPTER V

SUMMARY AND RECOMMENDATIONS

5.1 Model Summary

The model developed for this work seeks to achieve two main objectives. The first objective is to overcome prevalent simplifying assumptions in the area of drilling geomechanics modelling; and the second objective is to exploit recent technological advances to improve rock failure prediction. Concerning the first objective, the first assumption overcome by the model is the assumption of linear elasticity. The model replaces this assumption by implementing the plastic flow rule and strain hardening within the finite element model formulation. The second assumption is uniform wellbore shape, which is cancelled through the use of image logs analysis and re-meshing functions that can reflect the actual shape of the wellbore on the FEM mesh. The third assumption is the applicability of generalized and conventional failure criteria in predicting rock failure spatially around the wellbore. This is replaced by machine learning algorithms that are trained based on the outcome of image logs analysis and the output of the FEM. The final assumption is that the limits of the drilling window are static parameters. The model eliminates this assumption by the means of employing image logs analysis and re-meshing the wellbore accordingly to reproduce the limits of the drilling window, which can essentially reflect the effect of different drilling events as they occur on the wellbore rock strength. As for the second objective, the developed model works to take advantage of technological advances, which provide new forms of data that are not conventionally

utilized for drilling engineering applications. An example of these is LWD tools that can provide real-time density-based images of the wellbore, which can accurately describe the state of the wellbore and be easily reflected on the model mesh.

The new modelling scheme proposed in this work integrates a physics-based model, which is the geomechanics FEM, and a data science model, which is the trained machine learning algorithm. This scheme ensures that training can be performed using datasets that are smaller than those needed for purely data driven models. It also ensures that the predictions made by the integrated model are not merely predictions of failure, but can also provide a better understanding of the mechanisms leading to rock failure in specific scenarios. The training process of the machine learning algorithms can be engineered through features selection to achieve different goals and make different predictions such as the determination wellbore stability mud weights and the prediction of cavings shapes and volumes.

Comparing the performance of the trained and integrated FEM-ML model against the performance of conventional failure criteria such as the Mohr-Coulomb, Mogi, and Lade criterion by modelling true-triaxial tests of sandstone cores shows that the integrated model more than doubles the accuracy of rock failure prediction. To achieve this comparison, both the integrated model and the conventional failure criteria are evaluated as binary classification algorithms using the confusion matrix parameters. The integrated model also showed superior performance when trained based on field data. However, the improvement observed from the model trained based on field data was not as high as the improvement produced by the model trained based on lab experiments. The main reason

for this discrepancy is the quality of the field dataset available. The experiments dataset contained high resolution CT-scans to describe the rock failure; on the other hand, the field dataset contained low resolution resistivity logs from historical offset wells. As demonstrated from modelling the experimental work, the integrated model will provide better performance when higher quality well image logs are available.

In terms of field implementation, the integrated model allows for going beyond the standard practice of performing drilling geomechanics analysis in the pre-drilling phase only. The structure of the integrated model allows for it to be implemented in real-time, where it can respond to different drilling events as they occur. The real-time implementation produces updated limits to the drilling window, which reflect the effect of the different drilling events. The field implementation of the model as demonstrated using historical offset wells data show that the size of the drilling window can shrink considerably to where it can cause significant drilling difficulties. As shown in the example well used for demonstrating the real-time application of the integrated model, these difficulties can include wellbore instabilities, pack off, tight hole, fluid losses, excessive cuttings and fill, stuck pipe incidents, and lose of tools in hole.

5.2 Recommendations for Future Work

Considering that the work performed here proposes a new modelling scheme where a physics model is integrated with a trained machine learning algorithm, this scheme can be further improved by improving or replacing the physics model. In this work, the physics model used is an elasto-plastic finite element geomechanics model. This

is sufficient for determining the strength limits and predicting rock failure around the wellbore. However, other physics models can be employed in its place to perform a more complex analysis and produce different predictions. For example, in cases where the time dependency of wellbore failure needs to be examined or the effects of drawdown and injection need to be considered, the elasto-plastic finite element geomechanics model can be replaced by a coupled model that introduces the time variable and the mechanical effects of porous flow through the diffusivity equation in the finite element formulation. In this example, the coupled geomechanics and fluid flow in porous medium FEM can be integrated with a machine learning algorithm by introducing new features to the training process, which will mostly describe the time dependency of rock failure and the injection or production flow rates. Another example is the case where it is needed to discern the effects of the rheological properties of the drilling fluid on rock failure. In these cases, the elasto-plastic finite element geomechanics model can be replaced by a geomechanics model coupled with a wellbore hydraulics model, which is similar to what is described by Li et al. (2020), where a commercial software is used. In this example, the integration can be performed similarly between the geomechanics-hydraulics coupled model and the machine learning algorithm.

REFERENCES

Al-Ajmi A. M. and Zimmerman R. W. 2005. Relationship between the parameters of the Mogi and Coulomb failure criterion. *Int J Rock Mech Min Sci* **42**(3):431–9.

AlBahrani, H. and Al-Yami, A. 2018. Drillstring Vibrations and Wellbore Quality: Where Drillstring Design Meets Geomechanics. Paper SPE-193253-MS presented at the Abu Dhabi International Petroleum Exhibition & Conference, Abu Dhabi, UAE, 12-15 November. https://doi.org/10.2118/193253-MS

AlBahrani, H. and Morita, N. 2020. Understanding the Concurrence of Drilling Induced Fractures and Breakouts and Its Implications on Drilling Performance Using a 3D FEM and Field Imaging Data. Paper IPTC-20171-MS presented at the International Petroleum Technology Conference, Dhahran, Kingdom of Saudi Arabia, 13-15 January. https://doi.org/10.2523/IPTC-20171-MS

Bagala, S., McWilliam, I. A. G., Rourke, T. et al. 2010. Real -Time Geomechanics: Applications to Deepwater Drilling. Paper SPE-137071-MS presented at the SPE Deepwater Drilling and Completions Conference, Galveston, Texas, USA, 5-6 October. https://doi.org/10.2118/137071-MS

Barton, C. A., Moos, D., Peska, P. et al. 1997. Utilizing Wellbore Image Data To Determine The Complete Stress Tensor: Application To Permeability Anisotropy And Wellbore Stability. *Society of Petrophysicists and Well-Log Analysts* **38** (6): 21-33.

Cecílio, D. L., Devloo, P. R. B., Gomes, S. M. et al. 2019. An adaptive finite element model for wellbore breakout simulation. *Int J Numer Anal Methods Geomech*. **43** (4): 814– 828. https://doi.org/10.1002/nag.2895

Chen, X., Tan, C. P., and Detournay, C. 2001. Wellbore Behavior in Fractured Rock Masses. Paper presented at the 38th U.S. Symposium on Rock Mechanics (USRMS), Washington, D.C., 7-10 July.

Chen, X., Tan, C. P., and Detournay, C. 2002. The Impact of Mud Infiltration on Wellbore Stability in Fractured Rock Masses. Paper SPE-78241-MS presented at the SPE/ISRM Rock Mechanics Conference, Irving, Texas, 20-23 October. https://doi.org/10.2118/78241-MS

Chen, X., Tan, C. P., and Detournay, C. 2003. A Study on Wellbore Stability in Fractured Rock Masses with Impact of Mud Infiltration. *Journal of Petroleum Science and Engineering*. **38** (3–4): 145-154. https://doi.org/10.1016/S0920-4105(03)00028-7

Cheng, W., Jiang, G., Zhou, Z. et al. 2019. *Geotech Geol Eng.* **37** (Jan.): 2873–2881. https://doi.org/10.1007/s10706-019-00802-7

Daneshy, A. A. 1973. A Study of Inclined Hydraulic Fractures. *Society of Petroleum Engineers Journal* **13** (2): 61-68. SPE-4062-PA. http://dx.doi.org/10.2118/4062-PA.

Drucker D. C. and Prager W. 1952. Soil Mechanics and Plastic Analysis or Limit Design. *Quart Appl Math* **10** (1):157–65.

Du, X., Jin, Y., Wu, X. et al. 2020. Classifying Cutting Volume at Shale Shakers in Real-Time Via Video Streaming Using Deep-Learning Techniques. *SPE Drilling & Completion* Preprint. SPE-194084-PA. https://doi.org/10.2118/194084-PA

Feng, Y., Arlanoglu, C., Podnos, E. et al. 2015. Finite-Element Studies of Hoop-Stress Enhancement for Wellbore Strengthening. *SPE Drilling & Completion* **30** (1): 38-51. SPE-168001-PA. https://doi.org/10.2118/168001-PA

Furui, K., Fuh, G.-F., Abdelmalek, N. A. et al. 2010. A Comprehensive Modeling Analysis of Borehole Stability and Production-Liner Deformation for Inclined/Horizontal Wells Completed in a Highly Compacting Chalk Formation. *SPE Drilling & Completion* **25** (4): 530-543. SPE-123651-PA. https://doi.org/10.2118/123651-PA

Gaede, O., Karpfinger, F., Jocker, J., Prioul, R., Comparison between analytical and 3D finite element solutions for borehole stresses in anisotropic elastic rock, International Journal of Rock Mechanics and Mining Sciences, Volume 51, 2012, Pages 53-63, ISSN 1365-1609, https://doi.org/10.1016/j.ijrmms.2011.12.010.

Hamid, O., Hall-Thompson, B., Omair, A. et al. 2016. Application of Real-Time Geomechanics on a Horizontal Well. Paper SPE-183125-MS presented at the Abu Dhabi International Petroleum Exhibition & Conference, Abu Dhabi, UAE, 7-10 November. https://doi.org/10.2118/183125-MS

Huang, C., B. Akbari, and S. Chen. 2018. Quick approximate elastoplastic solutions of wellbore stability problems based on numerical simulation and statistical analysis. *J. Nat. Gas Sci. Eng.* **51** (Mar): 147–154. https://doi.org/10.1016/j.jngse.2018.01.005.

Huang, C. 2016. A Numerical Investigation of Wellbore Stability Problems Using an Elastoplastic Model. LSU Master's Thesis.

Jaeger, J.C., Cook, N.G.W. and Zimmerman, R. W. 2007. Deformation and Failure of Rock. In *Fundamentals of Rock Mechanics*, fourth edition, Chap. 4, 102. London: Blackwell Publishing.

James, G., Witten, D., Hastie, T., Tibshirani, R., 2017. *An Introduction to Statistical Learning*. Springer

Kanfar, M. F., Chen Z., and Rahman, S. S. 2015. Effect of Material Anisotropy On Time-Dependent Wellbore Stability, *International Journal of Rock Mechanics and Mining Sciences*, **78** (Sep.): 36-45. https://doi.org/10.1016/j.ijrmms.2015.04.024.

Kanfar, M. F., Chen, Z., and Rahman, S. S. 2015. Risk-Controlled Wellbore Stability Analysis in Anisotropic Formations, *Journal of Petroleum Science and Engineering*, **134** (Oct.): 214-222. https://doi.org/10.1016/j.petrol.2015.08.004

Karpatne, A. Atluri, G., Faghmouset G. H. et al. 2017. Theory-guided data science: a new paradigm for scientific discovery from data. *IEEE Trans. Knowl. Data Eng.* **29** (10): 2318–2331. https://doi.org/10.1109/TKDE.2017.2720168

Kumar, R. B., Vardhan, H., Govindaraj, M. et al. 2013. Regression analysis and ANN models to predict rock properties from sound levels produced during drilling. *International Journal of Rock Mechanics and Mining Sciences*, **58** (Feb.): 61-72. https://doi.org/10.1016/j.ijrmms.2012.10.002

Li, X., El Mohtar, C. S., and Gray, K. E. 2020. Modeling Progressive Wellbore Breakouts with Dynamic Wellbore Hydraulics. *Society of Petroleum Engineers Journal*. **25** (10): 541-557. SPE-199887-PA. https://doi.org/10.2118/199887-PA

Ma, T., Peng, N., Tang, T. et al. 2018. Wellbore Stability Analysis by Using a Risk-Controlled Method. Paper ARMA-2018-329 presented at the 52nd U.S. Rock Mechanics/Geomechanics Symposium, Seattle, Washington, 17-20 June.

Magini, L., Molteni, D., Zausa, F. et al. 2012. Collaborating on Real-Time Geomechanics across Organizational Boundaries. Paper SPE-151204-MS presented at the IADC/SPE Drilling Conference and Exhibition, San Diego, California, USA, 6-8 March. https://doi.org/10.2118/151204-MS

Meng, F. F., Fuh, G.-F., and Spencer, J. A. 2010. Managing Geomechanics Principles for Proper Handling of Wellbore-Stability Problems in Shale and Depletion Effects on Formations While Drilling. Paper SPE-135057-MS presented at the SPE Annual Technical Conference and Exhibition, Florence, Italy, 19-22 September. https://doi.org/10.2118/135057-MS

Mogi K. 1971. Fracture and Flow of Rocks Under High Triaxial Compression. *J Geophys Res* **76**(5):1255–69.

Morita, N. 2004. Well Orientation Effect on Borehole Stability. Paper SPE-89896-MS presented at the SPE Annual Technical Conference and Exhibition, Houston, Texas, 26-29 September. https://doi.org/10.2118/89896-MS

Morita, N. and Gray, K. E. 1980. A Constitutive Equation for Nonlinear Stress-Strain Curves in Rocks and Its Application to Stress Analysis Around A Borehole During Drilling. Paper SPE-9328 presented at the SPE Annual Technical Conference and Exhibition, Dallas, 21-24 September. http://dx.doi.org/10.2118/9328-MS

Morita, N. and Nagano, Y. 2016. Comparative Analysis of Various Failure Theories for Borehole Breakout Predictions. Paper SPE-181395-MS presented at the SPE Annual Technical Conference and Exhibition, Dubai, UAE, 26-28 September. https://doi.org/10.2118/181395-MS

Morita, N. and Ross, C. K. 1993. Core-Based Horizontal or Highly Inclined Well Stability Analysis for Unconsolidated Formations. Paper SPE-26332-MS presented at the SPE Annual Technical Conference and Exhibition, Houston, Texas, 3-6 October. https://doi.org/10.2118/26332-MS

Mylnikov, D., Tatur, O., Sabirov, A. et al. 2018. Integrated Approach to Real-Time Drilling Optimization Based on Methods of Geosteering, Petrophysics and Geomechanics. Paper presented at the SPE Russian Petroleum Technology Conference, Moscow, Russia, 15-17 October. https://doi.org/10.2118/191526-18RPTC-MS

Owen, D.R.J. and Hinton, E. 1980. *Finite Elements in Plasticity: Theory and Practice*. Swansea, United Kingdom: Pineridge Press Ltd.

Papamichos, E., Walle, L. E., Berntsen, A. N. et al. 2019. Sand Mass Production in True Triaxial Borehole Tests. Paper ARMA-2019-0238 presented at the 53rd U.S. Rock Mechanics/Geomechanics Symposium, New York City, New York, 23-26 June.

Phan, D. T., Liu, C., AlTammar, M. J. et al. 2020. Application of Artificial Intelligence to Predict Time-Dependent Safe Mud Weight Windows for Inclined Wellbores. Paper IPTC- presented at the International Petroleum Technology Conference, Dhahran, Kingdom of Saudi Arabia, 13-15 January. https://doi.org/10.2523/IPTC-19900-MS

Plumb, R. A. and S. H. Hickman. 1985. Stress-induced Borehole Enlargement: A Comparison between the Four-Arm Dipmeter and the Borehole Televiewer in the Auburn Geothermal Well. *J. Geophys. Res.* **90**: 5513-5521. https://doi.org/10.1029/jb090ib07p05513

Richardson, R. M. 1983. Hydraulic Fracture in Arbitrary Oriented Boreholes: An Analytic Approach. In *Hydraulic Fracturing Stress Measurements*, Chap. 3, 167-175. Washington, D.C.: National Academy Press.

Salehi, S. and Nygaard, R. 2010. Finite-element Analysis of Deliberately Increasing the Wellbore Fracture Gradient. Paper ARMA- presented at the 44th U.S. Rock Mechanics Symposium and 5th U.S.-Canada Rock Mechanics Symposium, Salt Lake City, Utah, 27-30 June.

Skea, C., Rezagholilou, A., Behnoud Far, P. et al. 2018. An approach for wellbore failure analysis using rock cavings and image processing. *Journal of Rock Mechanics and Geotechnical Engineering*. **10** (5): 865-878.

Soroush, H., Rasouli, V., and Tokhmechi B. 2018. A combined Bayesian–wavelet–data fusion approach for borehole enlargement identification in carbonates. *International Journal of Rock Mechanics and Mining Sciences*. **47** (6): 996-1005. https://doi.org/10.1016/j.ijrmms.2010.06.015

Watson, D., Brittenham, T. and Moore, P. 2003. *Advanced Well Control*. **10**, 98–112. Richardson, Texas: Textbook Series, SPE.

Wessling, S., Dahl, T., and Pantic, D. M. 2011. Challenges and Solutions for Automated Wellbore Status Monitoring - Breakout Detection as an Example. Paper SPE-143647-MS presented at the SPE Digital Energy Conference and Exhibition, The Woodlands, Texas, USA, 19-21 April. https://doi.org/10.2118/143647-MS

Zervos, A., Papanastasiou, P., and Cook, J. 1998. Elastoplastic Finite Element Analysis of Inclined Wellbores. Paper SPE-47322-MS presented at the SPE/ISRM Rock Mechanics in Petroleum Engineering, Trondheim, Norway, 8-10 July. https://doi.org/10.2118/47322-MS

Zhang, H., Yin, S. and Aadnoy, B. S. 2018. Finite-Element Modeling of Borehole Breakouts for In Situ Stress Determination. *International Journal of Geomechanics*. **18** (12): 04018174-1:15. https://doi.org/10.1061/(ASCE)GM.1943-5622.0001322

Zhang, X., Jeffrey, R. G., and Bunger, A. P. 2011. Hydraulic Fracture Growth from a Non-Circular Wellbore. Paper ARMA- 11-333 presented at the 45th U.S. Rock Mechanics / Geomechanics Symposium, San Francisco, California, 26-29 June.

Zienkiewicz, O.C., 1977. *The Finite Element Method*, pp 169-171. McGraw-Hill, London.

Zoback, M.D., 2007. *Reservoir Geomechanics*, Cambridge University Press.