DATA DRIVEN CASING FAILURE MITIGATION DURING DRILLING AND COMPLETION, OPERATIONS USING A NEW MACHINE LEARNING WORKFLOW ON DATA SET FROM THE GRANITE WASH FORMATION

A Dissertation

by

CHRISTINE IKRAM NOSHI

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,	Dr. Mahmood Amani
Committee Members,	Dr. Albertus Retnanto
	Dr. Hadi Nasrabadi
	Dr. Mahmoud El-Halwagi
Head of Department,	Dr. Jeff Spath

Spring 2022

Major Subject: Petroleum Engineering

ABSTRACT

This study seeks to minimize the likelihood of casing failure using data analysis and machine learning algorithms. The study focuses on the design of a data-driven workflow that can address several challenging aspects that pertain to casing failure including: (1) identification of potential risk factors amongst different exposures through the adoption of risk analysis techniques (case-control study design), (2) evaluation of the type and magnitude of the impact for each risk factor, determined through the application of different association measurements, (3) identification of the levels within each potential risk factor that impose the highest risk on casing failure, (4) acknowledgement of the depths most susceptible to casing failure implemented through non-parametric and semi-parametric survival analysis techniques, (5) prediction of the overall probability of casing failure given the information for pre-defined risk factors via the application of multiple classification learning algorithms , and finally (6) have a scheme for mitigating casing failure. To account for the rigidity of machine learning algorithms and to allow for more control from the user's end, risk assessment techniques were implemented, particularly, semi-quantitative probability-impact risk assessment matrices (PI-RAMs). These matrices were based on results obtained from frequency and survival analysis. PI-RAMs were used as feedback means to the initial predictions obtained from conventional ML algorithms to not only provide a better intuition of the overall risk, but also the contribution of each risk factor.

One major limitation to the proposed framework is the compliance of field engineers with the recommendations, due to cost-related concerns. This motivated the focus on yet another perspective of mitigating failure, that is the enhancement of casing integrity evaluation techniques, specifically fatigue failure induced from thermal stress. A steam injection simulated case has been used for the design and validation of the proposed data-driven model. A correct estimation of fatigue life of different casing parts must be made. The accuracy of fatigue life estimation is contingent on

the accuracy of local strain estimations. Two classes of factors impact local strains: 1) active/direct (temperature changes, casing material, etc.) and (2) passive/indirect factors (such as cement cracks or leaks). Although physics-based models can account for direct factors, they still fail to account for indirect factors, leading to false conclusions on casing fatigue life and abusive consumption of casing parts beyond their capabilities.

The proposed data-driven estimator takes as input all direct and indirect factors, and outputs the corresponding local strains that reflect those effects. Then, using the casing material properties, along with estimated strains as input for Manson's Eq. and estimating the fatigue life of those casing parts. Based on estimated fatigue life, the model can give recommendations on changing casing parts that are abused throughout any process (such as steam injection, or hydraulic fracturing). This would, ultimately, prevent or reduce the chances of the occurrence of casing failure. Using the proposed model, engineers can have a better understanding of the casing durability and ability to withstand the downhole conditions and practices.

DEDICATION

To my dad, Ikram, whom I owe everything I have become to, thank you for all your sacrifices, late nights waiting for me, money you have spent, effort you have put in, worrying about me and trying to encourage me along the way. This is for you. I love you.

To my mom and best friend in heaven, Nadia. I miss you so much. Thank you for encouraging me and always being *there* for me. I finally did it.

To my fiancé and husband to be, Patrick. Thank you for being there for me, for helping and encouraging me all through *thin* and *thin*. You are an *amazing* person. I love you more than words could tell.

ACKNOWLEDGEMENTS

The completion of this study could not have been possible without the help, support, and encouragement of Dr. Mahmood Amani, our beloved graduate advisor. I would like to also extend my gratitude to Dr. Alberus Retnanto and Dr. Hadi Nasrabadi for sitting on my panel and taking the time to read my dissertation.

A debt of gratitude goes to Dr. Mahmoud El Halwagi for his constant support, encouragement, and advice throughout *tough* times. You have been a great mentor; without you none of this would indeed be possible.

I would also like to extend my sincere gratitude to Dr. Maria Barrufet for taking the time to listen to my concerns and helping make this a reality.

I would like to thank my awesome boss, Rachel Shaffer for encouraging me to complete my studies and for being so patient throughout the process. I owe you.

ACRONYMS

AI	Artificial Intelligence
ANN	Artificial Neural Network
BBN	Bayesian Belief Network
BHT	Bottom Hole Temperature
DL	Dogleg
DLS	Dogleg Severity
EDA	Exploratory Data Analysis
e_v	External Non-Uniform Pressure
ft	Feet
GA	Genetic Algorithm
IQR	Interquartile Range
KNN	K-Nearest Neighbors
LK	Likelihood
MD	Measured Depth
ML	Machine Learning
NPV	Net Present Value
N	Sample Size
NPT	Non-Productive Time
OLAP	Online Analytical Processing Tools
OD	Outer Diameter
POOH	Pull out of Hole
PCA	Principal Component Analysis
PGM	Probabilistic Graphical Model
R^2	R Square

RBF	Radial Basis Function Neural Network
RIH	Run in Hole
ROC	Receiving Operating Characteristic Curve
ROP	Rate of Penetration
RPM	Rotation per Minute
SVR	Support Vector Regression
SVM	Support Vector Machine
TOC	Top of Cement
TVD	True Vertical Depth
WOB	Weight on Bit

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

LIST OF TABLES

Page

CHAPTER I
INTRODUCTION AND LITERATURE REVIEW

Reader Guide: In this chapter, the discussion starts with an introduction of the problem of casing failure which would be the focus of this study (Section 1.1). In this section, I briefly discuss: (1) the functionality of casing strings, (2) the importance of proper casing design, (3) downhole challenges and severe conditions imposed on casing strings, (4) criticality of casing failure, (5) previous contributions in identifying primary causes of casing damage, as well as, failure modes, (6) previous contributions in mitigating, or ultimately, avoiding casing failure including physics-based and data-driven approaches, and (7) the limitations of previous solutions and implemented techniques. Next, the motivation and main objectives of the work implemented and presented in this dissertation are highlighted (Section 1.2). Then, the framework/tool proposed for tackling the challenge of mitigating casing failure is presented (Section 1.3). In this section, I introduce the major contributions of this study along with an overview of my already published work that is related to the content of this dissertation. Following that, the outline of this dissertation is laid out (Section 1.4). In this section, I highlight the key steps taken in the following chapters in order to achieve the objectives of this study. Finally, a description of the data set compiled for the study is provided, along with the field from which the data was collected, as well as targeted formations (Section 1.5).

1.1 Problem Statement

Casing strings are an indispensable component in the design of any well and serve numerous purposes in oil and gas wells including supporting the oil and gas production, preventing loose formation from interfering into the borehole activities and, ultimately, ensuring normal operations during drilling and completion processes (Hua Tong et al. 2016). They also constitute a significant portion of the total well cost. The cost of casing pipes can constitute up to 20-30% of the total cost of the well (Alade 2018).

Safe and economical sizing of casing is, therefore, an important task in oil and gas well design (Junior et al. 2015). If casings are not sized correctly, then such casings are prone to failure at their early stage of operation. Design of casing string calls for knowledge of the operating conditions imposed on the casing as well as the concepts related to pipe properties (Bowers 1955; Halal et al. 1996; Akpan 2005; Edaigbini 2015; Thattil 2017).

Throughout the life of the well, several challenges and severe conditions are posed on casing strings owing to the existing and induced downhole stresses (Fleckenstein et al. 2001; Li and Samuel 2016; Feng and Gary 2017). Variations in in-situ stress associated to well location, reservoir characteristics, diagenesis, and local geo-stress distribution, all contribute to existing downhole stresses. Conversely, drilling and completion induced stresses are instigated from, well stimulation operations, well configurations, and production related stresses. These added stresses undermine casing integrity and put it at risk of casing failure.

Casing failure has always been the thorny problem in the process of oil field development (Dai et al. 2018). Over the past four decades, an escalating trend of casing failures have been observed; indicating that the current casing design for unconventional and horizontal wells might not be adequate. This has resulted in a higher-than-expected failure rate. According to Davies et al. 2014, recent worldwide statistics from countries such as Canada, China, Netherlands, Norway, United Kingdom, and the US on conventional and unconventional wells show that roughly 26,600 wells out of 380,000 wells have had at least one type or another of integrity failure. In addition, out of 14,297 wells in the US Gulf of Mexico, a whopping 45% of tubular failures were recorded during the 1980s. According to the Journal of Petroleum Technology (JPT), information collected from client operators shows that in certain US shale and tight oil fields, between 20%–30% of horizontal wells are integrally compromised to some degree due to hydraulic fracturing operations (Jacobs

2020). In their study, Salehi et al. 2009 reported that: in a carbonate oilfield in Iran, forty-eight casing collapses were noticed due corrosion, reservoir compaction, and geo-mechanical effects. They added that the incidents continued to escalate over the years. Another study by Xi et al. 2018 stated that, out of 101 wells drilled in Weiyuan shale play, a failure rate of 34% was noted.

There is a multitude of factors attributing to different casing failure modes and mechanisms. Local buckling along the casing wall, not extending to its center, was described as a failure mode by Wang et al. 2014. Whereas, in the case of columnar buckling, the center of the casing is completely bent and deformed. Yin et al. 2018a, b, c indicated that casing failure is induced during volume fracturing, formation deformation, and the consequent rock slippage. This is especially true for formations that are renowned for causing wellbore stability problems, such as shale. Additionally, both hydraulic and natural fracture slips tend to buckle the casing in the lateral section of shale gas horizontal wells during fracturing operations. Chipperfield et al. 2007; Zeng and Yao 2015; Hou et al. 2016, argued that huge displacement volumes with numerous stimulation stages along with high pumping pressures create a complex stress field leading to eventual deformation, leap, and ultimate shearing of casing pipes.

Factors such as rock strata displacement along bedding plane or sharply inclined faults were known to cause shear failure (Wang et al. 2011). Casing shear failure due to fracture slip during fracturing operations was reported by Yin et al. 2018a, b, c. Alterations in stress and pressure triggered by heating and injection was found to induce casing shear owing to formation shear (Dusseault et al. 2004). Han et al. 2006a, b reported that the presence of thick fragile shale lithology coupled with excessive vertical heterogeneity were amongst the underlying causes of casing shear failure.

Vudovich et al. 1988 summarized that casing failure types are interconnected with casing failing in one or more of the failure types, namely, collapse or burst, which is credited to radial stresses.

Casing collapse is principally classified into elastic, plastic, transitional, and yield. Differentiating these mechanisms are based on the slenderness ratio which is a ratio of casing diameter to its thickness, according to industry standards. Wang et al. 2014 suggested that collapse failure is an outcome of various mechanical loading from the casing itself, sand, and cement. Kiran et al. 2017 also noted that existence of voids and cement channels at the casing-cement interface could cause up to 60% drop in casing collapse resistance. Huang and Gao 2015 attributed the mechanism to imbalanced external loads surpassing the casing yield strength which alter its circular alignment to an oval configuration. Bastola et al. 2014, concluded that the increase in initial ovality would ultimately decrease the pipe's resistance causing collapse. Abdideh and Khah 2018 reported the phenomena of collapse as an abnormal formation displacement on the casing string leading to its ultimate collapse.

According to Wilson 2018, metal loss in casing strings, due to corrosion, can cause potential leakage. An average-metal-loss value above seventy percent leads to the probability of casing failure reaches unity and zero when it is under thirty percent. Zhang et al. 2016 confirmed that, during directional drilling, casing wear still lingers to be a leading problem as it causes degradation in casing strength, casing deformation, potentially casing failure. The characteristics of surface wear can be categorized into surface fatigue wear, abrasive wear, adhesive wear, and corrosive wear (Andersson, 2011). Casing wear instigated by drill string rotation may be classified as typical abrasive and adhesive wear (Best 1986). During extended reach drilling, Gao et al. 2010 and Shen and Beck 2014, noticed that casing wear occurs as a direct result from the contact between the stationary casing and the rotating drill string. Mao et al. 2018 indicated that the rate of wear differed with increasing drill pipe rotational speeds, non-linearly. The nonlinear behavior was attributed to the combination of corrosive, abrasive wear, and erosive wear.

According to Barton, 2003, several reports of critical elbow failures due to erosion have appeared on drilling units, production platforms, and different subsea equipment in the past. A study by Bai and Bai 2018 reported that, sand erosion due to sand particulates is the most widespread cause of erosion challenges in casing systems. Computational fluid dynamics (CFD) simulations showed that pipe erosion is more rampant at 45° angle in elbows (Ogunsesan et al. 2019).

According to Lin et al. 2016, corrosion develops pits and cracks on both the outer and inner walls of the casing, when subjected to corrosive conditions. Burst and collapse loads acting on the corroded casing will trigger further stress concentration and ultimately reduce the casing strength. Strength degradation can substantially diminish casing life, and even cause well failure. Yuan et al. 2012 analyzed five–eight years of operational casing using FEA. The authors advised using higher strength casing such as P110 and T-95 for production casing.

Bai and Bai 2018 recognized that fatigue is an irreversible, gradual, and localized structural damage that arises and continuously accumulates when a material is structurally exposed to cyclic loading (Gao and Hsu, 1998). Those cyclic could either be repeated, fully reversed, or fluctuating loads with either low or high cycles (Liu et al., 2015; Cirimello et al. 2018; Chen et al. 2018a, b). Lim et al. 2012 described that, during drilling in stormy weather in offshore operations; currents and wave, and heave motions of the sea are transported down the riser to the wellhead, to the casing structure causing fatigue failures at essential casing joints and welds. Furthermore, casing pipe fatigue failure could happen either during production, due to alternating temperatures and cyclic loading of pressure or during stimulation, due to temperature differentials between the reservoir fluids and stimulation fluids (Kiran et al. 2018). Liu et al. 2018 adds that casing joint fatigue has been reported during multistage hydraulic fracturing with varying temperature between stages.

Three categories of threading connections are specified by API, long and short rounded threaded couplings, buttress with asymmetrical trapezoidal thread couplings, and extreme-line connection thread without couplings. However, many well integrity challenges surround casing connections and seals, especially, in shale gas horizontal, HPHT, and deep-water wells. In order to reduce/remove voids existing in the cement slurry, operators rotate the casing in the lateral section during cementing jobs. However, this rotation puts the casing connection under excessive rotating-bending loads causing connection failures (Hamilton et al. 2019). Conversely, Jellison et al. 1998 suggested that local buckling and shear failure were the main culprits behind connection failure.

The aftermaths of casing failure are catastrophic and can include the possibility of blowouts, damage of adjacent wells, decreased production, environmental pollution, injuries or fatalities, loss of assets, negative financial implications pertinent to the cost of material as well as the non-productive time associated with repair (Payne et al. 1993). The aforementioned impacts along with today's focus on protection of the environment and the ambitions of the industry to operate safely in environmentally sensitive areas have urged the need for improved cross-disciplinary tools and work processes within the area of casing failure and risk management (Liu et al. 2012).

To date, various techniques have been implemented to avoid or mitigate casing failures. Most of the attempts were physics-based solutions that tackled one or few aspects on a small scale. The solutions were either analytical (Li and Samuel 2016; Lin et al. 2016; Cirimello et al. 2017), experimental (Han et al. 2006a, b; Ferla et al. 2009; Jinan et al., 2012; Zhang et al. 2012; Shen et al. 2014), or numerical in nature (Fleckenstein et al. 2001; Han et al. 2006a, b; Yuan et al. 2012; Lavrov et al. 2015; Hu et al. 2016; Feng and Gary 2017; Patel et al. 2018; Ogunsesan et al. 2019). However, former studies – based on conventional physics-based approaches – were found to

perform unreliably and did not attain wide-scale execution (Maharaj 1996; Dall'Acqua et al. 2013). This is mainly due to the dynamic complexity of the underlying problem.

Recently, with the advancement of data science and machine learning algorithms, attention has been geared towards the integration of data-driven solution in numerous oil and gas applications including reservoir characterization (Anifowose 2013; Akande et al. 2015; Hoeink and Zambrano 2017; Tiwari et al. 2018; Nanjo and Tanaka 2019), production optimization (Cao et al. 2016; Ounsakul et al. 2019), drilling (Zhao et al. 2017; Pollock et al. 2018; Yang et al. 2019) and well completion (Prislin et al. 2017; Ounsakul et al. 2019). Those efforts have been successful in resolving various challenging aspects that were, otherwise, unachievable using conventional physics-based approaches.

Yet, contributions in the area of casing failure have been minimal. To date, few attempts were published in the topic of predicting casing failure by (Noshi et al. 2018a, b; Noshi et al. 2019; Song and Zhou 2019; Tang 2019; Tan et al. 2020). The authors attempted to implement classical classification models to predict the probability of casing failure based on various risk factors deduced from historical cases. In addition, the authors managed to identify some features, or risk factors, that had the greatest impact on the occurrence of casing damage. This leaves drillers and drilling engineers with very little information on mitigating the risks that the casing is exposed to during drilling, completions, and production operations.

Implementation of data-driven solutions, nevertheless, has been a real challenge especially in the area of casing failure, in terms of, gathering enough failure data which has tied the hands of the industry into finding a proper solution. The lack of enough data has been due to a multitude of reasons, such as, reputation concerns, legal issues, competitors, partners, investors, avoiding legal actions and respecting customer privacy. This, in turn, has led the cases of casing failure to be

overlooked, under-reported or handled rather reactively instead of proactively. This urges the need to develop a common working platform for drilling engineers and the other disciplines involved in well planning, improving cross communication between different disciplines.

The study introduces the building blocks of this methodology through the development of a data-driven "casing failure mitigation" tool that has the capability of: (1) quantification of impact type and magnitude for potential risk factors, (2) identification of the levels within each potential risk factor that impose the highest risk on casing failure, (3) acknowledgement of the depths that are susceptible to casing failure, (4) adjustment of high risk casing design specifications in the direction of reducing the overall probability of casing damage and, ultimately, mitigating casing failure.

1.2 Dissertation Objectives and Related Published Work

In this dissertation, the primary objective is to develop a framework that can explore, analyze, and tackle several challenging aspects that pertain to the problem of casing failure in an attempt to better understand, also, mitigate the occurrence of that problem in the future. Motivated by the importance of the challenging problem of casing failure, the lack of a solid physical foundation capable of explaining, in full, the reasoning behind such a damaging incident and the deficiency of the know-how needed for its mitigation, a data-driven approach was adopted to be the cornerstone of the proposed workflow.

The proposed data-driven workflow will have the capability of: (1) quantification of impact type and magnitude for potential risk factors, (2) identification of the levels within each potential risk factor that impose the highest risk on casing failure, (3) acknowledgement of the depths that are susceptible to casing failure, (4) prediction of the overall probability of casing failure given the information for pre-defined risk factors and, (5) adjustment of high risk casing design

specifications in the direction of reducing the overall probability of casing damage and, ultimately, mitigating casing failure. This will be achieved through:

- Designing a workflow to be used for analyzing various exposures and conditions, then, extracting the potential risk factors that highly impact the occurrence of casing failure. The designed workflow will be based on a classical risk analysis study design, that is, case-control design in order to handle the relative rareness of the event of casing failure in the provided data set compiled from the Granite Wash formation (Chapter 2).
- Quantification of the impact type and magnitude of already identified potential risk factors for a better assessment of the inter-relationship between the various risk factors and the occurrence of casing failure. To that end, various association measurement techniques are integrated in the previously designed workflow (Chapter 2).
- Designing a workflow to be used for analyzing and assessing the significance of the subcategories within the potential risk factors, in addition to, identification of the depths that are most susceptible to casing failure/damage. This workflow will be based on another widely used statistical family of techniques known as survival analysis techniques. Those techniques will be modified in accordance with the focus and goal set for the proposed workflow (Chapter 3).
- Implementation and testing of various supervised machine learning algorithms, in particular, classification algorithms in order to predict the possibility of casing failure given the initial design specifications based on the pre-defined potential risk factors. This step will be part of the two-step "correction-prediction" procedure adopted in the proposed workflow in order to mitigate and, ultimately, avoid casing failure (Chapter 4).
- Integration of semi-quantitative risk assessment techniques in the designed workflow as a second step in the two-step "correction-prediction" procedure adopted in the workflow

(Chapter 4). This will give the tool the capability to correct and adjust high risk design specifications in the direction of reducing the overall risk of casing failure. This, in turn, will give the developed tool the allowance to handle casing failure proactively rather than actively.

- Validation of the developed "casing failure mitigation" tool against real cases of casing failure from the historical data set compiled from the Granite Wash formation in the Anadarko Basin (Chapter 4).

Following, a concise overview is given of my work already published related to the content of this dissertation. The three papers published are listed below, together with their abbreviated abstracts, for a quick orientation:

1) Noshi, C. I., Noynaert, S. F., and Schubert, J. J. 2018. Casing Failure Data Analytics: A Novel Data Mining Approach in Predicting Casing Failures for Improved Drilling Performance and Production Optimization. Presented at the SPE Annual Technical Conference and Exhibition, Dallas, Texas, USA, 24–26 September. SPE-191570-MS. https://doi.org/10.2118/191570-MS.

Abstract - Recent casing failures in the Granite Wash play in the western Anadarko Basin have sparked deep concerns to operators in North Texas and Oklahoma. Hydrostatic tests made in the field show that present API standards do not assure adequate joint and bursting strength to meet multi-staged hydraulically fractured well requirements. Past and present literature has been packed with numerous casing failures incidents. Despite the extensive documentation and recommendations, a mounting trend of failure is still on the rise. To find possible solutions for these failures, this study is a continuation of an on-going effort to minimize the likelihood of failure using machine learning algorithms. The study applied both descriptive visual representations such as Mosaic and Box Plots and predictive algorithms including Artificial Neural Networks (ANN)

and Boosted Ensemble trees on eighty land-based wells in the Anadarko basin, of which twenty possessed casing and tubing failures. The study used a predictive analytics software and python language to evaluate twenty-six different features compiled from drilling, fracturing, and geologic data. Comparison of different classification models showed the superiority of artificial neural network (ANN) compared to other models, in terms of, prediction accuracy (67% E), overall performance and model balance (33.33% FP, 33.33% FN) (Noshi et al. 2018a).

2) Noshi, C. I., Noynaert, S. F., and Schubert, J. J. 2018. Failure Predictive Analytics Using Data Mining: How to Predict Unforeseen Casing Failures? Presented at the Abu Dhabi International Petroleum Exhibition & Conference, Abu Dhabi, UAE, 12–15 November. SPE-193194-MS. https://doi.org/10.2118/193194-MS.

Abstract - Despite numerous studies in the subject matter, industry has yet to resolve casing failure issues. A more interdisciplinary approach is taken in this study integrating eighty land-based wells, from the panhandle, using a data-driven approach to predict the reasons behind casing failure. Principal component Analysis (PCA) was used for dimensionality reduction. Supervised and unsupervised approaches were selected respectively based on the response. The algorithms used in this study included Support Vector Machine, Boot strap, Random Forest, Naïve Bayes, XG Boost, and K-Means Clustering. Nine models were then compared against each other to determine the winner. Features contributing to casing failure were identified based on best algorithm performance, that was, SVM. Results showed that: lesser amounts of proppant along with an increased timing from drilling to fracturing operations had higher casing failure rate chances. While sudden failure occurred immensely at very low proppant concentration. Reduced proppant mass accompanied with a lower casing setting depth increased the likelihood of casing failure. A lower frac start month and a shallower casing setting depth caused high casing failure likelihood.

Moreover, the probability of failure is higher when acid is pumped while the probability of failure decreases with decreasing hole size. A higher mean temperature and bottom hole temperature increased the probability of success. The probability of failure is high when the pipe shrinkage length is between four and seven ft long. The probability of success decreased with increased hole size and decreased bottom hole temperature. The probability of failure decreased as base water increased and bottom hole temperature was not a controlling parameter in that case. The probability of failure decreased in hole sizes between 6 – 6.5 in and the time between drilling and fracturing was not a controlling parameter in this case. Finally, the probability of failure was reduced with cement presence and a higher bottom hole temperature. Box plots showed that fracturing during the first six month has a more significant probability of failure (Noshi et al. 2018b).

3) Noshi, C. I., Noynaert, S. F., and Schubert, J. J. 2019. Data Mining Approaches for Casing Failure Prediction and Prevention. Presented at the International Petroleum Technology Conference, Beijing, China, 26–28 March. IPTC-19311-MS. https://doi.org/10.2523/IPTC-19311-MS.

Abstract - The study assembled comprehensive data from eighty land-based wells during drilling, fracturing, workover, and production operations. Twenty wells suffered from casing failure while the remaining sixty offset wells were compiled from well reports, fracturing treatment data, drilling records, and recovered casing data. Results of the survival analysis showed that the following conditions (subgroups/levels) are the most risky compared to their companions within their corresponding risk factors: (a) drilling during winter or spring seasons, (b) implementation of acid treatment for depths greater than 10,000 ft, (c) having no cement support, (d) witnessing dogleg bending stress ≥ 95,600 lb./ft for depths greater than 10,000 ft, (e) bottomhole temperature ≥ 166

degree Fahrenheit, (f) lateral section shrinkage ≥ 10 in/100 ft (g) max inclination ≤ 95.5 degree for depths greater than 8,500 ft, (h) dogleg severity ≥ 15 degree per hundred foot for depths greater than 8,500 ft, (j) frequency of severe dogleg ≥ 13. Based on Risk Analysis, fracturing during spring turned out to increase the risk of casing failure by over 200%, an increase of one unit in the lateral section shrinkage increases the risk by 15% and an increase of one degree of maximum inclination increases the risk by 30%. On the other hand, cementing reduces the risk of casing failure by nearly 54%, while having a casing thickness greater than 0.65 in (P110) tends to reduce the risk of casing failure by nearly 90% (Noshi et al. 2019).

1.3 Novel Casing Failure Mitigation Tool Developed in The Study

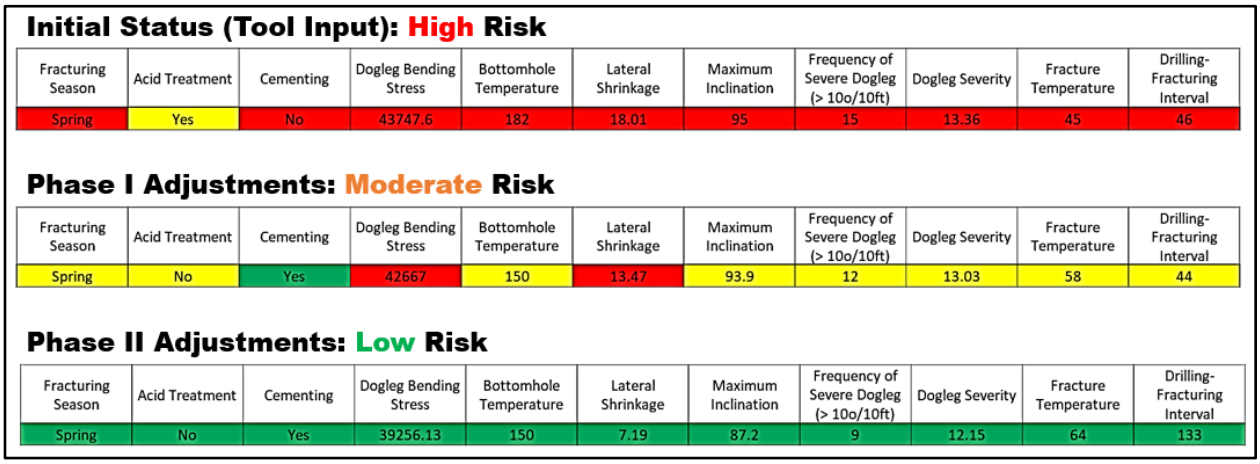
The major concern of this study is to design a novel tool, based on a data-driven workflow, that leverages several statistical techniques to automatically mitigate future casing failures. In other words, we seek a tool that, given a set of input features, not only could predict the occurrence of casing failure but also could give guidance on how an engineer should proceed, in terms of design adjustments, to avoid the occurrence of casing failure in the future.

In addition to mitigating casing failure, the developed workflow has the capability of addressing additional, yet important, challenging aspects of the problem of casing failure including: (1) quantification of impact type and magnitude for potential risk factors, (2) identification of the levels within each potential risk factor that impose the highest risk on casing failure, (3) acknowledgement of the depths that are susceptible to casing failure.

The proposed “casing failure mitigation” tool is based on a two-step “prediction-correction” procedure; where we add “feedback” to the initial predictions provided by the conventional machine learning algorithms that would serve as a guide that could help drilling engineers adjust their design and ultimately mitigate casing failure. This is accomplished through the integration of

semi-quantitative risk assessment techniques, in particular, probability-impact risk assessment matrices (PI-RAMs).

A case in point is the example illustrated in Figure 22. First, design specifications initially suggested by a drilling engineer for a particular well are fed to the "casing failure mitigation" tool. Next, design specifications are evaluated in terms of the risk imposed on casing failure (Figure 22a). In case of high risk, design specifications are automatically adjusted so that imposed risk is reduced and ultimately casing failure is avoided or mitigated (Figure 22c).



Initial Status (Tool Input): High Risk

Fracturing Season	Acid Treatment	Cementing	Dogleg Bending Stress	Bottomhole Temperature	Lateral Shrinkage	Maximum Inclination	Frequency of Severe Dogleg (> 10o/10ft)	Dogleg Severity	Fracture Temperature	Drilling-Fracturing Interval
Spring	Yes	No	43747.6	182	18.01	95	15	13.36	45	46

Phase I Adjustments: Moderate Risk

Fracturing Season	Acid Treatment	Cementing	Dogleg Bending Stress	Bottomhole Temperature	Lateral Shrinkage	Maximum Inclination	Frequency of Severe Dogleg (> 10o/10ft)	Dogleg Severity	Fracture Temperature	Drilling-Fracturing Interval
Spring	No	Yes	42667	150	13.47	93.9	12	13.03	58	44

Phase II Adjustments: Low Risk

Fracturing Season	Acid Treatment	Cementing	Dogleg Bending Stress	Bottomhole Temperature	Lateral Shrinkage	Maximum Inclination	Frequency of Severe Dogleg (> 10o/10ft)	Dogleg Severity	Fracture Temperature	Drilling-Fracturing Interval
Spring	No	Yes	39256.13	150	7.19	87.2	9	12.15	64	133

Figure 22. (a) Application of semi-quantitative probability-impact risk assessment matrices (PI-RAMs) on a case with initial status of "high" risk. (b) Using PI-RAMs, adjustments were done to the features of interest leading to reduction in risk status to "medium". (c) Further adjustments to the features of interest using PI-RAMs led to reduction of risk status to "low".

The tool design is based on four major statistical families of techniques (highlighted in "blue" in Figure 23), those are, risk analysis, survival analysis, supervised machine learning (ML) algorithms and semi-quantitative risk assessment techniques.

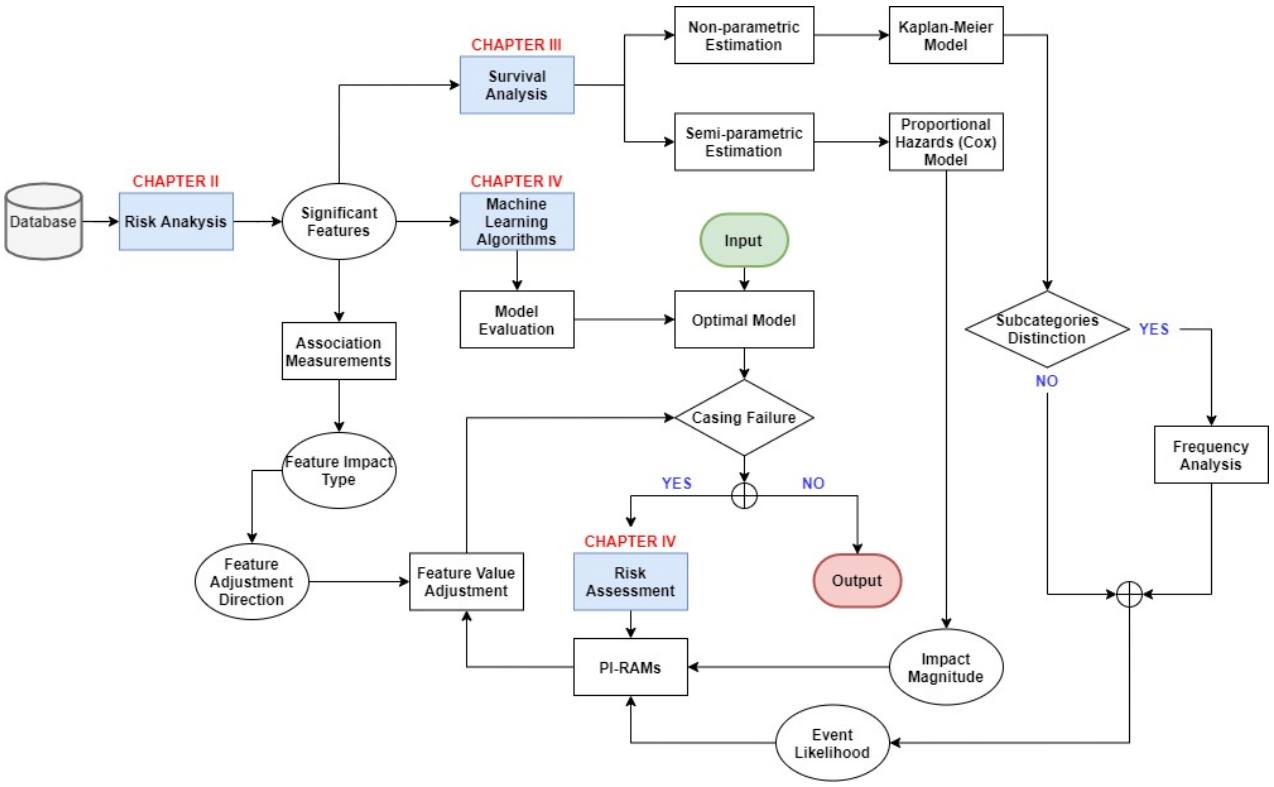


Figure 23. The developed workflow for the "automated" casing failure mitigation tool. Rectangles highlighted in "blue" correspond to the four major steps involved in the casing failure mitigation process which will be explained in detail throughout the thesis. Rectangles highlighted in "green" and "red" correspond to the input and output for the tool, respectively. Rectangle highlighted in "grey" correspond to the data set used for constructing the tool.

Supervised ML algorithms (Chapter 4) along with semi-quantitative risk assessment techniques (Chapter 4) are used as the cornerstone for the two-step "prediction-correction" procedure; where ML algorithms provide initial prediction of the probability of casing failure occurrence based on the input design specifications. Following that, the semi-quantitative risk assessment techniques (e.g., PI-RAMs) are used to automatically adjust the design specifications according to their impact type and the overall case evaluation until risk is reduced below pre-defined threshold.

Regarding the risk analysis (Chapter 2) and the survival analysis (Chapter 3), they constitute an integral part of the tool as they provide the necessary information to be later used as a basis for the

construction of ML predictive models, as well as risk assessment matrices (PI-RAMs). That information include: (1) identification of potential risk factors that are strongly associated with casing failure, (2) identification of the subgroups within each risk factor that impose the highest impact, (3) evaluation of the type and magnitude of the impact for each risk factor, (4) determination of the likelihood of the occurrence of the different scenarios within each risk factor. (Detailed discussion will be provided in the following chapters 2 through 4).

This study not only presents a methodology aiming to form the foundation for a new standard for casing risk assessment in the Anadarko Basin, but it can be applied to any geological area with different scenarios and can be developed into a more generalized tool in the future.

1.4 Dissertation Outline

In the remainder of **Chapter 1**, the geology and the data set used in this study along with some descriptive statistics are explained in (Sections 1.5).

Chapter 2 is dedicated to risk analysis family of techniques where I identify potential risk factors contributing to the problem of casing failure (Sections 2.1). In addition, I seek to identify the type and measure the magnitude of those potential risk factors. First, the various variable classes involved in the analysis and the different inter-relationship scenarios that might exist between them are identified (Sections 2.2 and 2.3). Then, the two possible study designs (i.e., case-control and cohort) are discussed, in some detail, emphasizing the key distinctions between them (Section 2.4). Afterwards, I proceed to explain the different sampling approaches involved in the study designs (Section 2.5).

Following that, I discuss the different association measurements used for testing the correlation between various exposures and the outcome (Section 2.6). Furthering the investigation, I seek

evaluating the type and magnitude each potential risk factor has on the occurrence of casing failure (Section 2.7). Due to dependency between the various exposures, or features, and the integral rule it has on the output of the analysis, different models with different combinations of exposures are tried out. In order accomplish that, another well-established family of techniques, known as subset selection techniques, is sought out (Section 2.8). To determine the best model with the optimal set of features, I set some criteria including Akaike's Information Criteria (AIC) and Bayesian Information Criteria (BIC) (Sections 2.9 and 2.10).

Chapter 3 is dedicated to survival analysis methods and builds forward on the results obtained in chapter 2 as I further my investigation to another concern of this study. Based on the information provided for the potential risk factors, I attempt to identify the depths that are susceptible and most vulnerable to casing failure. In addition, I seek to find the subgroups, or levels, within each risk factor that impose the highest risk. For that purpose, another classic statistical analysis technique, known as survival analysis, is implemented.

Discussion starts with laying the foundation of classical survival analysis techniques, as well as some modifications done by the author to the conventional analysis scale (Section 3.1). Then, I briefly discuss the various survival/hazard functions used for describing event distribution (Section 3.2). Afterwards, I proceed with discussing various methods used for estimating survival distributions. First, I discuss non-parametric estimators, in particular, Kaplan-Meier estimator (Section 3.3.1). Then, I move to semi-parametric estimators, in particular, proportional hazards (Cox) model, where some limitations in the non-parametric counterparts are addressed (Section 3.3.2).

In **chapter 4**, I take one final step towards the construction of the "automated casing failure mitigation" tool based on the information drawn from the analyses conducted in preceding

chapters. To achieve that goal, I discuss one family of machine leaning techniques, that is, supervised learning algorithms. Discussion starts with laying out the key distinctions between conventional physics-based and modern machine learning algorithms (Section 4.1). For the purposes of this study, I only investigate learning algorithms for classification problems such as logistic regression (Section 4.2.1), decision trees (Section 4.2.2), random forest (Section 4.2.3), support vector classifier (Section 4.2.4), support vector machine (Section 4.2.5) and neural networks (Section 4.2.6). Then, I discuss different criteria for evaluating the performance of the models in terms of prediction accuracy, model balance and overall performance (Section 4.3).

To add another level of controllability to serve as feedback means to the model initial predictions, discussion is extended to another area of statistical analysis, that is, risk assessment (Sections 4.4 and 4.5). However, discussion is limited to semi-qualitative risk assessment techniques, particularly, probability-impact risk assessment matrices, PI-RAM, (Section 4.6). I explain the basis for those techniques and how they are formulated. I, then, present some history cases from the oil and gas industry where implementation of those techniques showed a huge potential (Section 4.7). Then, I end the discussion with demonstration of the presented methods on the history data set (Section 4.8).

In **chapter 5**, results obtained from all the analyses conducted in the preceding chapters are discussed, highlighting key findings and major takeaways. Discussion starts with a layout of the major contributions achieved throughout the study (Section 5.1). Then, attention is geared towards evaluating the proposed solution(s) through the implementation of SWOT analysis (Section 5.2). Afterwards, a set of recommendations is provided for the reader in order to efficiently utilize and benefit from the proposed workflow (Section 5.3). Finally, we end our discussion with addressing

some of the aspects in our study that could be improved upon in the future for more refined/enhanced results (Section 5.4).

Chapter 6 provides a bird's-eye view of the achievements of this study and the major takeaways. The starting point is the results obtained from risk analysis and the potential risk factors identified throughout the process. Then, I show how those initial results were used as a primer for survival analysis; where more scrutinized examination was done to identify the subgroups/levels within the various risk factors that had the highest impact on the occurrence of casing failure. Following that, I showcase the importance of those results when fitting classification learning algorithms for casing failure prediction. And, finally, I show how integrating survival analysis with frequency analysis help establishing the risk assessment matrices that were used as a feedback step following initial predictions provided by ML algorithms which served as the basis for the two-step "correction-prediction" procedure used to establish the "casing failure mitigation" tool.

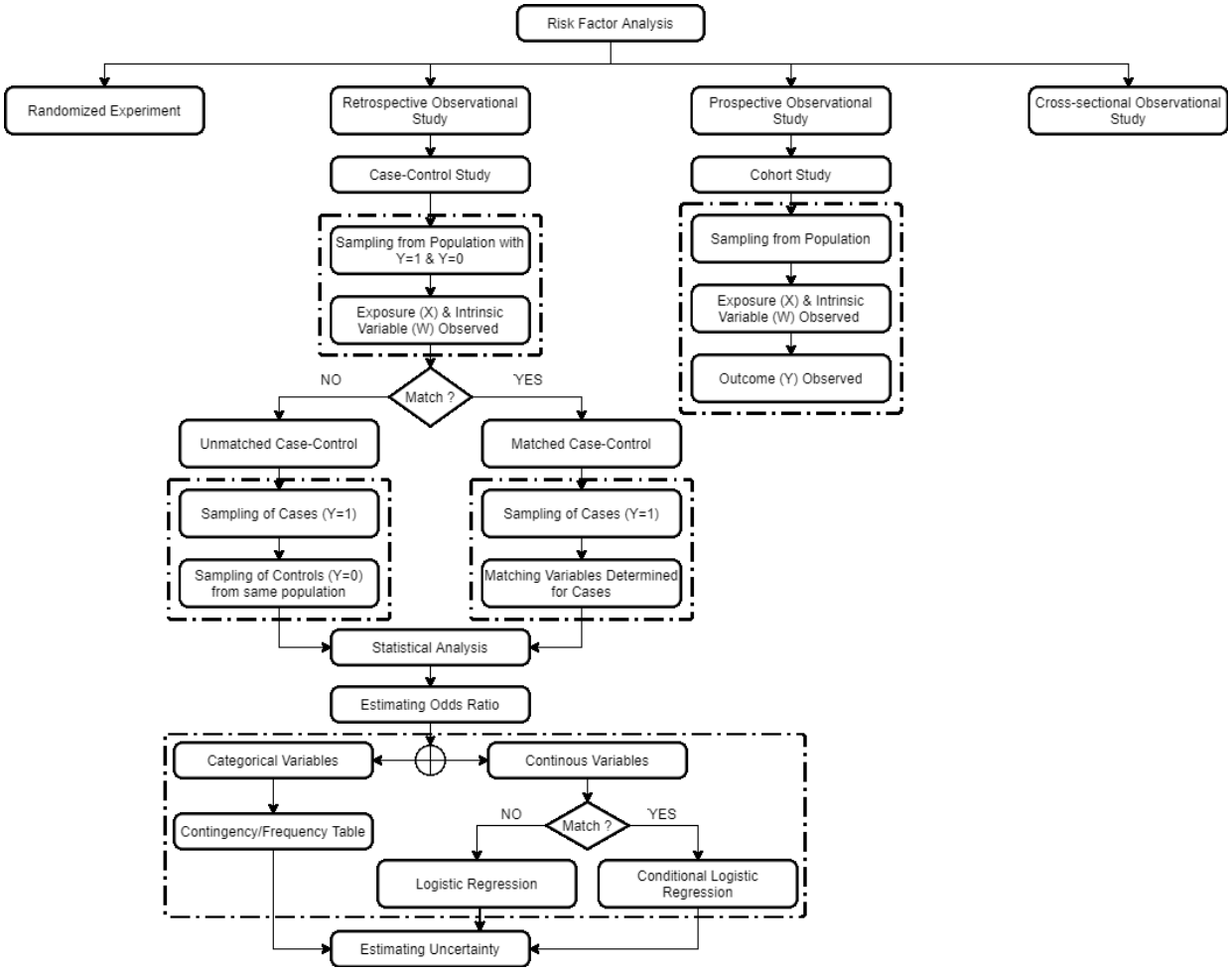


Figure 24. Workflow adopted in this thesis for general steps involved in risk factor analysis. Risk factor analysis comprises of four families of techniques, as shown at the top. Yet, two families are investigated in this study: prospective and retrospective designs. Key difference is the direction of inference, as highlighted in the top dotted rectangles. Within retrospective design, two sampling approaches are adopted, and the key distinction is shown in the dotted rectangles at the middle. Association measurements used for determining the potential risk factors and steps involved in the process is highlighted in the bottom dotted rectangle.

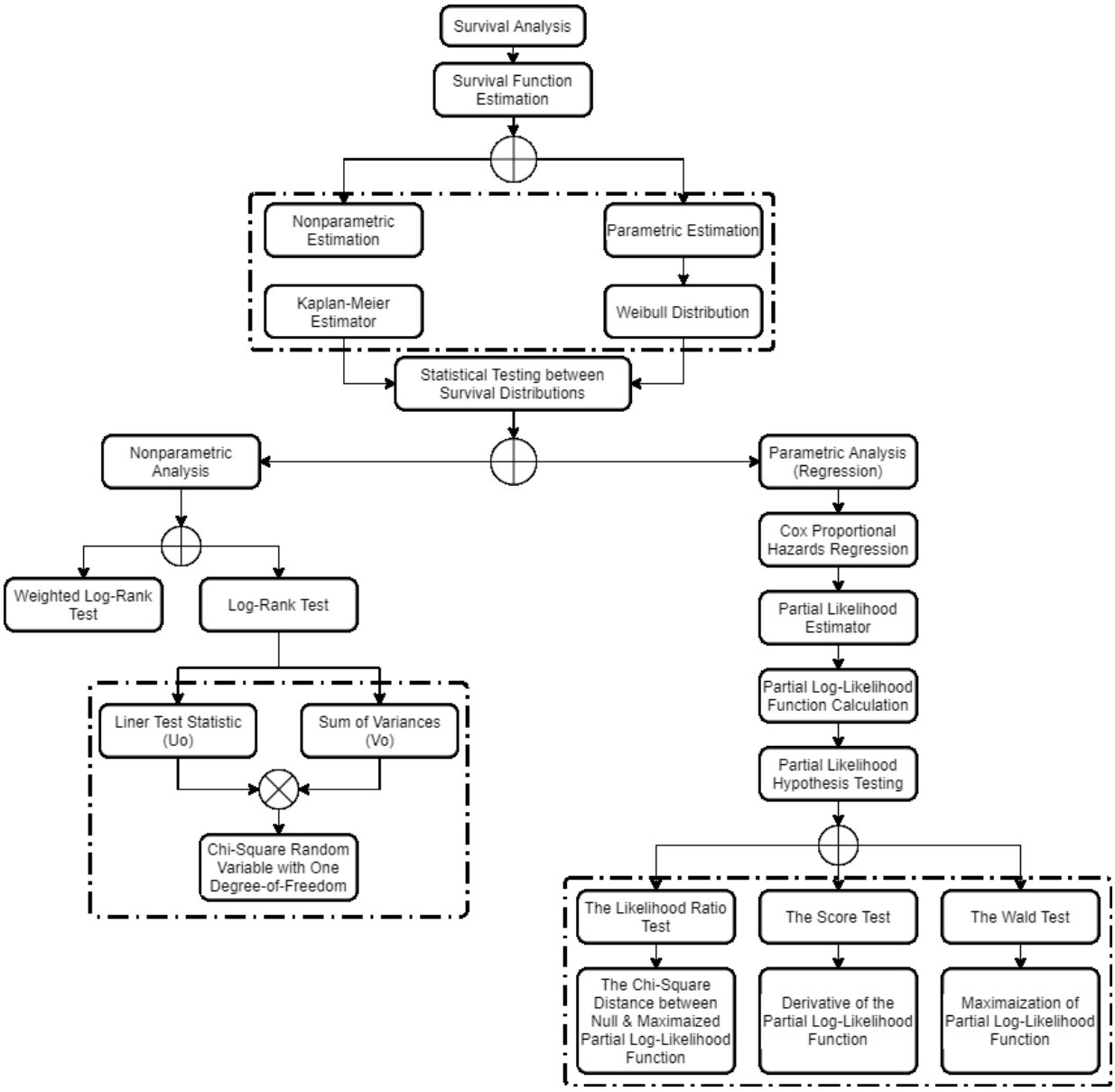


Figure 25. Workflow adopted in this thesis for general steps involved in survival analysis. The top dotted rectangle highlights the two major families of techniques used to estimate the survival/hazard distribution of the subcategories within each potential risk factor. Depending on the selected family of techniques, the key steps involved in the process are highlighted in the bottom dotted rectangles.

1.5 Data Set Used in This Study

The data set used in this study incorporated information extracted from an extensive geographical area the Granite Wash play located in Texas and Oklahoma. The High Plains Prospect is situated in the Western Anadarko Basin of the Panhandle with three major formations targeted, namely, the Cleveland Sandstone, Granite Wash, and the Marmaton (**Figure 26**).

The area has a complex tectonic setting with abundant faults and folds exhibiting a complex stratigraphic setup, vastly inconsistent lithologies, and congruently diverse reservoir properties accompanied by erratic formation types and varied porosity and permeability values (Moore 1979).

The Cleveland formation is a tight gas sandy formation spanning both the Lipscomb and Ochiltree Counties in the Texas Panhandle and permeability ranges from 0.03–1.1 millidarcy. The formation is composed of clean sand blocks with thin shaley laminations. It is exceptionally fine grained with a very high-water saturation owing to the presence of fine sandstone with interbedded shales. In this formation, essentially all the wells were detected to have drilling-induced fractures. The productive sandstone is composed of an average by total rock volume of three percent clay and thirteen percent cement.

The second formation is the Granite Wash tight gas play, composed of a succession of very different stacked pay zones, each of which should be developed differently. Lithology is composed of gravels, clean sands, and shale. The reservoirs are considered depleted due to production since 2002.

The third formation is the Marmaton group which encompasses Roger Mills County, Oklahoma and Wheeler County, Texas. The group contains interbedded gas-bearing sandstones and

conglomerates that thin into shale formations topped by limestones that taper laterally into shales in the north-eastern direction (basinward) in Hemphill and Roger Mills Counties with the units gradually updipping.

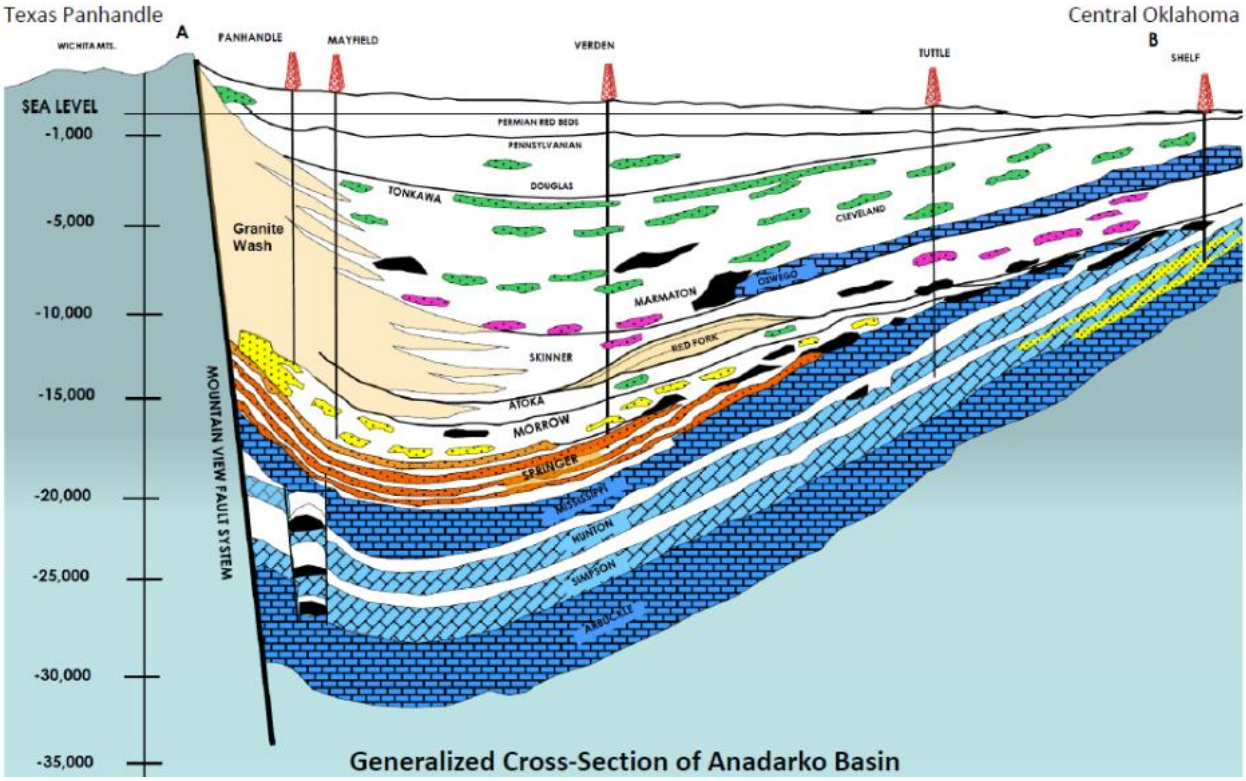


Figure 26. Describing the different wells drilled in the three formations presented in this study, namely The Marmaton, Cleveland, and Granite Wash (After Nostra Terra Oil & Gas Company PLC. 2018).

The study area spreads across seven counties: Ochiltree, Lipscomb, Roberts, Hemphill, Wheeler, Harper, and Roger Mills (***Figure 27***). In Wheeler County, the horizontal producing intervals include Marmaton Wash "B", "C", "D" Wash produces gas and condensate, while the Cottage Grove Wash produces oil and gas. Reservoirs in Wheeler County have as much as eighteen percent rock volume comprised of chlorite and a highly variable clay content with high S_w (Moore 1979). Ochiltree County is situated in the Texas Panhandle containing three prolific horizons. These

include the Cleveland, Mississippian, and Morrowan formations.

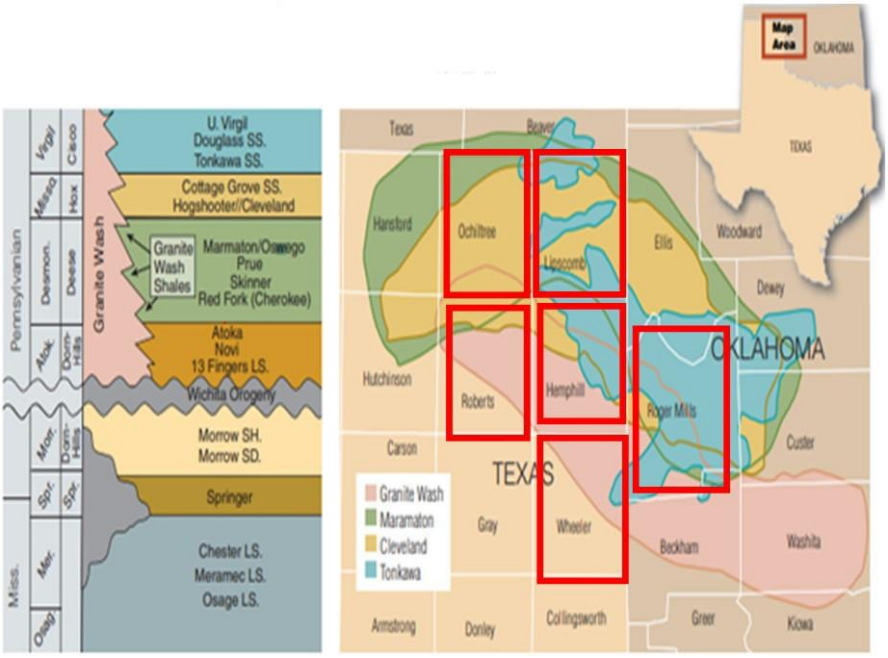


Figure 27. The Western Anadarko Basin showing the 7 different counties and 3 formations (After Karis 2015).

A total of eighty wells were originally gathered with twenty-six different variables in a spreadsheet (Kim 2017). The first step included cleaning and arranging the data in a suitable format. Feature engineering commenced with a cut off value of thirty-six percent; any column with missing data values of more than or equal thirty-six percent was removed from the data set. The cut-off threshold value was based on a trade-off between the amount of missing data per well and the number of wells with missing data. Having a higher value for the cut-off threshold would have led to a loss of a measurable number of wells. On the other hand, having a lesser value for the cut-off threshold would have led to having a high amount of missing data. According to the selected cut-

off value, two off-set wells were removed from the data, totaling to seventy-eight total wells used in this study. The data compiled was then divided into two parts: the wells with noted casing failures and offset wells with no obvious integrity issues.

Three neighboring wells for each of the failed well were selected for the analytical process. For comparison purposes, the chosen offset wells were drilled within five years of the failed wells within proximity to ensure similar application of drilling and completion technologies and practices (Kim 2017).

The data compiled for the offset wells were collected from a variety of public data repositories including Texas Railroad Commission (RRC), Frac Focus Chemical Disclosure Registry, Oklahoma Corporation Commission Oil and Gas Conservation Division, and Drillinginfo were used as a reference database. Different Completion and frac job documents were deduced from the aforementioned data repositories including W2 report, direction survey Gyro, direction survey MWD, W15 (casing design reports), Hydraulic fracturing fluid information disclosure. Various information was extracted from the collected files including well information, completion information, tubing records, treatment records, formation records, casing records, and other frac-related information (Kim 2017). The preceding information is displayed as a workflow for data collection and information gathering as shown in **Figure 28**.

After gathering all the required well information, the data went through a preprocessing step starting with data formatting. This included the construction of subset data frames and merging multiple datasets. Then, data type check including high level (whether continuous or categorical), low level (subcategories/levels within categorical variables) data type check, data dimension check (making sure that all input/output variables are of same dimension), and metadata check (making sure variables' names and definitions are unified). The clean data set then went through a data

engineering procedure starting with data scaling and standardization either basic scaling or z-score standardization. Then levels and sub-categories were defined based on data factorization. The preceding information is displayed as a workflow for data collection and information gathering as shown in **Figure 29**. For detailed discussion, the reader can refer to Endel et al. 2015; Furche et al. 2016.

Initial statistical analysis of the provided history data set; using basic contingency tables, showed that: the rate of casing failure was highest during the winter season with five out of seven wells failing (71.40%), followed by the fall season with only two out of twelve wells failing (16.67%). The spring season had seven out of twelve wells failing (58.33%), and finally, the summer season with six out of fourteen wells failed (42.86%). Concerning the "cement" variable, the data showed an increased rate of casing failure with lack of cement. As for the "acid" variable, the presence of an acid job was accompanied with increased risk of casing failure and well damage. In addition, an increased shrinkage of the lateral section, a longer lateral length, and a higher average DLS showed an increased probability of casing failure. Whereas a smaller casing thickness and a reduced amount of proppant used contributed to higher casing failure probability.

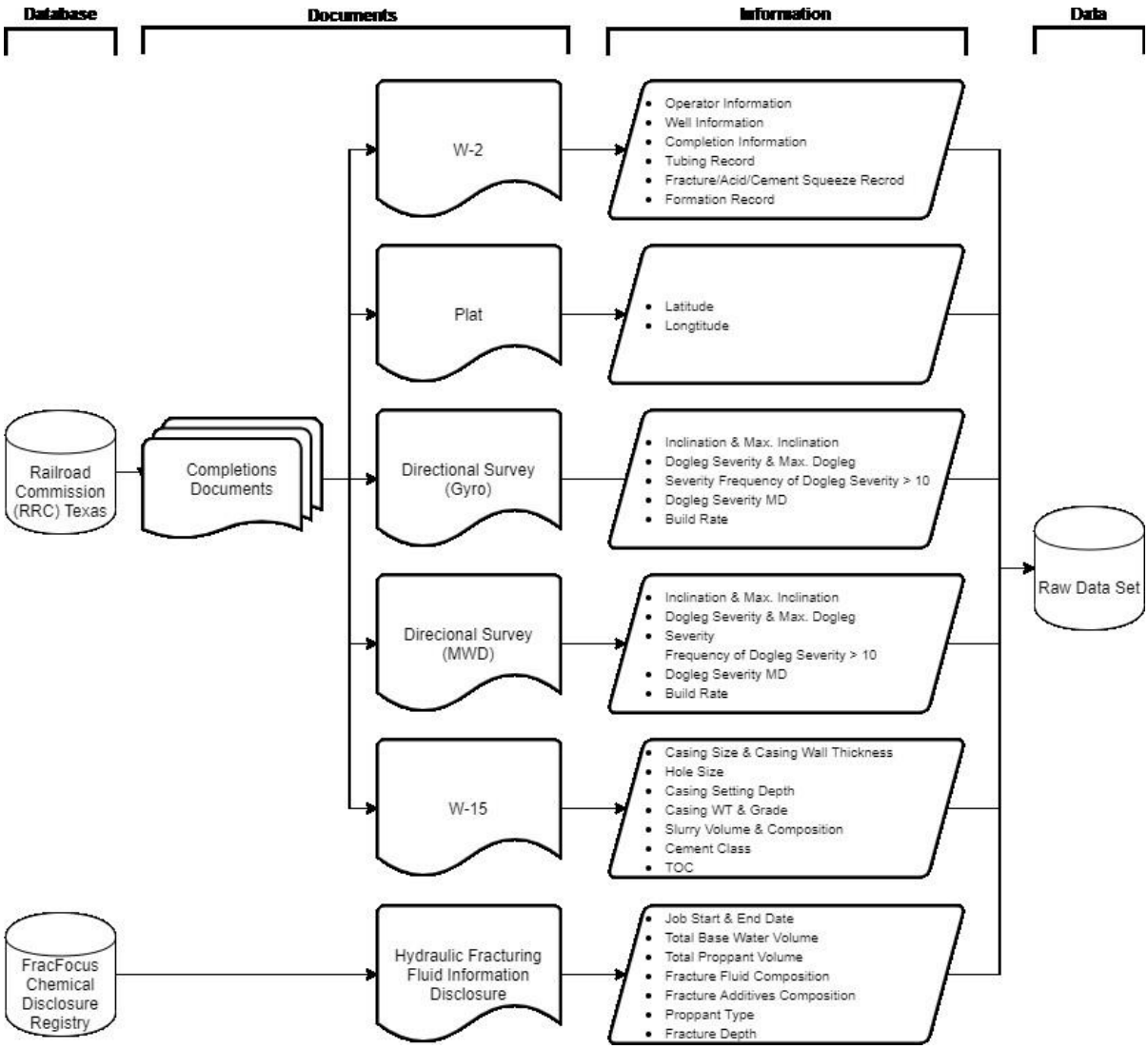


Figure 28. Workflow adopted in this thesis for data collection and information gathering. The data repositories included Texas Railroad Commission and Fracfocus Registry. Several files were extracted from the repositories including directional survey, hydraulic fracturing fluid information disclosure, field description, Casing design (W-15), etc.

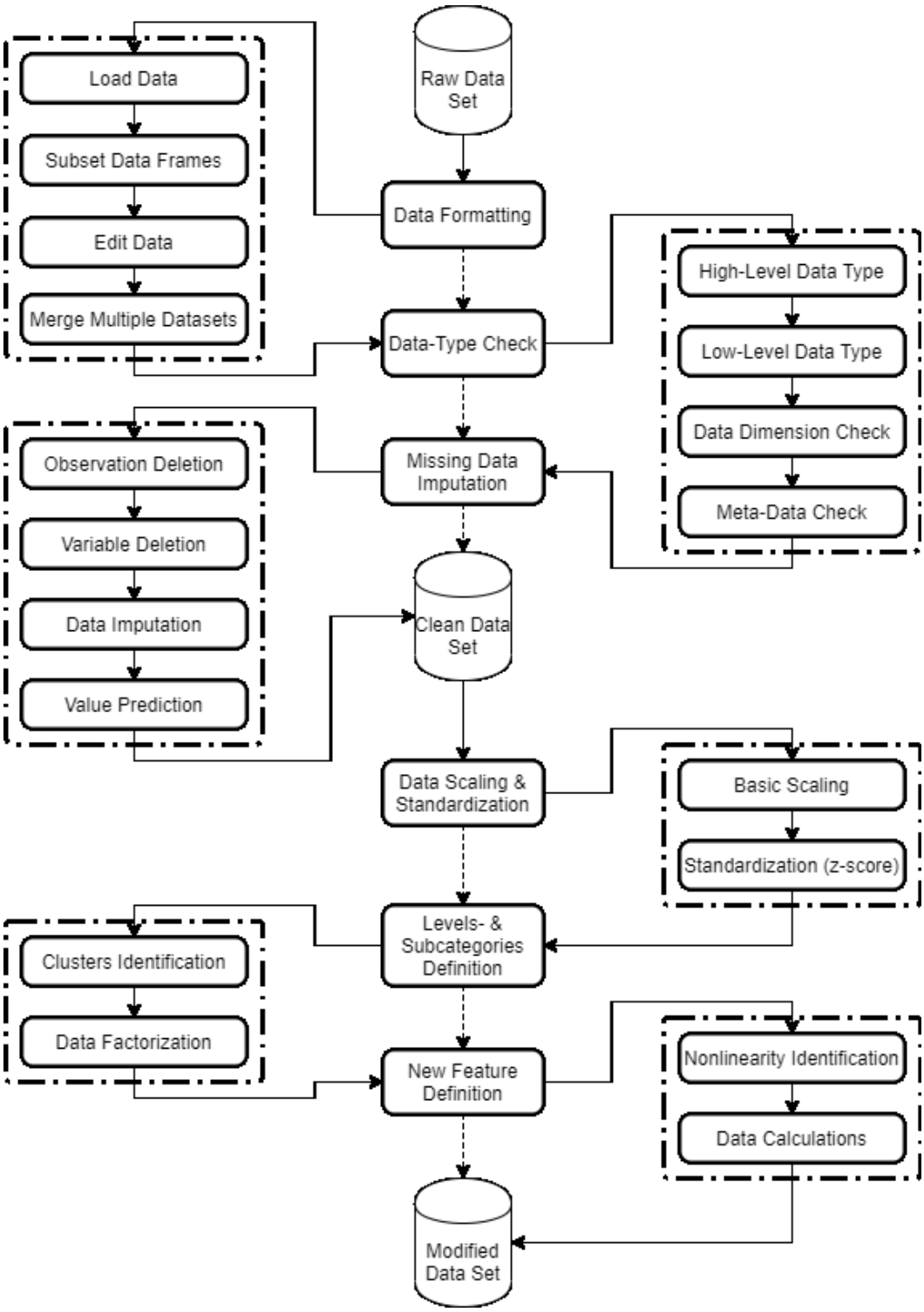


Figure 29. Workflow adopted in this thesis for data wrangling, cleaning, preprocessing and feature engineering. Six major steps are involved in the process (highlighted in the rectangles in the middle column), namely, data formatting, data-type check, data imputation, data scaling, variable categorization and new feature definition. Steps involved in each of those processes are highlighted in their corresponding dotted rectangles.

CHAPTER II
RISK ANALYSIS

Reader Guide: In this chapter, I address some of the first and foremost concerns of this study; (1) Identifying potential risk factors amongst different exposures initially provided in the data set, (2) evaluating the type of the impact each risk factor has on the occurrence of casing failure and (3) measuring the magnitude of that impact. To that end, I seek to design a workflow based on a well-established family of techniques, known as risk analysis techniques (Section 2.1). This workflow will, later, be integrated in the design and construction of "casing failure mitigation" tool.

Discussion starts with considering the various variable classes involved in the analysis and the different inter-relationship scenarios that might exist between them (Section 2.2 and Section 2.3). Then, the foundation for the two study designs (i.e., case-control and cohort) is laid out and discuss, in great detail, the key distinctions between them (Section 2.4). Afterwards, different sampling approaches involved in the study designs are explained (Section 2.5). Following that, I discuss the different association measurements used for testing the correlation between various exposures and the outcome and decide on the significant risk factors (Section 2.6). Then, I further my investigation into evaluating the type and magnitude each potential risk factor has on the occurrence of casing failure (Section 2.7).

Due to dependency between the various exposures, or features, and the integral rule that effect has on the output of the analysis, different models with different combinations of exposures are tested and evaluated. In order to accomplish that, I seek another well-established family of techniques, known as subset selection techniques (Section 2.8). To determine the best model with the optimal set of features, some criteria are discussed including Akaike's Information Criteria (AIC) and Bayesian Information Criteria (BIC) with slight modification to align with model design (Section 2.9). Following that, I implement techniques discussed in previous two sections on the provided data set in order to settle on the best model with the optimal set of features to continue with in subsequent analyses (Section 2.10). Finally, I highlight the work done in this chapter to address the aforementioned concerns, as well as, the major achievements, or takeaways, from the analyses conducted (Section 2.11).

2.1 Introduction

Literature is rife with contributions devoted to investigating and explaining the causing effects of casing failure in numerous historical cases. To date, the majority of those contributions were based on physics-based approaches either analytical (Li and Samuel 2016; Lin et al. 2016; Cirimello et al., 2017), experimental (Han et al. 2006a, b; Ferla et al. 2009; Jianiun et al. 2012; Zhang et al. 2012; Shen et al. 2014), or numerical in nature (Fleckenstein et al. 2001; Han et al. 2006a, b; Yuan et al. 2012; Lavrov et al. 2015; Hou et al. 2016; Feng and Gary 2017; Patel et al. 2018; Ogunsesan et al. 2019). Although, those contributions have provided a valuable insight into identification of some of the potential causing effects of casing failure, they failed to provide enough information on how to mitigate, or avoid, the occurrence of casing failure in the future, hence, did not attain wide-scale execution (Maharaj 1996; Dall'Acqua et al. 2013). This is partly due to their inability to quantify the impact; whether the type or magnitude, those potential risk factors had on the probability of the occurrence of casing failure.

This chapter represents the first step taken by the author towards designing a data-driven based tool that is capable of automatically mitigating and, ultimately, avoiding future casing failures. In this chapter, I address the concern of impact quantification for various potential risk factors as I seek to identify the impact type, as well as its magnitude for various potential risk factors considered in this study. This serves as a continuation of previous efforts to not only identify but also explain the impact of possible causing effects of casing failure. To achieve that, I seek to design a data-driven workflow based on traditional risk analysis statistical techniques. This workflow will give the allowance to (1) identify potential risk factors and (2) quantify their impact. Later, this workflow will be integrated in the design of the "casing failure mitigation" tool which is the pinnacle of this study.

Conventionally, risk-factor analysis includes two principal analytical study designs, those are, (1) cohort and (2) case-control designs (Keogh et al. 2014). In cohort studies, individuals are sampled, and their exposure status is determined initially; then the outcome is observed. In this study, individuals refer to all drilling/producing wells (failed and not failed), exposure status refers to all different sorts of exposure/treatment including – but not limited to – casing records (e.g. casing size, casing grade, etc.), fracture records (e.g. fracture length, fracture depth, base water total volume, etc.). As for the "final" status, it simply refers to the outcome which is – in this case – the casing failure. In case-control studies, cases, and non-cases (also known as controls) are sampled, and the "outcomes" being compared are the covariates (including exposure). Cases are usually a subset of the "individuals" that did actually experience the outcome. In this study, cases refer to the drilling/producing wells that encountered casing failure. On the contrary, non-cases, or simply controls, refer to wells that did not encounter casing failure.

Cohort and case-control designs can be distinguished primarily by the direction of inference: cohort studies reason forwards from the exposure to final outcome (Mann 2003). That is why cohort studies are also known as prospective observational studies. while case-control studies start with the final outcome and go back to find all possible causes (Porta 2014). That is why case-control studies are also known as retrospective studies. Graphical representation of both designs is provided in Figure 32.

2.2 System Description

The starting point for this discussion is having a population of study individuals (drilling/producing wells) that are assumed to be statistically independent. This assumption is critical since the primary objective is to understand the effect of exposures, e.g., treatments or conditions, on the final outcome, i.e., casing failure (Figure 30). Exposures are usually represented by a random variable,

X, and the outcome by a random variable, Y.

Another class of interest is "intrinsic" variables, typically represented by the random variable, W (Borgan et al. 2018). Depending on the study setting, these variables may affect X or Y or both; an aspect that will be further investigated in the subsequent section. In this study, both X and W are explanatory variables. This means that both X and W are independent variables which control, or explain, the variation in the outcome, Y.

For a variable to be considered as an exposure, X, it must be relevant, even if not realizable, to ask: how would the outcome of an individual have changed had their exposure been different from what it is, other things being equal? A typical example of variables that could be classified as exposure is the acid treatment, cement treatment or casing design, etc. Those are the variables that would be of interest to study their impact on the outcome (casing failure).

By contrast, intrinsic variables represent properties of individuals (drilling/producing wells) that are immutable, or fixed, when studying the possible effect on Y of changing the exposure X. A typical example of intrinsic variable would be the field where history wells are located or the type of formation, for instance.

Figure 30. The effect of an exposure, ***X***, on an outcome, ***Y***. The arrow represents statistical dependence, i.e., direction of inference, from ***X*** to ***Y***.

2.3 Interrelationships in Observational Studies

The below figure (Figure 31) showcases four different types, or scenarios, of inter-relationship between exposure, X, intrinsic variable, W, and the outcome, Y, which could exist in cohort, as well as, case-control studies (Keogh et al. 2014).

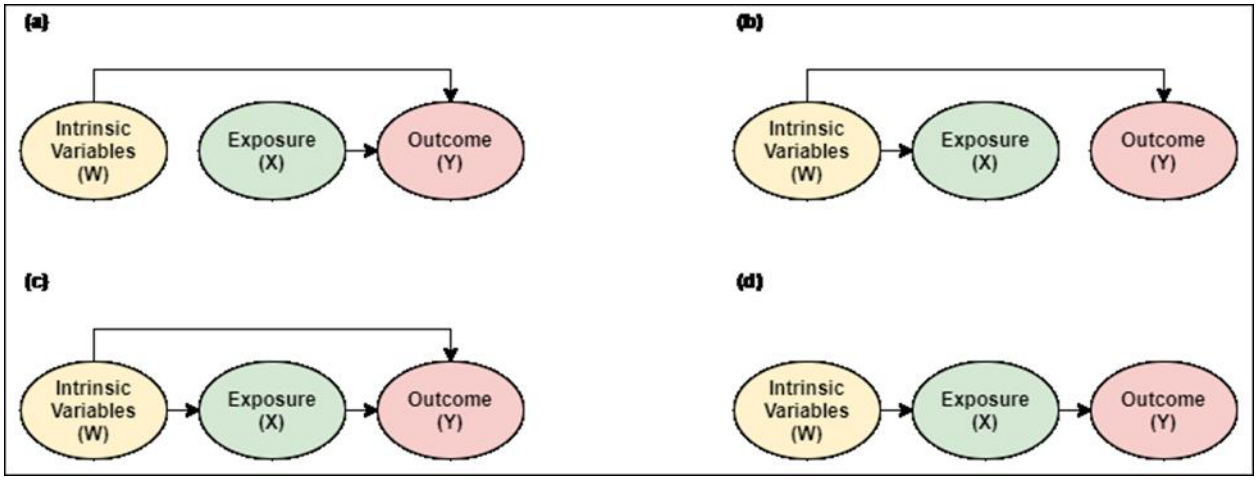


Figure 31. General specification: relationships between exposures, **X**, intrinsic variables, **W**, and outcome, **Y**, in observational studies (Reproduced from Keogh et al. 2014).

In Figure 31-a, the assumption is that the intrinsic variable has a potential impact on the outcome but not the exposure. Hence, W does not interfere in the effect of X on Y. In this case, ignoring W causes no systematic error, or bias, in the estimated effect of X on Y, though controlling for W may yield some improvement in precision for predicting the outcome, Y. A typical example would be to have the "formation type" as the immutable/fixed intrinsic variable and the "dogleg severity measured depth" as the exposure with the outcome of interest set to be "casing failure". Even though the "formation type" has no potential effect on the "dogleg severity measured depth", it still has an impact on the outcome that is the risk of casing failure.

In Figure 31-b, the assumption is that intrinsic variable affects both the outcome as well as, the exposure. Yet, there is no effect of exposure on the outcome given the intrinsic variable. This case, however, is rare especially in the context of this study since all the exposures of interest are somehow connected to the outcome.

Figure 31-c follows the same assumption as Figure 31-b except for the inclusion of an arrow from X to Y, which implies that the exposure has an impact on the outcome. In this case, neglection of the intrinsic variable would lead to an erroneous/biased estimate of the effect of X on Y given W. A case in point is having the "formation type" as the intrinsic variable, while having the "fracture design" as the exposure. It is obvious, then, that the intrinsic variable has an impact on both the outcome and the exposure. Meanwhile, the exposure has an impact on the outcome.

Figure 31-d, is analogous to Figure 31-a, with the only distinction that the assumption is reversed as we have W affects X but not Y. That is, the outcome is independent of the intrinsic variable given the specified exposure. Similar to the case explained in Figure 31-b, this case is quite rare in the context of this study and all the intrinsic variables are connected to the outcome.

For the purpose of the study and based on the former discussion, only cases 1 and 3; represented by Figure 31-a and Figure 9-c, respectively, were encountered during the analysis to be presented later in subsequent sections.

2.4 Analysis Approaches in Observational Studies

As mentioned previously, there are two major approaches or designs though which the system – described in previous sections – can be investigated/analyzed: (1) cohort, or prospective, study and (2) case-control, or retrospective, study. Following is a brief discussion of the case-control design and its key components.

2.4.1 Foundation of Cohort Study

In cohort studies, individuals are sampled, and their exposure status is determined initially; then the outcome is observed (Armenian 2009). In this study, individuals refer to all drilling/producing wells (failed and not failed), exposure status refers to all different sorts of exposure/treatment including – but not limited to – casing records (e.g., casing size, casing grade, etc.), fracture records (e.g., fracture length, fracture depth, base water total volume, etc.). As for the "final" status, it simply refers to the outcome which, in this case, is the casing failure (Figure 32-b).

For instance, if cohort design were to be implemented, the starting point would have been to drill/produce from new wells and try out different sorts/variations of exposures (e.g., different casing designs, different acid treatments, different fracture designs, etc.). Then, the outcome; whether this drilling/producing well would fail or not, is observed.

2.4.2 Foundation of Case-Control Study

In case-control study, the starting point is to sample individuals observed to have a specific outcome ($Y = 1$), referred to as *cases* (drilling/producing wells that encountered casing failure). Then, a suitable number of *controls* are chosen accordingly, often one (preferably three) control(s) for each case (Armenian 2009). Controls are defined as members of the population at risk of failure that are not cases yet ($Y = 0$). In the context of this study, controls are wells that did not experience casing failure during their lifetime. Exposures and intrinsic variables are then determined retrospectively on the chosen wells; cases and controls (Figure 32-c).

So, the essence of a case-control study is that we start with the outcome and look backward to find the exposure (i.e., explanatory/independent variable) of interest. For example, if case-control design were to be followed to study the association between an exposure and outcome, then the

starting point would be having drilling/producing wells that had either failed or not (i.e., historical case studies). Then, based on a comparison between exposures for failed (cases) and those for non-failed (controls) wells, the relationship between exposures and the outcome is determined.

Another core characteristic is that the ratio of cases-to-controls is not necessarily the same as in the true population. In the population, that ratio could be very small in case the outcome is rare, yet in the case-control study this ratio can be as high as one to one. The privilege of having this sort of control lies in the ability of leading a balanced analysis of different exposures which ultimately will lead to a sound conclusion on the potential causes.

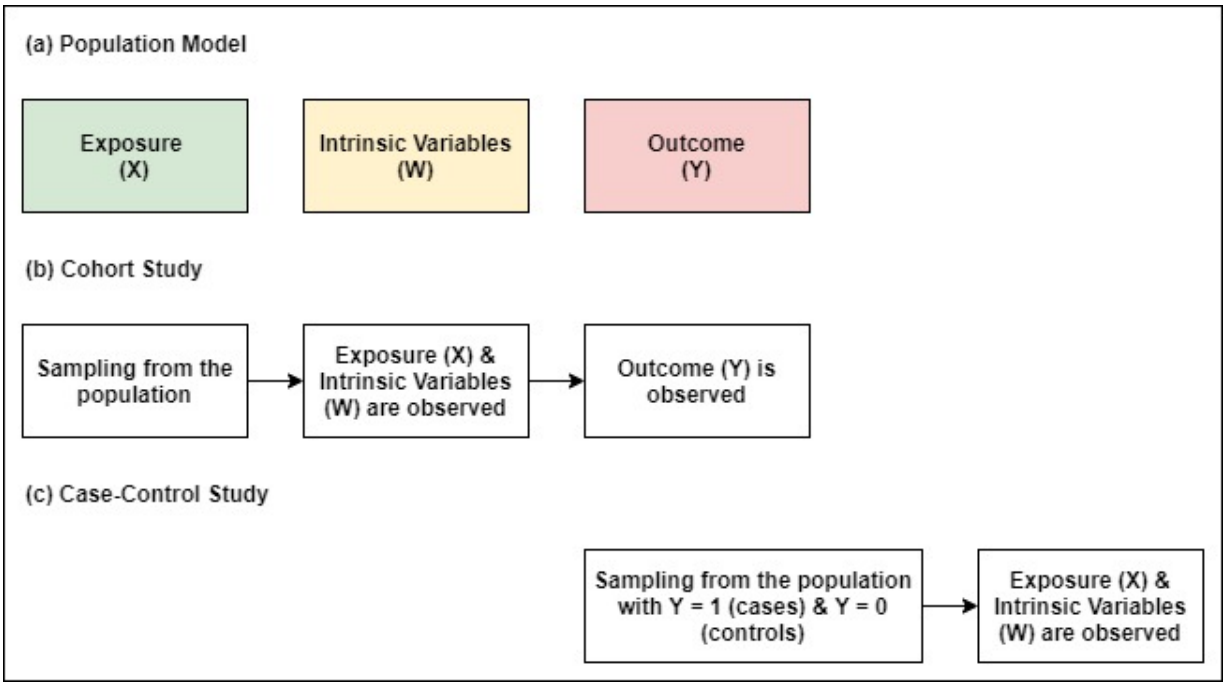


Figure 32. (a) key elements of population model before the sampling; (b) cohort study design where exposures are determined then outcome is observed; (c) case-control study design where outcome is determined then exposures are defined (Reproduced from Keogh et al. 2014).

2.4.3 Key Distinctions Between Cohort and Case-Control Designs

Even though cohort and case-control designs might share some similarities, there are key distinctions between the two designs that are critical to their application, hence, the decision on which to be adopted in the study.

One major distinction is the sampling design. For cohort design, since there is no prior knowledge of the wells that are going to fail given a specific set of exposures, then the study may end up having only the wells that failed or the opposite. This may lead to biased results and ultimately a false conclusion. For instance, if we were to study the relationship between casing grade and casing failure using the cohort design, and assuming that we have 1,000,000 wells to be drilled that we don't know which one of them would fail using a specific casing grade (e.g., Grade P110). Because of the large number of wells, we ought to choose – at random – a subset (e.g., 100 wells) to be drilled using that specific casing grade (i.e., P110). The problem here is the possibility that those 100 wells may all fail, for instance. This would lead to a conclusion that the grade P110 is imposing a high risk on casing failure. However, in those 1,000,000 wells, there might be another 100 wells that – if P110 casing were to be used – they wouldn't fail. In this case, we would have concluded that the grade P110 has no impact on casing failure since we have equal number of wells that failed and those who did not fail using same casing grade.

This problem could be solved by using the case-control design. Since, for case-control design, historical case studies are used, then we already know which wells did fail and those that did not. Accordingly, when sampling, we could choose a balanced number of failed wells (cases) and non-failed wells (controls) for analysis. This way we can eliminate any biases that may result from sampling and have a more balanced/fair conclusion.

Another problem with cohort study design is that since there is no prior knowledge of the outcome

(casing failure) for each selected individual (drilling/producing well), there is a need for sampling a huge number of wells in order for the sample to be representative of the outcome and the true population. The term “representative” reflects the fact that sampled wells should comprise all different sorts of exposure that exists in the true population. This, however, might not be feasible when the outcome is relatively rare, such as the one we are interested in this study.

This problem can be handled by following case-control design. With prior knowledge of the outcome (wells that encountered casing failure), there is no necessity for having an enormous sample size in order for the sample to be representative of the true population and the outcome. Only enough to accrue a sufficient number of cases and a comparable number of controls. Thus, the case-control design is ideal for studying rare outcomes.

Another key distinction that gives the edge to case-control studies is the utilization of the odds ratio (OR). Odds ratio approximates the relative risk by comparing exposed to unexposed in a cohort study. In this study, the term “exposed” refers to the drilling/producing wells that experienced the exposure of interest (e.g., specific acid treatment, cement treatment, etc.), while the term “unexposed” refers to the remaining wells that did not experience that specific exposure. For instance, if the impact of acid treatment on casing failure were to be tested, then the odds ratio would compare the wells that had acid treatment with those who did not. Based on the reasons mentioned previously, case-control design shows superiority over the cohort design for several reasons. Accordingly, the author has chosen to adopt this design in subsequent analysis.

2.5 Classification of Case-Control Study

Although the basic design principle of case-control study is that cases (wells experienced casing failure) should be representative of all cases in the population and controls (wells that did not

experience casing failure) should be representative of the source population of cases, this is not ideal. There still could be some bias, notably differences in the representation of cases and controls relative to their source populations.

This might lead to what is known as “confounding” effect. In the context of this study, the term “confounding” indicates the misrepresentation of the sample to source population. To minimize the effect of confounding, controls are generally individually- or group-matched to cases using known risk factors that are *not* of primary interest. The “matching” concept will be explained in detail in the sequel.

Before we proceed with the measurement techniques used to determine the inter-relationship between the outcome and different exposures, it is imperative to discuss another important trait of case-control design, that is, concerned with sampling cases and their corresponding controls.

Within the standard framework of case-control study, there are two main approaches of sampling (Borgan et al. 2018): (1) unmatched sampling and (2) matched sampling (Figure 33). In the unmatched case-control study, a shared control group for all cases is selected essentially at *random* (Figure 33-a). Contrarily, in the matched case-control study – as the name suggests – controls are selected case by case in such a way that they are constrained to match individual cases in certain specified respects, that is, so that to each case is attached one or more controls (Figure 33-b).

In this study, the latter approach has been followed. For each well that experienced failure (case), three corresponding neighboring wells with no experience of casing failure (controls) were selected. Each case and its corresponding controls were matched based on specific variables including their field, their location, targeted formation(s). So, for each well that experienced casing failure, three other wells that did not experience any casing failure were selected as controls, yet

they were drilled in the same field, within small vicinity and targeted the same formation(s). This matching has the potential of conducting a fair and balanced analysis when it comes to deciding the potential causes of casing failure among the different exposures in place.

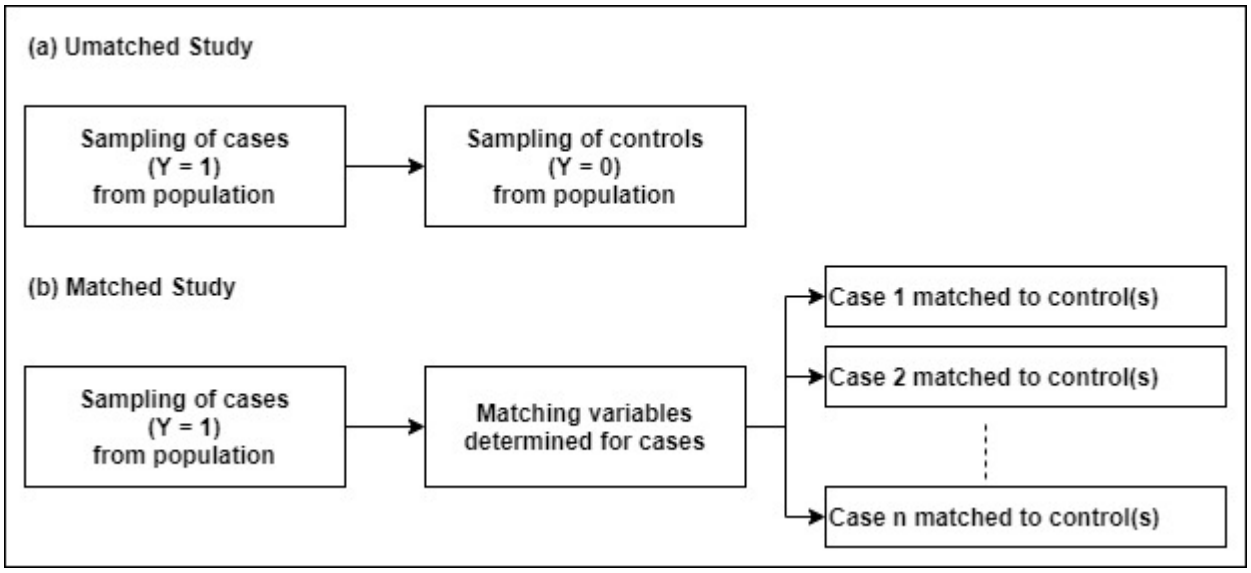


Figure 33. (a) An unmatched case-control study; (b) a matched case-control study, in which the cases are matched to one or more controls (Reproduced from Borgan et al. 2018).

2.6 Analysis Measurement Techniques

Now, after introducing the study design (i.e., case-control design) adopted for the subsequent analysis; discussing the motivation behind that selection and key components of that design, it is important to introduce the variety of techniques used for testing the association between different exposures and the outcome. Hence, the focus of this section is two-fold. First, a comprehensive review of several measures of testing the association of exposures (e.g., acid treatment, cementing, fracture design, etc.) to the outcome (casing failure) is presented; showing how association can be

estimated for binary exposures and outcome. Second, a layout of logistic regression analysis of matched case-control data is presented where multi-level and continuous exposures are handled.

The simplest type of exposure is binary, meaning that any individual (drilling/producing well) is either exposed or unexposed to the feature of interest. The association analysis, then, would require calculating the "prevalence of exposure" in the case group and in the control group and to examine whether the exposure prevalence differs by case and control status. However, in most cases, the exposures might have more than two levels or even continuous. In turn, this would require a more formal measure of association and certainly something more complex in order to consider those possibilities.

Before we start with our discussion of association measurements, it is important to introduce the various notations that will be used throughout the analysis. Exposures are represented by a random variable, normally a vector, X, while the outcome is represented by a random variable, also a vector, Y. For a binary exposure (taking values 0 and 1), X taking a value 0 represents *no exposure* to a risk factor and 1 as *exposure*. For a binary outcome, Y taking a value 0 represents *no failure* and 1 as a *failure*. In our analysis, each individual (drilling/producing well) falls into one of *four* types indexed first by exposure status and then by outcome status. The study population, then, defines probabilities as follows:

Table 1 Probabilities associated with binary explanatory and binary response variables.

	$\boldsymbol{Y = 0}$	$\boldsymbol{Y = 1}$
$\boldsymbol{X = 0}$	P_{00}	P_{01}
$\boldsymbol{X = 1}$	P_{10}	P_{11}

where,

$$P_{xy} = Pr(X = x, Y = y) \quad (1)$$

where X is the exposure of interest, Y is the outcome (casing failure) and,

$$P_{00} + P_{01} + P_{10} + P_{11} = 1 \quad (2)$$

There are two ways in which the population in question may be investigated: (1) from a prospective or (2) retrospective standpoint. Since the design adopted for the analysis is case-control design, the focus will be on the latter (i.e., retrospective). In this setting, the probability of interest is the probability of exposure, X, conditional on the outcome, Y. Mathematically, conditional probabilities for exposure, X, given outcome, Y, is defined as

$$Pr(X = x, Y = y) = \frac{\pi_{xy}}{\pi_{0y} + \pi_{1y}} \quad (3)$$

In other words, in case-control study, the interest is in evaluating the change in the exposure, X, given the outcome, Y. Case in point, assuming that we are interested in measuring the association between having an acid treatment, as the exposure, and the outcome, that is casing failure. Then, we would be interested to calculate the probability of wells that had acid treatment ($X = 1$) given that they experienced casing failure ($Y = 1$) and compare it with the wells that had acid treatment ($X = 1$) given that they did not experience casing failure ($Y = 0$). Similar comparison is conducted for the wells that did not have acid treatment ($X = 0$). In the following section, I briefly discuss the different ways of comparison, formally known as measurements of association.

2.6.1 Measurements of Association

The discussion in this section is limited to binary exposures and outcomes, however, later in this chapter, a more general discussion of association measurements will be presented in Section 2.6.2.

There are numerous ways to measure the association between an exposure, X, and the risk of the

event, i.e., the outcome, Y, in the population. They include: (1) risk difference, (2) relative risk, (3) relative risk difference and (4) odds ratio. The four methods are briefly introduced, however, only the odds ratio (OR) was selected for subsequent analysis for reasons that will be stated later.

2.6.1.1 Risk Difference (RD)

It is defined as the difference between the event incidence rates in exposed and unexposed individuals. For exposed subjects ($X = 1$) and unexposed subjects ($X = 0$), risk difference is defined as

$$RD = \lambda_e - \lambda_u \tag{4}$$

where λ_e is the outcome rate in the exposed group, while λ_u is the outcome rate in the unexposed group.

The risk difference is a useful measure when the association of exposure with the actual number of cases is of interest. A positive risk difference indicates a positive association between exposure and the outcome, a negative risk difference indicates a negative association between exposure and the outcome, and a risk difference of zero indicates no association between exposure and the outcome.

2.6.1.2 Relative Risk (RR)

It measures the proportional increase in the outcome, Y, associated with exposure, X (Porta 2014). It is defined by

$$RR = \frac{\lambda_e}{\lambda_u} = \frac{Pr(Y = 1|X = 1)}{Pr(Y = 1|X = 0)} = \frac{\pi_{11}/(\pi_{10} + \pi_{11})}{\pi_{01}/(\pi_{00} + \pi_{01})} \tag{5}$$

A relative risk greater than one indicates a positive association between exposure and the outcome,

a relative risk less than one indicates a negative association between exposure and the outcome and a relative risk equal to one indicates no association between exposure and the outcome.

2.6.1.3 Relative Risk Difference (RRD)

It measures the difference in outcome incidence at two different levels of exposure, relative to the incidence rate at some baseline level of exposure, often no exposure (Porta 2014). For λ_{e1} and λ_{e2} denoting incidence at two different levels of exposure, and λ_u the incidence at no exposure (i.e., baseline level), relative risk difference is defined as

$$RRD = \frac{\lambda_{e1} - \lambda_{e2}}{\lambda_u} = \frac{P_{e1}/[1 - P_{e1}] - P_{e2}/[1 - P_{e2}]}{P_u/[1 - P_u]} \tag{6}$$

When the second level of exposure, λ_{e2}, is the baseline level, λ_u, the RRD reduces to

$$RRD(t) = RR(t) - 1 \tag{7}$$

2.6.1.4 Odds Ratio (OR)

It compares the two probabilities $Pr(Y = 1|X = 1)$ and $Pr(Y = 1|X = 0)$. To achieve this, first, it is important to define the "odds" of an arbitrary event A with probability $Pr(A)$ to be the ratio $Pr(A)/[1 - Pr(A)]$. The odds of the outcome $Y = y$ given exposure $X = x$ are therefore

$$\frac{Pr(Y = y|X = x)}{1 - Pr\,(Y = y|X = x)} \tag{8}$$

It is evident that these odds in Eq. (8) are calculated from a prospective study but not from a retrospective, or case-control, study. The cornerstone of the analysis of case-control studies is that the ratio of the odds of $Y = 1$ given $X = 1$ and of $Y = 1$ given $X = 0$ calculated from the prospective study are the same as the formerly corresponding ratio of odds in the case-control study of $X = 1$ given $Y = 1$ to that given $Y = 0$. That is

$$\frac{Pr(Y=1|X=1)/Pr(Y=0|X=1)}{Pr(Y=1|X=0)/Pr(Y=0|X=0)} = \frac{Pr(X=1|Y=1)/Pr(X=0|Y=1)}{Pr(X=1|Y=0)/Pr(X=0|Y=0)} \quad (9)$$

Using the notation in Table 1, the odds ratio, (Edwards 1963), can be written as

$$e^{\psi} = \frac{P_{11} \cdot P_{00}}{P_{10} \cdot P_{01}} \quad (10)$$

where ψ is the *log odds ratio*. The four rates: π_{11}, π_{10}, π_{01} and π_{00} are defined in **Error! Reference source not found.**

Introducing notations for case-control study:

Table 2 Retrospective (case-control) sampling: separate samples from subpopulations Y = 0, 1 with relevant conditional probabilities.

	$\boldsymbol{Y=0}$	$\boldsymbol{Y=1}$
$\boldsymbol{X=0}$	P_{00}	P_{01}
$\boldsymbol{X=1}$	P_{10}	P_{11}
$\boldsymbol{Pr(X=x\|Y=y)}$	$P_{10}/(P_{10}+P_{00})=\theta_0$	$P_{11}/(P_{11}+P_{01})=\theta_1$

Now, using the notation introduced for the case-control study in Table 2.1(c), the odds ratio can also be written as

$$e^{\psi} = \frac{\theta_1/(1-\theta_1)}{\theta_0/(1-\theta_0)} \quad (11)$$

In many applications of case-control studies, as in this study, the probabilities of the occurrence of the outcome ($Y=1$) are small, as shown in Section 2.4.2 and Section 2.4.3. This is the main driver for adopting such a study design as it has the capability of conducting a balanced and unbiased analysis of relatively rare events (e.g., casing failure).

2.6.2 Logistic Regression

In this section, the discussion held in previous section is extended to include a more general view of how association of exposures and outcome is measured regardless of the levels that an exposure might have (Wright 1995; Menard 2002). The methods outlined are based on the *log odds ratio*, (Edwards 1963) introduced previously, using data from a case-control study. Assuming that observations are available on n independent individuals (drilling/producing wells, in our study). For the analysis of a case-control study we suppose initially that n_1 individuals, the cases, are selected at random from the subpopulation with $Y = 1$ and n_0 individuals, the controls, are selected at random from the subpopulation with $Y = 0$, generating the data shown in **Error! Reference source not found.**. By design, $n_0 \,/\, n_1$ is kept as a small integer as possible.

Table 3 Summary of data from a case-control study of n individuals; $r = r_0 + r_1$.

	$\boldsymbol{Y = 0}$	$\boldsymbol{Y = 1}$	**Total**
$\boldsymbol{X = 0}$	$n_0 - r_0$	$n_1 - r_1$	$n - r$
$\boldsymbol{X = 1}$	r_0	r_1	r
Total	n_0	n_1	n

Since the odds ratio (OR) are considered an integral part of the formulation of logistic regression, it is essential to start this discussion with describing a method for the estimation of the odds ratios. The frequencies r_1, r_0 in **Error! Reference source not found.**, which are the numbers of exposed cases and controls respectively, are values of independent binomially distributed random variables. This leads to the joint probability (Cramer 2002).

$$Pr(R_1 = r_1, R_0 = r_0) = \binom{n_1}{r_1} \theta_1^{r_1}(1-\theta_1)^{n_1-r_1} \times \binom{n_0}{r_0} \theta_0^{r_0}(1-\theta_0)^{n_0-r_0} \quad (12)$$

where θ_1 is the probability of exposure $X = 1$ for a case and θ_0 is that for a control. It is the comparison of the proportions of exposed individuals in the case group and in the control group that informs us of a possible association between exposure and outcome.

It is convenient to begin the formal analysis by considering the null hypothesis, that of *no difference* between the case and control groups in terms of exposure: $\theta_1 = \theta_0$. A sufficient statistic for the unknown common value is the total frequency $R = R_0 + R_1$, which is a binomially distributed random variable of index $n = n_0 + n_1$.We therefore condition on its observed value as the only way to obtain a test distribution not depending on a nuisance parameter. The test statistic may be either r_1 or r_0; we will choose the former.

Then, under the null hypothesis, the corresponding random variable R_1 has the distribution derived by dividing Eq. (14) by the marginal binomial probability for R:

$$Pr(R_1 = r_1, R = r) = \frac{\binom{n_1}{r_1}\binom{n_0}{r_0}}{\binom{n}{r}} \quad (13)$$

The same argument shows that for a non-null situation the conditional probability has the generalized hypergeometric form

$$Pr(R_1 = r_1, R = r) = \frac{\binom{n_1}{r_1}\binom{n_0}{r_0} e^{r_1\psi}}{\sum_{k=0}^{n_1}\binom{n_1}{k}\binom{n_0}{r-k} e^{k\psi}} \quad (14)$$

Where the odds ratio, e^{ψ}

$$e^{\psi} = \frac{\theta_1/(1-\theta_1)}{\theta_0/(1-\theta_0)} \tag{15}$$

The maximum likelihood estimate for the odds ratio is obtained by equating to zero the derivative with respect to ψ of the log of the likelihood, Eq. (16). The *estimated* odds ratio $\hat{\psi}$ is therefore obtained by solving

$$r_1 - \frac{\sum_k \binom{n_1}{r_1}\binom{n_0}{r_0} e^{k\hat{\psi}}}{\sum_k \binom{n_1}{k}\binom{n_0}{r-k} e^{k\hat{\psi}}} = 0 \tag{16}$$

Conditionally on n_1 and n_0 the frequencies in **Table 4** follow binomial distributions and the ratios r_1/n_1 and r_0/n_0 provide unbiased estimates of the conditional probabilities θ_1 and θ_0 respectively. It follows that a simple, and intuitive, estimate of the log odds ratio, (Edwards 1963), defined as

$$\hat{\psi} = log\frac{r_1(n_1 - r_1)}{r_0(n_1 - r_1)} \tag{17}$$

2.6.2.1 Logistic Model for Cohort-Study Data

Logistic regression is a useful way to estimate odds ratios (OR) after adjustment for potential confounding variables. This section expands the definition of x to include not only exposures of interest but also adjustment variables and effect-modifying variables. It describes logistic regression in the very simple situation where all explanatory variables are binary, and the data come from a cohort or case-control study design.

Let $x_E = 1$ for exposed subjects and $x_E = 0$ for unexposed subjects, and $x_A = 1$ for subjects belong to first level/category of the binary variable and $x_A = 0$ for subjects belong to the second level/category of the binary variable.

For $x = x_E$, the logistic model, (Cramer 2002), is defined as

$$logit\ (P_{x,\Delta t}(t)) = \alpha + \beta x_E \qquad (18)$$

The logistic model permits the estimation of the crude odds ratio $OR(t) = e^{\beta}$ based on regression coefficient estimates, since $logit(p) = log\ [p/\ (1-p)]$.

Table 4 Coefficient interpretation in a binary exposure logistic model.

$\boldsymbol{logit\ (P_{x,\Delta t}(t))}$			
$\boldsymbol{x_E = 1}$	$\boldsymbol{x_E = 0}$	**Difference**	**Odds Ratio (OR)**
$\alpha + \beta$	α	β	e^{β}

Adding the second variable x_A to the model permits estimation of a variable-adjusted odds ratio. The regression model, (Cramer 2002), will be adjusted as follows

$$logit\ (P_{x,\Delta t}(t)) = \alpha + \beta_A x_A + \beta_E x_E \qquad (19)$$

Table 5 Coefficient interpretation in a binary exposure logistic model.

Binary Variable	$\boldsymbol{logit\ (P_{x,\Delta t}(t))}$		**Difference**	**Odds Ratio (OR)**
	$\boldsymbol{x_E = 1}$	$\boldsymbol{x_E = 0}$		
$\boldsymbol{(x_A = 0)}$	$\alpha + \beta_E$	α	β_E	e^{β_E}
$\boldsymbol{(x_A = 1)}$	$\alpha + \beta_A + \beta_E$	$\alpha + \beta_A$	β_E	e^{β_E}

2.6.2.2 Logistic Model for Case-Control Study Data

Under the case-control design, there is no direct estimate of $P_{x,\Delta t}(t)$ in the population, but letting

$$S = \begin{cases} 1 & Sampled\ in\ the\ case/control\ sample \\ 0 & otherwise \end{cases} \quad (20)$$

one can estimate the related quantity

$$P^{CC}_{x,\Delta t}(t) = P(t \le T \le t + \Delta t | T \ge t, X = x, S = 1) \quad (21)$$

using sample proportions. Let π denote the ratio of case ($D = 1$) to control ($D = 0$) sampling probabilities for population members. Assuming there is no selection bias, π does not depend on exposure, and here we also assume it does not depend on other variables. The ratio π is typically far greater than one, as cases in the population are rare and usually have a much higher probability of being sampled than disease-free controls. The following equations show how logistic models, (Cramer 2002), for $P^{CC}(t)$ are related to logistic models for $P_{x,\Delta t}(t)$:

$$logit\left[P^{CC}_{x,\Delta t}(t)\right] = log\left[\frac{P(D = 1|x_E, x_A, S = 1)}{P(D = 0|x_E, x_A, S = 1)}\right] \quad (22)$$

$$= log\left[\overline{\frac{P(D = 1, x_E, x_A, S = 1)}{P(D = 0, x_E, x_A, S = 1)}}\right] \quad (23)$$

$$= log\left[\frac{P(S = 1|x_E, x_A, D = 1) \cdot P(D = 1|x_E, x_A).\ P(x_E, x_A)}{P(S = 1|x_E, x_A, D = 0) \cdot P(D = 0|x_E, x_A).\ P(x_E, x_A)}\right] \quad (24)$$

$$= log\ \pi + log[P(D = 1|X_E = x_E, X_A = x_A)] \quad (25)$$

$$= log\ \pi + log[P_{x,\Delta t}(t)] \quad (26)$$

Table 6 shows that, as for the odds ratio from a 2x2 table, log odds ratio coefficients in logistic models based on $P^{CC}_{x,\Delta t}(t)$ from case-control studies have the same interpretations as the analogous coefficients in logistic models for $P_{x,\Delta t}(t)$ in cohort studies. It is only the intercept term that does not have a population interpretation, for it depends not only on the baseline risk of outcome in the

population but also on the often-unknown ratio of sampling probabilities, π.

Table 6 Coefficient interpretation in a binary exposure logistic model fit to case-control data.

	$\boldsymbol{logit}\left(\boldsymbol{P}^{\boldsymbol{CC}}_{\boldsymbol{x},\Delta\boldsymbol{t}}(\boldsymbol{t})\right)$			
Binary Variable	$\boldsymbol{x_E = 1}$	$\boldsymbol{x_E = 0}$	**Difference**	**Odds Ratio (OR)**
Level/Category 1 ($\boldsymbol{x_A = 0}$)	$\log\pi + \alpha + \beta_E$	$\log\pi + \alpha$	β_E	e^{β_E}
Level/Category 2 ($\boldsymbol{x_A = 1}$)	$\log\pi + \alpha + \beta_A + \beta_E$	$\log\pi + \alpha + \beta_A$	$\beta_E + \gamma$	$e^{\beta_E+\gamma}$

2.7 Numerical Demonstration

Based on the odds ratio (OR) and following the case-control design, the different exposures in the data set had been analyzed and potential risk factors were identified. The complete results obtained from risk analysis are summarized in Appendix (A). A snapshot of the results obtained for one model is presented in Table 7. The risk analysis laid out in Table 7 can be broken down into three parts/sections: (1) model features, (2) model coefficients and (3) feature impact.

We start with the first part/section; that is "*model features*". The type of exposures, or features, that are considered for risk analysis is extremely critical for the final conclusions. The dependency between different exposures, or features, plays an integral rule in defining the association/impact different exposures have on the overall outcome.

Accordingly, special attention had been paid towards the selection of the optimal set of exposures to be considered for subsequent risk analysis. Conventionally, this challenge is tackled by trying out different combinations of the exposures, or features, initially considered. Then, risk analysis is

applied on all those combinations. Results obtained from the risk analysis are compared, and the best model is selected based on a pre-determined criterion. Detailed discussion of the techniques used for comparing different models, as well as the criteria for selecting the optimal model will be presented in the following sections (Sections 2.8 and 2.9).

The second part/section of the analysis is "*model coefficients*". As shown from the formulation of odds ratio model; presented in previous section, there is a set of parameters/coefficients, each corresponds to one of the exposures/features considered for the analysis, that needs to be calculated in order to have a model that is capable of estimating risk impacts. The estimates provided in the second part of Table 7 are the values of the coefficients corresponding to this particular set of exposures/features used for constructing that model, in particular. The estimates for the coefficient are heavily dependent on the set of exposures/features selected as a base for the model. This urges the need for having a systematic way of choosing the right set of exposures/features as a base for the model of interest, which – as mentioned earlier – will be the focus of the next section.

The third part of the analysis deals with "*exposure/feature impact*". The importance of applying risk analysis is two-fold: (1) defining the impact type, and (2) defining the impact magnitude.

First, we consider the impact "type". When analyzing the impact any exposure might have on the outcome, there are three impact types; either the exposure has slight/no impact on increasing the risk of the outcome, or the exposure increases the risk of the having the outcome, or the exposure reduces the risk of having the outcome (Table 7, Section 3, Column 3-6). As for the impact "magnitude", how far the odds ratio is from unity defines the magnitude of the impact an exposure has on either increasing or reducing the risk of the outcome (Table 7, Section 3, Column 7).

For an exposure with slight/no impact on increasing risk of the outcome, the odds ratio takes a value of unity. For instance, according to Table 7, the "*measured depth*" and "*dogleg severity measured depth*" have slight/no impact on the risk of experiencing "*casing failure*" with less than 1%.

For an exposure that increases the risk of the outcome, the odds ratio takes a value greater than unity. For instance, according to Table 7, the "*fracturing season*", "acid treatment", "*maximum inclination*" and "*lateral section shrinkage*" increase the risk of "*casing failure*" by nearly 73%-293%, 35%, 30% and 15% respectively.

Finally, for an exposure that reduces the risk of the outcome, the odds ratio takes a value less than unity. For instance, according to Table 7, the having "*cement*" support, increasing "*casing thickness*", increasing "*drill-frac time interval*" and reduce the risk of "*casing failure*" by 53%, 100% and 9%, respectively.

Table 7 Results from the application of risk analysis on the data set using odds ratio (OR). The definition of odds ratio (OR) follows case-control design.

AIC = 93.772

Model Features:

CEMENT	FRAC_SEASON	DL_SEVERE_MD
LATERAL_SHRINKAGE	MAX_INCL	BHT
DL_FREQ_10PLUS	DRILL_FRAC_INTERVAL	ACID
DL_BEND_STRESS	CSG_THICKNESS	

Coefficients:

Risk Factor	Level	Estimate	Std. Error	p-value
(Intercept)		-2.15E+01	1.77E+01	0.0225
CSG_THICKNESS		-6.77E+00	1.90E+01	0.0722
DL_BEND_STRESS		6.76E-06	2.70E-05	0.0802
FRAC_SEASON	**Spring**	1.37E+00	1.00E+00	0.0171
	Summer	5.50E-01	1.08E+00	0.0610
	Winter	1.53E+00	1.20E+00	0.0204
MD		-1.04E-04	4.02E-04	0.0795
ACID	**Yes**	3.16E-01	8.82E-01	0.0720
CEMENT	**Yes**	-6.25E-01	8.51E-01	0.0463
BHT		-3.49E-02	3.55E-02	0.0325
LATERAL_SHRINKAGE		1.47E-01	2.09E-01	0.0483
MAX_INCL		2.65E-01	1.98E-01	0.0181
DL_SEVERE_MD		2.32E-04	3.79E-04	0.0541
DRILL_FRAC_INTERVAL		1.17E-02	1.07E-02	0.0277
DL_FREQ_10PLUS		-8.85E-02	7.97E-02	0.0266

Risk Factor	Level	OR	LL	UL	Risk	Impact
CSG_THICKNESS		0.001145	7.22E-20	1.81E+13	▼	100%
DL_BEND_STRESS		1.000007	0.999954	1.000060	▼	1%
FRAC_SEASON	**Spring**	3.937058	0.55419	27.96954	▲	293%
	Summer	1.733069	0.209701	14.32291	▲	73%
	Winter	4.607039	0.437345	48.53103	▲	360%
MD		0.999896	0.999109	1.000683	▲	1%
ACID	**Yes**	1.371966	0.243384	7.733821	▲	37%
CEMENT	**Yes**	0.535253	0.101012	2.836260	▼	53%
BTH		0.965693	0.900787	1.035276	▼	4%
LATERAL_SHRINKAGE		1.157817	0.768677	1.743959	▲	15%
MAX_INCL		1.303575	0.884043	1.922200	▲	30%
DL_SEVERE_MD		1.000232	0.99949	1.000975	▲	0%
DRILL_FRAC_INTERVAL		1.011722	0.990679	1.033213	▲	0%
DL_FREQ_10PLUS		0.915278	0.782973	1.069940	▼	9%

2.8 Subset Selection

In the previous section, I have discussed the problem of the dependency between different exposures, or features, and their rule in defining the association/impact different exposures have on the overall outcome. The way this issue was addressed was through trying out different combinations of the exposures, or features, initially considered. Then, a model is fit to each combination. Results obtained from the different models are compared, and the best model is selected based on pre-determined criteria.

In this section, I briefly address the various methods that are used for comparing the performance of different models, the key distinctions between the different methods, as well as the criteria for selecting the optimal model. The first part of the discussion will be devoted to subset selection techniques. Subset selection comprises of two broad families of techniques for model selection: the best subset and stepwise model selection procedures.

2.8.1 Best Subset Selection

To perform best subset selection, a separate least squares regression best subset is fitted for each possible combination of the p predictors. That is, we fit all p models that contain exactly *one* predictor, all $\binom{p}{2} = p(p-1)/2$ models that contain exactly *two* predictors, and so forth.

Identification of the best model amongst the resulting models is implemented. The problem of selecting the best model from among the 2^p possibilities is usually broken up into two stages. First stage identifies the best model (on the training data) for each subset size, in order to reduce the problem from one of 2^p possible models to one of $p + 1$ possible models. Then, in order to select a *single best model*, we must simply choose among these $p + 1$ options. Implementation of conventional RSS and R^2 statistics for selection of the best model will always end up with a model involving *all* the variables, because the RSS of these $p + 1$ models decreases monotonically, and

the R^2 increases monotonically, as the number of features included in the models increases. Therefore, second stage, cross-validated prediction error estimators, such as AIC or BIC are used instead to select among $\mathcal{M}_0$, $\mathcal{M}_1$, ... , $\mathcal{M}_p$ models.

2.8.2 Stepwise Selection

For computational reasons, best subset selection cannot be applied with very large number of predictors p. Best subset selection may also suffer from statistical problems when p is large. The larger the search space, the higher the chance of finding models that look good on the training data, even though they might not have any predictive power on future data. Thus, an enormous search space can lead to overfitting and high variance of the coefficient estimates.

For both reasons, stepwise methods, which explore a far more restricted set of models, are attractive alternatives to best subset selection. Accordingly, stepwise selection methods were selected for subsequent analysis.

2.8.2.1 Forward Stepwise Selection

Forward stepwise selection is a *computationally efficient alternative* to best forward stepwise selection subset selection. While the best subset selection procedure considers all 2^p possible models containing subsets of the p predictors, forward stepwise considers a much *smaller* set of models.

Forward stepwise selection begins with a model containing no predictors, and then adds predictors to the model, one-at-a-time, until all the predictors are in the model. In particular, at each step the variable that gives the greatest additional improvement to the fit is added to the model.

Unlike best subset selection, which involved fitting 2^p models, forward stepwise selection involves fitting one null model, along with $p - k$ models in the k^{th} iteration, for $k = 0, \ldots, p - 1$. This

amounts to a total of $1 + p\ (p + 1)/2$ models, which constitutes a substantial difference. Accordingly, forward stepwise selection has computational advantage over best subset selection. However, forward stepwise does not guarantee to find the best possible model out of all 2^p models containing subsets of the p predictors especially with high-dimensional setting (where $n < p$).

2.8.2.2 Backward Stepwise Selection

Like forward stepwise selection, backward stepwise selection provides an efficient alternative to best subset selection. However, unlike forward stepwise selection, it begins with the full least squares model containing all p predictors, and then iteratively removes the least useful predictor, one-at-a-time.

Like forward stepwise selection, the backward selection approach searches through only $1 + p\ (p + 1)/2$ models, and so can be applied in high-dimensional settings (where p is too large) to apply best subset selection. Also, like forward stepwise selection, backward stepwise selection is not guaranteed to yield the best model containing a subset of the p predictors. Backward selection requires that the number of samples, n, is *larger* than the number of variables, p, (so that the full model can be fit). In contrast, forward stepwise can be used even when $n < p$, and so is the only viable subset method when p is very large.

2.9 Choosing the Optimal Model

As discussed in previous section, subset selection techniques are used for comparing the performance of different models fit to different combinations of exposures, or feature. Here, we discuss the different criteria that could be used to determine which of these models is best. In order to select the best model with the least test error, we need to estimate this test error. There are *two* common approaches:

1) **Indirect estimation of test error**: making an adjustment to the training error to account

for the bias due to overfitting.

2) **Direct estimation of the test error**: using either a validation set approach or a cross-validation approach.

Here, only indirect test error estimator approach is considered for discussion.

2.9.1 Akaike Information Criterion (AIC)

The AIC criterion, (Akaike 1973), was initially defined for evaluating and comparing regression models with different number of predictors. In case of classification model, a modified version of AIC could be used and is defined as follows

$$AIC = -2(L_d - L_S) + 2d \tag{27}$$

where L_d is the log-likelihood of the model with d predictors (for $d \ll n$), while L_S is the log-likelihood of the "saturated" model fitted on all n predictors. d is the number of predictors subset from the total number of n predictors. The term $-2(L_d - L_S)$ is defined as the "deviance".

Deviance is a goodness-of-fit statistic for a statistical model that is used to compare two different models. For classification model, deviance can be defined as the log-likelihood ratio of the full (saturated) model compared to the reduced model. Accordingly, AIC is – essentially – deviance penalized for model complexity. The "lower" the residual deviance, the "lower" the AIC, the "better" the model.

2.9.2 Bayesian Information Criterion (BIC)

BIC is derived from a Bayesian point of view, (Schwarz 1978), but similar to AIC, with the only difference that an "empirical" estimate of the model complexity; the term "$2d$" in Eq. (27), is used as the penalty. For a classification model with d predictors, the BIC is, then, given by

$$BIC = -2(L_d - L_S) + 2 * 2^{entropy} \quad (28)$$

Like AIC, the BIC will tend to take on a "small" value for a model with a "low" deviance, and so generally we select the model that has the "lowest" BIC value.

2.10 Numerical Demonstration

In our analysis, Akaike Information Criterion (AIC) was selected as the criterion for identification of the best model to proceed with. Results of the AIC values for the different models analyzed are presented in Figure 34. Based on a tradeoff between the AIC reduction and the model size, *model 10* has been selected as the "best" compared to other models analyzed.

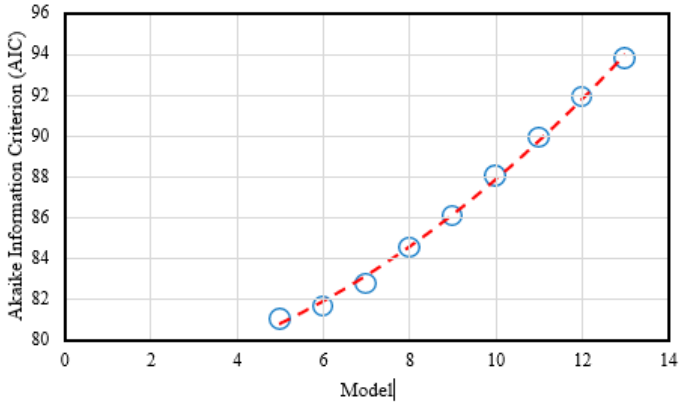


Figure 34. AIC results for the best models of each size for the collected Granite Wash data set.

Graphical representation of the risk analysis results for the best model selected is presented on a radar plot (Figure 35). In this plot, impact values of the different exposures/features are presented. A dotted polygon was set at value 1 corresponding to the "no" impact zone, where odds ratio is equal to unity. Any exposures that lie inside the polygon are considered to have positive impact (i.e., reduce the risk of casing failure). On the other hand, any exposures that lie outside the polygon are considered to have negative impact (i.e., increase the risk of casing failure).

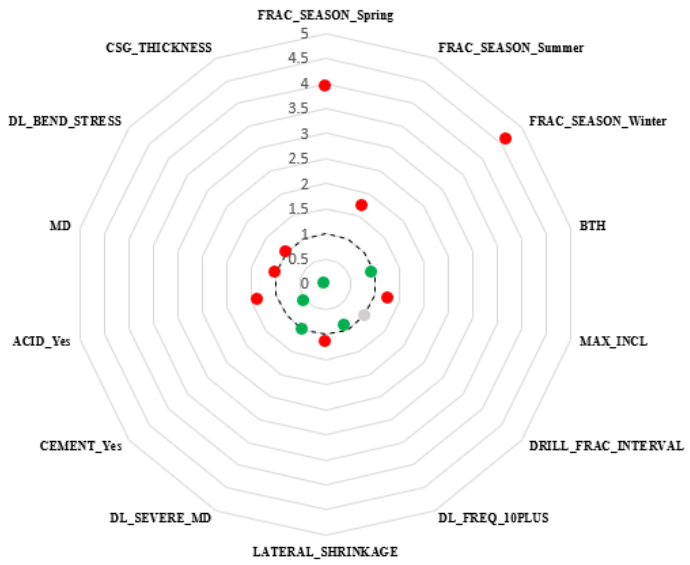


Figure 35. Risk factor analysis results for best selected model (based on AIC).

For completeness and comparison purposes, results of risk analysis implemented on all the different models with various combinations of the exposures/features are summarized in Appendix A. A graphical representation of the risk impacts corresponding to each exposure/feature for each model is presented in Figure 36.

In Figure 36, twelve plots corresponding to the twelve different exposures that were settled on based on the analysis are presented. For each plot, the x-axis indexes, or lists, the different model numbers, while the y-axis represents the impact value. The color of the plot corresponds to the "type" of the impact that exposure has on the outcome (casing failure). The red corresponds to increased risk, the green corresponds to reduced risk and the grey corresponds to slight/no risk.

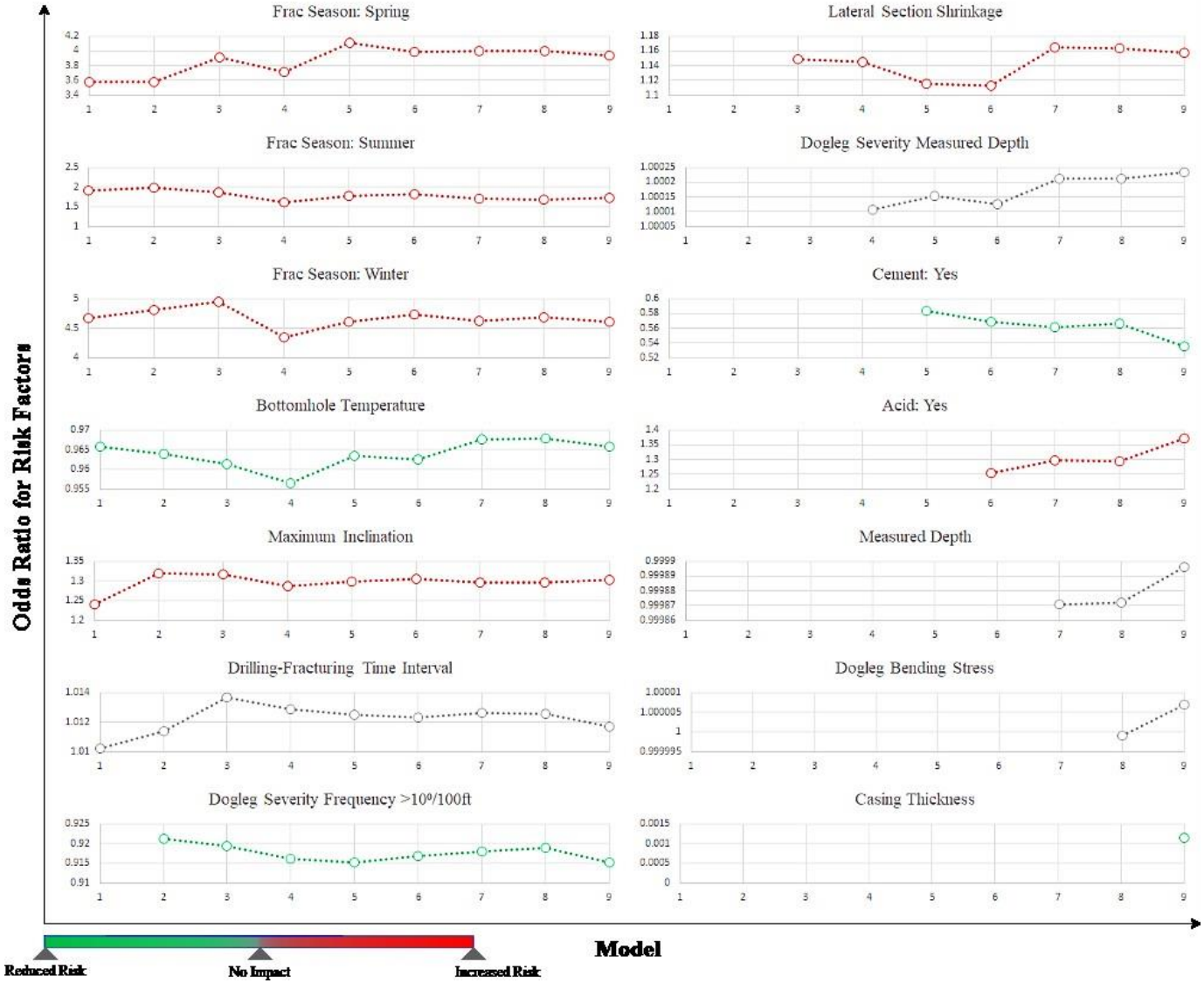


Figure 36. Odds ratio (OR) for various risk factors in a variety of model combinations.

2.11 Conclusions

In this chapter, I addressed one of the limitations of previous contributions done in the area of casing failure, that is, impact quantification of potential risk factors. To that end, I designed a data-driven workflow based on risk-factor analysis family of techniques that gives the allowance to evaluate the impact type, as well as the magnitude for potential risk factors. Quantification of various risk factors' impacts adds an invaluable insight to mitigating and, ultimately, avoiding future casing failures. This could be achieved through integration of such information in the design of the proposed "casing failure mitigation" tool, to be discussed in Chapter 4.

The data-driven workflow, designed in this chapter, is based on three main components: (1) definition of proper study design, (2) definition of proper sampling approach and (3) definition of suitable association measurement.

Regarding study designs, two major designs were initially discussed: cohort and case-control study design. Based on the direction of inference, sampling design and analysis biases, case-control design was adopted. In case-control study design, knowing the outcome (casing failure), we move backwards to test the different exposures, defines the potential risk factors and measure their impact.

Regarding sampling techniques, two approaches were presented and briefly discussed: (1) unmatched and (2) matched approach. Motivated by the structure of each analysis approach and the information provided in the data set, matched analysis approach was selected to proceed with. Following that approach, each case (well experienced casing failure) was matched with three controls (well that did not experience casing failure) based on different features (e.g., the field, the location, the formation, etc.).

To conduct the comparison between cases and controls, several association measurements were introduced and briefly addressed including risk difference (RD), relative risk (RR), relative risk difference (RRD) and odds ratio (OR). Odds ratio was, then, selected as the criterion for measuring the association/impact of different exposures on the outcome. Using the odds ratio, I managed to identify the potential risk factors, the type of impact they have (positive, negative or slight/no) and measure the magnitude of their impact.

Based on the initial results of the analysis, only thirteen features – out of the initial twenty-six features – were considered as potential risk factors of casing failure. Those potential risk factors were generally related to casing records, fracture records and drilling conditions.

Owing to the implementation of odds ratio (OR), I also managed to define the type of impact each potential risk had on the occurrence of casing failure; either positive, negative, or slight-to-no impact. Results showed that, for instance, "*fracturing season*" and "*maximum inclination*" had a negative impact on casing failure, while "*cementing*", "*casing thickness*" and "*drill-frac time interval*" had a positive impact. As for, "*measured depth*" and "*dogleg severity measured depth*" both had nearly no impact on the occurrence of casing failure.

In addition to evaluating risk type, using odds ratio (OR), we managed to measure the magnitude of the impact each potential risk factor had on the overall probability of casing failure. For instance, fracturing during spring turned out to increase the risk of casing failure by over 200%, an increase of one unit in the lateral section shrinkage increases the risk by 15%. On the other hand, cementing reduces the risk of casing failure by nearly 54%, while having a casing thickness greater than 0.65 in (P110) tends to reduce the risk of casing failure by nearly 90%. Those results were beneficial for generating probability-impact risk assessment matrix (PI-RAM) that was addressed later in the study.

Due to the dependency between different exposures, several combinations of the presented exposures/features were tested. A model was fit and analyzed for each combination. Then, the results from the different models were compared using subset selection techniques. Based on computational and statistical reasons, the stepwise family of techniques were selected for fitting and comparing the performance of the different models.

To define the best model to continue with, different criteria were discussed including AIC and BIC. Since the aforementioned criteria were, initially, designed for regression models, they had to be modified in order to suit the type of the machine learning algorithms implemented in the analysis (classification models). But it was the AIC criterion that, eventually, was chosen to identify the best model with optimal set of features/exposures.

CHAPTER III
SURVIVAL ANALYSIS

Reader Guide: In chapter 2, having initial set of exposures and using the designed workflow based on risk-factor analysis, I managed to identify the potential risk factors that are strongly associated with the occurrence of casing failure in the provided data set. In addition, I managed to identify their impact on the outcome; whether positive, negative or slight. Moreover, I managed to quantify their "overall" impact on the outcome and rank them accordingly.

In this chapter, I further my investigation by addressing the following concerns: (1) having a granular view of the impact each risk factor has on casing failure, in other words, testing the impact of the various subcategories within each potential risk factor, and (2) identifying the depths throughout the well that are most vulnerable to casing failure based on the provided information for different risk factors. To tackle those two challenges, I am designing another workflow that is based on another classical statistical analysis technique, known as "survival analysis".

Discussion starts with an overview of survival analysis family of techniques and their applications, along with a brief layout of the modifications/adjustments proposed by the author in order to align with this chapter's objectives (Section 3.1). Then, I present various survival/hazard functions used for describing an event, e.g., casing failure, distribution (Section 3.2). Those functions are important in terms of describing the behavior in which survival probability changes with respect to analysis scale (i.e., measured depth).

Following that, I proceed with discussing various methods used for estimating the aforementioned survival distributions (Section 3.3). I start with non-parametric estimators, in particular Kaplan-Meier estimator (Section 3.3.1). Due to their flexibility, non-parametric estimators provide the perfect solution for analyzing the significance of various subcategories within potential risk factors. I, then, move to semi-parametric estimators, in particular proportional hazards (Cox) model; where I address some of the limitations in the non-parametric counterparts (Section 3.3.2). Finally, I highlight the work done in this chapter, as well as the major achievements from the analyses conducted (Section 3.4).

3.1 Introduction

In Chapter 2, I highlighted one of the major limitations previous contributions – particularly in the area of casing failure – had, that is, their inability to provide a viable measure to mitigate, or even, avoid casing failure. One reason that contributed to such limitation was the inability to quantify the impact that various potential risk factors have on the occurrence of the casing failure. I tackled that challenge by designing a data-driven workflow, based on risk analysis techniques, that has the capability of evaluating the impact type, as well as the magnitude of the various risk factors.

Another reason that contributed to this limitation was the inability to tackle the impact of potential risk factors throughout the well, in addition to, how variation in a particular risk factor might affect the probability of casing failure accordingly. In this chapter, I tackle that challenge through designing another data-driven workflow based on "survival analysis" techniques. This workflow will give the allowance to evaluate the significance of the different subcategories within each potential risk factor, in addition, track their impact throughout the well length.

Survival analysis is a well-established statistical methodology that is of interest to researchers in many different fields including biology, sociology, technical reliability, econometrics, etc. The motivation behind the application of survival analysis is, normally, to analyze event history and predict the occurrence of that event/outcome with respect to time (Miller 1997). In other words, survival analysis is concerned about the time needed until a specific event/outcome occurs.

In survival analysis, different methodologies are used to describe the occurrence of any event/outcome such as survival curves and hazard rates, and to analyze the dependence on covariates. Survival and hazard terminologies are interchangeable. In essence, the two approaches complement each other, and both could be used to convey the same output. Traditionally, survival-based methodologies were used for positive outcomes, while hazard-based methodologies were

used when negative outcome is the case.

Although techniques involved in survival analysis normally include time scale in their construction and formulation, yet – motivated by our own research interests – the analysis scale has been tweaked. So, the data would, now, be presented and analyzed with respect to "reached measured depth" instead of the conventional analysis scale that is "time" (Figure 16). The author believes that this slight modification has the potential for resolving the objectives set for this chapter.

For classical survival analysis, where "time" is the analysis scale, individuals considered for the study are assumed to have the same starting point, usually the date when the study begins. In the context of this study, the definition of the starting point is different. Since the analysis scale has been adjusted to reflect the "measured depth" reached before the occurrence of the casing failure, the starting point is now defined in terms of spatial coordinates. This point was selected to be the datum/well surface. Hence, the depth will be measured from surface until the point of failure (Figure 37).

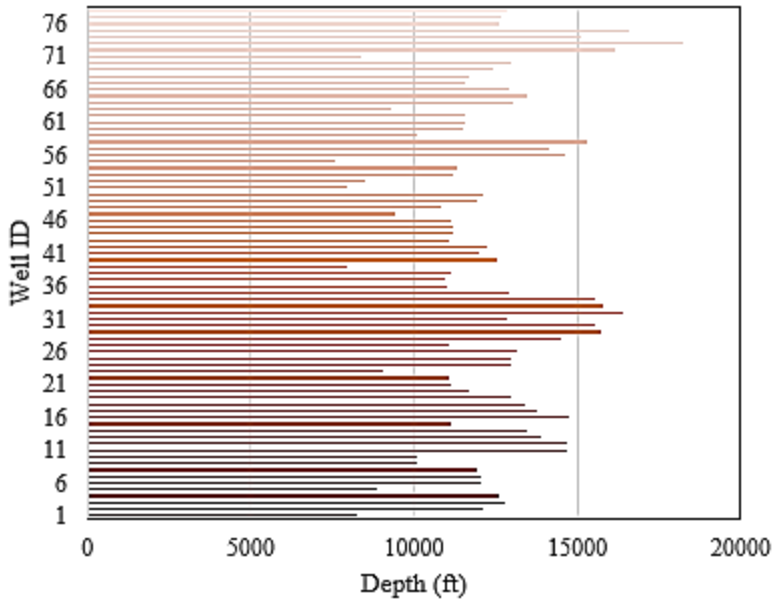


Figure 37. Line plot in the proposed analysis scale (x-axis); that is the measured depth reached until casing failure, for the eighty wells considered in the study compiled from the Granite Wash formation (y-axis).

Before we proceed with the discussion, it is important to introduce some essential elements and key notes that will be used in subsequent analysis. Based on the previously implemented risk analysis (Chapter 2), conclusions were drawn regarding the potential risk factors among different exposures initially considered and the association, or impact, they had on the outcome of interest (casing failure). Those risk factors were selected as a starting point for survival analysis.

Risk factors deduced from the risk analysis included both numerical variables, as well as categorical variables. Numerical risk factors have been categorized into two or more classes/levels for reasons that has to deal with some of the implemented survival estimators, e.g. Kaplan-Meier estimator (detailed discussion is provided in Section 3.3.1). The number of classes/levels was purely defined based on the statistical inference from the frequency analysis conducted on the data set that was compiled from Granite Wash formation. A snapshot of the results obtained from

frequency analysis is presented in Figure 38. The remaining results from the frequency analysis are summarized in Appendix (B). As for the final risk factors and their categories, that will be used for proceeding analysis, they are summarized in Table 8.

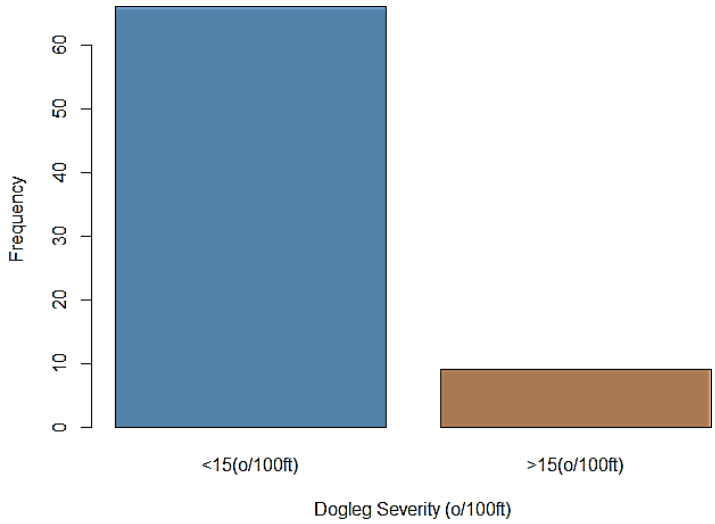


Figure 38. Graphical representation of the two statistically significant categories of "dogleg severity" risk factor; represented by blue and brown colors, obtained from the application of frequency analysis on "dogleg severity".

Table 8 Description of the different risk factors considered in survival analysis.

Variable	Description	Codes/Values
MD	Measured Depth	ft
CSG_THICKNESS	Casing Wall Thickness	in
FRAC_TEMP	Fracture Temperature	ºF
BHT	Bottomhole Temperature	ºF
DRILL_FRAC_INTERVAL	Drilling-Fracturing Time Interval	Days
DL_SEVERITY	Dogleg Severity	º/100ft
DL_FREQ_10PLUS	Frequency of Doglegs > 10º/100ft	Dimensionless
DL_SEVERE_MD	Measured Depth of Maximum Dogleg	ft
MAX_INCL	Maximum Inclination	º
DL_BEND_STRESS	Dogleg Bending Stress	lbf
LATERAL_SHRINKAGE	Shrinkage of Lateral Section	ºF/100ft
FRAC_SEASON	Fracturing Season	1 = Fall
		2 = Spring
		3 = Summer
		4 = Winter
ACID	Acid Usage	1 = Yes
		0 = No
CEMENT	Cement Usage	1 = Yes
		0 = No
FAIL_DEPTH	Measured Depth of Failure	ft
STATUS	Censoring	1 = Failure
		0 = Survived

3.2 Representations of Survival Distribution

In this section, I briefly address the basic functions that are used for defining a survival distribution. Those functions are critical in terms of describing the behavior in which survival/hazard probabilities change with respect to analysis scale (measured depth). These functions include: (1) the survival function, (2) the hazard function, (3) the cumulative distribution function. As mentioned in Section 3.1, all definitions, formulations and methodologies will be adjusted to align with the one of the research concerns, that is, defining the depths that are most vulnerable to casing failure based on analyzing the provided information for the different risk factors and estimating survival, or conversely the hazard, probability accordingly.

3.2.1 The Survival Function, $\boldsymbol{S(d)}$

The survival function defines the expected proportion of individuals (drilling/producing wells) that have not yet experienced the event/outcome (casing failure) by the specified measured depth, d (Figure 39-a). It should be noted that survival function does not consider the impact of different exposures/risk factors on the overall probability. In other words, the survival function provides the *unconditional* probability that the event of interest has not happened by depth d, formally,

$$S(d) = pr(D > d) \qquad (29)$$

where D is a random variable that denotes the survival depth. Like any probability distribution, the survival function ranges between 0 and 1. It takes the value 1 at depth d = 0 (the surface, in our case), and tends to decrease, or remain constant, as the measured depth, d, increases; since more and more individuals (drilling/producing wells) will experience the event of interest (casing failure), and never drops below 0.

3.2.2 The Hazard Function, $\boldsymbol{h(d)}$

The hazard function is, simply, the instantaneous failure rate. In other words, it is the probability that, given that a subject has survived up to depth d, it is expected to fail in the next small interval of depth, $[d, d + \delta)$, divided by the length of that interval, δ, (Figure 39-b). Mathematically, this may be expressed as (Maxim 2008)

$$h(d) = \lim_{\delta \to 0} \frac{pr(d < D < d + \delta \mid D \geq d)}{\delta} \qquad (30)$$

As mentioned before (Section 3.1), hazard and survival functions, are complementary. The core aspect that differentiates both functions is the perspective from which each function deals with the problem of interest. The survival function assumes an individual (drilling/producing well) will survive up to a certain measured depth. On the other hand, the hazard function assumes that an individual (drilling/producing well) will fail after a certain depth. Based on the outcome/event of interest or the researcher's convenience, it is decided which notion to adopt in subsequent analysis.

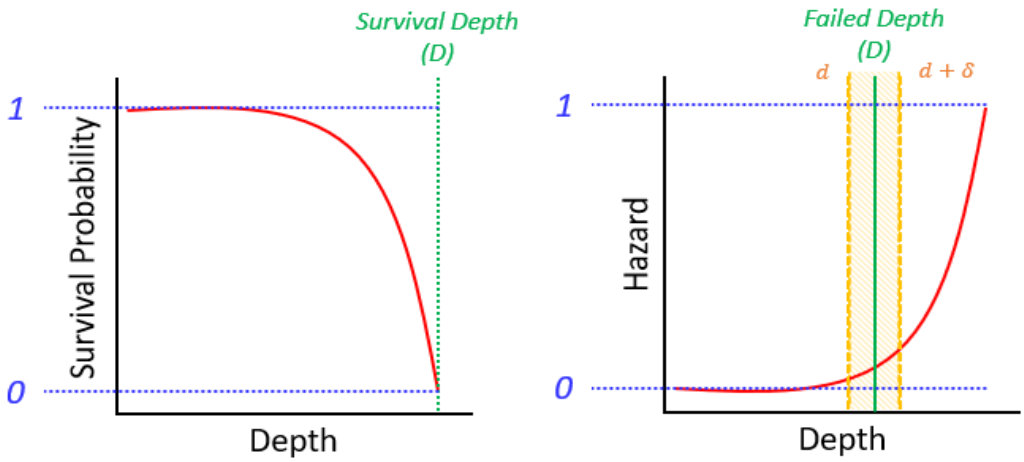


Figure 39. (left**)** Survival function with high initial probability (~ 1) that decreases exponentially until hits "zero" probability threshold at the depth of failure, D. (Right) Hazard function with low (~zero) initial probability that increases exponentially until reaches unity at the depth of failure, D.

3.2.3 The Probability Density Function, $\boldsymbol{f(d)}$

Now, after introducing the two key functions (survival and hazard functions) used for describing any survival distribution, it is equally important that we link the two together. This will give the allowance for calculating one function given that the other is known. This sort of conduit is known as the probability density function.

Before we proceed with the definition of the probability density function, it is imperative to introduce another simple concept; that is the cumulative distribution function (also known as cumulative risk function). Cumulative distribution function can be considered as the complement of the survival function, (Miller 1971), and is defined as

$$F(d) = pr(D \leq d) \tag{31}$$

This leads us to the definition of the probability density function, (Ord 1971), which is simply the rate of change of the cumulative distribution function, or minus the rate of change of the survival function, Formally

$$f(d) = -\frac{\partial}{\partial d} S(d) = \frac{\partial}{\partial d} F(d) \tag{32}$$

The *hazard function* can, then, be related to the *probability density function* and *survival functions*, (Miller 1971), through the following equation

$$h(d) = \frac{f(d)}{S(d)} \tag{33}$$

This means that the hazard at depth d is the probability that an event/outcome (casing failure) will occur in the neighborhood of depth, d, divided by the probability that the individual (drilling/producing well) will survive until the measured depth, d. It is this relationship that allows us to compute the *survival function*, $S(d)$, corresponding to a *hazard function*, $h(d)$.

3.3 Survival Curve Estimation

The upcoming sections build on the definitions and functions that were introduced in previous sections (Section 3.2.1 and Section 3.2.2). In the upcoming sections, I proceed with the discussion of different techniques that are used for estimating survival, or conversely hazard, probabilities at certain measured depths.

Generally, there are three broad families of estimators that are used for the aforementioned purposes. Here, the discussion is limited to only two: (1) non-parametric estimators and (2) semi-parametric estimators. Non-parametric estimators, by definition, provide estimates of survival probabilities based purely on the history data set. On the contrary, semi-parametric estimators have a more formal way of conveying the survival probabilities based on the assumption that survival functions follow parametric distribution.

For completeness and comparison purposes, the foundation of both families of techniques will be

laid out in detail and results of the analysis using both approaches will be presented, as well.

3.3.1 Nonparametric Survival Curve Estimation

Here, I discuss, in detail, the first family of survival curve estimators; the non-parametric estimators. The focus will be on the Kaplan-Meier estimator, since it is one of the most widely used non-parametric estimators of the survival function (Kaplan 1958). This estimator is, basically, the product over the various failure depths, d_i, of the conditional probabilities of surviving up to the next failure depth. Mathematically, it is defined by

$$S(d) = \prod_{d_i \leq d} \left(\frac{n_i - l_i}{n_i} \right) \tag{34}$$

where n_i is the number of wells at risk – yet did not fail – at depth d_i, and l_i is the number of the wells who already failed at that measured depth.

To illustrate how survival curves are initiated using Kaplan-Meier estimator, a trivial example is provided based the baseline survival curve, known as the "null" model. In this case, the impact of the different risk factors (e.g., acid treatment, cementing, etc.) is totally ignored and the focus is only on final outcome; that is the wells at risk of failure, n_i, and wells that already failed, l_i (column 2 and 3 in Table 9). We use those information as inputs for the estimator, Eq. (34), and calculate the survival probability at each of the known failure depths (column 4 in Table 9). The output of those calculations is the corresponding survival curve (Figure 40).

Table 9 Basic statistics for Kaplan-Maier survival estimator with 95% confidence intervals.

					Confidence Interval	
Depth	**Wells at risk**	**Wells failed**	**Survival Function**	**Std. dev.**	**lower 95%**	**upper 95%**
d_i	n_i	l_i	$S(d)$	σ	-	-
7927	60	1	0.983	0.0165	0.9515	1
7954	59	1	0.967	0.0232	0.9223	1
8280	58	1	0.95	0.0281	0.8964	1
8400	57	1	0.933	0.0322	0.8723	0.999
8846	56	1	0.917	0.0357	0.8493	0.989
9052	55	1	0.9	0.0387	0.8272	0.979
9297	54	1	0.883	0.0414	0.8057	0.968
9403	53	1	0.867	0.0439	0.7848	0.957
10074	52	1	0.85	0.0461	0.7643	0.945
10100	51	1	0.833	0.0481	0.7442	0.933
11065	47	1	0.816	0.0502	0.7228	0.92
11096	45	1	0.797	0.0523	0.7013	0.907
11129	44	1	0.779	0.0542	0.6801	0.893
11537	37	1	0.758	0.0566	0.655	0.878
12830	23	1	0.725	0.0631	0.6117	0.86
12920	22	1	0.692	0.0683	0.5707	0.84
12938	21	1	0.659	0.0725	0.5315	0.818
14705	11	1	0.599	0.0873	0.4506	0.797
16558	2	1	0.3	0.2164	0.0728	1

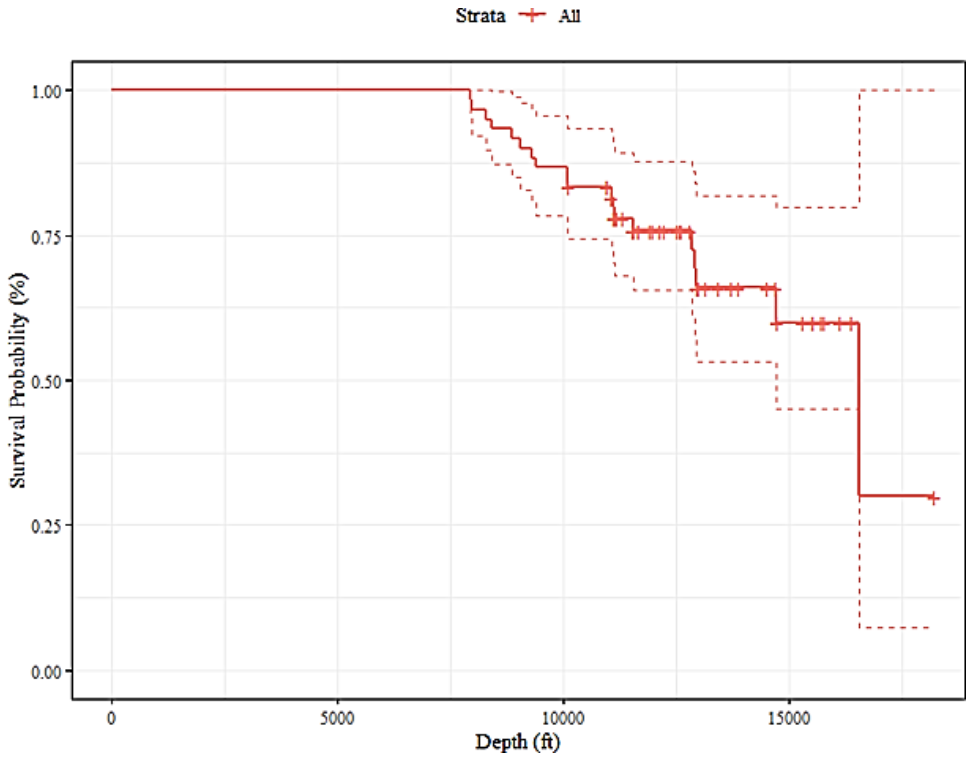


Figure 40. Survival probability developed in this study of null model with respect to measured depth. The estimates of survival probabilities at each measured depth are represented in "solid" red line. The confidence levels: upper and lower, are represented by the top and bottom "dotted" red lines, respectively.

In Figure 40, survival probability, i.e., the probability of not experiencing casing failure, is represented on the y-axis, while the reached (drilled) measured depth is represented on the x-axis. According to Figure 40, the survival probability is "inversely" proportional to the reached measured depth. Since the history wells were either deviated or horizontal, the probability of casing failure was expected to increase with increasing the depth due to multiple reasons including: (a) increased dogleg severity and maximum inclination, (b) increased temperature differentials, (c) increased shrinkage of lateral section and (d) inexistence of proper cement support for some cases. Those reasons led to increased stresses on the casing string, hence, increased risk of casing failure.

The previous example (explained in Table 9 and Figure 40), however, neglected all the potential risk factors (i.e. null model) and the sole purpose of introducing it was to familiarize the reader with the process of constructing survival curves using non-parametric estimators, more specifically Kaplan-Meier estimator. The actual analysis is introduced in the upcoming section, where the impact of various risk factors/covariates (e.g., acid treatment, cementing, dogleg severity, etc.) on survival probability will be considered.

3.3.1.1 Comparing Survival Curves

In this section, we address the objective of adopting this family of techniques, in particular, and survival analysis, in general. As mentioned in the beginning of this chapter that one of the research's interests – and the focus of this chapter – is to define the places/depths throughout the well that are most vulnerable to casing failure. This require estimating the probability of survival, or more conveniently, the probability of failure at various depths. To that end, we have adopted the classical "survival" analysis techniques with little adjustment to the analysis scale to align with our interests.

In the previous section, we discussed who survival curves can be constructed using one of the two major estimation approaches; that is non-parametric estimators (Kaplan-Meier estimator). Yet, we totally ignored the impact of different risk factors on the survival curves. In this section, we discuss – in detail – how the impact of different risk factors can be tested and incorporated in the construction of survival curves.

To study the impact of various risk factors on survival curves, each risk factor is considered separately. Then, each risk factor is categorized into two – or more – levels/groups. For each level/group, a survival curve is constructed. Survival curves for all levels/groups within each risk factor are compared. Based on that comparison, it can be concluded for every depth, the impact

each of those levels/groups have on the survival curve. Also, it can be decided which of those groups have the highest impact.

We start our discussion with a simple case; assuming a two-level risk factor (acid treatment, for instance). The goal – as mentioned – is to test the equivalence of the two groups (i.e., the well had an acid treatment, 1, or not, 0). Typically, we are interested in testing a null hypothesis, H_0, that the two-population means are *equal* versus an alternative hypothesis, H_A, that the means are *not equal* (two-sided test) or that the mean for one group is *greater than* that for the other group (one-sided test). In other words, we initially assume that both groups (Yes Acid and No Acid) have the same impact on survival curves, and we want to examine this assumption (i.e., null hypothesis). If our initial assumption (i.e., null hypothesis) is true, then, we should expect the alignment of the two survival curves corresponding to the two groups and we end up with one survival curve as we had in the case in previous section. If the initial assumption is false, then we should expect two different survival curves for the two groups. Based on the condition we have (having acid treatment, for instance), one can follow the survival curve that corresponds to that condition and estimate the failure probability at the depth of interest. If the conditions change with depth, then one can move between the different survival curves and estimate the failure probability, accordingly.

The way we test the null hypothesis is through computing a test statistic from the observed data and reject the null hypothesis if the test statistic exceeds a particular constant. The significance level (known as, p-value) of the test is defined as the probability that we reject the null hypothesis when the null hypothesis is in fact true.

Regarding the non-parametric tests of equivalence of two survival functions. We adopt the *log-rank* test, χ^2, (Nathan 1966) which is defined as follows

$$\frac{{U_a}^2}{V_a} \sim \chi_1^2 \qquad (35)$$

where,

$$U_a = \sum_{i=1}^{N} (d_{ai} - e_{ai}) = \sum d_{ai} - \sum e_{ai} \qquad (36)$$

$$V_a = var(U_a) = \sum v_{ai} \qquad (37)$$

e_{ai} is defined as the expected mean, while d_{ai} is the observed mean and v_{ai} is defined as the expected variance. N is the total number of wells.

3.3.1.2 Numerical Demonstration

The log-rank test, or simply Chi-square, is calculated for each level/group (Table 10; column 6) within each risk factor using Eqs. (35)-**Error! Reference source not found.**. Inputs for those equations can be found in Table 10; columns 3-5. Then, Chi-square values for all levels/groups within each risk factor are added together to form the overall Chi-square value for each risk factor. The Chi-square then are compared with pre-determined thresholds (Table 11; column 3). Based on that comparison, the null hypothesis is either accepted or rejected. According to results shown in Table 11; column 4, every level/group within each risk factor has a distinct and different impact on the survival curves. This conclusion is critical when deciding which survival curve to follow when estimating failure probability at the depth of interest given specific set of values for the different risk factors.

Table 10 Data and calculations of the log-rank test statistic.

Feature	Class	N	$\sum d_{ai}$	$\sum e_{ai}$	χ_1^2	p-value
Fracture Season	**Fall**	12	2	3.58	0.701	
	Spring	17	7	5.24	0.588	
	Summer	20	5	7.42	0.787	
	Winter	11	5	2.76	1.827	
					3.9	0.3
Acid	**No**	43	13	12.19	0.0532	
	Yes	17	6	6.81	0.0953	
					0.2	0.7
Cement	**No**	28	10	7.94	0.534	
	Yes	32	9	11.06	0.383	
					1	0.3
Dogleg Bending Stress (lbf)	**< 95800**	54	17	16.49	0.0156	
	> 95800	6	2	2.51	0.1027	
					0.1	0.7
Measured Depth (ft)	**< 13500**	41	13	10.86	0.42	
	> 13500	19	6	8.14	0.561	
					1.1	0.3
Bottomhole Temperature (ºF)	**< 166**	31	11	8.74	0.584	
	> 166	29	8	10.26	0.498	
					1.1	0.3
Shrinkage of Lateral Section (in/100F)	**< 10.7**	4	2	0.75	2.0856	
	> 10.7	56	17	18.25	0.0857	
					2.2	0.1
Maximum Inclination (degree)	**< 95.5**	50	16	15.56	0.0124	
	> 95.5	10	3	3.44	0.0561	
					0.1	0.8
Measured Depth of max. Dogleg	**< 10900**	49	16	14.11	0.254	
	> 10900	11	3	4.89	0.733	
					1.1	0.3
Frequency of Dogleg > 10	**< 13**	30	10	8	0.502	
	> 13	30	9	11	0.365	
					0.9	0.3
Dogleg Severity (º/100ft)	**< 15**	51	16	15.94	0.00019	
	> 15	9	3	3.06	0.000994	
					0	1
Fracture Temperature (ºF)	**< 49.5**	18	8	4.81	2.111	
	> 49.5	42	11	14.19	0.716	
					2.9	0.09
Drilling-Fracturing Time Interval (Days)	**< 75**	58	18	18.465	0.0117	
	> 75	2	1	0.535	0.4042	
					0.4	0.5

Table 11 Data and calculations of the log-rank test statistic.

Feature	χ_1^2	$max(\chi_1^2)$	Reject H_0 ?
Fracture Season	3.9	7.815	Yes
Acid	0.2	3.841	Yes
Cement	1	3.841	Yes
Dogleg Bending Stress (lbf)	0.1	3.841	Yes
Measured Depth (ft)	1.1	3.841	Yes
Bottomhole Temperature (F)	1.1	3.841	Yes
Shrinkage of Lateral Section (ft/100F)	2.2	3.841	Yes
Maximum Inclination (degree)	0.1	3.841	Yes
Measured Depth of max. Dogleg	1.1	3.841	Yes
Frequency of Dogleg > 10	0.9	3.841	Yes
Dogleg Severity (º/100ft)	0	3.841	Yes
Fracture Temperature (ºF)	2.9	3.841	Yes
Drilling-Fracturing Time Interval (Days)	0.4	3.841	Yes

Based on the former discussion, a conclusion was drawn that *different* levels/groups within each risk factor of interest have *different* impacts on the outcome (casing failure). Now, it is imperative to have a more focused/granular view on how changing between levels/groups within each risk factor affect the overall survival/failure of the drilling/producing well.

To that end, I seek to construct survival curves for each level/group within each and every risk factor. Then, combine survival curves for all levels/groups within a risk factor in a single plot for comparison. The relative positioning of survival curves with respect to each other will reflect the relative impact each level/group has on the survival curve compared to other levels/groups within the same risk factor.

Case in point, considering the "fracturing season" as one of the potential risk factors. The goal,

then, is to differentiate the impacts of the different levels/groups; winter, spring, fall and summer, on the outcome (casing failure). In Figure 41, survival probability, i.e., probability of not experiencing casing failure, is represented on the y-axis, while the reached measured depth is represented on x-axis. The four categories within “fracturing season”; fall, spring, summer and winter, are represented by red, green, cyan and mauve, respectively.

As shown in Figure 41, survival curves corresponding to “winter” and “spring” seasons are, generally, in a lower position than those corresponding to “fall” and “summer”. Considering the fact that we are, essentially, comparing “survival” probabilities, this relatively *low* positioning indicates that drilling during “winter” or “spring” puts the well in a higher risk of failure than drilling during the other two seasons. This is understandable since, for relatively cold seasons, the temperature differential between the fracturing fluid and reservoir is high and the shrinkage in lateral section will be high, as well. This, in turn, induces additional stresses on the casing string, hence, imposes a higher risk of casing failure.

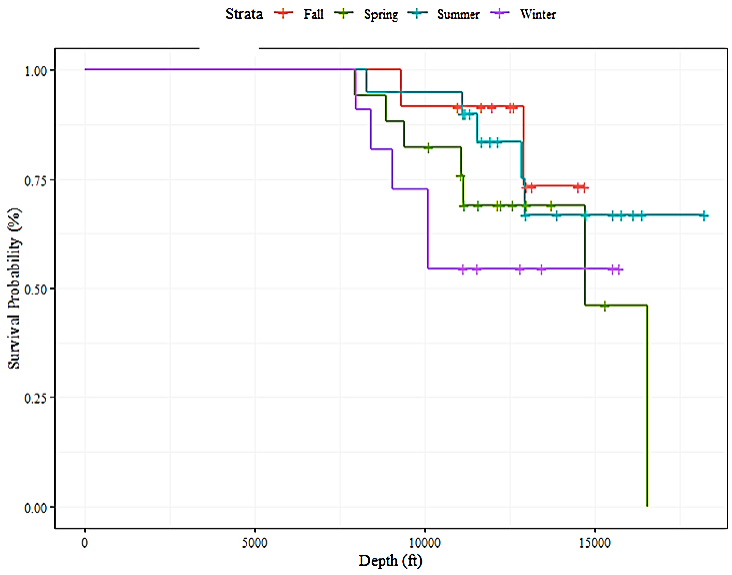


Figure 41. Graphical representation of survival curves developed in this study based on change in drilling season. Survival curve corresponding to "winter" is highlighted in mauve. Survival curve corresponding to "spring" is highlighted in green. Survival curve corresponding to "summer" is highlighted in cyan. Survival curve corresponding to "fall" is highlighted in red.

Similar to "fracturing season" risk factor, explained in **Figure 41**, the remaining potential risk factors were analyzed through survival curves. The following conclusions, for the remaining potential risk factors, have been reached regarding conditions that relatively *increase* the risk of casing failure compared to their companions (other levels/groups within same risk factor):

- Implementation of "acid" treatment at relatively lower depths (< 8,500 ft) imposes no increased risk on casing failure as opposed to not having acid treatment. However, for intermediate depths, ranging from 8,500 – 10,000 ft, as well as relatively high depths (> 10,000 ft) the risk imposed by having acid treatment becomes much higher than that imposed by not having acid treatment (~ 1.5-fold higher).
- Similar to acid treatment, having no "cement" support at lower depths (< 8,500 ft) imposes

no higher risk on casing failure compared to having a cement support. Yet, for intermediate depths (8,500 – 10,000 ft) and high depths (> 10,500 ft), having a cement support starts to play an integral rule in reducing the risk of casing failure; reaching nearly 2-fold less than the risk imposed by not having cement support.

- Regarding "dogleg bending stress", $\sigma_{bending}$, at low depths (< 8,500 ft) and intermediate depths (8,500 – 10,000 ft), relatively high $\sigma_{bending}$ (> 95.8k lb./ft^2) are needed to impose high risk on casing failure. Conversely, at high depths (> 10,000 ft), relatively lower $\sigma_{bending}$ could be sufficient to impose high risk on casing failure. This inverse proportionality between allowed $\sigma_{bending}$ and reached depth is due to the deviation of the well. For deviated and horizontal wells, inclination reaches its highest values at high depths. This, in turn, imposes high stresses on casing string. Accordingly, lower than usual $\sigma_{bending}$ will lead to higher risk of casing failure at those high depths where inclination is the highest.

- Considering "bottomhole temperature", BHT, at relatively low depths (< 8,500 ft) along with intermediate depths (8,500 – 10,000 ft), BHT has no differentiating rule on either increasing or decreasing casing failure. However, with high depths (> 10,000 ft), high BHT (> 170 ^{o}F) imposes higher risk on casing failure than relatively low BHT. This could be due to high temperature differentials between fracturing fluid and bottomhole which leads to high shrinkage of lateral section, hence, higher risk of casing failure.

- Similar to BHT, the variation in the magnitude of "lateral shrinkage" tends to have no noticeable impact on the risk of casing failure for low depths (< 8,500 ft), as well as intermediate depth (8,500 – 10,000 ft). Yet, for high depths (> 10,000 ft), change in lateral

section starts to play an integral rule on increasing/reducing the risk of casing failure. For such high depths, a small increase in "lateral shrinkage" (≤ 10 ft/100ft) doubles the risk of casing failure. This is due to the high stress already imposed on the casing string due to high inclinations and high temperature differentials, as well.

- Similar to "bending stress", "maximum inclination", α_{max}, possesses an inverse proportionality with respect to the measured depth. So, for low and intermediate depths, α_{max} of a value > 95.5º is needed to increase a high risk of failure on casing string. On the other hand, for high depths (> 10,000 ft), relatively lower α_{max} would be sufficient to nearly double the risk of failure. This is for the same reason explained in the previous two points.

- Regarding "frequency of severe dogleg", ω_{DLS}, for low depths (< 8,500 ft), ω_{DLS} has no indicative influence on increased/reduced risk of casing failure. For intermediate depths (8,500 – 10,000 ft), lower value ω_{DLS} of (ranging from 7 – 12) is indicative of increased risk of casing failure. As for high depths (> 10,000 ft), a higher than normal ω_{DLS} (> 13) would be indication of increased risk of casing failure. This direct proportionality between ω_{DLS} and depth is owing to the fact that, for deviated wells, inclination increases with respect to depth.

- Similar to "maximum inclination", "dogleg severity", DLS, is inversely proportional to reached measured depth. This means that, for low and intermediate depths (< 10,000 ft), the allowance for high DLS (≥ 15º/100ft) is high. Conversely, for high depths (> 10,000 ft), the allowance for high DLS (≥ 15º/100ft) is much less.

- Finally, for" fracture temperature", T_{frac}, it is only with high depths (> 10,000 ft) that low

T_{frac} ($\leq$ 49.5 ºF) imposes a high risk of casing failure for the same reasons as BHT.

Graphical representations of the results from the application of non-parametric survival curve estimators, i.e., Kaplan-Meier estimator, on the aforementioned potential risk factors are summarized in Appendix (D).

From previous discussion, it is obvious that non-parametric estimators (e.g., Kaplan-Meier estimator) show a huge flexibility, in terms of the relative positioning of the levels/groups with respect to depth. For instance, within a particular risk factor, Group A might impose a higher risk than another Group B within a specific range of depth, yet, in different depth range Group B might impose a higher risk than Group A. This sort of flexibility helps reflecting the reality of the situation that: it is not necessary to have one level/group within a risk factor that imposes the highest risk throughout the well.

However, one major disadvantage emerged during incorporating risk factors/covariates. Creating a model for each group/level within a particular risk factor becomes infeasible as the number of levels/groups increase or when the risk factor is continuous (e.g., dogleg severity). To incorporate a continuous variable, it must be categorized into several levels/groups and a model is created for each group. It is the categorization process that makes it infeasible. It is for this reason, I will be discussing another family of techniques, that is the semi-parametric survival curve estimators.

3.3.2 Semi-Parametric Survival Curve Estimation

In previous section (Section 3.3), it was mentioned that two broad families of techniques; parametric and non-parametric survival curve estimators, can be used for survival analysis. Non-parametric estimators were, first, introduced (Section 3.3.1). Detailed discussion of the different steps involved in the process, as well as, analysis outputs from the provided data set were

presented (Section 3.3.1.1 and Section 3.3.1.2). In this section, discussion is extended to the second family of techniques used for survival analysis. The motivation behind discussing semi-parametric estimators is address the limitations discussed in Section 3.3.1.2. First, like non-parametric estimators (e.g., Kaplan Meier estimator), semi-parametric estimators (e.g., Cox regression) allow for flexible baseline. However, unlike non-parametric estimators, semi-parametric estimators allow having different survival functions, corresponding to different levels/groups, within the same fitted model. Hence, they better handle risk factors, or covariates, of multiple levels, as well as continuous covariates.

3.3.2.1 Proportional Hazards Model (Cox Regression Model)

In essence, semi-parametric survival curve estimators decompose the hazard or instantaneous risk into (1) a *non-parametric baseline,* shared across all individuals (drilling/producing wells), and (2) a *relative risk,* which describes how risk factors/covariates impact the outcome (casing failure). Being one of the widely used semi-parametric models, Cox regression model will be the focus of this section. Here, we formulate the "relative risk" since it comprises the core of Cox model.

Similar to non-parametric estimators, the starting point is the definition of the alternative hypothesis that assumes difference between survival distributions of levels/groups within a risk factor/covariate. For simplicity, a two-level covariate is considered. The alternative hypothesis can be defined, in terms of survival function, (Cox 1972), as

$$H_A: S_1(d) = [S_0(d)]^{\psi} \qquad (38)$$

Equivalently, alternative hypothesis can be defined, in terms of proportional hazards, as

$$h_1(d) = \psi\, h_0(d) \qquad (39)$$

where ψ the *proportional hazards constant*, indexes the difference between the two survival (or, hazard) distributions.

Extension of the model to include other covariate information, represented by a vector z, yields

$$\psi = e^{z\beta} \tag{40}$$

Then, the Cox regression model (Cox 1972) can be defined as

$$\lambda_n(d) = \underbrace{\lambda_0(d)}_{basline}\ \underbrace{exp(\beta^T z_n)}_{relative\ risk} \tag{41}$$

where $\lambda_0(d)$, the baseline risk, and z_n is a vector of risk factors/covariates for the individual (drilling/producing well), *n*.

In order to use Cox regression model for estimating the hazard, it is imperative to calculate the relative risk. In order to accomplish that, we implement "Partial Likelihood" function, which is similar to the "Likelihood" function used for the parametric survival curve estimators. Detailed discussion on "Partial Likelihood" can be found in Appendix (C).

3.3.2.2 Comparing Survival Curves (Using Partial Likelihood Function)

Here, discussion of the partial likelihood starts by considering the simple case of comparing two groups of survival data. The partial likelihood allows for the usage of an unspecified baseline survival distribution to define the survival distributions of subjects based on their covariates.

The partial likelihood differs from a likelihood in two ways. First, it is a product of expressions, one for each failure time, while censoring times do not contribute any factors. Second, the factors

of a partial likelihood are conditional probabilities.

Consider now the first failure depth d_1. The set of all wells "at risk" for failure at this depth is denoted by R_1: (Just before the first failure, this set is comprised of all the wells.) Among the wells in the risk set R_1, all are at risk of failure (i.e., of experiencing the event), and one of them, say well i, does fail. The probability that well i is the one who fails is the hazard, $h_i(d) = \psi\, h_0(d)$, for that well divided by the sum of the hazards of all the wells:

$$p_1 = \frac{h_i(d_1)}{\sum_{k \in R_1} h_k(d_1)} = \frac{h_0(d_1)\psi_i}{\sum_{k \in R_1} h_0(d_1)\psi_k} = \frac{\psi_i}{\sum_{k \in R_1} \psi_k} \qquad (42)$$

After the event at d_1, that well drops out of the risk set R_1, as do any censored observations that occur after d_1 up to and including the second failure depth d_2, resulting in a new (and smaller) risk set R_2. We then repeat this calculation to obtain p_2, and so on up to the last failure time.

The partial likelihood, (Edwards 1992), is the product

$$\prod_{n=1}^{N} \left(\frac{\lambda_0(t) \cdot \exp(\beta^T z_n)}{\sum_{k=1}^{N} \lambda_0(t) \cdot \exp(\beta^T z_n)} \right)^{\delta_n} \left(\sum_{k=1}^{N} \lambda_0(t) \cdot \exp(\beta^T z_n) \right)^{\delta_n} S_n(t)$$

Assuming that there are N failure depths. In each factor the baseline hazard cancels out of the numerator and denominator, so that it plays no role in the final partial likelihood. The maximum partial likelihood estimate is the value of ψ that maximizes this function. After defining the value of relative risk, ψ; that maximizes the "*partial likelihood*" function, we use the "*partial hazards*" model, or simply, Cox model for estimating hazard functions for different levels/groups within various risk factors/covariates.

3.3.2.3 Numerical Demonstration

The way Cox model handles partial hazards is divided into stages. First, the model picks one of

the levels/groups within the risk factor and sets it as a "reference" point for subsequent calculations. Then, the model calculates the relative impact other levels/groups have on the outcome compared to the reference point. This process is, then, repeated for all risk factors/covariates.

The output of the analysis is represented in terms of "*hazard ratio*". A hazard ratio of value equal to unity indicates *no* impact. A value of hazard ratio greater than 1 indicates an *increased* risk and a value of hazard ratio less than 1 indicates a *reduced* risk.

Results obtained from the application of Cox model on the data set are summarized in **Table 12** and **Figure 42**. As expected, the conclusions drawn from the application of semi-parametric estimators (Cox model) match those obtained from the implementation of non-parametric estimators (Kaplan-Meier estimator). What gives the edge to semi-parametric estimators over non-parametric estimators, however, is the manner both techniques handle multiple covariates. For non-parametric estimators, as shown before, the survival function corresponding to each covariate needs to be fit on a separate model. This has shown to be infeasible and quite cumbersome for complex covariates (e.g., multi-level covariates and continuous covariates). Unlike non-parametric estimators, semi-parametric estimators have the ability to handle all risk factors/covariates of interest on a single model fit.

Table 12 Basic Statistics of the statistical model using 10 selected features.

Feature/Covariate	Class	Coef.	exp(co.)	SE	z	p
DL_BEND_STRESS	**>95800(lbf)**	1.0343	2.8133	0.9375	1.103	0.2699
CEMENT	**Yes**	-0.3324	0.7172	0.6116	-0.543	0.5868
MD	**>13500(ft)**	-0.5052	0.6034	0.6888	-0.733	0.4633
BTH	**>166(F)**	-0.7635	0.4660	0.6570	-1.162	0.2452
LATERAL_SHRINKAGE	**>10.7(in/100F)**	-2.3007	0.1002	0.9399	-2.448	0.0144
MAX_INCL	**>95.5(o)**	0.5716	1.7710	0.8556	0.668	0.5041
DL_SEVERE_MD	**>10900(ft)**	-0.1813	0.8342	0.9662	-0.188	0.8511
DL_FREQ_10PLUS	**>13**	-0.6780	0.5076	0.5935	-1.142	0.2533
FRAC_TEMP	**>49.5(F)**	-1.0852	0.3378	0.535	-2.028	0.0425
DRILL_FRAC_INTERVAL	**>75(days)**	1.8279	6.2209	1.2375	1.477	0.1397

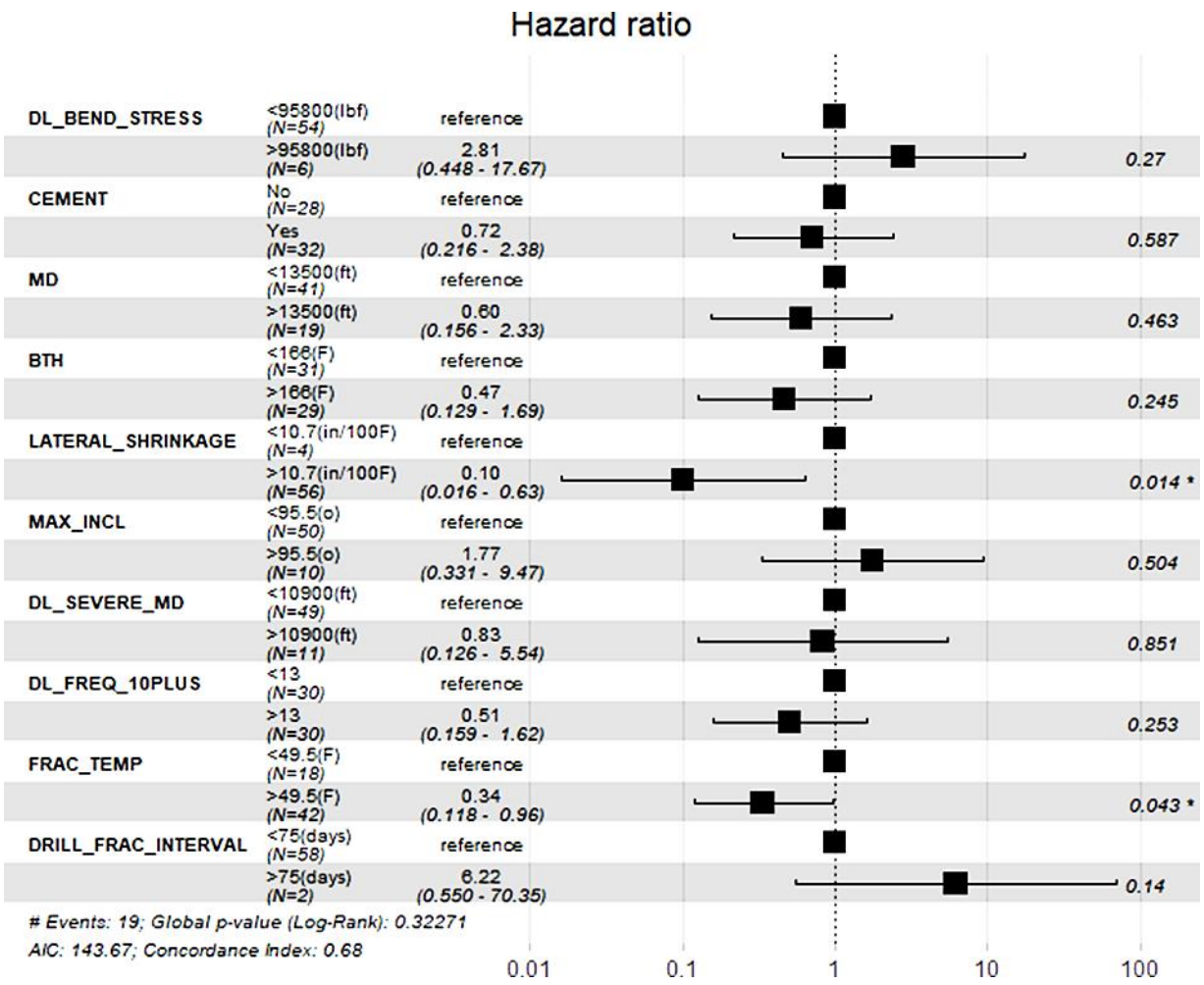


Figure 42. Hazard ratio envelope for the statistical model using 10 selected features.

3.4 Conclusions

In this chapter, investigation continued addressing some of the causations of a major limitation previous contributions had concerning their inability to provide a viable measure to mitigate casing failure. Two major drivers of that limitation are: (1) the inability to "quantify" the impact that various potential risk factors have on the occurrence of the casing failure and (2) the inability to tackle the impact of potential risk factors throughout the well length, in addition, how variation in a particular risk factor might affect the probability of casing failure accordingly.

I chapter 2, I tackled the first challenging aspect by designing a data-driven workflow, based on risk analysis techniques, that has the capability of evaluating the impact type, as well as the magnitude of the various risk factors.

In this chapter, I tackled the second challenging aspect through designing another data-driven workflow based on "survival analysis" techniques. Using the designed workflow, I managed to evaluate/test the significance of the different subcategories within each potential risk factor. In addition, I managed to track their impact throughout the well length which was otherwise not feasible. The information drawn from this chapter and the previous chapter hold a great potential for developing an advisory system for drillers and drilling managers and ultimately an automated casing failure mitigation system. Conclusions drawn from the analysis would be of great value when adjusting design specifications of high risk; using the proposed "casing failure mitigation" tool that would be based on "correction-prediction" procedure (Chapter 4).

Although the designed workflow was based on survival analysis, yet, motivated by the research interests, little modification was made to the analysis scale (conventionally, time scale) so it now corresponds to reached "measured depth".

First, several survival/hazard functions that are used to draw conclusion on survival distribution were introduced. To provide a systemic way of constructing survival distributions, two broad families of survival curve estimators were discussed: (1) non-parametric estimators, e.g., Kaplan Meier estimator, and (2) semi-parametric estimators, e.g., Cox model.

Due to their simplicity and their graphical representations, non-parametric estimators were used for tracing the changes in the impact of subgroups/levels within each potential risk factor throughout the well length. That served our first goal for this chapter.

Results of the analysis showed that the following conditions (subgroups/levels) are the riskiest compared to their companions within their corresponding risk factors with respect to depth:

(1) Regarding "fracturing season", it's been noticed that for depths less than 8,500 ft, no significant difference in risk impact between different seasons. Yet, fracturing during "winter" or "spring" still puts the well at a higher risk of failure than the other two seasons: with an increase of risk by 20% and 7%, respectively, for depths ranging from 8,500 – 10,000 ft, while 30% and 20%, respectively, for depths greater than 10,000 ft.

(2) Regarding "acid treatment", it has been noticed that: implementation of "acid" treatment at depths less than 8,500 ft imposes no increased risk on casing failure as opposed to not having acid treatment. However, for depths ranging from 8,500 to 10,000 ft, as well as depths greater than 10,000 ft, the risk imposed by having acid treatment becomes much higher (7.5% and 12% increase, respectively) than that imposed by not having acid treatment. In addition, the risk constantly increases with increasing reached depth.

(3) Regarding "cementing", it has been noticed that: having no "cement" support at depths less than 8,500 ft imposes no higher risk on casing failure compared to having a cement support. Yet, for depths ranging from 8,500 to 10,000 ft, as well as depths greater than 10,000 ft,

having a cement support starts to play an integral rule in reducing the risk of casing failure; reaching nearly 30% less than the risk imposed by not having cement support for depths greater than 10,000 ft.

(4) Regarding "dogleg bending stress", it has been noticed that at depths ranging from 8,500 to 10,000 ft, as well as, depths < 10,000 ft, relatively high DL bending stress (> 95.8k lb./ft^2) are needed to impose high risk on casing failure (~ 7.5% increase). Conversely, at depths > 10,000 ft, relatively lower DL bending stress (less than 95.8k lb./ft^2) could be sufficient to impose a higher risk of casing failure (~ 12-13% increase).

(5) Regarding "bottomhole temperature", it has been noticed that for depths < 8,500 ft, relatively higher BHT (> 166 ^{o}F) is needed in order to have a mere increase in risk of casing failure (~ 3.2% increase). In contrast, for depths ranging from 8,500 to 10,000 ft, as well as, those > 10,000 ft, relatively less BHT (< 166 ^{o}F) could be enough to impose a significant increase in risk of casing failure (~ 5.1-12.8% increase).

(6) Regarding "lateral shrinkage", it has been noticed that the variation in the magnitude of "lateral shrinkage" tends to have no noticeable impact on the risk of casing failure for depths < 8,500 ft. Yet, for depths ranging from 8,500 to 10,000 ft, as well as, those > 10,000 ft, slight changes in lateral section (less than 10 ft/100ft) could play an integral rule in increasing the risk of casing failure (~ 26% increase).

(7) Regarding "maximum inclination", as well as "DL severity", it has been noticed that – similar to "bending stress" – "maximum inclination" and "DLS" possesses an inverse proportionality with respect to the measured depth. So, for depths < 8,500 ft and those ranging from 8,500 – 10,000, a relatively high value of max. inclination, as well as DLS is needed in order to have a noticeable impact on increasing risk. Yet, for depths > 10,000 ft,

relatively lower values would be sufficient to nearly increase the risk by 9-12%.

(8) Regarding "frequency of severe dogleg", it has been noticed that it follows a direct proportionality with respect to the measured depth. So, for depths ranging from 8,500 – 10,000, a relatively lower value of DLS frequency (< 13) could be indicative of increased risk (~ 4.6% increase). And, for depths > 10,000 ft, relatively higher values (> 13) would be a strong indication of increased risk (~ 16.3%).

(9) Finally, for" fracture temperature", it has been noticed that, it follows direct proportionality with respect to the measured depth, for the same reasons as BHT.

However, one limitation was noticed: with increased number of features/risk factors or subgroups/levels within risk factors, some complexities were introduced in the analysis. That's when semi-parametric came into play as semi-parametric estimators had the ability to incorporate all risk factors in one model, unlike the non-parametric estimators. Besides, semi-parametric compromises for multi-level, as well as continuous variables.

CHAPTER IV
CASING FAILURE MITIGATION TOOL DESIGN

Reader Guide: In this chapter, the information provided by the analyses conducted in chapters 2 and 3 are leveraged to construct the proposed "automated casing failure mitigation" tool, which will be based on a two-step "prediction-correction" procedure. Based on the input information (initial design specifications), the tool provides an initial prediction of the probability of casing failure, then automatically adjust, or correct, the design specifications in order to reduce the risk below a pre-defined threshold.

To tackle the first goal of the tool, that is concerned with prediction, a family of machine learning algorithms; formally known as supervised learning algorithms, is implemented (Section 4.1). In general, machine learning can be viewed as an alternative to the conventional engineering approach for the design of an algorithmic solution (Mitchell 1997; Goodfellow et al. 2016; Brynjolfsson 2017; Simeone 2018b).

We start our discussion with investigating learning algorithms for classification problems such as logistic regression (Section 4.2.1), decision trees (Section 4.2.2), random forest (Section 4.2.3), support vector classifier (Section 4.2.4), support vector machine (Section 4.2.5) and neural networks (Section 4.2.6). In addition, we discuss different criteria for evaluating the performance of the models in terms of prediction accuracy, model balance and overall performance (Section 4.3).

To add a level of controllability to serve as feedback to the model predictions, the discussion is extended to another area of statistical analysis: risk assessment (Sections 4.4 and 4.5). The discussion will be limited, however, to semi-qualitative risk assessment techniques, particularly, probability-impact risk assessment matrices, PI-RAM, (Section 4.6). The basis for those techniques, as well as their formulation will then be laid out. History cases from the oil and gas industry will be presented to showcase the huge potential for such methodology (Section 4.7). Afterwards, application of the presented methods on the historical data set will be demonstrated (Section 4.8). Finally, highlights of the work done in this chapter, as well as the major takeaways from the analysis conducted will be pointed out (Section 4.9).

4.1 Introduction

In previous chapters, I addressed the main drivers of a major limitation previous contributions had concerning their inability to provide a viable measure to mitigate casing failure. Those were: (1) the inability to "quantify" the impact that various potential risk factors have on the occurrence of the casing failure and (2) the inability to tackle the impact of potential risk factors throughout the well length, in addition, how variation in a particular risk factor might affect the probability of casing failure accordingly.

I chapter 2, I tackled the *first* challenging aspect by designing a data-driven workflow based on "risk-factor" analysis techniques. The developed workflow had the capability of evaluating the impact type, as well as the magnitude of the various risk factors (Section 2.7). In chapter 3, I tackled the *second* challenging aspect through designing another data-driven workflow based on "survival analysis" techniques. Using the designed workflow, I managed to evaluate/test the significance of the different subcategories within each potential risk factor, in addition, I managed to track their impact throughout the well length which was otherwise not feasible (Section 3.3.1.2 and Section 3.3.2.3).

The information drawn from the two previous chapters hold a great potential for developing an advisory system for drillers and drilling managers and ultimately an automated casing failure mitigation system. Consequently, in this chapter, the information provided from the analyses conducted in Chapter 2 and Chapter 3 are leveraged to construct the proposed automated "casing failure mitigation" tool, which will be based on a two-step "prediction-correction" procedure (overview provided in Section 1.3).

Based on the input information (initial design specifications), the tool will provide an initial

prediction of the probability of casing failure, then automatically adjust, or correct, the high-risk design specifications in order to reduce the overall risk of casing failure below a pre-defined threshold. This tool, hence, will serve as an advisory system for drillers and drilling managers to automatically mitigate, or avoid, future casing failures which was otherwise not feasible.

What gives the edge to the proposed tool is that, unlike previous attempts that followed physics-based approach, this tool is based on data-driven workflow. Figure 43 highlights the key distinctions between the two approaches. Conventional engineering design flow starts with the acquisition of domain knowledge required for the problem of interest. Then, a mathematical model that capture the physics of the study is constructed. Based on that model, an optimized algorithm is produced that offers performance guarantees under the assumption that the given physics-based model is an accurate representation of reality (Figure 43-a). In contrast, the machine learning approach substitutes the step of acquiring domain knowledge with the task of collecting a sufficient data set (known as the "training" set) that captures the behavior of interest. The training set is, then, fed to a learning algorithm to produce a trained "machine" that carries out the desired task, e.g., prediction of casing failure, in our case (Figure 43-b).

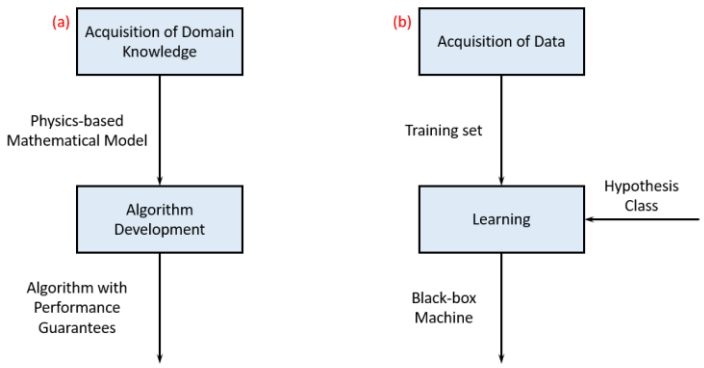


Figure 43. (a) Conventional engineering design flow; and (b) baseline machine learning methodology (Reproduced from Simeone 2018a).

Machine learning comprises of two major families of techniques: supervised learning and unsupervised learning techniques (Simeone 2018a). In supervised learning, the training set consists of pairs of input and desired output, and the goal is that of learning a mapping between input and output spaces (Shalev-Shwartz et al. 2014; Klimberg et al. 2016a; Arpit et al. 2017). In unsupervised learning, the training set consists of unlabeled inputs, that is, of inputs without any assigned desired output, and the goal is to discover properties of the mechanism generating the data or cluster the data points that are close to each other (Hastie et al. 2009). Based on the former discussion, along research interests, that is, prediction of casing failure, supervised learning techniques were selected for subsequent analysis.

As for the second goal of this chapter, we investigate some of the risk assessment techniques. Risk assessment is an inherent part of a broader risk management strategy to help eliminate any potential risk-related consequences (Rausand 2013; Zackmann 2014; Popov et al. 2016). The risk assessment process consists of a sequence of steps and sub-steps. The steps in the process are shown in Figure 44.

Figure 44. The six steps of a risk assessment (Reproduced from Rausand 2013).

Generally, risk assessment follows one of three major analytic approaches; qualitative, quantitative or semi-quantitative. In quantitative risk assessment, probabilities and consequences are determined purely qualitatively. In quantitative risk assessment, however, numerical estimates are provided for probabilities and consequences. Semi-quantitative risk assessment, in essence, is a quantitative approach, yet probabilities and consequences are quantified within *ranges*. Based on the provided data set, semi-quantitative approach was adopted for conducting the risk assessment.

Based on the former discussion, the focus of this chapter will be: (1) discussion of various supervised learning models, their key distinctions, major limitations and overall prediction performance, and (2) discussion of risk assessment techniques, history cases in oil and gas industry, construction and application of PI-RAMs on the provided data set.

4.2 Supervised Learning Techniques

In this section, I address the first of the proposed tool, that is, providing initial predictions of the

probability of casing failure given input values for pre-defined potential risk factors (Chapter 2). To that end, I will be implementing various machine learning algorithms to fulfill that goal. Following the research interest, that is, prediction of casing failure, supervised learning techniques were selected for the analysis.

The essence of the supervised learning is that the machine (i.e., the learning algorithm) is given a dataset, along with the right answers to a question corresponding to that data set. In our case, the question would be: would a casing failure occur given the conditions at hand? To determine the answer to that question, the learning algorithm has to learn the key characteristics/patterns within each data point in the dataset. Based on those key characteristics, the algorithm should be able to give the right answer (predict casing failure, in this case). To put that into mathematical terms, in supervised learning, the goal is to infer a function, f, that maps from the input data to the desired output. The input data are represented by vector, X, and the output is represented by vector, Y.

$$Y = f(X) + \varepsilon \qquad (43)$$

To find the optimal supervised learning technique that perfectly fits the study goal, *two* integral components need to be defined: (1) the outcome type and (2) the learning algorithm. Regarding the first component, supervised learning techniques can be categorized, based on the type of the outcome, into two major classes or groups: classification and regression techniques. The goal set (predict whether a casing failure would occur or not) falls into the classification learning domain.

As for the second component, Literature is rife with learning algorithms that fit similar purposes, such as logistic regression, supervised hierarchal clustering, decision trees, support vector machines (SVMs), bootstrap forest (bootstrapping), naïve-bayes, and xgboost. Following is brief

introduction to a variety of supervised learning techniques. Detailed discussion will be provided in the upcoming sections.

Logistic regression is somehow a modification to “linear regression” technique for classification problems. In logistic regression, rather than modeling the outcome, e.g., casing failure, directly (as in “linear regression”), Figure 45 (left), it models the probability of each of the categories of the outcome (e.g., “failure” or “no failure”). Then, use that as a basis for the classification, Figure 45 (right).

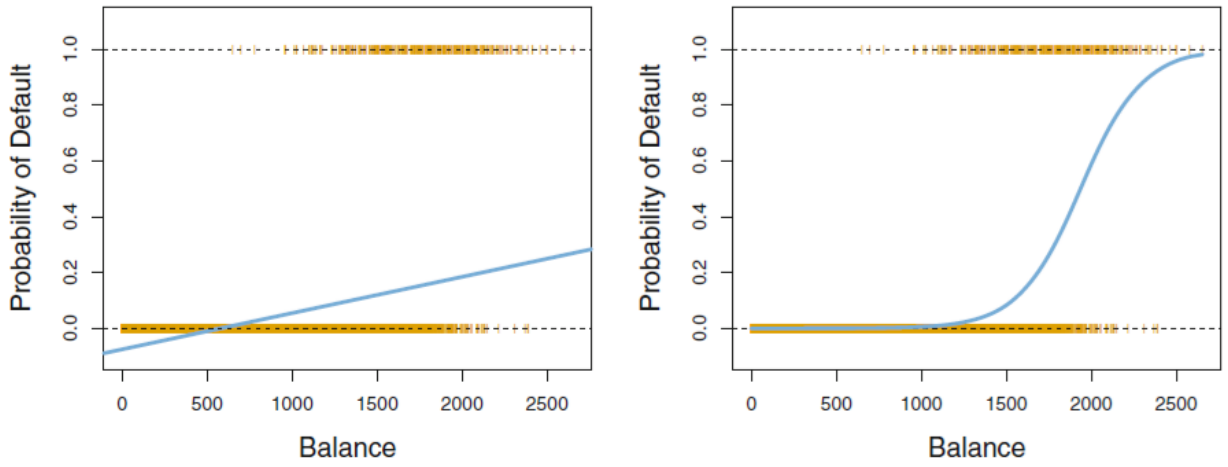


Figure 45. Example reproduced from Hastie et al. 2009; showing classification using the Default data. The orange ticks indicate the 0/1 values coded for default (No or Yes). Left: Estimated probability of default using linear regression. Right: Predicted probabilities of default using logistic regression.

Another family of techniques is tree-based methods (Gerrity 1979; Sprague 1980; Waller 1988). Those methods involve segmenting the input space into a number of regions, then assign a value (for regression trees) or a label (for classification trees) to an observation based on the region where

it belongs (**Figure 46**). Tree-based methods include simple decision trees, random forest and bootstrap forest. Tree-based methods are simple and useful for interpretation, yet, for some cases, they may not be competitive with other supervised learning approaches, such as SVMs, in terms of prediction accuracy (Misra et al. 2019).

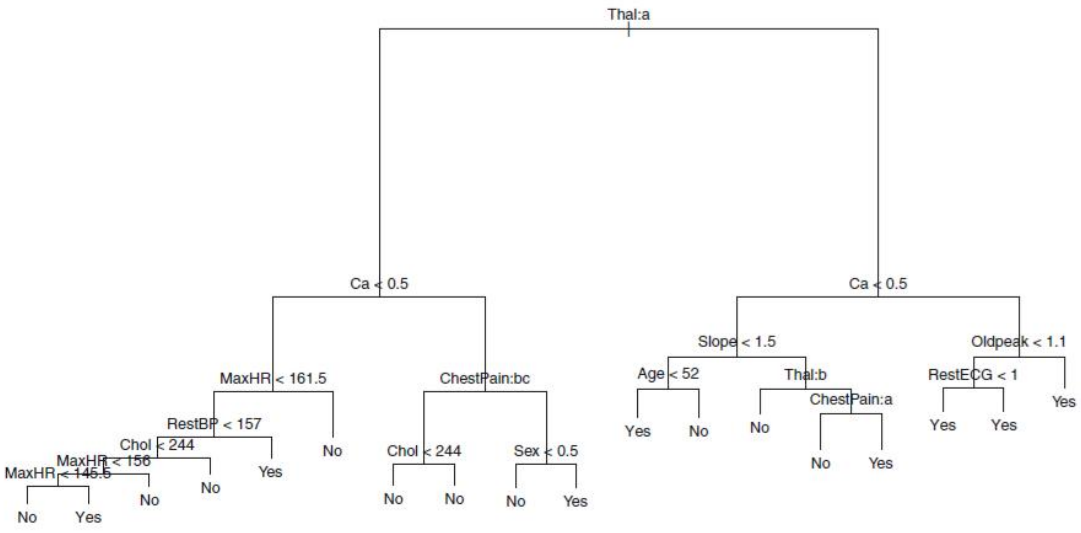


Figure 46. Example reproduced from Hastie et al. 2009; showing a snippet of unpruned classification tree for "Heart" data.

Support vector machine (SVM) is another approach for classification that was developed in the 1990s and is considered as one of the best classifiers in terms of accuracy (Cortes et al. 1995). Support vector machine, in essence, can be viewed as a generalization of a much simpler classifier known as the "maximal margin classifier". Unlike maximal marginal classifier, which only assumes that classes can be separated by a linear boundary, Figure 47 (Left), SVM has the ability to handle non-linear class boundaries through the implementation of kernels, Figure 47 (Right).

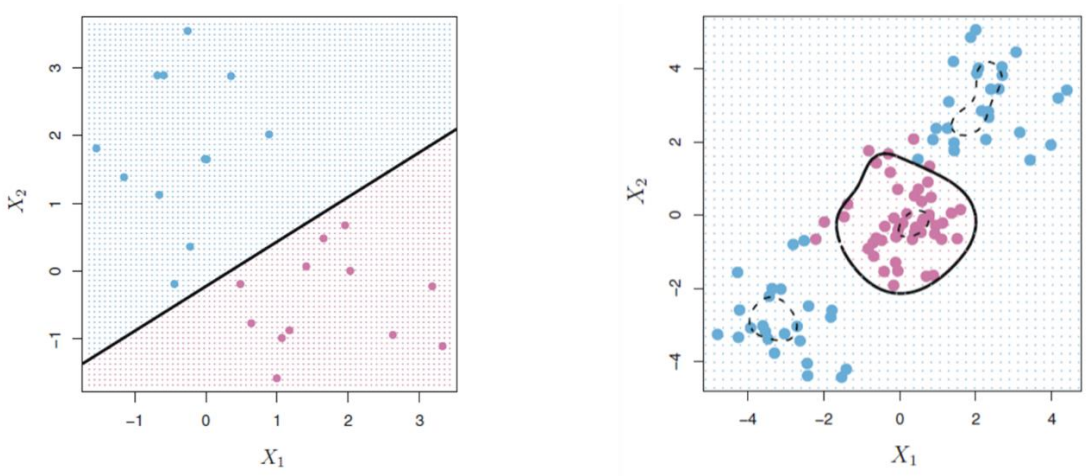


Figure 47. Example reproduced from Hastie et al. 2009. Left: Maximal marginal classifier with linear boundary. Right: An SVM with radial kernel.

Artificial neural networks (ANN), or simply, neural networks (NN), is one of the widely used supervised learning techniques (Hopfield 1982; Sarle 1994). A neural network is a two-stage model, typically represented by a network diagram (Figure 48). This network comprises of three major layers; input, hidden and output layer, and could be applied to regression or classification problems. The central idea is to extract linear combinations of the inputs as derived features (hidden layer), and then model the target (output layer) as a nonlinear function (known as "activation" function) of these features.

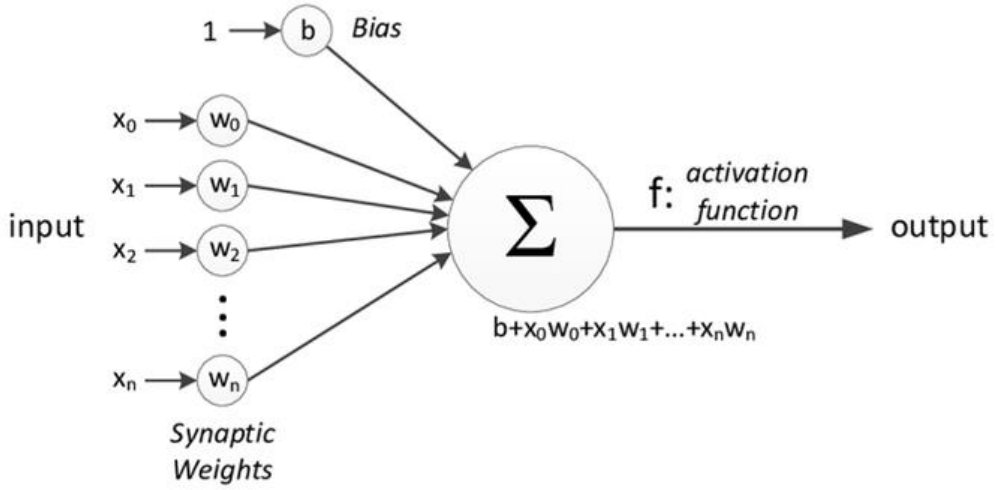


Figure 48. Schematic of feed-forward neural network (reproduced from Mohammed et al. 2017).

In the following sections, detailed discussion of select supervised learning algorithms, their intuition, mathematical formulation and key distinction. In addition, major results from the application of the learning techniques on the provided data set are presented.

4.2.1 Logistic Regression

As mentioned before (Section 4.1), logistic regression can be viewed as a modification to linear regression for classification problems. To further investigate that concept, we start with the definition of a simple linear regression model

$$P(Y = 1|X) = \beta_0 + \beta_1 X \qquad (44)$$

where $P(Y = 1|X)$ is the probability of having the outcome, Y, given the input, X. β_0 is the intercept, or formally, the expected value of the outcome assuming independence of the given input. β_1 is the slope.

For a classification problem, assuming a binary outcome (e.g., casing failure) represented by 1 or 0, using a typical linear regression model might lead to some serious errors in the overall prediction. For instance, having a very small value for input (close to zero) might lead to negative outcome. On the other hand, having a relatively large value for the input might lead to prediction of the outcome greater than 1. Both cases would be misrepresentations of the true outcome and not sensible since probabilities must take a value between 0 and 1.

To fix that problem, we model $P(Y = 1|X)$, or simply, $P(X)$ using a function that gives outputs between 0 and 1. For our discussion, we implement the logistic function, (Cramer 2002), that is defined as

$$P(Y = 1|X) = P(X) = \frac{e^{\beta_0+\beta_1 X}}{1 + e^{\beta_0+\beta_1 X}} \qquad (45)$$

Using that definition to construct the model, we managed to keep $P(X)$ confined between 0 and 1 regardless of the input value. In this case, having a low input value would yield a $P(X)$ close, but never below, zero. Also, having a high input value would yield a $P(X)$ close, but never greater than, one.

To fit that model to the given data set, in other words, to estimate the unknown coefficients β_0 and β_1, we use the "maximum likelihood function". This method was introduced in great detail, previously, in chapter 3. Here we touch on the key aspects of that method.

The intuition behind using the concept of "maximum likelihood" to fit our logistic regression model, **Eq. (45)**, is that: we seek estimates for the coefficients; β_0 and β_1, such that the *predicted* probability $P(X)$ for each individual corresponds as closely as possible to the individual's

observed status. In other words, we try to find the coefficients; β_0 and β_1, such that plugging these estimates into our model, **Eq. (45)**, yields a number close to "one" for all wells that experienced casing failure, and a number close to "zero" for all wells that did not.

So, initially, we have the "likelihood function" (Edwards 1992) defined as

$$l(\beta_0, \beta_1) = \prod_{i:y_i=1} p(x_i) \prod_{i`:y_{i`}=0} [1 - p(x_{i`})] \tag{46}$$

and the goal is to find the value of β_0 and β_1 that maximizes that likelihood function.

For binary outcome (e.g., casing failure) using multiple predictors/inputs, the logistic model, **Eq. (45)**, can be generalized as follows (Cramer 2002)

$$P(Y = 1|X) = P(X) = \frac{e^{\beta_0+\beta_1 X+\cdots+\beta_p X}}{1 + e^{\beta_0+\beta_1 X+\cdots+\beta_p X}} \tag{47}$$

4.2.2 Classification Decision Tree

As mentioned earlier (Section 4.1), decision trees, or generally, tree-based methods are useful both for the ease of interpretation and for the manner they handle qualitative predictors as they don't require the usage of dummy variables. In this section, we discuss the simplest of tree-based methods, that is, basic decision trees. This is important for laying the foundation for more complex tree-based methods, such as random forest which will be discussed next. For the purposes of our study, we will limit the discussion to classification trees only.

Constructing a decision tree is a two-step method. First, we split the predictor space; that is a set of possible features/risk factors, into distinct and non-overlapping regions. Then, for each

observation that falls into a particular region, we assign a label to that observation based on the most occurring class of the training observations in that particular region.

When growing a tree, we follow a recursive binary splitting approach to grow a very large and complex tree, based on the training data set, and we stop when each terminal node has fewer than a minimum number of observations.

Afterwards, we start searching for a subtree (one with less terminal nodes) that has the lowest test error rate. The process of cutting down a tree is formally known as "pruning". The reason why we implement the "pruning" when searching for the best subtree instead of stop growing the tree early on when the reduction in the error rate exceeds a pre-determined threshold is that: the latter approach might seem short-sighted as a seemingly worthless split early that can cause the tree to stop growing might be followed by a very good split that should have led to a huge reduction in error rate.

For classification trees, the criteria for making the binary splits are classification error rate, E, Gini index, G, and cross-entropy, D. The classification error rate is defined as the fraction of the training observations in a region that do not belong to the most common class in that particular region. It is, then, defined as (Hastie 2013)

$$E = 1 - \max_k(\hat{p}_{mk}) \qquad (48)$$

where $\hat{p}_{mk}$ is the proportion of training observations in the m^{th} region that are from k^{th} class.

The Gini index, on the other hand, measures the total variance across the K classes and it is defined as follows

$$G = \sum_{k=1}^{K} \hat{p}_{mk}(1 - \hat{p}_{mk}) \qquad (49)$$

The Gini index can be viewed as a measure of node purity as it takes on a small value if all $\hat{p}_{mk}$ are close to zero or one.

Finally, for the cross-entropy, (Hastie 2013) it is defined as

$$G = -\sum_{k=1}^{K} \hat{p}_{mk} \cdot log(\hat{p}_{mk}) \qquad (50)$$

Similar to the Gini index, cross-entropy takes on a small value if all $\hat{p}_{mk}$ are close to zero or one. So, it can also be considered as a measure of node purity.

When building a classification tree, either the Gini index or the cross-entropy are typically used to evaluate the *quality* of a particular split, since these two approaches are more sensitive to node purity than is the classification error rate. In case the *prediction accuracy* of the final pruned tree is the goal, then using classification error rate is preferable.

4.2.3 Random Forests

Random forests are considered an improvement over the basic decision trees, discussed in previous section. One major problem with the application of basic decision trees is that they suffer from high variance. Meaning that if we randomly split the training data into two groups, then fit a decision tree to both halves, the results could be a bit different. In contrast, a procedure with low variance will yield similar results if applied repeatedly to distinct data sets. Random forests tackle that problem by implementing a modified version of a methodology known as bootstrap aggregation, or simply, bagging.

Bagging follows the essence that averaging any set of observations would reduce the variance between the obtained results. Typically, different B training sets are taken from the population. Then, separate prediction model is fit to each training set, $\hat{f}^i(x)\ for\ i = 1,2,\dots,B$. Finally, the resulting predictions are averaged, $\hat{\hat{f}}_{avg}(x)$. This process can be summarized as follows (Hastie 2013)

$$\hat{\hat{f}}_{avg}(x) = \frac{1}{B}\sum_{b=1}^{B}\hat{f}^b(x) \qquad (51)$$

For classification bagged tree, for each tested observation, the class predicted by each of the B trees is recorded. Then, the overall prediction is determined by the majority vote (i.e., the most commonly occurring class among the B predictions).

The reasons why random forests are considered as a modified version of bagging is that they decorrelates the trees. This means that when building a tree, for each split, a random sample of predictors is chosen as a split candidates from the full set of predictors. The importance of having random split candidates is that if there are some strong predictors, then following the typical bagging procedure we will end up with nearly the same set of split candidates for the majority of the trees. This would lead eventually to having similar trees and the predictions from those trees will be highly correlated.

4.2.4 Support Vector Classifier

As mentioned previously (Section 4.1), support vector machines family of techniques is one of the best classifiers in terms prediction accuracy. Support vector machines include the basic maximal margin classifier, support vector classifier and, of course, support vector machine. The essence of all those techniques is to find a classification boundary (hyperplane) that can perfectly separate the

different classes, hence, allowing for a more accurate prediction/labeling of new observation.

In this section, the focus will be on the support vector classifier. Support vector classifier is considered a modified version of the basic maximal margin classifier. Both classifiers are linear, Eq. (52) and both seek a "hyperplane" that can separate the different classes and help providing a better prediction/labeling. However, the key distinction between both techniques – and what gives edge to support vector classifier – is the way those hyperplanes are constructed.

$$f(x) = \beta_0 + \sum_{i=1}^{n} \alpha_i \langle x, x_i \rangle \qquad (52)$$

Maximal margin classifier suffers from two major. First, it requires that each observation to be on the right side of the hyperplane. This leads the constructed hyperplane to be extremely sensitive to a change in a single observation (Figure 49).

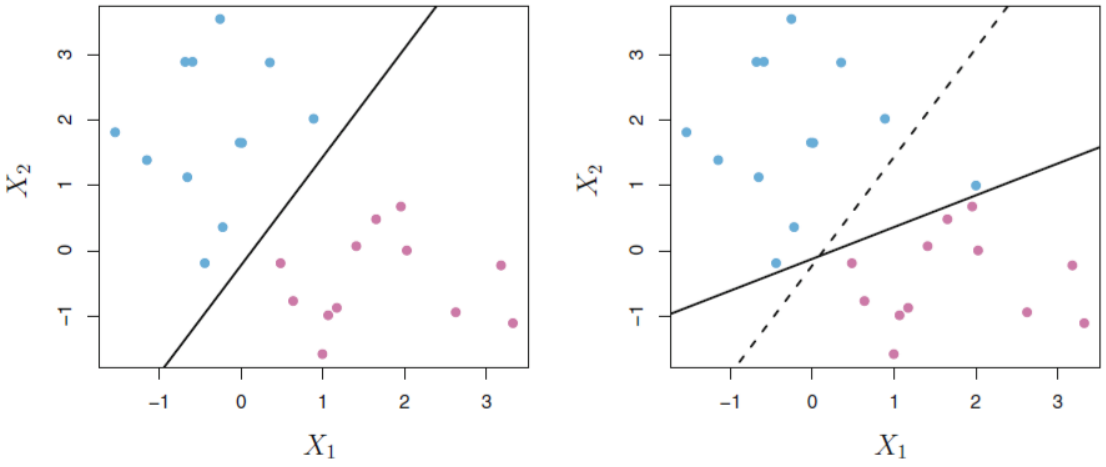


Figure 49. (Example reproduced from Hastie 2009) Left: Two classes of observations are shown in blue and in purple, along with the maximal margin hyperplane. Right: An additional blue observation has been added, leading to a dramatic shift in the maximal margin hyperplane shown as a solid line. The dashed line indicates the maximal margin hyperplane that was obtained in the absence of this additional point.

Second, since the perpendicular distance from the nearest observation to the hyperplane defines the level of confidence in the model predictions, for cases where this distance is extremely small, the level of confidence drops heavily affecting the overall performance of the classifier.

The way the support vector classifier handles those problems, is that it allows for a few training observations to be misclassified in order to do a better job in the overall classification of the remaining observation. This is formally known as “soft margin”. In other words, instead of seeking the largest possible margin so that every observation is not only on the correct side of the hyperplane but also on the correct side of the margin (as in maximal margin classifier), we instead allow some observations to be on the incorrect side of the margin, or even the incorrect side of the hyperplane (Figure 50).

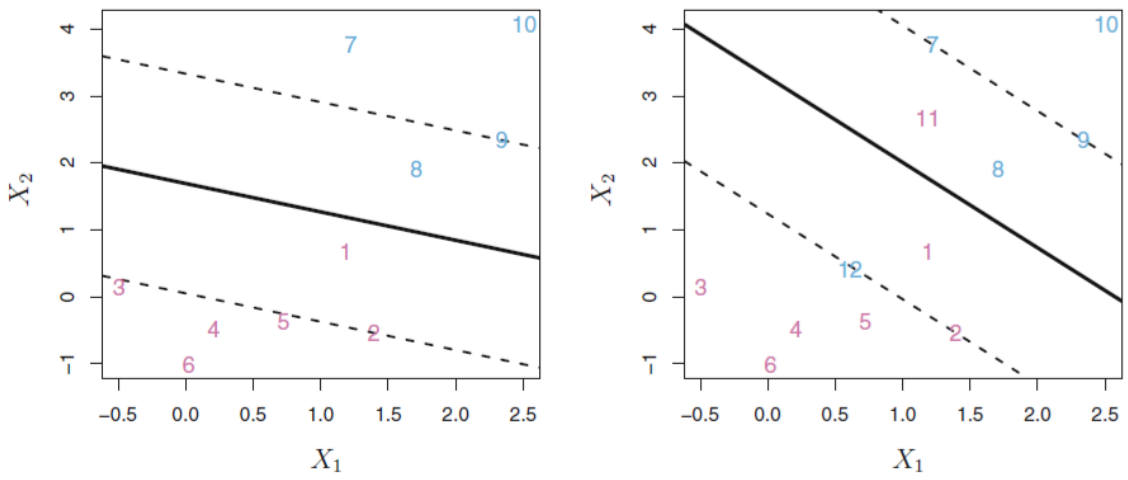


Figure 50. (Reproduced from Hastie 2013) Left: A support vector classifier was fit to a small data set. The hyperplane is shown as a solid line and the margins are shown as dashed lines. Purple observations: Observations 3, 4, 5, and 6 are on the correct side of the margin, observation 2 is on the margin, and observation 1 is on the wrong side of the margin. Blue observations: Observations 7 and 10 are on the correct side of the margin, observation 9 is on the margin, and observation 8 is on the wrong side of the margin. No observations are on the wrong side of the hyperplane. Right: Same as left panel with two additional points, 11 and 12. These two observations are on the wrong side of the hyperplane and the wrong side of the margin.

4.2.5 Support Vector Machine

The support vector machine (SVM) is considered an extension of the support vector classifier. SVM seeks to enlarge the predictor/feature space in order to accommodate a non-linear boundary between the classes, wherever needed, using "kernels". The kernel approach is simply an efficient computational approach for enacting this idea.

Similar to the support vector classifiers, SVMs (Cortes 1995) can be defined, in terms of the "kernel", K, as

$$f(x) = \beta_0 + \sum_{i=1}^{n} \alpha_i K\langle x, x_i \rangle \tag{53}$$

where $K\langle x, x_i \rangle$ is the "kernel". The kernel is what defines the shape of the classification boundary. It could be linear, Eq (54), polynomial, Eq. (55), or radial, Eq. (56). A graphical representation of the polynomial kernel and the radial kernel is provided in Figure 51.

$$K\langle x_i, x_{i`} \rangle = \sum_{j=1}^{p} x_{ij} \cdot x_{i`j} \tag{54}$$

$$K\langle x, x_i \rangle = \left(1 + \sum_{j=1}^{p} x_{ij} \cdot x_{i`j}\right)^d \tag{55}$$

$$K\langle x, x_i \rangle = exp\left(-\gamma \sum_{j=1}^{p} \left(x_{ij} \cdot x_{i`j}\right)^2\right) \tag{56}$$

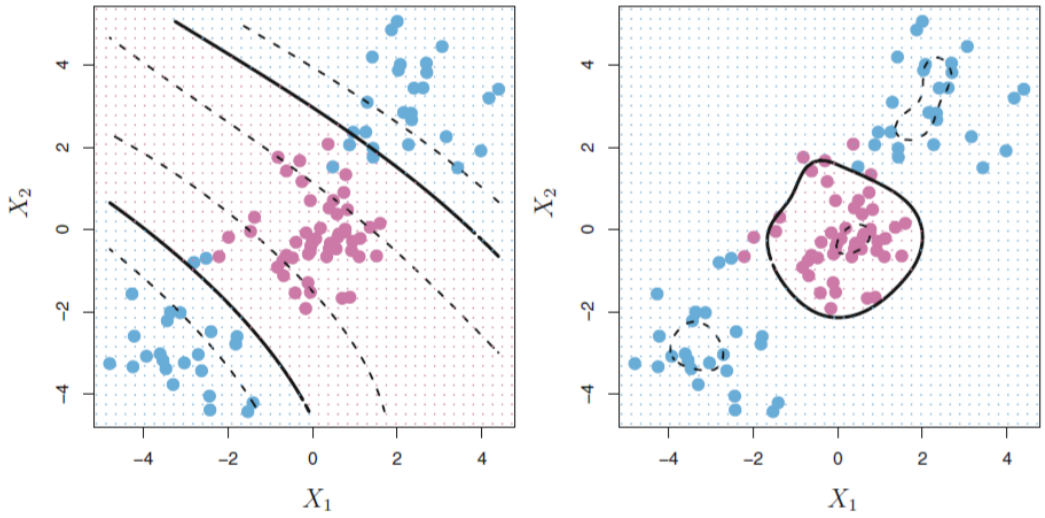


Figure 51. (Reproduced from Hastie 2009) Left: An SVM with a polynomial kernel of degree 3. Right: An SVM with a radial kernel is applied.

4.2.6 Artificial Neural Networks (ANNs)

Artificial neural networks, or simply, neural networks (NN) are popular machine learning techniques that simulate, in a gross manner, the networks of nerve cells (neurons) in biological organisms. The basic structure of a neural network contains several computation units, called neurons. Those units are connected to each other through weights, W_i. The weights in neural networks are analogous to the strengths of synaptic connections in biological organisms.

The basic neural network function known as the "perceptron" (Rosenblatt 1960), is comprised of one layer, where it directly propagates computations from the input neurons (one layer) to a single output neuron through intermediate variables (i.e., weights). Learning occurs by changing the weights connecting the neurons. Then, weights are adjusted according to "feedback" on how well the predicted output for a particular input matches the annotated output label in the training data.

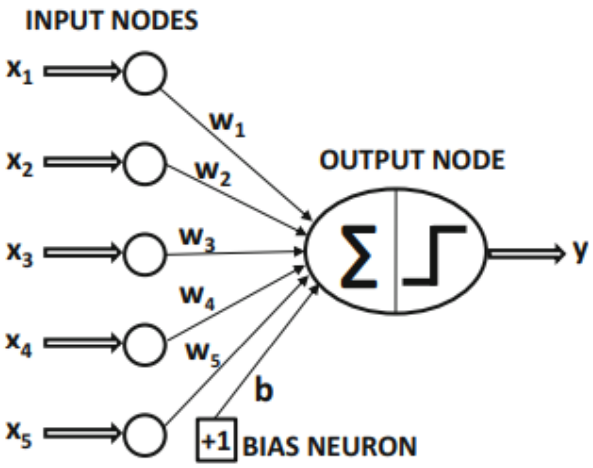


Figure 52. The basic architecture of the perceptron (Reproduced from Goodfellow 2016).

Mathematically, the output can be defined, in terms of the "perceptron", (Freund 1999) as

$$\hat{y} = sign(\bar{W} \cdot \bar{X} + b) = sign\left(\sum_{j=1}^{d} w_j x_j + b\right) \tag{57}$$

where $\hat{y}$ is the predicted output; either +1 or -1. $\bar{X}$ is the vector of input features. $\bar{W}$ is the vector of the weights/edges corresponding to the input features. d is the number of input features, or equivalently, neurons in input layer and b is the bias neuron. $\sum_{j=1}^{d} w_j x_j$ is basically the linear function that is computed at the output.

The error in the outcome is, then, minimized through the adoption of least-square algorithm applied for all training instances, (Hastie 2013) as follows

$$\min_{\bar{W}} \sum_{(\bar{X},y)\in D} (y - \hat{y})^2 = \min_{\bar{W}} \sum_{(\bar{X},y)\in D} \left(y - sign(\bar{W} \cdot \bar{X})\right)^2 \tag{58}$$

where y is the true outcome. $(\bar{X}, y)$ is the input-output pair for an instance in the training data set. D is the training data set. This minimized function is known as the "loss function". There is a wide variety of loss functions including sigmoid function and tanh function.

Based on the calculated error, the weights are updated through the following relation (Hastie 2013)

$$\bar{W} \Leftarrow \bar{W} + \propto (y - \hat{y})\bar{X} \qquad (59)$$

where $\propto$ is known as the "learning rate". $(y - \hat{y})\bar{X}$ is an approximation of the gradient.

A generalization of the simple one-layer "perceptron" is the multi-layer NNs. Multilayer neural networks contain additional, intermediate, computational layers other than the input and the output layers. Those additional layers are referred to as "hidden" layers because the computations performed are not visible to the user (**Figure 53**). The architecture of multilayer NNs is referred to as feed-forward networks, because successive layers feed into one another in the forward direction from input to output. This architecture assumes that all nodes in one layer are connected to those of the next layer. Once the number of layers and the number/type of nodes in each layer have been defined, the only remaining detail is the loss function that is optimized in the output layer.

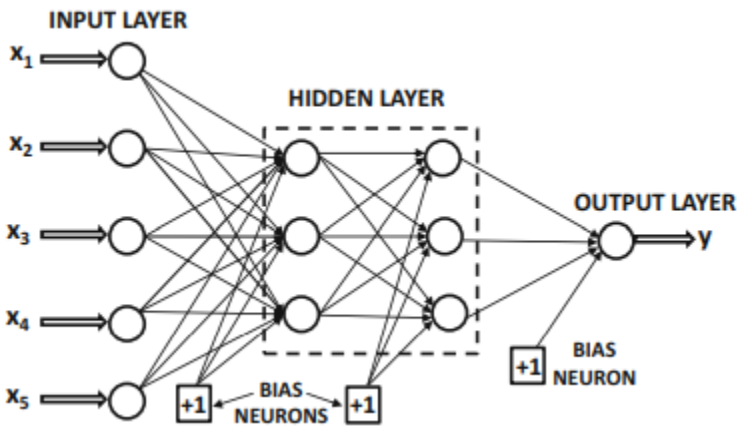


Figure 53. The basic architecture of a feed-forward network with two hidden layers and a single output layer (Reproduced from Goodfellow 2016).

4.3 Numerical Demonstration

In the previous sections, detailed discussion of a variety of supervised machine learning algorithms; their key distinctions, major advantages, and limitations, was provided. Motivated by the research interest, as well as the type of the outcome (i.e., prediction of casing failure) to be analyzed, the discussion was limited to classification problems.

Several learning algorithms were tested on the provided data set and the results were compared in order to reach a solid conclusion on the best/optimal model (in terms of overall performance and prediction accuracy) to be adopted for future analysis. The implemented learning algorithms fall into four major families of techniques: (1) regression, (2) tree-based, (3) support vector machines and (4) deep learning.

Several criteria were considered for evaluating and comparing the different models. For evaluating the prediction accuracy of the model, the “classification error rate” was used for determining the

overall accuracy. As for evaluating the model performance and balance, two different criteria were used: (1) the “false negative” rate and (2) the “false positive” rate.

False negative (FN), formally known as “type II” error, is defined as rejecting a true null hypothesis, while false positive (FP), formally known as “type I” error, is defined as accepting a false null hypothesis. In our case, the “false negative” would correspond to the number of instances when the model did not predict a casing failure while there was failure in reality. In contrast, the “false positive”, in our case, would be the number of instances where the model predicted a casing failure while it was not true in reality (**Figure 54**).

Table of error types		Null hypothesis (H_0) is	
		True	**False**
Decision about null hypothesis (H_0)	**Don't reject**	Correct inference (true negative) (probability = $1-\alpha$)	Type II error (false negative) (probability = β)
	Reject	Type I error (false positive) (probability = α)	Correct inference (true positive) (probability = $1-\beta$)

Figure 54. Tabularized relations between truth/falseness of the null hypothesis and outcomes of the test.

Ideally, the best model would have a zero “false positive”, as well as zero “false negative”. In reality, however, the chance of having similar condition is quite low due to many reasons including – but not limited – sampling biases, model biases, noise, etc. Yet, there are many variations of

"false positive" and "false negative" that might be acceptable, depending on the "type" of the outcome.

For our case, the outcome, that is, casing failure is negative. Accordingly, having a – somewhat – "conservative" model might seem a plausible solution for mitigating that outcome. In other words, we seek a model that lies on the pessimistic side of reality, hence, predicts more casing failure cases than what there are in reality. This, in turn, would reduce the chances of facing a casing failure in the future. Statistically speaking, we seek a model that does not have a high overall accuracy but also a relatively *low* "false negative" and relatively *high* "false positive".

The values corresponding to the three different criteria based on testing different supervised learning techniques on the data set are presented in **Table 13**. Regarding the model prediction accuracy, based on the overall accuracy, the best three models are: (1) artificial neural networks, (2) support vector machine and (2) support vector classifier, sharing the same overall accuracy of 66.67%.

Regarding the model overall performance, based on the "false negative" and "false positive" rate, the best three models are: (1) artificial neural network, (2) support vector machine and (3) logistic regression. Artificial neural network (ANN) has the lowest "false negative" of 33.33% and highest "false positive" of 33.33%. Following that, the support vector machine (SVM) with "false negative" of 33.33% and "false positive" of 50%. Finally, logistic regression with "false negative" of 50% and "false positive" of 33.33%.

Based on the "overall accuracy" alone, one might assume that artificial neural network and support vector machines are both equally good since both share the same overall accuracy (66.67%). However, with considering the other two criteria, ANN would be a better of a choice than SVM

since ANN seems more conservative than SVM.

Besides the previous case, there are many instances where “false positive” and “false negative” can have a huge impact on the final conclusion/evaluation of a model performance. For instance, considering the case of support vector classifier (with tuning) and support vector machine (with tuning), although both might share the same overall accuracy, yet there is one major issue that give the edge to SVM.

From **Table 13**, it is noticed that support vector classifier (with tuning) has a 100% false negative rate and 0% false positive rate. This means that the model is heavily skewed, and it always predicts casing failure. Hence, choosing that model to use in future cases is a risky choice, although it might have a comparable “overall accuracy”.

Although, supervised learning algorithms show a good potential for predicting the outcome of interest (i.e., casing failure), they are basically “black box”. The Majority of machine learning algorithms, although are robust in terms of prediction accuracy, they don’t give the allowance for more control over the problem from the user’s end. In the next section, we tackle that issue by following of the risk assessment approaches, that is, probability-impact risk assessment matrix.

Table 13 Results from testing various supervised learning algorithms of the data set.

Supervised Learning Algorithm	Overall Accuracy (%)	False Positive (Type I Error) (%)	False Negative (Type II Error) (%)
Regression:			
Linear Regression	61.11	8.33	100
Logistic Regression	61.11	33.33	50
Tree-based Methods:			
Decision Tree	61.11	16.67	83.33
Random Forest (Without tuning)	55.56	0	100
Random Forest (With tuning)	61.11	16.67	83.33
Support Vector Machines:			
Support Vector Classifier (Without tuning)	55.56	33.33	66.67
Support Vector Classifier (With tuning)	66.67	0	100
Support Vector Machine (Without tuning)	55.56	50	33.33
Support Vector Machine (With tuning)	66.67	0	100
Deep Learning:			
Artificial Neural Network (Without tuning)	66.67	8.33	83.33
Artificial Neural Network (With tuning)	66.67	33.33	33.33

4.4 Risk Assessment

Classical risk management includes three main phases: (1) a *hazard assessment* including a hazard analysis (hazard characterization and frequency analysis) and a consequence analysis (consequence scenario and severity of consequences), (2) a *risk assessment* including – but not limited to – risk estimation and tolerance criteria, and (3) a proper *risk management plan* through mitigation and feedback. These phases must be sequential but also iterative (**Figure 55**).

The hazard assessment gathers, organizes and summarizes all data relevant to risk assessment and management. It includes qualitative and quantitative characteristics of the hazard, addresses uncertainties and provides a range of forecasts based on plausible scenarios. The hazard is characterized in terms of probability of a measurable physical parameter, exceeding a certain threshold, during a period of time. A recurrence interval is then defined. However, if there is no past experience with a hazard, there is no basis for any forecast. Such a conventional probabilistic approach cannot be applied to non-recurrent hazards. This approach can be considered – somehow – qualitative or semi-quantitative in nature.

In this section, a brief review of risk analysis, as well as survival analysis and their application in identification of critical risk factors and their impact on failure probability, respectively. A particular attention is paid to the likelihood and the consequences of variation in each of potential risk factors. A Probability-Impact Risk Assessment Matrix (PI-RAM) is proposed, to be used as part of a global Risk Management (RM) process during the different phases of an E&P project (drilling and field development).

In the following sections, several historical case studies are presented showcasing the application of the proposed technique in various fields in Oil and Gas (O&G) industry. Following that, a

detailed discussion with demonstration of the proposed workflow on the dataset of interest is provided.

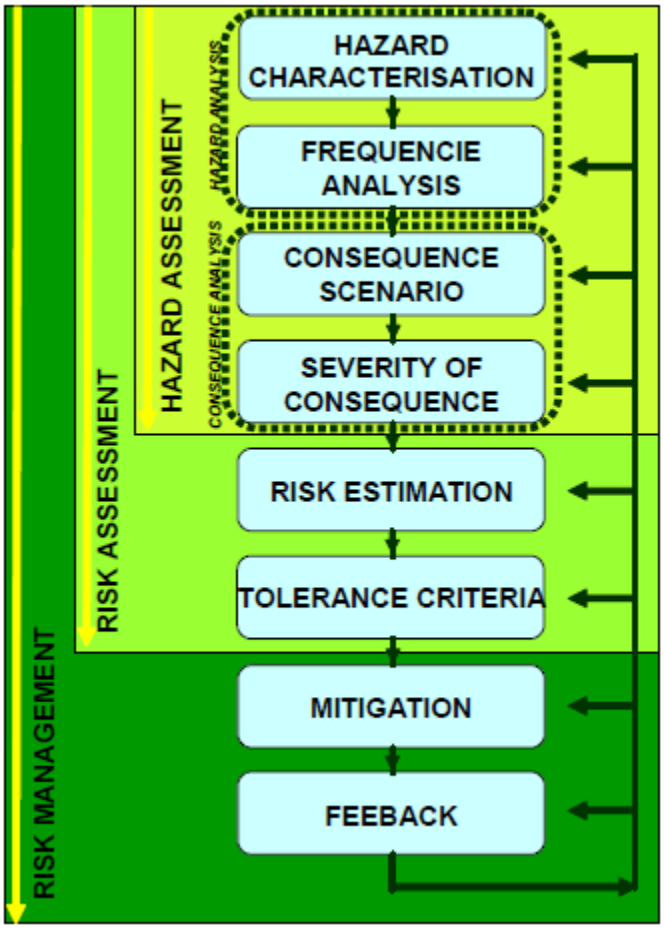


Figure 55. Risk Management flow chart (Reproduced from Cauquil 2009).

4.5 Risk Management Process

Risk Management (RM) generally includes the processes concerned with conducting risk

management planning, identification, analysis, responses, and monitoring and control on a project; most of these processes are updated throughout the project. The objectives of Project Risk Management are to increase the probability and impact of positive events and decrease the probability and impact of events adverse to the project (PMI 2008; Conroy and Sulton 1998; Raz and Michael, 2001; Kayis et al. 2007). The steps involved in the PRM can be summarized as follows:

- Risk Identification
- Risk Assessment/Analysis
- Risk Prioritization
- Risk Response Planning/Management
- Risk Monitoring and Control

The focus of this study is on the second step; that is Risk Assessment (RA). This step can be carried out fully quantitatively, fully qualitatively or semi-quantitatively. Of course, the fully quantitative approach is the most accurate amongst its peers. However, it requires a tremendous amount of data that most of the cases don't have. This, in turn, qualifies the semi-quantitative approach to be a plausible option for conducting risk assessment. This approach requires the construction of Risk Matrix (RM).

4.6 Probability-Impact Risk Assessment Matrix (PI-RAM)

"Risk" can be defined as the statistical probability of damage, loss, or other negative occurrences. Risk statistics can be based only on *historical* information, as it is not possible to calculate the probability of something that has not yet taken place. For this reason, risk assessments are inherently *quantitative*, meaning that they are capable of being measured or expressed in numerical terms based on historical data that have taken place over a period of time.

After *risk levels* have been established for a particular operational situation, a matrix-driven approach is used, based on variables such as frequency and severity, to define what preventative or mitigating measures should be put in place. The *matrix-driven quantitative risk assessment* is a two-step process that (1) builds a quantitative risk assessment matrix using historical data and then (2) performs an evaluation of that matrix.

The process starts with defining risk criteria. Risk criteria are normally defined based on the operational situation. In this study, risk criteria include tubing record (e.g., tubing OD, tubing surface area, tubing grade, tubing design), casing record (e.g. casing OD, casing surface area, casing grade, casing design, casing setting depth, hole size), formation record (e.g. formation composition, corresponding measured depth), fracture record (e.g. fracture depth, total base water volume, total proppant volume), acid record, cement record, as well as, directional survey (e.g., dogleg severity, dogleg severity measured depth, inclination).

Following the definition of appropriate risk criteria, the assessment process begins by gathering historical data on various events for each of the defined criteria, relevant to the specific area of operation (e.g., casing failure).

The frequency versus severity of these events is, then, plotted on a color-coded risk matrix (**Figure 56**), also known as "*Probability-Impact Matrix*", to indicate the actual risk levels in the area of operation (PMI, 2008; Cox, 2008; INCOSE, 2011). In Figure 56, the risks are color-coded, typically using *red* to indicate risks having high probability and high impact; *orange* or *yellow* to indicate risks with intermediate combinations of impact and probability; and *green* to indicate risks having either low impact or low probability or both.

Frequency and consequence categories can be developed in a qualitative or quantitative manner (

Table 14 and **Table 15**). Qualitative schemes (i.e., low, medium, or high) typically qualitative criteria and examples of each category will ensure consistent risk classification. Multiple consequence classification criteria may be required to address different types of consequences. However, in this study, the focus is on one consequence; that is casing failure.

Once assignment of consequences and probabilities is complete, a Risk Matrix (RM) can be used as a mechanism for assigning risk (**Eq. (***60)*), using a risk categorization approach.

$$R = I \times P \qquad (60)$$

Each cell in the matrix corresponds to a specific combination of probability; derived from uncertainty of risk occurrence, and consequence/impact; that is the effect of the contingency, (Hillson and Hulett 2004) and can be assigned a priority number or some other risk descriptor (**Table 16**). An organization must define the categories that it will use to score risks and, more importantly, how it will prioritize and respond to the various levels of risks associated with cells in the matrix".

It is not always possible to get good historical data on all types of incidents, so as a second step in the process, it is imperative to validate the initial quantitative matrix results against a well-defined historical case.

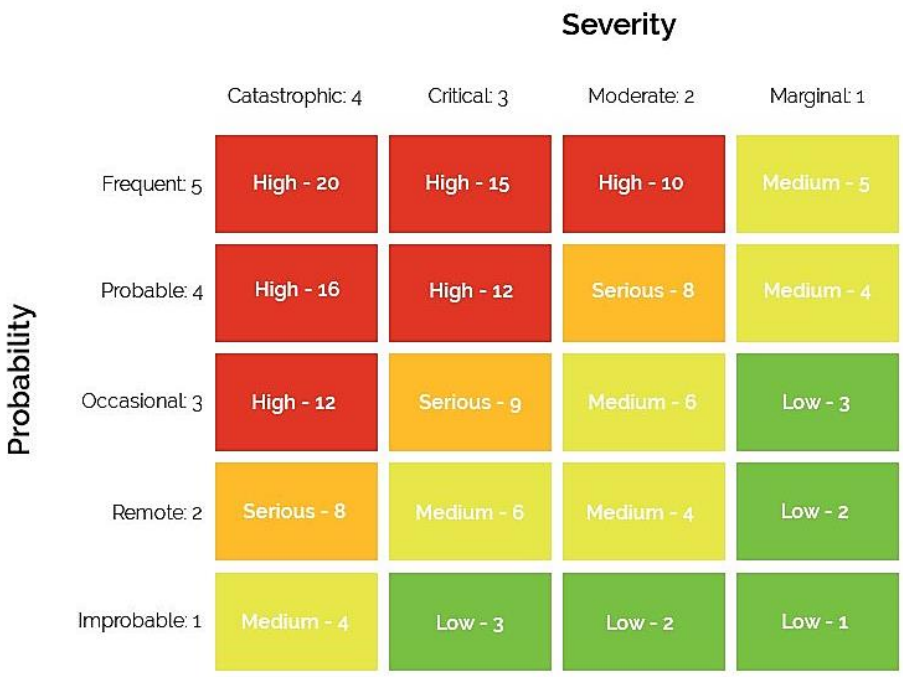


Figure 56. A typical risk matrix.

Table 14 Likelihood score risk.

Likelihood Level	**Score**
Very Low	0 – 20
Low	21 – 40
Medium	41 – 60
High	61 – 80
Very High	81 – 100

Table 15 Impact analysis.

Magnitude of Impact	Impact Definition	Score	Rating
High Impact / High Probability	Very high	1	A
High Impact / Medium Probability Medium Impact / High Probability	High	2	B
Medium Impact / Medium Probability	Medium	3	C
Medium Impact / Low Probability Low Impact / Medium Probability	Low	4	D
Low Impact / Low Probability	Insignificant	5	E

Table 16 Calculation of the exposure risk.

No.	Risk	Occurrence Likelihood		Impact		Degree of Risk Exposure	
		Probability	Score	Probability	Risk	Rating	Score
1	RT1	Very Low	20	Very High	5	E	12.5
2	RT2	Low	40	High	4	D	22
3	RT3	Medium	60	Medium	3	C	31.5
4	RT4	High	80	Low	2	B	41
5	RT5	Very High	100	Very Low	1	A	50.5

4.7 Historical Case Studies

4.7.1 Case (1): ADNOC EISM

In this section, a demonstration of the application of PI-RAM in Oil & Gas industry is presented. The analysis provided in the sequel is reproduced from a study shared by ADNOC (ADNOC CoP, 2004; Alkendi, 2006). In this study, an Environmental Impact Severity Matrix (EISM) was designed to rate the severity of the environmental impacts (EI).

The study investigates several environmental aspects (EA) whether accidental (i.e., non-recurrent); as in the case of an oil spill, or recurrent; as in the case of exhaust emissions and consumption of raw material. The study also discusses the method used for identifying, evaluating and managing the environmental impacts (EIs) of those environmental aspects (EAs).

The motivation behind the choice of EISM – according to the authors – was two-fold. First and foremost was the simplicity of the proposed technique, which made it a viable option for handling the different EAs considered in their study. Second, the allowance that EISM gives for the involvement of a broad range of non-specialist stakeholders; from environmental engineers to instrument engineers, and from operators to asset managers.

One complication of the implementation of EISM, however, was the multitude of parameters that contributed to the magnitude of an EI including the sensitivity and the resilience of the impacted resources, the value of the impacted resources to the ecosystem and the community, the level of damage, exposure of people, breach of regulation, duration of the event causing the impact, and the time it takes the environment to heal. In the study, the parameters were summed in *two* sets: impact effect and impact duration, for simplicity.

Impact Effect is basically the magnitude of change to the environment caused by the EA, without taking into consideration its duration or persistence. It accounts for the value of the subject environmental resource, the magnitude of the impact, human exposure, etc. (**Figure 58**). Meanwhile, *Impact Duration* is the sum of the duration of the subject environmental aspect and how long it will take the environment to recover after ceasing that aspect (**Figure 57**).

This disassociation between the level of change to the environment (i.e., Impact Effect) and its persistence (i.e., Impact Duration) allowed for a *more flexible and consistent* evaluation. The value of disassociating the impact effect and impact duration comes to light when applying this matrix to projects, because it can distinguish between the *temporary impacts* that come with construction and the *long-term impacts* that come with operating the facility.

In the study, the risk assessments (RAs) implementing this method (i.e., EISM) were conducted in *two* stages. First, usage of the EISM (assuming that the scenario will happen). This step yielded the impact *severity* which was used – in second step – in combination with the probability to rate the risk itself.

One challenging aspect of the proposed risk assessment methodology was the long-term operations as most impacts are linked to the expected age of the facility. This means that as long as the facility is running this aspect will be reoccurring. Consequently, the EISM was at risk of being reduced from a 5×5 matrix to a 5×1 vector, or less. This challenge, however, was resolved by overlooking the duration of the operation focusing only on the persistence of that impact after ceasing the aspect.

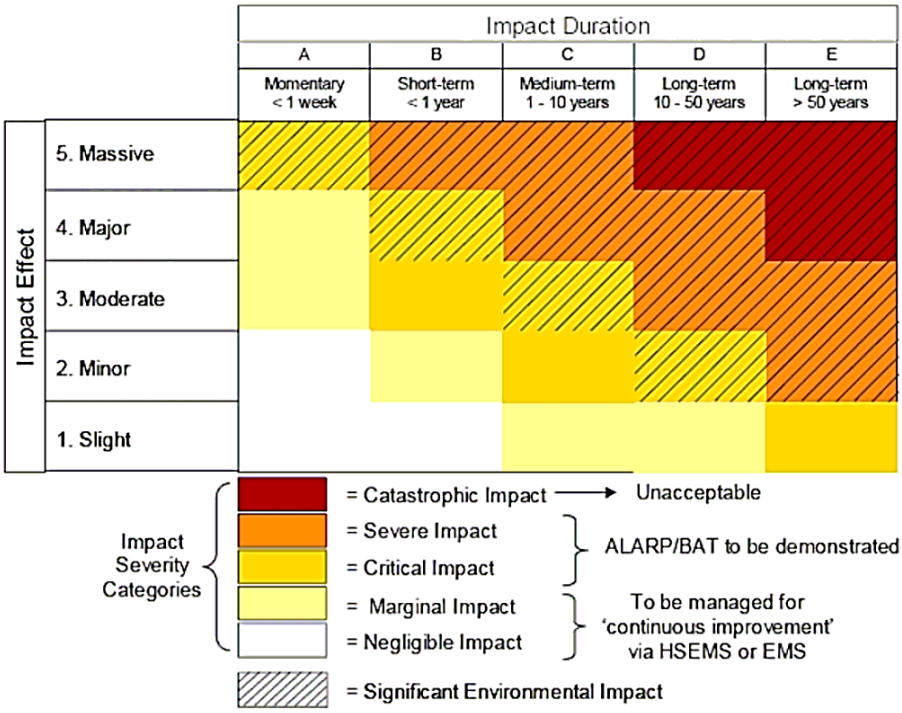


Figure 57. ADNOC Environmental Impact Severity Matrix (Reproduced from Alkendi 2006).

One case that was explored in the study and analyzed through the proposed Environmental Risk Assessment (ERA) method was evaluating risk associated with oil spill in a sensitive area. Two scenarios were designed for comparison reasons.

First scenario assumed an oil spill hits a habitat area (e.g., mangrove forest). Because no intervention was made to clean up the spill and because the forest was expected to protect the spill from any wearing actions, it was likely for the spill to stay for several years. Consequently, the forest was expected to degrade and lose its value as a habitat and a nursery for many marine organisms. In this scenario, the impacted area happened to be relatively small (less than 1 km2). The effect was, then, rated moderate. And because it was likely to stay for over 10 years, therefore, the impact was rated *Severe*. This outcome is the consequence parameter in the risk matrix, which

also takes into consideration the probability.

The second scenario, on the contrary, assumed immediate clean up. Accordingly, the evaluation was different. Removal of the oil minimized the direct impact and accelerate recovery to less than 10 years. Impact severity rating was, therefore, reduced to *Critical* status.

Effect	Area	Air	Water	Waste	Biodiversity	Resources	Regulation	Society
Massive (5)	▪ Extending over a wide area (>100 km2)	▪ Severe acute impact on the receptor(s) potentially leading to fatality. ▪ Significant deterioration of air quality on a regional or global level.	▪ Regional contamination or depletion of potable groundwater. ▪ Regional contamination of marine environment.	▪ Severe uncontrolled generation, handling, storage or disposal of priority hazardous waste leading to catastrophic contamination of valued groundwater, soil and/or loss of non-renewable natural resources.	▪ Catastrophic effect on regionally important attributes of the ecological environment is observable or measurable over a wide area.	▪ Significant damage to archaeological, cultural or natural resources regional/international importance.	▪ Constant very high exceedence of statutory or prescribed limits (assessed not to be ALARP) ▪ Immediate intervention by third parties or governmental body	▪ High profile community outrage.
Major (4)	▪ Extending over 10 km2 from site	▪ Acute impact on the receptors. ▪ Significant deterioration of air quality on a national level.	▪ Serious local contamination or depletion of potable groundwater. ▪ Serious local or considerable national contamination of marine environment.	▪ Serious uncontrolled generation, handling, storage or disposal of hazardous waste leading to serious contamination of valued groundwater, soil and/or loss of non-renewable natural resources.	▪ A serious effect on locally important attributes of the ecological environment is observable or measurable locally.	▪ Significant damage to archaeological, cultural or natural resources of national importance.	▪ Constant exceedence of statutory or prescribed limits (assessed not to be ALARP) ▪ Significant alarm raised by third parties or governmental body with instruction to mitigate	▪ Serious community concern and complaints.
Moderate (3)	▪ Extending over 1 km2 area.	▪ Significant deterioration of air quality on a local level. ▪ Strong odour and irritation caused by deterioration of air quality.	▪ Considerable local contamination or depletion of potable groundwater. ▪ Considerable local contamination of marine environment.	▪ Considerable uncontrolled generation, handling, storage or disposal of haz-waste, or serious uncontrolled for non-haz-waste, leading to considerable land contamination putting valued groundwater at risk, and natural resources.	▪ A considerable effect on significant attributes of the ecological environment is observable or measurable.	▪ Significant damage to archaeological, cultural or natural resources of local importance.	▪ Multiple intermittent exceedence of statutory or prescribed limits (assessed not to be ALARP) ▪ Concern raised from third parties or governmental body	▪ Considerable community concern and potentially a single community complaint
Minor (2)	▪ Extending over 10,000 m2 area.	▪ Significant deterioration of air quality on a limited level. ▪ Some odour and irritation caused by deterioration of air quality	▪ Limited degradation of potable groundwater quality or reduction of its volume ▪ Limited contamination of marine environment.	▪ Limited uncontrolled generation, handling, storage or disposal of waste leading to limited loss of natural resources and/or land contamination without putting valued groundwater at risk.	▪ A limited effect on significant attributes of the ecological environment is observable or measurable.	▪ Limited impact on natural resources or damage any archaeological or cultural resources.	▪ Multiple, constant exceedence of statutory or prescribed limits but assessed to be ALARP ▪ No complaint from third parties or governmental body	▪ Some community concern raised.
Slight (1)	▪ Measurable above background ▪ Confined within fenceline	▪ Slight change of ambient air quality over a limited area. ▪ Some local influence.	▪ Slight degradation of quality or reduction of volume of potable groundwater. ▪ Some contamination of marine environment.	▪ Controlled generation, handling, storage or disposal of waste.	▪ An effect on any attribute of the ecological environment is observable or measurable above background.	▪ Some impact on natural resources or damage to archaeological or cultural resources.	▪ Some exceedence of statutory or prescribed limits but assessed to be ALARP. ▪ No third party or government concern.	▪ Possible community focus.

Figure 58. Impact Effect Comparison Guide (Reproduced from Alkendi 2006).

4.7.2 Case (2): Gas-Hydrate Hazard

In this section, another demonstration of the application of PI-RAM in Oil & Gas industry is presented. The analysis provided in the sequel is reproduced from a study presented by Cauquil, Total SA (Cauquil 2009). In this study, a gas-hydrate (GH) hazard-consequence risk assessment matrix was proposed, to be used as part of a global risk management process (RMP) during the different phases of a Deepwater E&P project (drilling and field development).

Unlike previous case, this study handles a *non-recurrent* geohazards; that is gas hydrate (GH). Normally, the hazard is characterized in terms of probability of a measurable physical parameter, exceeding a certain threshold, during a period of time (known as recurrence interval). However, for non-recurrent hazards (e.g., GH) where there is *no past experience* with a hazard, there is *no basis* for any forecast. Accordingly, conventional probabilistic approaches cannot be applied to non-recurrent geohazards like Gas Hydrates (GH).

The study proposes *knowledge-based* evaluation workflow for the GH hazard analysis (and the possible consequences). As for any PI-RAM based workflow, the adopted approach is qualitative or semi-quantitative. The limitation of implementing a fully quantitative approach was the lack of extensive and accurate data, as well as the spatial variability of factors and parameters.

Generally, for hazard assessment, the concept of "*Knowledge*" implies: (1) capture of the location of the hazardous feature, (2) the potential magnitude of the hazard event (i.e., Impact) and the rate of occurrence/recurrence of the hazard event (i.e., Likelihood).

In this study, the *location* was related to the vertical (depth & thickness) and spatial extension of the GH formation and was determined by the data acquisition/interpretation. The *magnitude* of the GH hazard meant the type and concentration of GH in place and – also – was determined by the

data acquisition/interpretation. As for the third – and final – aspect that is the *likelihood* of the hazard, although important, was not relevant for GH hazard assessment since it is non-recurrent. In other words, at a specific location, there was no rate of occurrence/recurrence of GH, as GH were either present or not.

The GH hazard assessment was conducted on two stages. First, a Deterministic Hazard Assessment (DHA) was implemented – based on geological knowledge and field observations – to define the Maximal Credible Event (MCE) as the largest possible event that could be produced by a geohazard feature. At a second stage, a probabilistic analysis using geostatistics was founded based on the DHA in order to spatially extrapolate the GH occurrence.

The review of GH proxies provided a conceptual geological model based on available indices. Based on those geological models, the qualitative likelihood of GH occurrence based on the proxy types defined previously was proposed (**Figure 59**) relying on the number of GH indices collected over the area of interest. Based on those information, GH hazard-consequence risk matrix was proposed (**Figure 60**).

INDIRECT INDICES			DIRECT INDICES		UNCERTAINTY	LIKELIHOOD	LEVEL
Model	Geophy 1	Geophy 2	Seafloor	In situ			
Yes	Yes	Yes	Yes	Yes	Very low	Certain	1
Yes	Yes	Yes	Yes	No	Low	Almost certain	2
Yes	Yes	Yes	No	No	Medium	Very likely	3
Yes	Yes	No	No	No	High	Likely	4
Yes	No	No	No	No	Very high	Unlikely	5

Figure 59. Gas Hydrate occurrence likelihood levels (Reproduced from Cauquil 2009).

LIKELIHOOD			CONSEQUENCE			RISK	
Proxies Type 1, 2, 3, 4 & 5 Occurrence of GH is proven by all indirect and direct indices GH qualification, quantification and stability modelling is possible (well-bore stabilty and long term dissociation / dissolution models)	CERTAIN	1	GH proven to be present and having a strong impact on facilities if no mitigation / remediation measures applied Extensive damages to the entire installation (field layout) with escalation processes (chain processes) Complete loss and / or full disruption of activity (production) Major asset and environmental impact	CATASTROPHIC	A	VERY HIGH	VH
Proxies Type 1, 2, 3 & 4 only GH may be present from a thermodynamic stand point and their occurrence is suspected from conventional geophysical indices (Geophy 1 & 2) GH quantification (concentration) may be estimated from dedicated data processing (Vp, inversion...) but without in-situ calibration Proxy Type 4 shows that free gas possibly related to GH dissociation / dissolution affect the seafloor - if sampled, seafloor GH characterisation (C1, C2...) may be used to refine modelling and quantification	ALMOST CERTAIN	2	Extensive damages to the particular installation with impact on it functional integrity but without collateral impact on remote installation Minor asset and environmental impact	MAJOR	B	HIGH	HR
Proxies Type 1, 2 & 3 only GH may be present from a thermodynamic stand point and their occurrence is suspected from conventional geophysical indices (Geophy 1 & 2) GH quantification (concentration) may be estimated from dedicated data processing (Vp, inversion...) but without in-situ calibration	VERY LIKELY	3	Moderate damages to the particular installation with no impact on it functional integrity Damages must be handled in order to avoid escalation No asset and environmental impact	MEDIUM	C	MODERATE	MR
Proxies Type 1 & 2 only GH may be present from a thermodynamic stand point and their occurrence is suspected from conventional geophysical indices (Geophy 1: BSR...)	LIKELY	4	Not serious enough to cause damages	MINOR	D	LOW	LR
Proxy Type 1 only GH may be present from a thermodynamic stand point but neither indirect nor direct indices are observed	UNLIKELY	5	Not to be considered	NONE	E	INEXISTANT	NR

Figure 60. Gas Hydrate hazard-consequence risk matrix (Reproduced from Cauquil 2009).

4.8 Numerical Demonstration

In this section, I apply the proposed workflow (PI-RAM) on the dataset of interest. Generally, risk analysis (RA) can be applied to multiple outcomes simultaneously (Shen et al. 2014). However, in this study, the focus of the presented RA is on a *single* outcome that is casing failure. The results of the analysis are presented in two-dimensional PI-RAMs: (1) Likelihood, i.e., frequency of the incidence occurrence, and (2) consequence, i.e., impact on risk probability of casing failure if it happens.

Figure 61 showcases a snippet of the calculations involved in the construction of risk assessment matrices for potential risk factors, in particular, "fracturing season". Results of the RA and final risk matrices (RMs) for all the remaining potential risk factors are summarized in Appendix (E). In Figure 61, three plots are presented: probability-impact risk assessment matrix (Figure 61; *left*), risk values for different variations of "fracturing season" risk factor (Figure 61; *top-right*) and frequency analysis of the different seasons within "fracturing season" risk factor (Figure 61; *bottom-right*).

Regarding the "*probability*", it was arranged and valued according to the categorization of each risk factor, separately. Frequency analysis (Section 3.1) was used as a basis for determining the occurrence probability of different sublevels/categories within each potential risk factor (Figure 61; *bottom-right*). Survival analysis (Chapter 3), in particular non-parametric survival estimators, e.g. Kaplan-Meier estimator (Section 3.3.1), was used as a basis for defining the significance of sublevels/categories within each potential risk factor.

As for the "*consequence*", it was arranged and valued according to the reached measured depth. Survival analysis (Chapter 3), in particular semi-parametric survival estimators, e.g. Cox model

(Section 3.3.2), was used as a basis for defining the impact magnitude of different variations within each risk factor with respect to reached measured depth.

Upon defining the two parameters; "*probability*" and "*consequence*", and using Eq. (60), the resulting risk values were assessed, then, classified on a three-level severity scale; slight, moderate and severe, (Figure 61; *left*). The construction of the severity scale was based, purely, on the researcher's judgement and knowledge of historical casing failure case studies (Figure 61; *top-right*).

As shown in the risk assessment matrix for "fracturing season", (Figure 61; *left*), fracturing during *any* season at relatively low depths (< 8500 ft) imposes no risk at casing failure. However, starting from intermediate depths (> 8500 ft), the time (or generally, season) during which the fracturing operation is conducted starts to play a rule in defining the risk imposed on casing failure. For intermediate depths; ranging from 8,500-10,000 ft, as well as high depths; greater than 10,000 ft, fracturing during "spring" or "winter" imposes the highest risk compared to the two other seasons. This could be partially explained by the high temperature difference between the reservoir and fracturing fluid assuming that the temperature of fracturing fluid resembles that of well surface. This high temperature difference leads, in turn, to high shrinkage of lateral section, accordingly, a high bending stress on the casing string.

The reasons why risk imposed by fracturing during "spring" or "winter" is classified as *moderate* (represented by "yellow" color) for intermediate depths, as opposed to *severe* (represented by "red" color) at high depths, is that the bottomhole temperature tends to increase with increasing depth. This means that, for high depths (> 10,000 ft), the temperature difference between fracturing fluid and reservoir will be *higher* than that for intermediate depths (8,500-10,000 ft). Accordingly,

shrinkage of lateral section will be *greater* for high depths than intermediate depths. This, in turn, leads to *higher* bending stress and *higher* overall risk of casing failure at high depths than intermediate depths.

The importance of having probability-impact risk assessment matrices was to serve as "feedback" for the initial predictions obtained from conventional machine learning algorithms, as shown in Section 1.3. In other words, PI-RAMs were used as basis for adjusting the input design specifications in case they were initially classified as high risk by any of the conventional classification machine learning algorithms (Section 4.2). For validation, a series of already identified failure cases were reassessed using the developed PI-RAMs (Figure E-21).

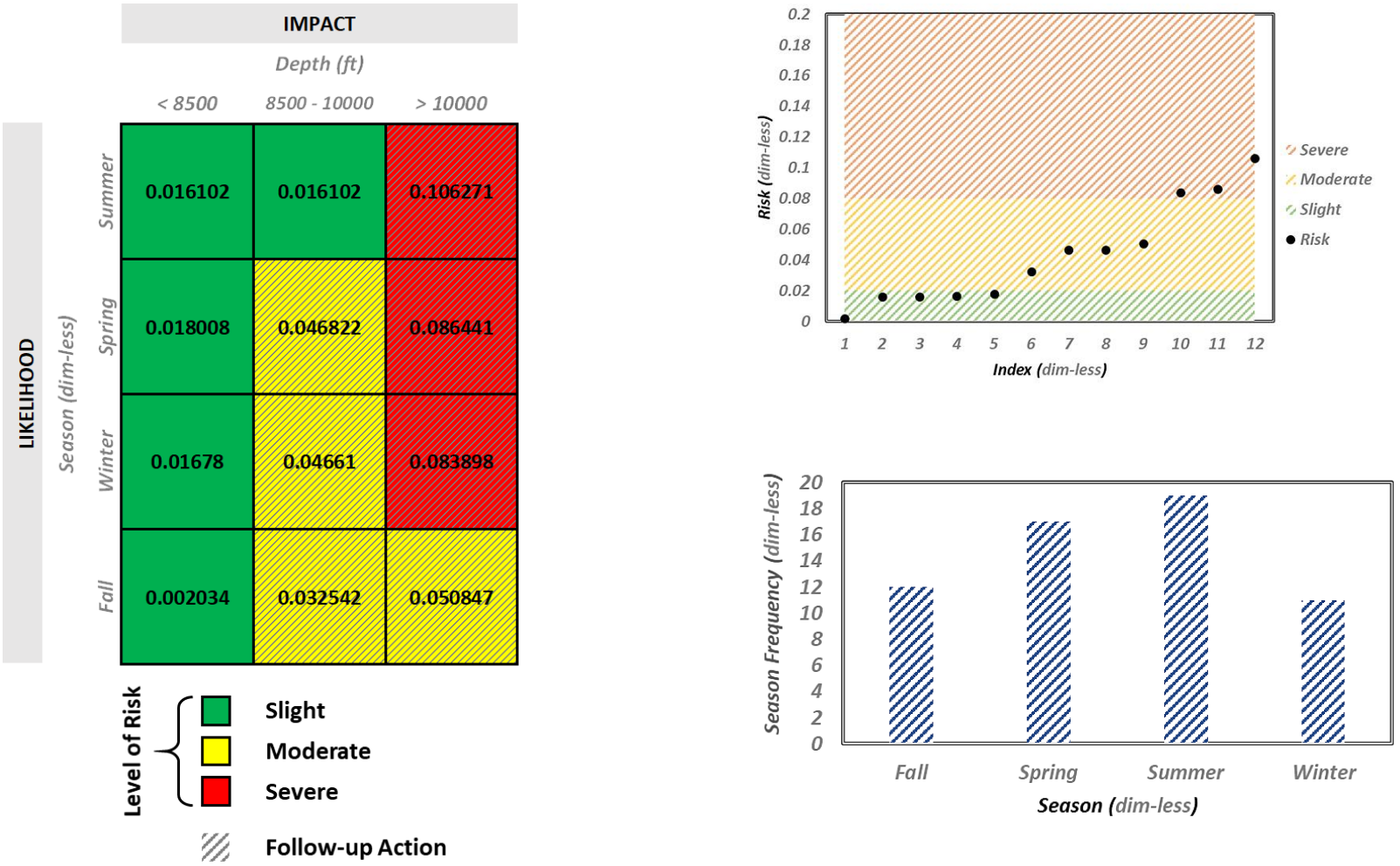


Figure 61. (left) Risk Assessment Results developed in this study for "Season" Risk Factor. (Top-right) Calculated Risk Values for Different Variations of "Season" Risk Factor. (Bottom-right) Summery of Frequency Analysis Used for Determining Likelihood of Various Combinations of "Season" Risk Factor.

4.8.1 Validation Case (1)

Now, after I have explained the two-step "prediction-correction" procedure adopted in the proposed casing failure mitigation tool, it is important that I test/validate the performance of the developed tool against multiple historical cases that experienced casing failure. Figure 62 showcases one of the history wells that experienced casing failure. Figure 62 comprises of three levels; layout of initial design specifications (Figure 62; *top level*), layout of adjusted design specifications after first round of iterations (Figure 62; *middle level*) and layout of the final design specifications after the second round of iteration (Figure 62; *bottom layer*). The number of features considers were eleven features based on the results obtained from risk analysis (Chapter 2). Each feature was color-coded according to the risk imposed by its current value. Three levels of risk severity were considered: high (represented by "red" color), moderate (represented by "yellow" color) and low (represented by "green" color)

Initial design specifications included: (1) acid treatment, (2) no casing cementing, (3) dogleg bending stress of 43.7k lb./ft^2, (4) maximum inclination of 95 deg., (5) dogleg severity of 13.36 deg./100ft, (6) fracture temperature of 45 ^{o}F and (7) drilling-fracturing time interval of 46 days.

The aforementioned initial design specifications were fed to the tool for evaluation, in particular, the optimal classification model. Preliminary evaluation had shown a probability of casing failure and classified the case as a "*high*" risk; indicated by "red" color (Figure 62a). Following that, the tool started on adjusting the design specifications' inputs in the direction of reducing the risk of casing failure using feed from PI-RAMs. After the first iteration, the risk was reduced to "*moderate*" severity class; indicated by "yellow" color (Figure 62b) with suggestion of: (1) no implementation of acid treatment, (2) application of casing cementing, (3) reduction of dogleg bending stress to 42.6k lb./ft^2, (4) reduction of inclination to 93.9 deg., (5) reduction of dogleg severity to 13

deg./100ft, (6) increase of fracture temperature to 58 ^{o}F and finally (7) increasing the drilling-fracturing time interval to 44 days.

Further iteration over the design specifications has reduced the risk of casing failure to "*low*" ; indicated by "green" color (Figure 62c) with the following suggestions: (1) no implementation of acid treatment, (2) application of casing cementing, (3) reduction of dogleg bending stress to 39.2k lb./ft^2, (4) reduction of inclination to 87.2 deg., (5) reduction of dogleg severity to 12.15 deg./100ft, (6) increase of fracture temperature to 64 ^{o}F and finally (7) increasing the drilling-fracturing time interval to 133 days.

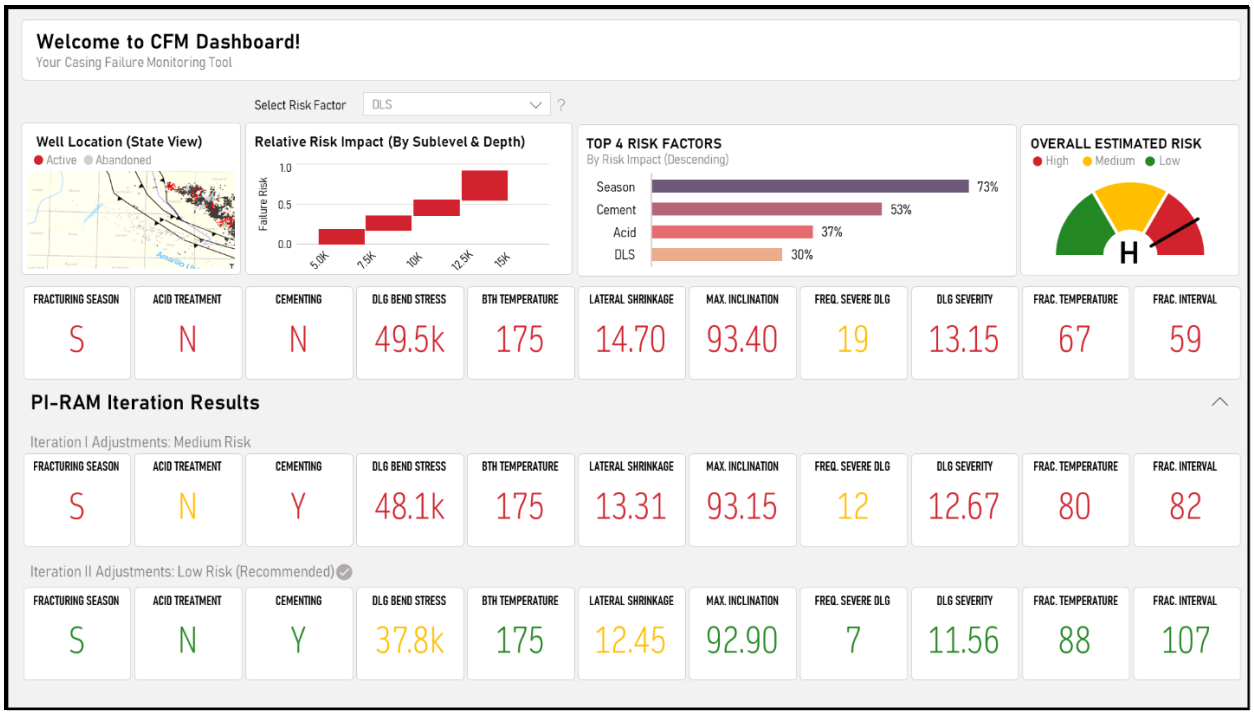


Figure 62. A snapshot of the CDM dashboard for a historical case from Granite Wash data set with an initial status of high risk. Using PI-RAMs, adjustments were done to the features of interest leading to reduction in risk status to medium after the first iteration. Further adjustments to the features of interest using PI-RAMs led to reduction of risk status to low.

4.8.2 Validation Case (2)

Initial design specifications included: (1) no acid treatment, (2) no casing cementing, (3) dogleg bending stress of 49.5k lb./ft^2, (4) maximum inclination of 93.4 deg., (5) dogleg severity of 13.15 deg./100ft, (6) fracture temperature of 67 ^{o}F and (7) drilling-fracturing time interval of 59 days.

The aforementioned initial design specifications were fed to the tool for evaluation, in particular, the optimal classification model. Preliminary evaluation had shown a probability of casing failure and classified the case as a "*high*" risk; indicated by "red" color (Figure 48a). Following that, the tool started on adjusting the design specifications' inputs in the direction of reducing the risk of casing failure using feed from PI-RAMs. After the first iteration, the risk was still classified as "*high*" severity class; indicated by "red" color (Figure 48b) with suggestion of: (1) no implementation of acid treatment, (2) application of casing cementing, (3) reduction of dogleg bending stress to 48.1k lb./ft^2, (4) reduction of inclination to 93.15 deg., (5) reduction of dogleg severity to 12.67 deg./100ft, (6) increase of fracture temperature to 80 ^{o}F and finally (7) increasing the drilling-fracturing time interval to 82 days.

Further iteration over the design specifications has reduced the risk of casing failure to "*low*" ; indicated by "green" color (Figure 48c) with the following suggestions: (1) no implementation of acid treatment, (2) application of casing cementing, (3) reduction of dogleg bending stress to 37.8k lb./ft^2, (4) reduction of inclination to 92.9 deg., (5) reduction of dogleg severity to 11.56 deg./100ft, (6) increase of fracture temperature to 88 ^{o}F and finally (7) increasing the drilling-fracturing time interval to 107 days.

Initial Status (Tool Input): High Risk

Fracturing Season	Acid Treatment	Cementing	Dogleg Bending Stress	Bottomhole Temperature	Lateral Shrinkage	Maximum Inclination	Frequency of Severe Dogleg (> 10)	Dogleg Severity	Fracturing Temperature	Drilling-Fracturing Interval
Summer	No	No	49479.43	175	14.7	93.4	19	13.15	67	59

Phase I Adjustments: High Risk

Fracturing Season	Acid Treatment	Cementing	Dogleg Bending Stress	Bottomhole Temperature	Lateral Shrinkage	Maximum Inclination	Frequency of Severe Dogleg (> 10)	Dogleg Severity	Fracturing Temperature	Drilling-Fracturing Interval
Summer	No	Yes	48037.23	175	13.31	93.15	12	12.67	80	82

Phase II Adjustments: Low Risk

Fracturing Season	Acid Treatment	Cementing	Dogleg Bending Stress	Bottomhole Temperature	Lateral Shrinkage	Maximum Inclination	Frequency of Severe Dogleg (> 10)	Dogleg Severity	Fracturing Temperature	Drilling-Fracturing Interval
Summer	No	Yes	37853.5	175	12.45	92.9	7	11.56	88	107

Figure 63. (a) Application of semi-quantitative probability-impact risk assessment matrices (PI-RAMs) on a case from Granite Wash data set with initial status of "high" risk. (b) Using PI-RAMs, adjustments were done to the features of interest leading to reduction in risk status to "medium". (c) Further adjustments to the features of interest using PI-RAMs led to reduction of risk status to "low".

4.8.3 Validation Case (3)

Initial design specifications included: (1) no acid treatment, (2) no casing cementing, (3) dogleg bending stress of 48.9k lb./ft^2, (4) maximum inclination of 95 deg., (5) dogleg severity of 18.97 deg./100ft, (6) fracture temperature of 19 ^{o}F and (7) drilling-fracturing time interval of 82 days.

The aforementioned initial design specifications were fed to the tool for evaluation, in particular, the optimal classification model. Preliminary evaluation had shown a probability of casing failure and classified the case as a "*high*" risk; indicated by "red" color (Figure 49a). Following that, the tool started on adjusting the design specifications' inputs in the direction of reducing the risk of casing failure using feed from PI-RAMs. After the first iteration, the risk was reduced to "*moderate*" severity class; indicated by "yellow" color (Figure 49b) with suggestion of: (1) no implementation of acid treatment, (2) application of casing cementing, (3) reduction of dogleg

bending stress to 62.1 k lb./ft^2, (4) reduction of inclination to 93.4 deg., (5) reduction of dogleg severity to 14.27 deg./100ft, (6) increase of fracture temperature to 41 ^{o}F and finally (7) increasing the drilling-fracturing time interval to 93 days.

Further iteration over the design specifications has reduced the risk of casing failure to "*low*" ; indicated by "green" color (Figure 49c) with the following suggestions: (1) no implementation of acid treatment, (2) application of casing cementing, (3) reduction of dogleg bending stress to 42.8k lb./ft^2, (4) reduction of inclination to 93.2 deg., (5) reduction of dogleg severity to 12 deg./100ft, (6) increase of fracture temperature to 42 ^{o}F and finally (7) increasing the drilling-fracturing time interval to 200 days.

Initial Status (Tool Input): High Risk

Fracturing Season	Acid Treatment	Cementing	Dogleg Bending Stress	Bottomhole Temperature	Lateral Shrinkage	Maximum Inclination	Frequency of Severe Dogleg (> 10)	Dogleg Severity	Fracturing Temperature	Drilling-Fracturing Interval
Winter	No	No	84905.49	150	15.2	95	17	18.97	19	82

Phase I Adjustments: Moderate Risk

Fracturing Season	Acid Treatment	Cementing	Dogleg Bending Stress	Bottomhole Temperature	Lateral Shrinkage	Maximum Inclination	Frequency of Severe Dogleg (> 10)	Dogleg Severity	Fracturing Temperature	Drilling-Fracturing Interval
Winter	No	Yes	62117.67	150	14.31	93.4	13	14.27	41	93

Phase II Adjustments: Low Risk

Fracturing Season	Acid Treatment	Cementing	Dogleg Bending Stress	Bottomhole Temperature	Lateral Shrinkage	Maximum Inclination	Frequency of Severe Dogleg (> 10)	Dogleg Severity	Fracturing Temperature	Drilling-Fracturing Interval
Winter	No	Yes	42830.74	150	13.57	93.2	12	13.08	42	200

Figure 64. (a) Application of semi-quantitative probability-impact risk assessment matrices (PI-RAMs) on a case from Granite Wash data set with initial status of "high" risk. (b) Using PI-RAMs, adjustments were done to the features of interest leading to reduction in risk status to "medium". (c) Further adjustments to the features of interest using PI-RAMs led to reduction of risk status to "low".

This “prediction-correction” procedure is what gives the edge to the developed workflow in this study over other data-driven workflows presented/introduced in literature, specifically, in this research area (casing failure). Following this two-step procedure, the proposed tool can handle the risk of casing failure proactively rather than reactively. This, in turn, gives the allowance for drillers and drilling engineers to adjust their design specifications in order to avoid or mitigate potential casing failure.

4.9 Conclusions

In this chapter, I leveraged the information provided by the analyses conducted in Chapter 2 and Chapter 3 to construct the proposed automated “casing failure mitigation” tool based on a two-step “prediction-correction” procedure. Based on the input information (initial design specifications), the tool provides initial prediction of the probability of casing failure, then automatically adjust, or correct, the high-risk design specifications in the direction of reducing the overall risk of casing failure below a pre-defined threshold. This gives the allowance for drillers and drilling managers to successfully modify their design and ultimately mitigate casing failure which was otherwise infeasible.

Regarding the “prediction” part of the tool, machine learning techniques, particularly, classification algorithms were implemented to fulfill that task. Due to the type of the outcome, only classification algorithms were considered, such as, logistic regression, basic decision trees, random forests, support vector classifier, support vector machines and artificial neural networks.

The different models were trained and tested on the provided data set, and results were compared based on two key elements: (1) prediction accuracy and (2) model balance and overall performance. For the prediction accuracy, the “overall classification rate” criterion was used.

Regarding the model performance and balance, both “false positive rate” and “false negative rate” were implemented.

Comparison of different classification models, based on the aforementioned criteria, showed the superiority of artificial neural network (ANN) compared to other models, in terms of, prediction accuracy (67% E), overall performance and model balance (33.33% FP, 33.33% FN). Support vector machines (SVMs) came second (67% E, 50% FP, 33.33% FN), followed by logistic regression (61% E, 33.33% FP, 50% FN).

As for the “correction” part of the tool, I sought risk assessment techniques with a focus on semi-quantitative risk assessment matrices (PI-RAMs). Construction of PI-RAMs involved use of two key inputs, namely, exposure occurrence likelihood and exposure impact. Regarding likelihood estimation, non-parametric survival curve estimators along with frequency analysis were used. As for exposure impact, semi-parametric survival estimators along with risk analysis were implemented.

Upon the development of the “automated casing failure mitigation” tool, it was evaluated/validated against some of the historical wells that experienced casing failure. Results for three of the history wells that experienced casing failure were presented. Based on the initial design specifications fed to the tool, preliminary evaluation concluded a probability of casing failure and classified the case as a “high” risk. Following that, the tool started on adjusting the design specifications inputs in the direction of reducing the risk of casing failure. After the first iteration, the risk was reduced to “moderate” severity class. Further iteration over the design specifications has reduced the risk of casing failure to “low”.

This “prediction-correction” procedure is what gives the edge to the developed workflow in this study over other data-driven workflows presented/introduced in literature, specifically, in the

casing failure research area. Unlike previous contributions, physics-based or data-driven, that handled casing failure *reactively*, the developed workflow in this study can handle the casing failure rather *proactively*, hence, give the allowance for drillers and drilling managers to constantly adjust their designs and ultimately mitigate potential casing failures.

CHAPTER V
DATA DRIVEN PHYSICS-GUIDED CASING FATIGUE LIFE ESTIMATION

Reader Guide: In this chapter, I introduce a hybrid fatigue life estimator based on modern machine learning algorithms coupled with conventional analytical estimators. In section 5.1, I start the discussion by introducing casing fatigue failure as the mode of interest for sequel analysis. Then, in section 5.1.1, I discuss the general framework adopted in conventional fatigue life estimators, along with the most popular analytical models used in this regard. Following that, in section 5.1.2, I highlight the main shortcomings of the analytical models through history cases and provide key remarks on potential reasons in section 5.1.3. Afterwards, in section 5.2., I introduce the proposed hybrid model where I showcase the proposed framework along with key adjustments to address defects in most analytical models. Finally, in section 5.3, I test the proposed solution through a simulated case (using ANSYS) for a number of scenarios and validate results against conventional techniques.

5.1. Fatigue Failure

In their study, Gao and Hsu, 1998 defined fatigue as an irreversible, gradual, and localized structural damage that arises and continuously accumulates when a material is structurally exposed to cyclic loading.

During drilling in stormy weather in offshore operations; currents and wave, and heavy motions of the sea are transported down the riser to the wellhead, to the casing structure causing fatigue failures at essential casing joints and welds (Lim et al. 2012). Furthermore, casing pipe fatigue failure could happen either during production, due to alternating temperatures and cyclic loading of pressure or during stimulation, due to temperature differentials between the reservoir fluids and stimulation fluids (Kiran et al. 2018). Casing joint fatigue is another mode that has been reported numerously during multistage hydraulic fracturing with varying temperature between stages (Liu

et al. 2018). In all those cases, those cyclic variations induce thermal axial casing stress (compressive and tensile) depending on the direction of the change. As a result, two major effects may occur from those thermally induced stresses: 1) casing hot-yield and the resultant casing collapse failure, and 2) casing fatigue.

5.1.1. Conventional Fatigue Life Estimators

Typical steps involved in casing fatigue analysis workflow are as follows:

1. Definition of wellbore configuration (e.g., casing size, steel grade, connection type, cement top, etc.)
2. Definition of operations history, including cyclic operations (e.g., steam injection, plug & perf, etc.)
3. Thermal flow simulations (using commercial software) to obtain temperature and pressure profiles
4. Selection of critical locations (stress concentration points) such as connections
5. Calculations of local casing loads/stresses (mainly, thermal loads)
6. Calculations of elastic-plastic response at the critical locations (to determine strength)
7. Estimation of fatigue life (how many years to failure)

In case of low-cycle fatigue (LCF) operations, e.g., steam injection, multistage hydraulic fracturing, *strain-based methods* are used to implement steps 5–7 in the aforementioned fatigue workflow. One example of strain-based methods is *Manson's universal slope method*, which will be considered for subsequent discussion due to its simplicity.

According to Manson's equation, fatigue life, N, is correlated to total cyclic change of local strain, $\Delta\varepsilon_t$, as follows:

$$\Delta\varepsilon_t = K_f \frac{\sigma_u}{E} N^{-0.12} + {\varepsilon_f}^{0.6} N^{-0.6} \tag{61}$$

Based on local maximum stress, $\Delta\sigma_t$, various methods could be used for calculating the total cyclic change of local strain, $\Delta\varepsilon_t$.

In case of *elastic* behavior (i.e., local maximum stress, σ_t, is less than material yield stress, σ_y), then, the following equation could be used for calculating the total cyclic change of local strain, $\Delta\varepsilon_t$.

$$\Delta\varepsilon_t = K_f\,\Delta\varepsilon_p = K_f\,\alpha\,\Delta T \tag{62}$$

In case of *plastic* behavior (i.e., local maximum stress, σ_t, is greater than material yield stress, σ_y), then, Neuber's equation could be used for calculating the total cyclic change of local strain, $\Delta\varepsilon_t$.

$$\Delta\varepsilon_t = \frac{\Delta\sigma_t}{E} + 2\left(\frac{\Delta\sigma_t}{2K^*}\right)^{1/n^*} \tag{63}$$

Figure 44 is a schematic that summarizes steps involved in the calculation of total cyclic change of local strain, $\Delta\varepsilon_t$, and ultimately the fatigue life, N.

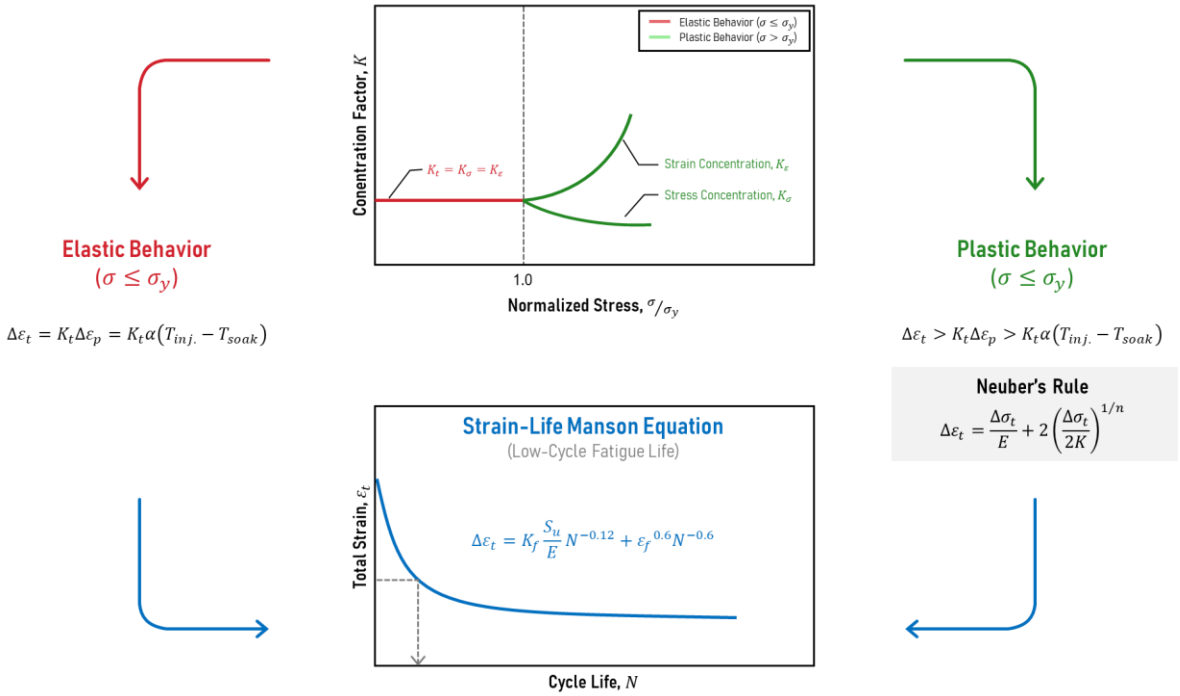


Figure 44. A schematic of steps involved in calculation of total cyclic change of local strain and fatigue life estimation.

From former discussions of fatigue workflow and from the schematic shown in Figure 44, it is evident that the correct calculation of total cyclic change of local strain at concentration points is key to the accurate estimation of fatigue life of various casing parts and, in turn, mitigation of casing failure in the future.

However, estimation of cyclic local strain could be challenging due to multiple factors that will be highlighted in the next section. Wrong cyclic strain calculations will lead to false estimation of casing fatigue life. This, in turn, would lead to abuse of casing material beyond its capability and increased possibility of failure.

In the next section, two instances are highlighted where conventional models failed to calculate the true induced strain leading to false estimations of different casing parts' fatigue life and, eventually, the occurrence of casing failure.

5.1.2. Instances of Failure

5.1.2.1. Case 1: Steam Injection Operation – Chevron (Wu et al. 2008)

In this study, authors analyzed casing failure data from Chevron's Bakersfield Cymric 1Y steam injection project. Their focus was on thermally induced tensile/compressive axial stresses in restricted/fixed casing due to temperature differentials during steam injection. The major failure mode noticed was low-cycle casing fatigue. Accordingly, authors integrated strain-based methods in their casing design workflow in order to accurately analyze their case.

Typically, casing fatigue failure results from alteration of casing axial compressive and tensile stresses. In cyclic steam-injection operation, casing undergoes axial compressive stress during steam periods and axial tensile stress during soak periods. Depending on temperature changes during those two periods (steam and soak), this alteration of stresses can lead to casing fatigue and, ultimately, failure. In the study, authors noticed that although pipe body was not affected by the alteration of compressive and tensile stresses during the steam-injection operation, pipe connections were impacted. The reason for that was attributed to stress concentration effect developed around those relatively weak/sensitive locations. Authors used strain-based methods to analyze stresses developed at connections and measure their fatigue life. Based on their calculations (using Manson Eq.), authors predicted fatigue life of ~930 cycles for casing connections. Yet, their predictions contradicted the observed values from history data which were around 81-108 cycles. Authors had named few factors that may have led to an unexpected increase in total cyclic change of local strain at casing connections including: 1) casing buckling, (2) higher

connections' fatigue factor and (3) cement leak/crack due to casing-cement contact pressure from thermal radial expansion of casing during steam period.

5.1.2.2. Case 2: Multi-stage Hydraulic Fracturing Operation – Halliburton (Liu et al. 2018)

Similar to Case 1, this study was concerned with thermally induced tensile/compressive axial stresses in restricted/fixed casing due to temperature variations. The only difference was the source of temperature variation, that is, multistage hydraulic fracturing. Fatigue was the dominant failure mode in that case, hence, authors integrated strain-based methods in their typical casing design workflow.

Based on their calculations (using Manson Eq.), authors predicted fatigue life of ~9216 cycles for casing connections. Similar to case 1, their predictions contradicted the observed values from history data which were around 15 cycles. Authors had named few factors that could lead to an increase in total cyclic change of local strain at casing connections including: 1) rupture of protective film due to localized plastic strain, (2) corrosion due to acid solution (possibly saturated with H2S gas). Those effects had proven, later, to further reduce the fatigue life to 1/235 of its initial value/estimate.

5.1.2.3. Remarks on Cases 1 & 2

Although authors in both Cases 1 & 2 managed to successfully account for fatigue effects in their casing design workflow; using various casing fatigue models (e.g., Manson Universal Slope Method and Energy-Based Fatigue Model), their conclusions were utterly false and far from observed values in history data.

The major challenging aspect of the implemented physics-based models was their inability to account for various effects that were later found to have a great impact on the durability of different casing parts. Those factors are passive in nature, hence, can't be directly included in the

implemented fatigue models. Examples of those factors were 1) corrosion due to acid solution, (2) pitting effects, (3) cement leaks/cracks besides any other casing body imperfections. Using the proposed modified model, those passive effects can be integrated and accounted for without the need for explicit mathematical formulation.

5.2. Proposed Fatigue Life Estimator

To mitigate casing failure that results from fatigue, correct estimation of fatigue life of different casing parts must be made. As shown earlier, the accuracy of fatigue life estimation is contingent on the accuracy of local strain estimations. Two classes of factors impact local strains: 1) active/direct (such as temperature changes, casing material, etc.) and (2) passive/indirect factors (such as cement cracks or leaks, casing-cement contact pressures, material imperfections/wrong field handling, etc.).

Although conventional models can account for direct factors (Figure 45-*Left*), they bound to fail to account for indirect factors, leading to false conclusions on casing fatigue life and abusive consumption of casing parts beyond their capabilities.

In this paper, a data-driven alternative is proposed that takes as input all direct and indirect factors, and outputs the corresponding total local strain that reflect those effects (Figure 45-*Right*). Then, using the casing material properties, along with estimated strains as input for Manson's Eq. and estimating the fatigue life of those casing parts. Based on estimated fatigue life, the model can give recommendations on changing casing parts that are abused throughout any process (such as steam injection, or hydraulic fracturing). This would, ultimately, prevent or reduce the chances of the occurrence of casing failure.

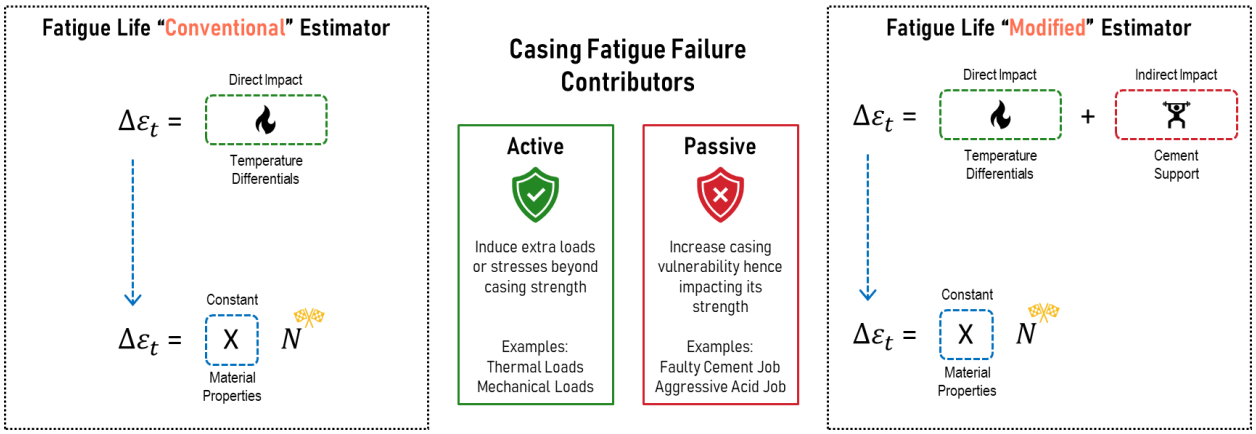


Figure 45. (Left) Schematic of the adopted framework for conventional analytical fatigue life estimators, (Right) schematic of the proposed framework for the data-driven fatigue life estimator.

5.3. Model Training

For model training, ANSYS Simulation was performed to analyze the effect of only one indirect factor, *cement job quality*, represented by cement support volume on maximum strain for a 7-inch L-80 production casing subjected to cyclic steam injection at 14000 ft. Multiple variations were considered for multiple temperature differentials and multiple volumes of cement that is supporting the casing. Figure 46 highlights the different steps involved in the stress-strain analysis using ANSYS Workbench©:

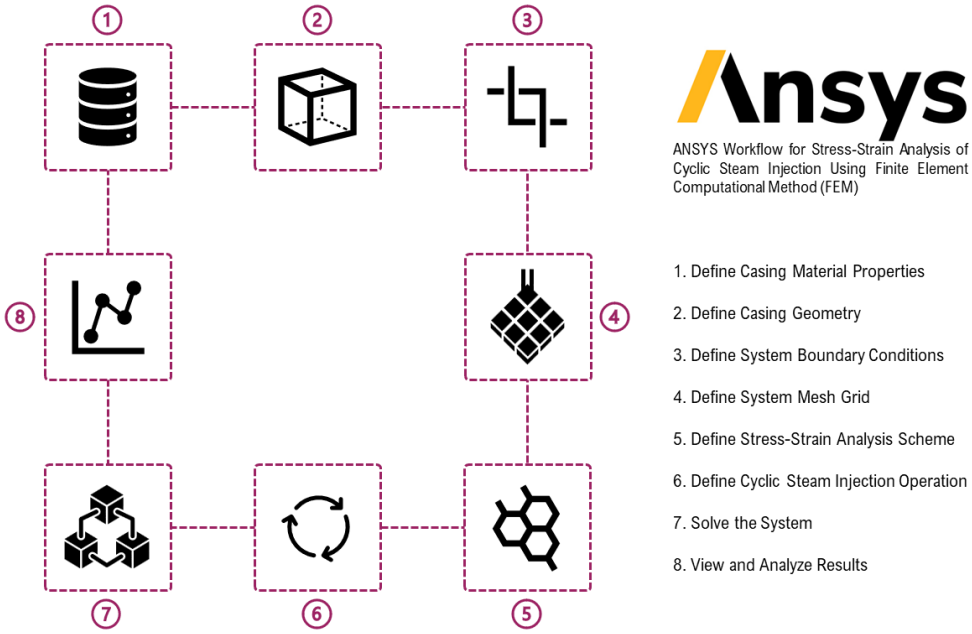


Figure 46. Adopted framework for stress-strain analysis of cyclic steam injection.

In the simulation, dry steam is injected through a 10-ft production casing with multiple temperature differentials (50-350 degrees) representing various scenarios of the cyclic operation. In addition, various supporting cement volumes were simulated foe the different scenarios for research purposes. Properties of the corresponding casing, steam, cement, and formation are summarized in Tables 17-19 below.

Table 17 Input Data for L-80 Production Casing.

Casing Properties		Value
Size – OD (*inch*)	:	7
Nominal Weight (*lb./ft*)	:	23
Burst (*psi*)	:	6,340

Casing Properties		Value
Collapse (*psi*)	:	3,830
Young's Modulus (*psi*)	:	30 x 10^6
Yield Strength (*psi*)	:	85,000
Tensile Strength (*psi*)	:	95,000
Length (*ft*)	:	10
Density (lb/ft^3)	:	473.28
Poisson Ratio	:	0.3
Thermal Conductivity (*W/mC*)	:	28
Thermal Expansion (C^{-1})	:	1.09^{-05}
Specific Heat (*J/Kg*)	:	453

Table 18 Input Data for Unconsolidated Sandstone Formation.

Formation Properties		Value
Young's Modulus (*MPa*)	:	1.2 x 10^6
Poisson Ratio	:	0.27
Density ($lb./ft^3$)	:	473.28
Thermal Conductivity (*W/mC*)	:	28
Thermal Expansion (C^{-1})	:	1.09^{-05}
Specific Heat Capacity ($1/^0F$)	:	453

Table 19 Input Data for Injected Steam during Cyclic Operation.

Steam Properties		Value
Density of Steam (Kg/m^3)	:	36.5107
Specific Entropy of Steam (*KJ/Kg K*)	:	5.81334
Specific Heat of Steam (*kJ/Kg*)	:	5.1192
Specific Enthalpy of Steam (*kJ/Kg*)	:	2771.89

The geometric design of the system of interest is depicted in Figure 47-a below. The 10-ft of 7" L-80 casing is modelled, followed by an 8-1/2" x 10-3/4" cement-filled annulus, and the formation on the outside boundary is approximately 30-in in diameter. As for system meshing, multiple factors were considered for improved mesh efficiency including mesh defeaturing, trap curvature, and grab proximity. Figure 47-b highlights the results of finning mesh sizing with element edge length of 5-in for casing, 1-in for the cement, and 2-in for the target formation.

Regarding applied loads and boundary conditions, the inner casing surface is subjected to a single load approximating the internal pressure caused by the 1000 psi steam injection (as shown in Figure 47-c). The applied internal pressure is constant along the 10-ft casing.

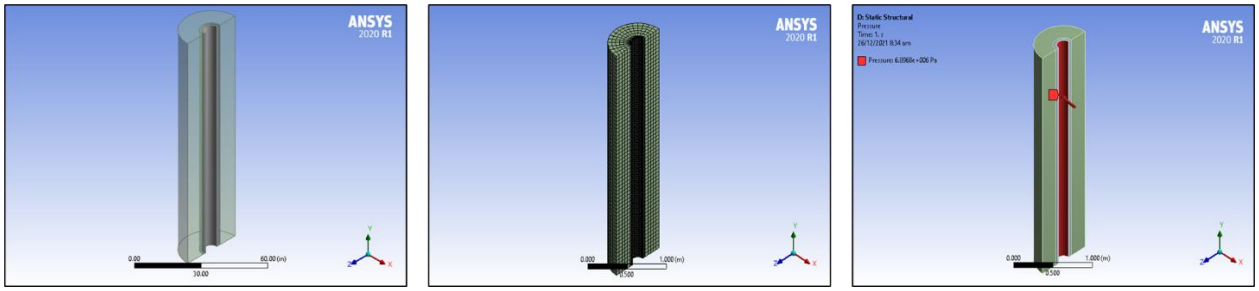


Figure 47. (a) Results of the geometric design of the system of interest, (b) results of the fine-meshing step of the system of interest, and (c) visual representation of the applied boundary conditions applied to the system of interest.

Table 20 and Fig. 48 represent the numerical and graphical representation of simulation results for one of the 20 different scenarios tested during the training phase, in particular for 100-degree temperature differential for multiple cement support volumes. While Figure 49 highlights results of all different scenarios tested.

Table 20 Max Strain Values at a 100-Degree Temperature Differential for Different Cement Support Volumes.

Cement Support (%)	Cement Thickness (in)	Total Strain (in/in)
100	1.125	0.0016955
80	0.900	0.0017978
60	0.675	0.0019452
40	0.450	0.0019689
20	0.225	0.0019882
0	0.000	0.0020564

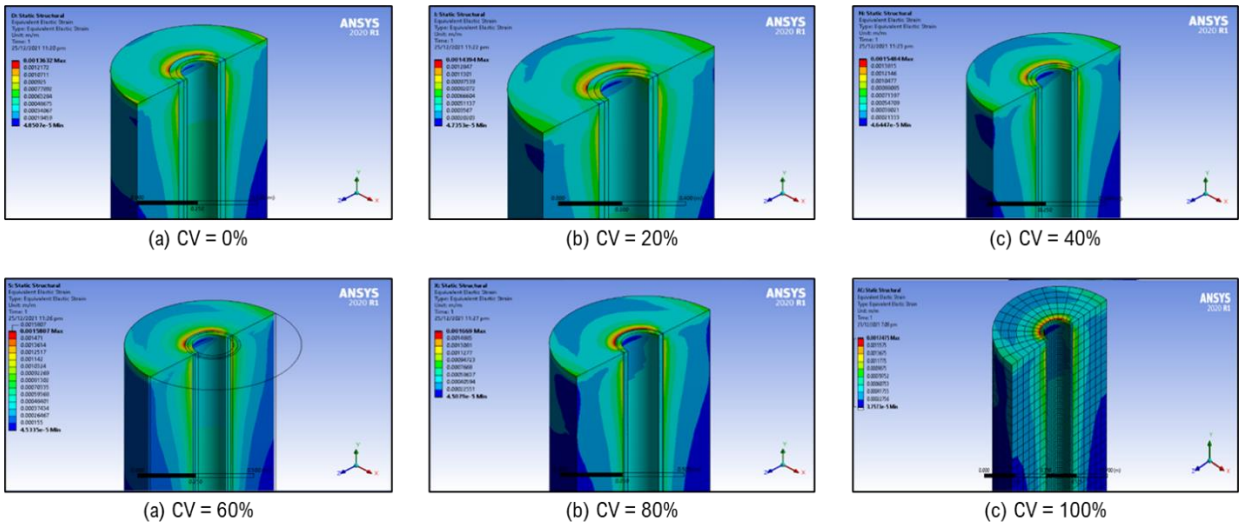


Figure 48. Results of ANSYS FEA stress-strain simulation for a wide range of cement support volumes for a temperature difference of 100-degrees.

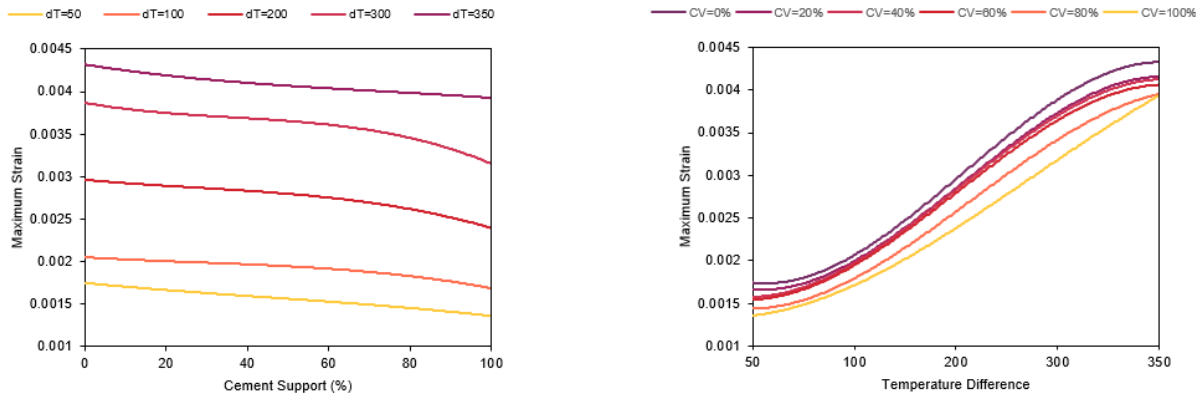


Figure 49. (Left) Maximum strain plot with temperature difference variation for multiple cement support volumes. (Right) Maximum strain plot with cement support volume variation for multiple temperature differences.

Results of simulation runs along with casing material properties were then coupled with analytical fatigue life estimators (e.g., Mansion's Eq.) to estimate the corresponding fatigue cycle life, as shown in Figure 50.

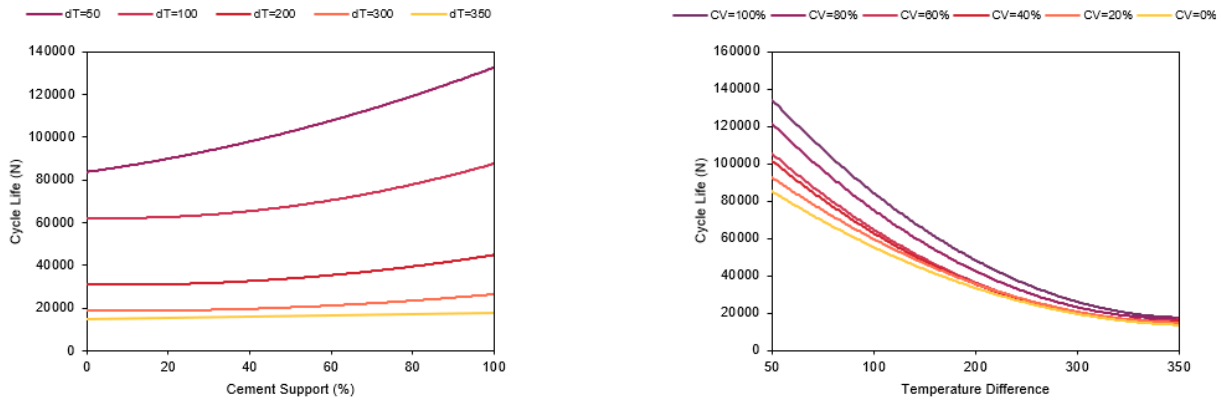


Figure 50. (Left) Fatigue cycle life plot with temperature difference variation for multiple cement support volumes. (Right) Fatigue cycle life plot with cement support volume variation for multiple temperature differences.

Results of conventional analytical models were then compared with those estimated from the proposed modified model as shown in Figure 51.

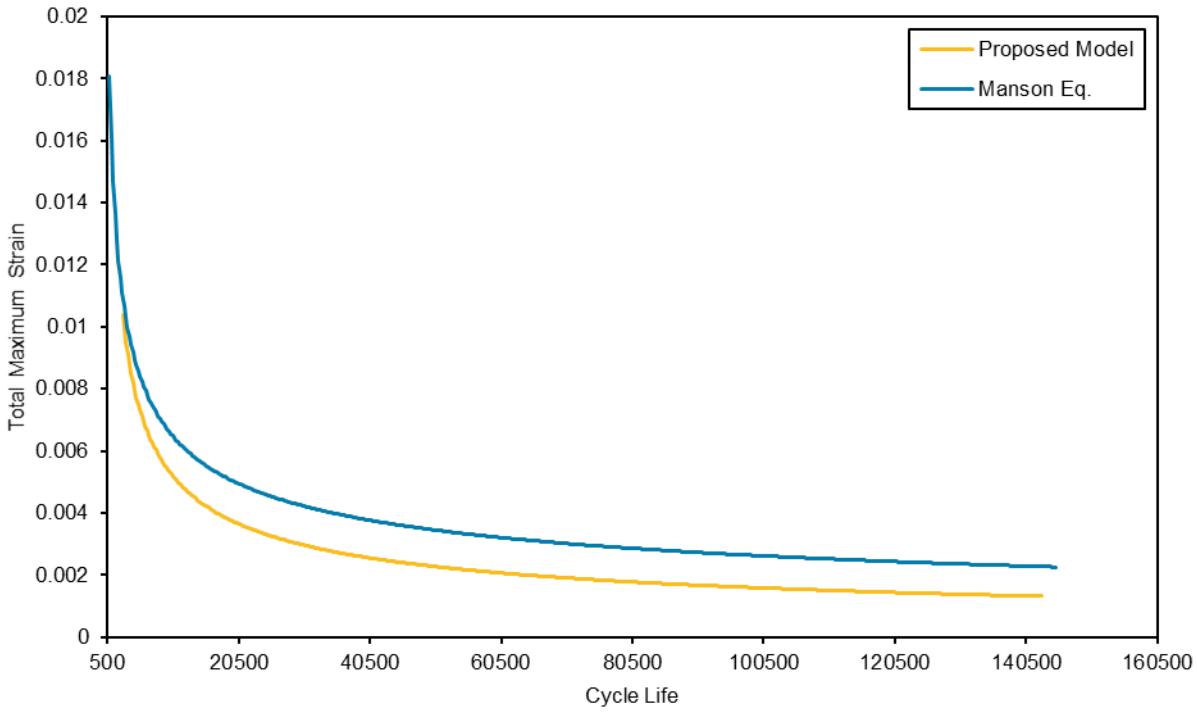


Figure 51. A plot of fatigue cycle life for various total local strain values using conventional analytical models (blue) as compared to the proposed modified data-driven model.

As shown above, there is a huge difference between results of the proposed modified data-driven model and typical analytical models. It's clear that the proposed model is more sensitive to changes in total local strain as compared to the conventional counterparts. For the proposed model, a total local strain half of that calculated using conventional models will be enough to keep the casing integrity for a certain amount of well overall lifetime. This is due to the consideration of other sources of vulnerability which in this case is the lack of supporting cement support

Using the proposed model would help drilling engineer better plan and design their casing and maintain their integrity without abusing the material or risking their failure.

CHAPTER VI
DISCUSSION

Reader Guide: In this chapter, I briefly discuss the results obtained from all the analyses conducted in the preceding chapters; highlighting key findings and major takeaways. The discussion starts with a layout of the major contributions achieved through the work presented in this dissertation (Section 6.1). Then, attention is geared towards evaluating the proposed solution(s) through the implementation of SWOT analysis (Section 6.2). Afterwards, I provide a set of recommendations for the reader in order to efficiently utilize and benefit from the proposed workflow (Section 6.3). Finally, I end the discussion with addressing some of the aspects in the study that could be improved upon in the future for more refined/enhanced results (Section 6.4).

6.1 Original Contribution

Contributions previously introduced in literature in the area of casing failure were devoted to investigating and explaining the causing effects of casing failure in numerous historical cases. To date, the majority of those contributions were based on physics-based approaches (analytical, experimental or numerical). Although, those contributions had provided a valuable insight into identification of some of the potential causing effects of casing failure, they failed to provide enough information on how to mitigate, or avoid, the occurrence of casing failure in the future, hence, did not attain wide-scale execution. This left drillers and drilling engineers with very little information on mitigating the risks that the casing is exposed to during drilling, completions, and production operations.

The main drivers of the major limitation previous contributions had concerning their inability to provide a viable measure to mitigate casing failure were: (1) the inability to “quantify” the impact that various potential risk factors have on the occurrence of the casing failure and (2) the inability

to tackle the impact of potential risk factors throughout the well length, in addition, how variation in a particular risk factor might affect the probability of casing failure accordingly.

In this study, I managed to tackle the aforementioned challenging aspects through developing data-driven workflows based on risk-factor, as well as survival analysis statistical techniques. In addition, I managed to develop a data-driven based "casing failure mitigation" tool that is capable of proactively handling the causing effects of casing failure through the implementation of a two-step "prediction-correction" procedure based on the integration of predictive analytics, in addition, semi-quantitative risk assessment techniques.

Application of the developed casing failure mitigation tool will give the allowance for drillers and drilling engineers to constantly and automatically adjust their initial design specifications in the direction of reducing the overall risk of casing failure and, ultimately, avoiding its occurrence in the future.

6.2 SWOT Analysis

SWOT analysis is a technique used to identify strengths, weaknesses, opportunities, and threats – conventionally – related to project planning and management. In this section, we use SWOT analysis for evaluating the solution(s) proposed in our study for mitigating casing failure.

6.2.1 Strengths

- Early study applying integrated machine learning to inform drilling and drilling engineers on how to proactively correct rather than reactively a set of features in drilling, completions, and production.
- Some results of this study's work (Noshi et al. 2018a, Noshi et al. 2018b; Noshi et al. 2019) was used in three studies predicting casing damage (Song and Zhou 2019; Tang 2019; Tan et

al. 2020).

- A generic tool that identifies the parameters that have the greatest impact on casing damage. For example, applying the tool to wells in the upper Granite Wash formation, the optimal dog leg severity was identified for the minimum risk of failure.
- Quantification of the impact each of the pre-defined risk factors has on the overall probability of casing failure.
- Tracking the variation in the impact that is caused by variation within potential risk factors throughout the length of the well.
- Closed-loop adjustment by iteratively modifying the key input parameters to prevent failure based on five statistical approaches developed in this study. The adopted "prediction-correction" procedure will inform drillers how to react to mitigate casing failure and well loss.
- Adaptive data-driven workflow that could be applied to much larger data sets with diverse set of features that might not, yet, be supported by physical evidence.

6.2.2 Weaknesses

- The provided data set of Granite Wash formation is relatively small with limited set of features which might lead to inconclusive results, in terms of, identifying all possible causes of casing failure.
- The limitation of the provided Granite Wash formation data set, in terms of, the data size imposes a high risk of overfitting during the training/testing phase of the implemented classification models.
- Regarding the construction of risk matrices, the thresholds selected for assigning risk values to one of the three pre-determined risk severity classes were set manually based on the author's knowledge of history failure cases. This might induce some errors for more complex cases

with much more diverse set of features.

- Implementation of non-parametric survival estimation techniques, e.g., Kaplan Meier estimator, might be infeasible when integrating a much broader set of features; either categorical with multiple levels/categories or continuous features.

6.2.3 Opportunities

- With the constant development of machine learning algorithms, there is an allowance for enhancing the tool prediction capability through the integration of more robust and efficient classification models.
- Model validation on a much broader set of casing failure cases from outside the Anadarko basin for model generalization.
- With the acquisition of real-time data, the proposed tool could be further modified to handle design specifications in real-time.

6.2.4 Threats

- Limitation in casing failure data as companies refrain from publishing in order to protect their reputation might hinder the progression of data-driven solutions.
- The tool shouldn't be used as a standalone model but should become an addition to proper casing handling, completion, and production practices.

6.3 Practical Recommendations

- When handling large data sets that have: (a) a much broader set of features, (b) categorical variables with multiple levels/categories or (c) continuous features, it is recommended that the user implement semi-parametric or parametric survival curve estimators instead of non-parametric estimators that were used for this study.

- When handling data sets that have features with less variation, frequency analysis might be troublesome and there might be a risk of class imbalance. This, in turn, might affect the statistical analysis. it is recommended, then, that the user pay attention to the weight of the different subgroups/clusters within each feature to guarantee the balance of the analysis.
- When constructing risk assessment matrices, it is recommended for the user to implement quantitative techniques rather than semi-quantitative techniques (adopted in the study). Unlike semi-quantitative techniques, fully quantitative techniques have the ability to assign continuous risk values to any variation within any feature/risk factor.
- When accounting for missing values in the provided data set using imputation techniques, it is important that the user select an optimal threshold so that they don't lose more wells (in case of high cutoff value) or have a great deal of missing inputs (in case of low cutoff value).
- When optimizing supervised learning algorithms, it is important that the user check the model balance and overall performance. Choosing an optimal model based solely on having high prediction accuracy might be misleading in some cases, as the model might be suffering from overfitting. This could ultimately lead to misclassification of future cases and affect the performance of the tool.

6.4 Future Work

- Enhancing the tool prediction specificity to account for the mode of casing failure (e.g., collapse, burst, tensile, etc.).
- Enhancing the tool adjustment, or follow-up, actions upon prediction of casing failure. For instance, it could provide the engineer with the specific cement composition, optimum number of acid gallon for stimulation, type of acid to use, etc.
- Integration of more geological and metallurgical features, in addition to other features in

drilling, completion and production.

- Integration of data sets from different fields with different conditions.

CHAPTER VII
CONCLUSIONS

Reader Guide: In this chapter, I highlight the work done throughout this dissertation along with the major takeaways concluded from the conducted analyses. I start with highlighting the contribution of this dissertation in terms of tackling of the problem of casing failure (Section 7.1). Afterwards, I present the outlines of the proposed solution, that is, an "automated casing failure mitigation" tool (Section 7.2). Following that, I share the key results obtained for each of the four major statistical techniques used as a basis for achieving the study's objectives including: risk analysis (Section 7.3), survival analysis (Section 7.4), predictive analytics (Section 7.5) and, finally, risk assessment (Section 7.6). Based on the analyses applied in this study, I present some practical insights that can benefit from the work conducted (Section 7.7).

7.1 Study Contribution

This study represents an on-going effort to minimize the likelihood of casing failure and attempts to find a possible solution. Contributions previously introduced in literature in the area of casing failure were devoted to investigating and explaining the causing effects of casing failure in numerous historical cases. To date, the majority of those contributions were based on physics-based approaches (analytical, experimental or numerical). Although, those contributions had provided a valuable insight into identification of some of the potential causing effects of casing failure, they failed to provide enough information on how to mitigate, or avoid, the occurrence of casing failure in the future, hence, did not attain wide-scale execution. This left drillers and drilling engineers with little information on mitigating the risks that the casing is exposed to during drilling, completions, and production operations.

The main drivers of the major limitation previous contributions had concerning their inability to provide a viable measure to mitigate casing failure were: (1) the inability to "quantify" the impact

that various potential risk factors have on the occurrence of the casing failure and (2) the inability to tackle the impact of potential risk factors throughout the well length, in addition, how variation in a particular risk factor might affect the probability of casing failure accordingly.

In this study, I managed to tackle the aforementioned challenging aspects through developing data-driven workflows based on risk-factor, as well as survival analysis statistical techniques. In addition, I managed to develop a data-driven based "casing failure mitigation" tool that is capable of proactively handling the causing effects of casing failure through the implementation of a two-step "prediction-correction" procedure based on the integration of predictive analytics, in addition, semi-quantitative risk assessment techniques.

Application of the developed casing failure mitigation tool will give the allowance for drillers and drilling engineers to constantly and automatically adjust their initial design specifications in the direction of reducing the overall risk of casing failure and, ultimately, avoiding its occurrence in the future. Additionally, this study not only presents a methodology aiming to form the foundation for a new standard for casing risk assessment in the Anadarko Basin, but it can be applied to any geological area with different scenarios and can be developed into a more generalized tool in the future.

7.2 Casing Failure Mitigation Automation

I propose a tool for automated casing failure mitigation (Sections 4.8 and 4.9). This tool is based on a two-step "prediction-correction" procedure; where I add "feedback" to the initial predictions provided by the conventional ML algorithms that would serve as a guide that could help drilling engineers adjust their design and ultimately mitigate casing failure. This is accomplished through the integration of semi-quantitative risk assessment techniques probability-impact risk assessment matrices (PI-RAMs).

First, design specifications initially suggested by a drilling engineer for a particular well are fed to the "casing failure mitigation" tool. Next, design specifications are evaluated in terms of the risk imposed on casing failure (Sections 4.8 and 4.9). In case of high risk, design specifications are automatically adjusted so that imposed risk is reduced and ultimately casing failure is avoided or mitigated. The tool is based on four major concepts, those are, risk analysis, survival analysis, supervised machine learning (ML) algorithms and semi-quantitative risk assessment.

Supervised ML algorithms (Chapter 4) along with semi-quantitative risk assessment techniques (Chapter 4) are used as the cornerstone for the two-step "prediction-correction" procedure; where ML algorithms provide initial prediction of the probability of casing failure occurrence based on the input design specifications. Following that, the semi-quantitative risk assessment techniques (e.g., PI-RAMs) are used to automatically adjust the design specifications according to their impact type and the overall case evaluation until risk is reduced below pre-defined threshold.

Regarding the risk analysis (Chapter 2) and the survival analysis (Chapter 3), they constitute an integral part of the tool as they provide the necessary information to be later used as a basis for the construction of ML predictive models, as well as risk assessment matrices (PI-RAMs). That information include: (1) identification of potential risk factors that are strongly associated with casing failure, (2) identification of the subgroups within each risk factor that impose the highest impact, (3) evaluation of the type and magnitude of the impact for each risk factor, (4) determination of the likelihood of the occurrence of the different scenarios within each risk factor.

7.3 Risk Analysis techniques

First, I focused on testing the association between the different features, or exposures, and the occurrence of casing failure. To that end, I implemented a well-established family of techniques, known as risk analysis techniques (Section 2.1). Based on the inferential direction of the study and

its computational superiority, “case-control” study design was adopted, where cases (failed wells) were, basically, compared with controls (non-failed wells) based on the different exposures they had experienced (Sections 2.2 through 2.4).

For a more balanced and reliable analysis, a “matched” analysis approach was followed, where each failed well (case) was matched with three other non-failed wells (controls) based on a set of common features including location, formation, etc. (Section 2.5).

As for association measurements, different techniques were briefly introduced from literature, such as, relative risk (RR), risk difference (RD), relative risk difference (RRD) and odds ratio (OR). However, it was the odds ratio (OR) technique that was selected for evaluating and interpreting the results of the analysis (Section 2.6).

Based on the initial results of the analysis, only thirteen features – out of the initial twenty-six features – were considered as potential risk factors of casing failure. The risk factors were related to casing design, fracture design and drilling conditions (Section 2.7).

Owing to the implementation of odds ratio (OR), we had also managed to define the type of impact each potential risk had on the occurrence of casing failure; either positive, negative, or slight-to-no impact. Results showed that, for instance, “frac season” and “maximum inclination” had a negative impact on casing failure, while “cementing” and “casing thickness” had a positive impact. And “measured depth” and “dogleg severity measured depth” had nearly no impact on the occurrence of casing failure (Section 2.7).

In addition to evaluating risk type, using odds ratio (OR), we managed to measure the magnitude of the impact each potential risk factor had on the overall probability of casing failure. For instance, fracturing during spring turned out to increase the risk of casing failure by over 200%, an increase of one unit in the lateral section shrinkage increases the risk by 15%. On the other hand, cementing

reduces the risk of casing failure by nearly 54%, while having a casing thickness greater than 0.65 in (P110) tends to reduce the risk of casing failure by nearly 90%. Those results were beneficial for generating probability-impact risk assessment matrix (PI-RAM) that was addressed later in the study (Section 2.7).

Due to the dependency between different exposures, or features, and their rule in defining the association/impact different exposures have on the overall outcome, several models were run and tested against various combinations of input features using subset selection techniques (Section 2.8). Due to computational reasons, as well as, statistical reasons, stepwise selection techniques were used for generating the models. Results obtained from the different models were compared, and the best/optimal model was selected based on pre-determined criteria including AIC, BIC, Cp, and R^2 (Sections 2.9 and 2.10).

7.4 Survival Analysis

Based on the results obtained from the risk analysis, the focus then shifted to another concern, that was, defining the depths that are venerable to casing failure. The motivation behind that was to give engineers a guidance on where they should expect to have failure so they can pay attention during drilling/production operations. To achieve that goal, we have extended our investigation to another classic statistical technique; known as survival analysis (Section 3.1).

Motivated by the research interests, little modification was made to the analysis scale (conventionally time scale) so it corresponded to reached measured depth (Sections 3.1 and 3.2). Two broad families of survival analysis techniques were discussed (Section 3.3): (1) non-parametric estimators, e.g., Kaplan Meier estimator (Section 3.3.1), and (2) semi-parametric estimators, e.g., Cox model (Section 3.3.2).

Non-parametric techniques, due to simplicity (since there is no need for formal equations/models) and their graphical representations, they helped trace the changes in the impact of subgroups/levels within each potential risk factor (Sections 3.3.1.1 and 3.3.1.2). That ultimately served our goal. However, with increased number of features/risk factors, as well as subgroups/levels within risk factors, some complexities were introduced in the analysis. That's when semi-parametric came into play. Semi-parametric estimators had the ability to incorporate all risk factors in one model, unlike the non-parametric estimators (Sections 3.3.2.1 and 3.3.2.2).

Then we moved one step further in the analysis, as we evaluated the impact each subgroup/level within each risk factor had throughout the entire well. This added another layer of understanding that was not provided by risk analysis. That proved to be of great value, especially, when constructing the PI-RAM.

Results of the analysis showed that the following conditions (subgroups/levels) are the most risky compared to their companions within their corresponding risk factors: (1) drilling during winter or spring seasons, (2) implementation of acid treatment, (3) having no cement support, (4) witnessing dogleg bending stress ≥ 95,600 lb. ft, (5) bottomhole temperature ≥ 166 ºF, (6) lateral section shrinkage ≥ 10 in/100 ft (7) max inclination ≤ 95.5º, (8) dogleg severity ≥ 15º/100 ft, (9) frequency of severe dogleg ≥ 13 (Section 3.3.2.3).

7.5 Predictive Analytics

Afterwards, the attention was geared towards a different – yet equally important – issue, that is, prediction of the probability of casing failure using potential risk factors defined from risk analysis (Section 4.1). Machine learning techniques, particularly, supervised learning algorithms were used to fulfill that task (Section 4.2). Due to the type of the outcome, only the learning algorithms used for classification purposes were discussed, such as, logistic regression (Section 4.2.1), basic

decision trees (Section 4.2.2), random forests (Section 4.2.3), support vector classifier (Section 4.2.4), support vector machines (Section 4.2.5) and artificial neural networks (Section 4.2.6).

The different models were trained and tested on the data set, and results were compared based on two major bases: (1) prediction accuracy and (2) model balance and overall performance. For the prediction accuracy, the "overall classification rate" criterion was used. Regarding the model performance and balance, both "false positive rate" and "false negative rate" were implemented (Section 4.3).

Based on the aforementioned criteria, artificial neural network (ANN) showed a superiority compared to other models, in terms of, prediction accuracy, overall performance and model balance. Support vector machines (SVMs) came second, followed by logistic regression (Section 4.3).

Although supervised learning algorithms showed huge potential for predicting the outcome of interest, that is, casing failure, yet they gave no allowance for engineer inputs/interventions. It is for that reason that we sought risk assessment techniques (Sections 4.4 through 4.6).

7.6 Risk Assessment

Risk assessment techniques had proven its efficiency in many applications in oil and gas industry (Section 4.7). In our study, we focused on one type of risk assessment techniques that is semi-quantitative techniques. The implemented technique was PI-RAM. Results obtained from risk analysis, survival analysis and frequency analysis were used for the construction of those matrices.

Using the PI-RAMs helped have a better intuition of not only the overall risk, but also the contribution of each risk factor (Section 4.9). This, in turn, could help experienced engineers have a clear view of changes to risk probabilities with changing conditions/designs.

Findings of the implemented workflow are in a close match with proven theories. This outcome has two important takeaways: (1) it gives validation to the proposed workflow, so, we can trust its performance and its judgement when handling much more complex situations with a much bigger data set and much more features that might, yet, not be supported by physical proofs, (2) even though the proposed data-driven workflow led to the same results achieved by physics-based approached, yet, there still a key distinction between both that we did not have to get through all the physics behind that phenomena to reach to the final conclusions.

7.7 Practical Conclusions

The analyses applied in this study revealed the following practical insights:

1. Based on the results of the risk analysis, only thirteen features – out of the initial twenty-six features – were considered as potential risk factors of casing failure. The risk factors were generally related to casing design (e.g., casing grade, casing surface area, etc.), fracture design (e.g., fracture length, fracture stages, etc.), well conditions (e.g., lateral shrinkage, bottomhole temperature, etc.) and drilling conditions (e.g., dogleg severity, maximum inclination, etc.).

2. Based on odds ratio (OR) analysis, fracturing during spring turned out to increase the risk of casing failure by over 200%, an increase of one unit in the lateral section shrinkage increases the risk by 15% and an increase of one degree of maximum inclination increases the risk by 30%. On the other hand, cementing reduces the risk of casing failure by nearly 54%, while having a casing thickness greater than 0.65 in (P110) tends to reduce the risk of casing failure by nearly 90%.

3. Results of the survival analysis showed that the following conditions (subgroups/levels) are

the riskiest compared to their companions within their corresponding risk factors:

a. Regarding “fracturing season”, it’s been noticed that for depths less than 8,500 ft, no significant difference in risk impact between different seasons. Yet, fracturing during “winter” or “spring” still puts the well at a higher risk of failure than the other two seasons: with an increase of risk by 20% and 7%, respectively, for depths ranging from 8,500 – 10,000 ft, while 30% and 20%, respectively, for depths greater than 10,000 ft.
b. Regarding “acid treatment”, it has been noticed that: implementation of “acid” treatment at depths less than 8,500 ft imposes no increased risk on casing failure as opposed to not having acid treatment. However, for depths ranging from 8,500 to 10,000 ft, as well as depths greater than 10,000 ft, the risk imposed by having acid treatment becomes much higher (7.5% and 12% increase, respectively) than that imposed by not having acid treatment. In addition, the risk constantly increases with increasing reached depth.
c. Regarding “cementing”, it has been noticed that: having no “cement” support at depths less than 8,500 ft imposes no higher risk on casing failure compared to having a cement support. Yet, for depths ranging from 8,500 to 10,000 ft, as well as depths greater than 10,000 ft, having a cement support starts to play an integral rule in reducing the risk of casing failure; reaching nearly 30% less than the risk imposed by not having cement support for depths greater than 10,000 ft.
d. Regarding “dogleg bending stress”, it has been noticed that at depths ranging from 8,500 to 10,000 ft, as well as, depths < 10,000 ft, relatively high DL bending stress (> 95.8k lb./ft^2) are needed to impose high risk on casing failure (~ 7.5% increase). Conversely, at depths > 10,000 ft, relatively lower DL bending stress (less than

95.8k lb./ft^2) could be sufficient to impose a higher risk of casing failure (~ 12-13% increase).

e. Regarding "bottomhole temperature", it has been noticed that for depths < 8,500 ft, relatively higher BHT (> 166 ^oF) is needed in order to have a mere increase in risk of casing failure (~ 3.2% increase). In contrast, for depths ranging from 8,500 to 10,000 ft, as well as, those > 10,000 ft, relatively less BHT (< 166 ^oF) could be enough to impose a significant increase in risk of casing failure (~ 5.1-12.8% increase).

f. Regarding "lateral shrinkage", it has been noticed that variation in the magnitude of "lateral shrinkage" tends to have no noticeable impact on the risk of casing failure for depths < 8,500 ft. Yet, for depths ranging from 8,500 to 10,000 ft, as well as, those > 10,000 ft, slight changes in lateral section (less than 10 ft/100ft) could play an integral rule in increasing the risk of casing failure (~ 26% increase).

g. Regarding "maximum inclination", as well as "DL severity", it has been noticed that – similar to "bending stress" – "maximum inclination" and "DLS" possesses an inverse proportionality with respect to the measured depth. So, for depths < 8,500 ft and those ranging from 8,500 – 10,000, a relatively high value of max. inclination, as well as DLS is needed in order to have a noticeable impact on increasing risk. Yet, for depths > 10,000 ft, relatively lower values would be sufficient to nearly increase the risk by 9-12%.

h. Regarding "frequency of severe dogleg", it has been noticed that it follows a direct proportionality with respect to the measured depth. So, for depths ranging from 8,500 – 10,000, a relatively lower value of DLS frequency (< 13) could be indicative of increased risk (~ 4.6% increase). And, for depths > 10,000 ft,

relatively higher values (>13) would be a strong indication of increased risk (~ 16.3%).

i. Finally, for" fracture temperature", it has been noticed that it follows direct proportionality with respect to the measured depth, for the same reasons as BHT.

4. Comparison of different classification models showed the superiority of artificial neural network (ANN) compared to other models, in terms of, prediction accuracy (67% E), overall performance and model balance (33.33% FP, 33.33% FN). Support vector machines (SVMs) came second (67% E, 50% FP, 33.33% FN), followed by logistic regression (61% E, 33.33% FP, 50% FN).

5. The developed automated casing failure mitigation tool could be used by drillers and drilling managers to automatically and constantly check and correct their design specification in order to mitigate, or avoid, potential casing failure.

6. The developed data-driven workflows (based on risk analysis) could be used to explore new causing effects in far more complicated cases where there is not yet a physical proof to draw a solid conclusion.

7. The developed data-driven workflows (based on survival analysis) could be used to explore the significance the variation in risk factors has on changing the overall impact of casing failure, in addition, expose the depths that are most vulnerable to casing failure based on the given downhole conditions.

REFERENCES

Abdideh, M. and Khah, S. H. 2018. Analytical and Numerical Study of Casing Collapse in Iranian Oil Field. *Geotech Geol. Eng.* **36** (3): 1723–1734. https://doi.org/10.1007/s10706-017-0428-0.

Akaike, H. 1973. Information Theory and An Extension of the Maximum Likelihood Principle. *Proc*., 2nd International Symposium on Information Theory, Tsahkadsor, Armenia, 2–8 September.

Akande, K. O., Taoreed O. O., and Sunday O. 2015. Investigating the Effect of Correlation-Based Feature Selection on the Performance of Support Vector Machines in Reservoir Characterization. *Journal of Natural Gas Science and Engineering* **22**: 515–522. https://doi.org/10.1016/j.jngse.2015.01.007.

Akpan, H. O. and Kwelle, S. O. 2005. Efficient Computational Method for Casing String Design. Presented at the Nigeria Annual International Conference and Exhibition, Abuja, Nigeria, 1–3 August. SPE-98790-MS. https://doi:10.2118/98790-MS.

Alkendi, M. Y. M. Sa. 2006. ADNOC Environmental Impact Severity Matrix, an Innovative Impact Rating Matrix. SPE International Health, Safety & Environment Conference, Abu Dhabi, UAE, 2–4 April. SPE-98852-MS. https://doi.org/10.2118/98852-MS.

Andersson, S. 2008. Wear simulation with a focus on mild wear in rolling and sliding contacts. Presented at the International Symposium on Friction, Wear and Wear Protection, Aachen, Germany, 9–11 April. https://doi:10.1002/9783527628513.ch1.

Anifowose, F., Labadin, J., and Abdulraheem, A. 2013. A Least-Square-Driven Functional Networks Type-2 Fuzzy Logic Hybrid Model for Efficient Petroleum Reservoir Properties

Prediction. *Neural Comput. & Applic.* **23**: 179–190. https://doi.org/10.1007/s00521-012-1298-2.

Arpit, D., Stanisław J., and Nicolas B. et al. 2017. A closer look at memorization in deep networks. *Proc.* 34th International Conference on Machine Learning, Sydney, Australia. (in press; posted 1 July 2009). arXiv:1706.05394.

Bai, Y. and Bai, Q. 2018. *Subsea Engineering Handbook*. Gulf Professional Publishing.

Barton, N. 2003. *Erosion in Elbows in Hydrocarbon Production Systems: Review Document*. UK: Sudbury press. http://hdl.handle.net/10068/635680.

Bastola, A., Wang, J., Mirzaee-Sisan, A., et al. 2014. Predicting hydrostatic collapse of pipes using finite element analysis. Presented at the 33rd International Conference on Offshore and Arctic Engineering, San Francisco, California, USA, June 8–13. OMAE2014-23690. https://doi.org/10.1115/OMAE2014-23690.

Best, B. 1986. Casing Wear Caused by Tool Joint Hard Facing. *SPE Drilling Engineering*, **1** (1): 62–70. SPE-11992-PA. https://doi.org/10.2118/11992-PA.

Borgan, Ø., Breslow, N., Chatterjee, N., et al. ed. 2018. *Handbook of Statistical Methods for Case-Control Studies*. CRC Press, 536 pages. ISBN: 9781498768580. Biometrical Journal **61** (6). https://doi.org/10.1002/bimj.201900212.

Bowers, C. N. 1955. Design of Casing Strings. Presented at the Fall Meeting of the Petroleum Branch of AIME, New Orleans, Louisiana, 2–5 October. SPE-514-G. https://doi.org/10.2118/514-G.

Cao, Z., Wang, Y., and Li, D. 2016. Bayesian Perspective on Geotechnical Variability and Site Characterization. *Engineering Geology* **203**: 117–125. https://doi.org/10.1016/j.enggeo.2015.08.017.

Cauquil, E. C. 2009. Risk Matrix for Non-Recurrent Geological Processes: Application to the Gas Hydrate Hazard. Offshore Technology Conference, Houston, Texas, 4–7 May. OTC-20014-MS. https://doi.org/10.4043/20014-MS.

Chen, Z., Zhu, W. and Di, Q. 2018, *Elasticity Solution for The Casing Under Linear Crustal Stress*. *Engineering Failure Analysis*, **84**: 185–195. https://doi.org/10.1016/j.engfailanal.2017.11.007.

Chipperfield, S. 2007. Shear Dilation Diagnostics: A New Approach for Evaluating Tight Gas Stimulation Treatments. SPE Hydraulic Fracturing Technology Conference, College Station, Texas, USA, 29-31 January. SPE-106289-MS. https://doi.org/10.2118/106289-MS.

Cirimello, P., Otegui, J., and Carfi, G. et al. 2017. Failure and Integrity Analysis of Casings Used for Oil Well Drilling. *Engineering Failure Analysis* **75**: 1–14. https://doi.org/10.1016/j.engfailanal.2016.11.008.

Cirimello, P., Otegui, J., and Sanchez, M. et al. 2018. Oil Well Drill Bit Failure During Pull Out: Redesign to Reduce Its Consequences. *Engineering Failure Analysis* **83**: 75–87. https://doi.org/10.1016/j.engfailanal.2017.09.020.

Cortes, C., Vapnik, V. 1995. Support-Vector Networks. *Mach Learn* **20** (3): 273–297. https://doi.org/10.1007/BF00994018.

Cox, D. R. 1972. Regression Models and Life-Tables. *Journal of the Royal Statistical Society: Series B (Methodological)* **34** (2): 187–202. https://doi.org/10.1111/j.2517-6161.1972.tb00899.x.

Cramer, J.S. 2020. The Origins of Logistic Regression. *Tinbergen Institute* **119**: 167–178, http://dx.doi.org/10.2139/ssrn.360300.

Dai, W., Noel, B., Alvord, C., et al. 2018. A Practical Approach to Casing Wear Prediction, Modeling and Mitigation on Challenging ERD Wells. SPE Annual Technical Conference and Exhibition, Dallas, Texas, USA, 24-26 September. SPE-191495-MS. https://doi.org/10.2118/191495-MS.

Dall'Acqua, D., Chartier, M., Nowinka, J. et al. 2018. Unified Basis for Thermal Casing/ Connection System Design. SPE Thermal Well Integrity and Design Symposium, Banff, Alberta, Canada, 27-29 November. SPE-193363-MS. https://doi.org/10.2118/193363-MS.

Davies, R., Almond, S., Ward, R., et al. 2014. Oil and Gas Wells and Their Integrity: Implications for Shale and Unconventional Resource Exploitation. *Marine and Petroleum Geology*, **56**: 239–254. https://doi.org/10.1016/j.marpetgeo.2014.03.001.

Dusseault, M.B., Maury, V., Sanfilippo, F. et al. 2004. Drilling around salt: risks, stresses, and uncertainties. Presented at the 6th North America Rock Mechanics Symposium (NARMS), Houston, Texas, 5–9 June. ARMA-04-647.

Edaigbini P.I. and Maikobi A.A. 2015. Casing Design for High Pressure/High Temperature Wells. *Innovative Systems Design and Engineering* **6**(3): 9–25.

Edwards, A. W. F. 1963. The Measure of Association in a 2×2 Table. *Journal of the Royal Statistical Society. Series A (General)*. **126** (1): 109–114. http://doi.org/10.2307/2982448.

Edwards, A. W. F. 1992. *Likelihood*. Johns Hopkins University Press. ISBN 9780521318716.

Endel F., Piringer, H. 2015. Data Wrangling: Making Data Useful Again. *IFAC-PapersOnLine* **48** (1): 111–112. https://doi.org/10.1016/j.ifacol.2015.05.197.

Feng, Y. and Gray, K.E. 2017. Parameters Controlling Pressure And Fracture Behaviors In Field Injectivity Tests: A Numerical Investigation Using Coupled Flow And Geomechanics Model. *Computers and Geotechnics* 87: 49–61. https://doi.org/10.1016/j.compgeo.2017.02.002.

Ferla, A., Lavrov, A. and Fjær, E. 2009. Finite-Element Analysis of Thermal-Induced Stresses Around A Cased Injection Well. *Journal of Physics: Conference Series* **181**: 12–51. https://doi.org/10.1088%2F1742-6596%2F181%2F1%2F012051.

Finkelstein, M., 2008. *Failure Rate Modelling for Reliability and Risk*. London: Springer Science and Business Media. https://doi.org/10.1007/978-1-84800-986-8.

Fleckenstein, W.W., Eustes, A.W., and Miller, M.G. 2001. Burst Induced Stresses in Cemented Wellbores. SPE/AAPG Western Regional Meeting, Long Beach, California 19–22 June. SPE-62596-MS. https://doi.org/10.2118/62596-MS.

Freund, Y. and Schapire, R.E. 1999. Large Margin Classification Using the Perceptron Algorithm. *Machine learning*, **37**: 277–296. https://doi.org/10.1023/A:1007662407062.

Furche, T., Gottlob, G., Libkin, L., et al. 2016. Data Wrangling for Big Data: Challenges and Opportunities. *Proc*. 19th International Conference on Extending Database Technology, Bordeaux, France, 15-18 March. http://doi.org/10.5441/002/edbt.2016.44.

Gao, D., Sun, L., and Lian, J. 2010. Prediction of casing wear in extended-reach drilling. Petrol. Sci. 7 (4): 494–501. https://doi.org/10.1007/s12182-001-0098-6.

Gao, L. and Hsu, C.T. 1998. Fatigue of concrete under uniaxial compression cyclic loading. Materials Journal 95 (5): 575–581.

Gerrity, T. P. 1979. Decision Support Systems: An Organizational Perspective. *Sloan Management Review*, **20** (2): 81.

Gideon, S. 1978. Estimating the Dimension of a Model. *The annals of statistics*, **6**(2): 461–464. https://doi.org/10.1214/aos/1176344136.

Halal, A. S., Warling, D. J., and Wagner, R. R. 1996. Minimum Cost Casing Design. SPE Annual Technical Conference and Exhibition, Denver, Colorado, 6–9 October. SPE-36448-MS. https://doi.org/10.2118/36448-MS.

Hamilton, K. and Pattillo, P.D. 2019. Developing an evaluation method for casing connections used in hydraulically fractured wells. Presented at SPE Hydraulic Fracturing Technology Conference and Exhibition, The Woodlands, Texas, USA, 5–7 February. SPE-194369-MS. https://doi.org/10.2118/194369-MS.

Han, H., Dusseault, M., and Ioannidis, M. et al. 2006a. Multiscale Pore Structure Characterization by Combining Image Analysis and Mercury Porosimetry. SPE Europe/EAGE Annual Conference and Exhibition, Vienna, Austria, 12–15 June. SPE-100353-MS. https://doi.org/10.2118/100353-MS.

Han, H., Dusseault, M., and Xu, B., et al. 2006a. Simulation of Tectonic Deformation and Large Area Casing Shear Mechanisms–Part B: Geomechanics. In 41st U.S. Symposium on Rock Mechanics, Golden, Colorado, 17-21 June. ARMA-06-1004.

Han, H., Dusseault, M.B., Ioannidis, M. et al. 2006b. Multiscale pore structure characterization by combining image analysis and mercury porosimetry. Presented at SPE Europec/EAGE

Annual Conference and Exhibition, Vienna, Austria, 12–15 June. SPE-100353-MS. https://doi.org/10.2118/100353-MS.

Haroutune K. A. 2009. Avoiding Information Bias in Exposure Assessment. In *The Case-Control Method: Design and Applications*, ed. Haroutune K. A., Chap. 4, 63–86. Oxford University Press.

Hastie, T., James, G., and Witten, D. et al. 2013. *An Introduction to Statistical Learning*, **112**: 18. New York: springer. https://doi.org/10.1007/978-1-4614-7138-7.

Hastie, T., Tibshirani, R. and Friedman, J., 2009. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. New York: Springer-Verlag. https://doi.org/10.1007/978-0-387-84858-7.

Hillson, D., Hulett, D. 2004. Assessing Risk Probability: Alternative Approaches. *Proc*. PMI Global Congress Europe, Prague, Czech Republic, April.

Hoeink, T. and Zambrano C. 2017. Shale Discrimination with Machine Learning Methods. 51st U.S. Rock Mechanics/Geomechanics Symposium, San Francisco, California, USA, 25-28 June. ARMA-2017-0769.

Hopfield, J. J. 1982. Neural Networks and Physical Systems with Emergent Collective Computational Abilities. *National Academy of Sciences*, **79**(8): 2554–2558. https://doi.org/10.1073/pnas.79.8.2554.

Hou, Z., Yang, C., and Wang, L., et al. 2016. Hydraulic Fracture Propagation of Shale Horizontal Well by Large-Scale True Triaxial Physical Simulation Test. *Rock and Soil Mechanics* **37** (2), 407–414. https://doi.org/10.16285/j.rsm.2016.02.013.

Hou, Z., Yang, C., Wang, L., et al. 2016. Hydraulic fracture propagation of shale horizontal well by large-scale true triaxial physical simulation test. *Rock Soil Mech*. **37** (2): 407–414. https://doi.org/10.16285/j.rsm.2016.02.013.

Hu, Z., Yang, Z., and Salakhutdinov, et al. 2016. Deep Neural Networks with Massive Learned Knowledge. *Proc*. Conference on Empirical Methods in Natural Language Processing, Austin, Texas, 1–5 November. https://doi.org/10.18653/v1/D16-1173.

Huang, W., Gao, D. 2015. A theoretical study of the critical external pressure for casing collapse. *J. Nat. Gas Sci. Eng.* **27**: 290–297. https://doi.org/10.1016/j.jngse.2015.08.063.

Jellison, M.J., Brock, J.N. 1998. The impact of compression forces on casing string designs and connectors. Presented at IADC/SPE Asia Pacific Drilling Technology, Jakarta, Indonesia, 7–9 September. SPE-47790-MS. https://doi.org/10.2118/47790-MS.

Junior, N.M., Carrasquila, A., Figueiredo, A. et al. 2015. Worn pipes collapse strength: experimental and numerical study. *J. Pet. Sci. Eng*. **133**: 328–334. https://doi.org/10.1016/j.petrol.2015.06.024.

Kaplan, E. L. and Meier, P. 1958. Nonparametric Estimation from Incomplete Observations. *J. Amer. Statist. Assoc.* **53** (282): 457–481. https://doi.org/10.2307/2281868.

Kayis, B., Koh, S., and Saad, S. et al. 2007. A Review of Techniques for Risk Management in Projects. *Benchmarking: An International Journal*, 14(1): 22–36. https://doi.org/10.1108/14635770710730919.

Keogh, R., and Cox, D. 2014. *Case-Control Studies*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139094757.

Kim, J., 2017. *Study of Casing Failure in Unconventional Wells Using Data Analysis*. MSc Thesis, Texas A&M University, College Station, Texas (August 2017).

Kiran, R., Salehi, S., and Hayatdavoudi, A. et al. 2018. Wellbore Strengthening Characteristics Evaluation of Sandstones Using Experimental and Analytical Studies. *Proc*. 52nd U.S. Rock Mechanics/Geomechanics Symposium, Seattle, Washington, 17-20 June. ARMA-2018-1129.

Klimberg, R. and McCullough, B.D., 2016. *Fundamentals of Predictive Analytics with JMP*, second edition. SAS Institute. ISBN: 9781629608013.

Lavrov, A., Todorovic, J. and Torsæter, M. 2015. Numerical Study of Tensile Thermal Stresses in A Casing-Cement-Rock System with Heterogeneities. *Proc*. 49th US Rock Mechanics/Geomechanics Symposium, San Francisco, California, 28 June–1 July. ARMA-2015-110.

Li, C. and Samuel, R. 2016. Casing Integrity: Modeling Strength Degradation. IADC/SPE Drilling Conference and Exhibition, Fort Worth, Texas, USA, 1–3 March. SPE-178791-MS. https://doi.org/10.2118/178791-MS.

Lim, T.K., Tellier, E. and Howells, H. 2012. Wellhead, Conductor and Casing Fatigue–Causes and Mitigation. *Proc*. 32nd International Conference on Ocean, Offshore and Artic Engineering, Nantes, France, 9–14 June. OMAE2013-11112.

Lin, T., Zhang, Q., Lian, Z. et al. 2016. Evaluation of casing integrity defects considering wear and corrosion–Application to casing design. *J of Natural Gas Science and Engineering* **29**: 440– 452. https://doi.org/10.1016/j.jngse.2016.01.029.

Liu, Z., Samuel, R., Gonzales, A. et al. 2018. Analysis of Casing Fatigue Failure During Multistage Fracturing Operations. Abu Dhabi International Petroleum Exhibition & Conference, Abu Dhabi, UAE, 12-15 November. SPE-193189-MS. https://doi.org/10.2118/193189-MS.

Liu, J. and Yu, X. 2012. Stress Analysis on the Combination of Casing-Cement Ring -Surrounding Rock Considering Fluid-Solid Coupling. *Electronic Journal of Geotechnical Engineering* **17**:1863–73.

Liu, R., Zhang, Z., Zhang, P. et al. 2015. Extremely-Low-Cycle Fatigue Behaviors of Cu and Cu–Al Alloys: Damage Mechanisms and Life Prediction. *Acta Materialia*, **83**: 341–356. https://doi.org/10.1016/j.actamat.2014.10.002.

Maharaj, E.A., 1996. A Significance Test for Classifying ARMA Models. *Journal of Statistical Computation and Simulation*, **54** (4): 305–331. https://doi.org/10.1080/00949659608811737.

Mann, C. J. 2003. Observational Research Methods. Research Design II: Cohort, Cross Sectional, And Case-Control Studies. *Emergency medicine journal*, **20**(1): 54-60. http://dx.doi.org/10.1136/emj.20.1.54.

Mantel, N., 1966. Evaluation of Survival Data and Two New Rank Order Statistics Arising in Its Consideration. *Cancer Chemother. Rep.*, 50: 163–170.

Mao, L., Cai, M., and Wang, G. 2018. Effect of Rotation Speed on the Abrasive-erosive-corrosive wear of steel pipes against steel casings used in drilling for petroleum. *Wear* **410**: 1-10. https://doi.org/10.1016/j.wear.2018.06.002.

Menard, S. 2002. *Applied Logistic Regression Analysis*, second edition. Thousand Oaks, CA: Sage. https://dx.doi.org/10.4135/9781412983433.

Miller, R. G. 1971. *Survival Analysis*, second edition. New York: John Wiley and Sons, Inc. ISBN 0-471-25218-2.

Misra, S., Ganguly, E., and Wu, Y. 2019. Generalization of Machine Learning Assisted Segmentation of Scanning Electron Microscopy Images of Organic-Rich Shales. *Machine Learning for Subsurface Characterization*, 315.

Moore, G. E. 1979. Pennsylvanian Paleogeography of the Southern Mid-Continent. In Pennsylvanian Sandstones of the Mid-Continent. *Tulsa Geological Society*, Special Publication **1**: 2–12.

Nanjo, T. and Tanaka, S., 2019. Carbonate Lithology Identification with Machine Learning. Abu Dhabi International Petroleum Exhibition & Conference, Abu Dhabi, UAE, 11-14 November. SPE-197255-MS. https://doi.org/10.2118/197255-MS.

Noshi, C. I., Noynaert, S. F., and Schubert, J. J. 2018. Casing Failure Data Analytics: A Novel Data Mining Approach in Predicting Casing Failures for Improved Drilling Performance and Production Optimization. Presented at the SPE Annual Technical Conference and Exhibition, Dallas, Texas, USA, 24–26 September. SPE-191570-MS. https://doi.org/10.2118/191570-MS.

Noshi, C. I., Noynaert, S. F., and Schubert, J. J. 2018. Failure Predictive Analytics Using Data Mining: How to Predict Unforeseen Casing Failures? Presented at the Abu Dhabi International Petroleum Exhibition & Conference, Abu Dhabi, UAE, 12–15 November. SPE-193194-MS. https://doi.org/10.2118/193194-MS.

Noshi, C. I., Noynaert, S. F., and Schubert, J. J. 2019. Data Mining Approaches for Casing Failure Prediction and Prevention. Presented at the International Petroleum Technology Conference, Beijing, China, 26–28 March. IPTC-19311-MS. https://doi.org/10.2523/IPTC-19311-MS.

Ogunsesan, O.A., Hossain, M., Iyi, D. et al. 2019. CFD Modelling of Pipe Erosion Due to Sand Transport. *Proc*. 1st International Conference on Numerical Modelling in Engineering. https://doi.org/10.1007/978-981-13-2273-0_22.

Ogunsesan, O.A., Hossain, M., Iyi, D. et al. 2019. CFD modelling of pipe erosion due to sand transport. Presented at the Numerical Modelling in Engineering. Springer, Singapore, pp. 274–289.

Oluwaseun, A. 2018. Relevant Information on Oil and Gas Casing Design. *Chem Eng J*. ***2***(8): 000179. https://doi:10.23880/ppej-16000179.

Ord, J. K. 1972. *Families of Frequency Distributions*. London: Griffin. https://cds.cern.ch/record/110249.

Ounsakul, T., Sirirattanachatchawan, T., Pattarachupong, W. et al. 2019. Artificial Lift Selection Using Machine Learning. International Petroleum Technology Conference, Beijing, China, 26–28 March. IPTC-19423-MS. https://doi.org/10.2523/IPTC-19423-MS.

Patel, M., 2018. *Data Structure and Algorithm with C*, first edition. Educreation Publishing.

Payne, M. L., Leturno, R. E., and Harder, C. A. 1993. Fatigue Failure of API 8-Round Casings in Drilling Service. Presented at the SPE Annual Technical Conference and Exhibition, Houston, Texas, 3–6 October. SPE-26321-MS. https://doi.org/10.2118/26321-MS.

PMI, S.C., 2008. *A Guide to The Project Management Body of Knowledge*. Newton Square, PA, USA: Project Management Institute, Inc.

Pollock, J., Stoecker-Sylvia, Z., Veedu, V., et al. 2018. Machine Learning for Improved Directional Drilling. Offshore Technology Conference, Houston, Texas, USA, 30 April - 3 May. OTC-28633-MS. https://doi.org/10.4043/28633-MS.

Popov, G., Lyon, B. K., & Hollcroft, B. (2016). *Risk assessment: A practical guide to assessing operational risks*. John Wiley and Sons.

Porta M, ed. 2014. *Dictionary of Epidemiology*, sixth edition. Oxford University Press. https://doi.org/10.1093/acref/9780199976720.001.0001.

Prislin, I. and Maroju, S., 2017. Mooring Integrity and Machine Learning. Offshore Technology Conference, Houston, Texas, USA, 1–4 May. OTC-27866-MS. https://doi.org/10.4043/27866-MS.

Rausand, M. and Haugan, S. 2013. *Risk Assessment: Theory, Methods, and Applications*, second edition. John Wiley and Sons. ISBN: 978-1-119-37723-8.

Raz, T. and Michael, E., 2001. Use and Benefits of Tools for Project Risk Management. *International journal of project management*, **19**(1): 9–17. https://doi.org/10.1016/S0263-7863(99)00036-8.

Salehi, S., Hareland, G., Dehkordi, K., et al. 2009. Casing Collapse Risk Assessment and Depth Prediction with A Neural Network System Approach. *Journal of Petroleum Science and Engineering*, **69**: 156-162. https://doi.org/10.1016/j.petrol.2009.08.011.

Sarle, W. S. 1994. Neural Networks and Statistical Models. *Proc*. 19th Annual SAS Users Group International Conference, Cary, NC: SAS Institute, USA.

Shalev-Shwartz, S. and Ben-David, S., 2014. *Understanding machine learning: From theory to algorithms*. New York, NY: Cambridge University Press.

Shen, Z., Beck, F.E. and Ling, K., 2014. The Mechanism of Wellbore Weakening in Worn Casing-Cement-Formation System. *Journal of Petroleum Engineering*. https://doi.org/10.1155/2014/126167.

Sihag, P., Tiwari, N.K. and Ranjan, S., 2018. Support Vector Regression-Based Modeling of Cumulative Infiltration of Sandy Soil. *ISH Journal of Hydraulic Engineering*, **26**(1): 44–55. https://doi.org/10.1080/09715010.2018.1439776.

Simeone, O. 2018. A Very Brief Introduction to Machine Learning with Applications to Communication Systems. *IEEE Transactions on Cognitive Communications and Networking*, **4**(4): 648-664. https://doi.org/10.1109/TCCN.2018.2881442.

Song, M. and Zhou, X. 2019. A casing damage prediction method based on principal component analysis and gradient boosting decision tree algorithm. Presented in SPE Middle East Oil and Gas Show and Conference, Manama, Bahrain.

Sprague Jr, R. H. 1980. A Framework for The Development of Decision Support Systems. *MIS quarterly*, **4**(4): 1-26. https://doi.org/10.2307/248957.

Tan, C., Wu, H., Liu, J. et al. 2019. Casing Damage Prediction Model Based on the Data-Driven Method. Presented in Proceedings of the Eurasia Conference on IOT, Communication and Engineering, Yunlin, Taiwan. https://doi.org/10.1155/2020/8315908.

Tang, Q., Wu, H., Teng, G. et al. 2019. Prediction of casing damage in unconsolidated sandstone reservoirs using machine learning algorithms. Presented in Proceedings of the International Conference on Computation, Communication and Engineering, Fujian, China.

Thattil, M.S. 2017. Casing Design for Casing/Liner while Drilling. *International Research Journal of Engineering and Technology*, **4**(7): 1–3.

Tong, H. and Tang, X.H. 2016. Oil Casing Introduction. *International Journal of Science and Research*. **5**(6): 696–698. http://dx.doi.org/10.21275/v5i6.NOV164332.

Vudovich, A., Chin, L., Morgan, D. 1988. Casing Deformation in Ekofisk. Presented at the Offshore Technology Conference, Houston, Texas, 2–5 May. OTC-5623-MS. https://doi.org/10.4043/5623-MS.

Walden, D. D., Roedler, G. J., and Forsberg, K. J. et al. ed. 2011. *INCOSE Systems Engineering Handbook: A Guide for System Life Cycle Processes and Activities*, Fourth Edition. John Wiley and Sons, Inc. ISBN: 9781118999400.

Waller, P. 1988. Expert Systems and Artificial Intelligence in Decision Support Systems. *Journal of the Royal Statistical Society*, **37**(3): 348–349. https://doi.org/10.2307/2348181.

Wang, H., Liao, X., Zhao, X. et al. 2014. Advances of Technology Study of Stimulated Reservoir Volume in an Unconventional Reservoir. *Special Oil Gas Reservoirs* **21** (2): 8–15.

Wang, P., Zhu, Q., Bu, X. 2011. Finite Element Analysis on Casing Failure Based on ADINA. *Syst. Eng. Procedia* **1**: 42–47. https://doi.org/10.1016/j.sepro.2011.08.008.

Wilson, A. 2018. Risk-Based Statistical Approach to Predict Casing Leaks. *J. Pet. Technol.* **70** (06): 71–72. SPE-0618-0071-JPT. https://doi.org/10.2118/0618-0071-JPT.

Wright, R. E. 1995. Logistic regression. *Reading and understanding multivariate statistics*, 217–244. American Psychological Association.

Wu et al. 2008. Casing Failures in Cyclic Steam Injection Wells. Paper presented at the IADC/SPE Asia Pacific Drilling Technology Conference and Exhibition, Jakarta, Indonesia, 2–5 August. SPE-114231-MS. https://doi.org/10.2118/114231-MS.

Xi, Y., Li, J., Liu, G. et al. 2018. Numerical Investigation for Different Casing Deformation Reasons in Weiyuan-Changning Shale Gas field During Multistage Hydraulic Fracturing. *J. Pet. Sci. Eng.* **163**: 691–702. https://doi.org/10.1016/j.petrol.2017.11.020.

Yang, S., Han, L., Feng, C., et al. 2018. Mechanical performance of casing in in-situ combustion thermal recovery. *Journal of Petroleum Science and Engineering*, **168**: 32–38.

Yin, F., Deng, Y., He, Y. et al. 2018b. Mechanical Behavior of Casing Crossing Slip Formation in Water flooding Oilfields. *J. Pet. Sci. Eng.* **167**: 796–802. https://doi.org/10.1016/j.petrol.2017.12.069.

Yin, F., Han, L., Yang, S. et al. 2018a. Casing Deformation from Fracture Slip in Hydraulic Fracturing. *J. Pet. Sci. Eng.* **166**: 235–241. https://doi.org/10.1016/j.petrol.2018.03.010.

Yin, F., Xiao, Y., Han, L. et al. 2018. Quantifying the Induced Fracture Slip and Casing Deformation in Hydraulically Fracturing Shale Gas Wells. J. *Nat. Gas Sci. Eng.*, **60**: 103–111. https://doi.org/10.1016/j.jngse.2018.10.005.

Yuan, Z., Schubert, J., Teodoriu, C. et al. 2012. HPHT gas well cementing complications and its effect on casing collapse resistance. SPE Oil and Gas India Conference and Exhibition, Mumbai, India, 28-30 March. SPE-153986-MS. https://doi.org/10.2118/153986-MS

Zachmann, K., 2014. Risk in Historical Perspective: Concepts, Contexts, and Conjunctions. *Risk-A Multidisciplinary Introduction*, 3–35. Springer, Cham. https://doi.org/10.1007/978-3-319-04486-6_1.

Zeng, Q. and Yao, J., 2015. Experiment of Shale Failure Mechanism Based on Particle Flow Theory. *Petrol. Drill. Tech.*, **43** (1): 33–37.

Zhang, P., Zhang, S. et al. 2012. Effect of Pressure Depletion on Stress Field and Casing Load Alteration in Mature Fields: A Case Study. Presented at the SPE Latin America and Caribbean Mature Fields Symposium, Salvador, Bahia, Brazil.15–16 March. SPE-184902-MS. https://doi.org/10.2118/184902-MS.

Zhang, S., Zhong, R., and Liu, Y. 2016. Numerical Modeling of Land Subsidence Resulting for Oil Production. *Proc*. 50th US Rock Mechanics/Geomechancis Symposium, Houston, Texas, USA, 26–29 June. ARMA-2016–382.

Zhao, Z., Xu, Z. and He, P. 2017, Impaction and Wear-Resistant Carbon Elastic Seal for HTHP Downhole Valve Application, SPE Kingdom of Saudi Arabia Annual Technical Symposium and Exhibition, Dammam, Saudi Arabia, 24-27 April. SPE-187975-MS. https://doi.org/10.2118/187975-MS.

APPENDIX A

This section is devoted to the results obtained from case-control analysis implemented using odds ratio (OR) as the association measurement.

MODEL [1]

AIC = 81.024

Model Features:

MAX_INCL	FRAC_SEASON	BHT
DRILL_FRAC_INTERVAL		

Coefficients:

Risk Factor	Level	Estimate	Std. Error	z-value	p-value
(Intercept)		-16.9326	15.34374	-1.104	0.270
FRAC_SEASON	**Spring**	1.27619	0.9615	1.327	0.184
	Summer	0.64597	0.98086	0.659	0.510
	Winter	1.54095	1.0995	1.402	0.161
BHT		-0.0348	0.02547	-1.366	0.172
MAX_INCL		0.21509	0.16845	1.277	0.202
DRILL_FRAC_INTERVAL		0.01019	0.00907	1.124	0.261

Risk Factor	Level	OR	LL	UL	Risk	Impact
(Intercept)		4.43E-08	3.85E-21	509494.4	-	-
FRAC_SEASON	**Spring**	3.582973	0.544252	23.5878	▲	260%
	Summer	1.907827	0.279004	13.04573	▲	90%
	Winter	4.669031	0.541146	40.28458	▲	367%
BTH		0.965802	0.918779	1.015232	▼	4%
MAX_INCL		1.239976	0.891299	1.725057	▲	24%
DRILL_FRAC_INTERVAL		1.010245	0.992445	1.028365	▼	1%

MODEL [2]

AIC = 81.654

Model Features:

DL_FREQ_10PLUS	FRAC_SEASON	MAX_INCL
DRILL_FRAC_INTERVAL	BHT	

Coefficients:

Risk Factor	Level	Estimate	Std. Error	z-value	p-value
(Intercept)		-21.3322	16.37178	-1.303	0.193
FRAC_SEASON	**Spring**	1.274856	0.968498	1.316	0.188
	Summer	0.684717	0.988129	0.693	0.488
	Winter	1.569822	1.127770	1.392	0.164
BHT		-0.03685	0.026252	-1.404	0.160
MAX_INCLs		0.276491	0.185307	1.492	0.136
DRILL_FRAC_INTERVAL		0.01135	0.009356	1.213	0.225
DL_FREQ_10PLUS		-0.08204	0.071231	-1.152	0.249

Risk Factor	Level	OR	LL	UL	Risk	Impact
(Intercept)		5.44E-10	6.30E-24	46933.21	-	-
FRAC_SEASON	**Spring**	3.578187	0.536119	23.88170	▲	257%
	Summer	1.983211	0.285928	13.75567	▲	98%
	Winter	4.805794	0.526972	43.82708	▲	380%
BTH		0.963818	0.915481	1.014708	▼	4%
MAX_INCL		1.318495	0.916943	1.895897	▲	31%
DRILL_FRAC_INTERVAL		1.011415	0.993037	1.030133	▲	1%
DL_FREQ_10PLUS		0.921239	0.801197	1.059267	▼	8%

MODEL [3]

AIC = 82.75

Model Features:

LATERAL_SHRINKAGE	FRAC_SEASON	MAX_INCL
DL_FREQ_10PLUS	DRILL_FRAC_INTERVAL	BHT

Coefficients:

Risk Factor	Level	Estimate	Std. Error	z-value	p-value
(Intercept)		-22.8954	16.58420	-1.381	0.167
FRAC_SEASON	**Spring**	1.363045	0.970678	1.404	0.160
	Summer	0.623232	0.989067	0.63	0.529
	Winter	1.598323	1.124400	1.421	0.155
BHT		-0.03951	0.026623	-1.484	0.138
LATERAL_SHRINKAGE		0.138172	0.151765	0.91	0.363
MAX_INCL		0.276077	0.186439	1.481	0.139
DRILL_FRAC_INTERVAL		0.01356	0.009862	1.375	0.169
DL_FREQ_10PLUS		-0.08408	0.072369	-1.162	0.245

Risk Factor	Level	OR	LL	UL	Risk	Impact
(Intercept)		1.14E-10	8.71E-25	14907.89	-	-
FRAC_SEASON	**Spring**	3.908075	0.583049	26.19513	▲	290%
	Summer	1.864947	0.268383	12.95917	▲	86%
	Winter	4.944734	0.545801	44.79727	▲	394%
BTH		0.961264	0.912391	1.012755	▼	4%
LATERAL_SHRINKAGE		1.148172	0.852751	1.545939	▲	14%
MAX_INCL		1.317949	0.914531	1.899325	▲	31%
DRILL_FRAC_INTERVAL		1.013653	0.994247	1.033437	▲	1%
DL_FREQ_10PLUS		0.919354	0.797776	1.059461	▼	9%

MODEL [4]

AIC = 84.53

Model Features:

LATERAL_SHRINKAGE	FRAC_SEASON	DL_SEVERE_MD
DL_FREQ_10PLUS	MAX_INCL	BHT
DRILL_FRAC_INTERVAL		

Coefficients:

Risk Factor	Level	Estimate	Std. Error	z-value	p-value
(Intercept)		-2.06E+01	1.73E+01	-1.191	0.234
FRAC_SEASON	**Spring**	1.31E+00	9.76E-01	1.346	0.178
	Summer	4.81E-01	1.04E+00	0.465	0.642
	Winter	1.47E+00	1.16E+00	1.271	0.204
BHT		-4.44E-02	2.87E-02	-1.545	0.122
LATERAL_SHRINKAGE		1.36E-01	1.50E-01	0.905	0.366
MAX_INCL		2.52E-01	1.92E-01	1.311	0.190
DL_SEVERE_MD		1.06E-04	2.29E-04	0.465	0.642
DRILL_FRAC_INTERVAL		1.28E-02	9.98E-03	1.279	0.201
DL_FREQ_10PLUS		-8.76E-02	7.36E-02	-1.19	0.234

Risk Factor	Level	OR	LL	UL	Risk	Impact
(Intercept)		1.19E-09	2.47E-24	575473.1	-	-
FRAC_SEASON	**Spring**	3.718162	0.549407	25.163	▲	271%
	Summer	1.617872	0.212585	12.31275	▲	61%
	Winter	4.350393	0.450515	42.00953	▲	335%
BTH		0.956593	0.904236	1.011981	▼	5%
LATERAL_SHRINKAGE		1.145313	0.853594	1.536729	▲	14%
MAX_INCL		1.28664	0.882666	1.875501	▲	28%
DL_SEVERE_MD		1.000106	0.999658	1.000555	▲	0%
DRILL_FRAC_INTERVAL		1.012847	0.993222	1.032859	▲	0%
DL_FREQ_10PLUS		0.916164	0.793158	1.058246	▼	9%

MODEL [5]

AIC = 86.078

Model Features:

CEMENT	FRAC_SEASON	DL_SEVERE_MD
LATERAL_SHRINKAGE	MAX_INCL	BHT
DL_FREQ_10PLUS	DRILL_FRAC_INTERVAL	

Coefficients:

Risk Factor	Level	Estimate	Std. Error	z-value	p-value
(Intercept)		-2.24E+01	1.75E+01	-1.277	0.201
FRAC_SEASON	**Spring**	1.41E+00	9.94E-01	1.42	0.155
	Summer	5.79E-01	1.05E+00	0.55	0.582
	Winter	1.53E+00	1.17E+00	1.307	0.191
CEMENT	**Yes**	-5.38E-01	8.08E-01	-0.666	0.506
BHT		-3.73E-02	3.08E-02	-1.21	0.226
LATERAL_SHRINKAGE		1.09E-01	1.54E-01	0.71	0.478
MAX_INCL		2.61E-01	1.94E-01	1.347	0.178
DL_SEVERE_MD		1.54E-04	2.44E-04	0.629	0.529
DRILL_FRAC_INTERVAL		1.24E-02	1.02E-02	1.215	0.224
DL_FREQ_10PLUS		-8.86E-02	7.39E-02	-1.198	0.231

Risk Factor	Level	OR	LL	UL	Risk	Impact
(Intercept)		1.94E-10	2.43E-25	154793.4	-	-
FRAC_SEASON	**Spring**	4.103800	0.584906	28.79298	▲	300%
	Summer	1.783528	0.226884	14.02024	▲	78%
	Winter	4.606826	0.465838	45.55840	▲	360%
CEMENT	**Yes**	0.584138	0.119982	2.843915	▼	58%
BTH		0.963390	0.906912	1.023386	▼	4%
LATERAL_SHRINKAGE		1.115431	0.82487	1.508341	▲	11%
MAX_INCL		1.298576	0.887981	1.899026	▲	29%
DL_SEVERE_MD		1.000154	0.999675	1.000633	▲	0%
DRILL_FRAC_INTERVAL		1.012484	0.992419	1.032955	▲	0%
DL_FREQ_10PLUS		0.915255	0.791773	1.057994	▼	10%

MODEL [6]

AIC = 88.009

Model Features:

CEMENT	FRAC_SEASON	DL_SEVERE_MD
LATERAL_SHRINKAGE	MAX_INCL	BHT
DL_FREQ_10PLUS	DRILL_FRAC_INTERVAL	ACID

Coefficients:

Risk Factor	Level	Estimate	Std. Error	z-value	p-value
(Intercept)		-22.5015	17.49441	-1.286	0.198
FRAC_SEASON	**Spring**	1.382646	0.996585	1.387	0.165
	Summer	0.597784	1.051459	0.569	0.570
	Winter	1.553488	1.172369	1.325	0.185
ACID	**Yes**	0.225756	0.86296	0.262	0.794
CEMENT	**Yes**	-0.56512	0.815023	-0.693	0.488
BHT		-0.03815	0.031091	-1.227	0.220
LATERAL_SHRINKAGE		0.107123	0.152429	0.703	0.482
MAX_INCL		0.266527	0.19481	1.368	0.171
DL_SEVERE_MD		0.000125	0.000268	0.466	0.641
DRILL_FRAC_INTERVAL		0.012214	0.0102	1.197	0.231
DL_FREQ_10PLUS		-0.08682	0.074595	-1.164	0.244

Risk Factor	Level	OR	LL	UL	Risk	Impact
(Intercept)		1.69E-10	2.17E-25	131603.3	-	-
FRAC_SEASON	**Spring**	3.985434	0.565152	28.10516	▲	298%
	Summer	1.818086	0.231523	14.27690	▲	81%
	Winter	4.727933	0.47504	47.05570	▲	372%
ACID	**Yes**	1.253270	0.230929	6.801600	▲	25%
CEMENT	**Yes**	0.568291	0.115029	2.807581	▼	56%
BTH		0.962571	0.905665	1.023052	▼	4%
LATERAL_SHRINKAGE		1.113071	0.825607	1.500627	▲	11%
MAX_INCL		1.305422	0.891097	1.912392	▲	30%
DL_SEVERE_MD		1.000125	0.9996	1.000650	▲	0%
DRILL_FRAC_INTERVAL		1.012289	0.992252	1.032731	▲	0%
DL_FREQ_10PLUS		0.916844	0.792134	1.061187	▼	9%

MODEL [7]

AIC = 89.901

Model Features:

CEMENT	FRAC_SEASON	DL_SEVERE_MD
LATERAL_SHRINKAGE	MAX_INCL	BHT
DL_FREQ_10PLUS	DRILL_FRAC_INTERVAL	ACID

Coefficients:

Risk Factor	Level	Estimate	Std. Error	z-value	p-value
(Intercept)		-2.23E+01	1.76E+01	-1.272	0.203
FRAC_SEASON	**Spring**	1.39E+00	9.98E-01	1.388	0.165
	Summer	5.30E-01	1.07E+00	0.494	0.622
	Winter	1.53E+00	1.17E+00	1.305	0.192
MD		-1.29E-04	3.93E-04	-0.328	0.743
ACID	**Yes**	2.59E-01	8.66E-01	0.299	0.765
CEMENT	**Yes**	-5.76E-01	8.15E-01	-0.707	0.480
BHT		-3.30E-02	3.46E-02	-0.953	0.341
LATERAL_SHRINKAGE		1.53E-01	2.07E-01	0.74	0.459
MAX_INCL		2.59E-01	1.97E-01	1.317	0.188
DL_SEVERE_MD		2.12E-04	3.73E-04	0.568	0.570
DRILL_FRAC_INTERVAL		1.25E-02	1.04E-02	1.202	0.230
DL_FREQ_10PLUS		-8.56E-02	7.45E-02	-1.149	0.251

Risk Factor	Level	OR	LL	UL	Risk	Impact
(Intercept)		1.98E-10	2.22E-25	176325.5	-	-
FRAC_SEASON	**Spring**	3.993101	0.564891	28.22644	▲	299%
	Summer	1.698039	0.207342	13.90618	▲	69%
	Winter	4.626766	0.463897	46.14597	▲	362%
MD		0.999871	0.9991	1.000642	▲	1%
ACID	**Yes**	1.295801	0.237163	7.079947	▲	29%
CEMENT	**Yes**	0.561902	0.11365	2.778134	▼	56%
BTH		0.967580	0.904164	1.035443	▼	4%
LATERAL_SHRINKAGE		1.165000	0.777287	1.746106	▲	16%
MAX_INCL		1.295481	0.881325	1.904258	▲	29%
DL_SEVERE_MD		1.000212	0.999482	1.000942	▲	0%
DRILL_FRAC_INTERVAL		1.012611	0.99212	1.033526	▲	0%
DL_FREQ_10PLUS		0.918006	0.793352	1.062247	▼	9%

MODEL [8]

AIC = 91.899

Model Features:

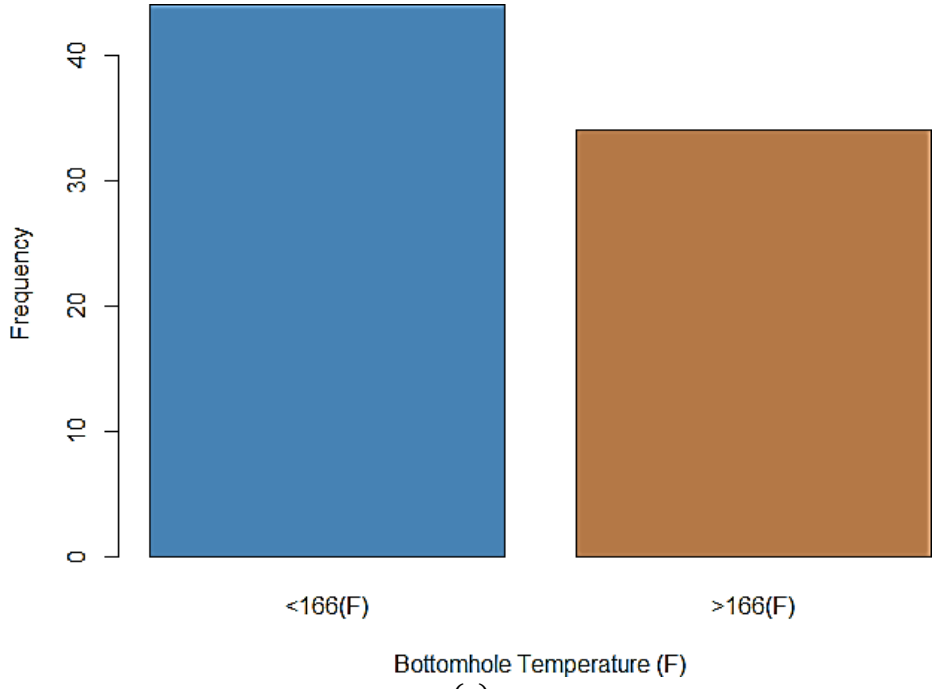
CEMENT	FRAC_SEASON	DL_SEVERE_MD
LATERAL_SHRINKAGE	MAX_INCL	BHT
DL_FREQ_10PLUS	DRILL_FRAC_INTERVAL	ACID
DL_BEND_STRESS		

Coefficients:

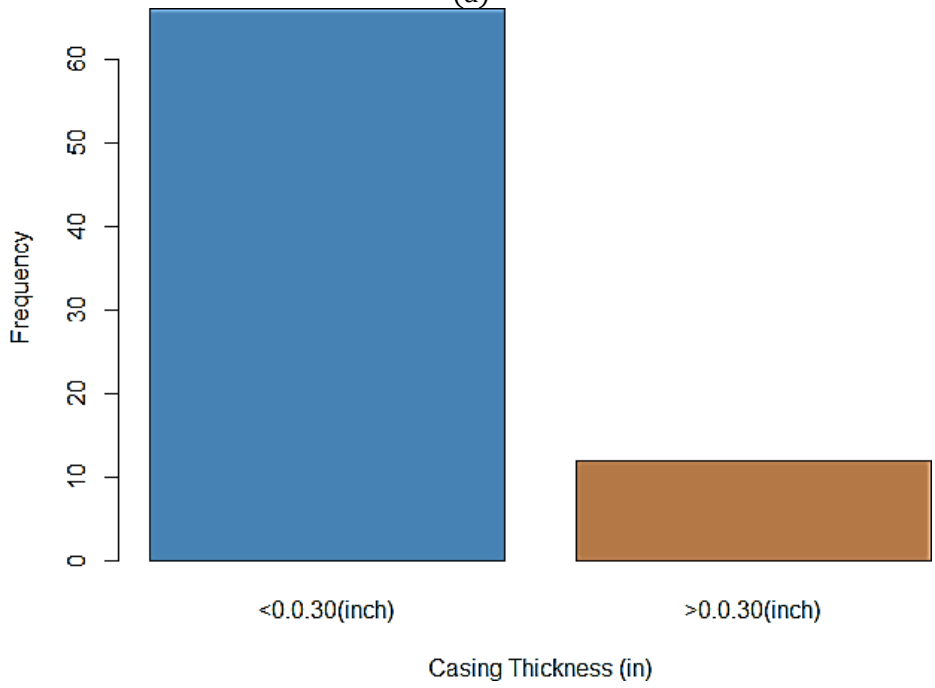
Risk Factor	Level	Estimate	Std. Error	z-value	p-value
(Intercept)		-2.24E+01	1.76E+01	-1.273	0.203
DL_BEND_STRESS		-7.71E-07	1.68E-05	-0.046	0.963
FRAC_SEASON	**Spring**	1.39E+00	9.98E-01	1.388	0.165
	Summer	5.27E-01	1.07E+00	0.49	0.624
	Winter	1.54E+00	1.20E+00	1.286	0.199
MD		-1.28E-04	3.95E-04	-0.323	0.747
ACID	**Yes**	2.57E-01	8.68E-01	0.296	0.767
CEMENT	**Yes**	-5.69E-01	8.30E-01	-0.686	0.493
BHT		-3.28E-02	3.48E-02	-0.942	0.346
LATERAL_SHRINKAGE		1.51E-01	2.09E-01	0.725	0.468
MAX_INCL		2.59E-01	1.97E-01	1.318	0.188
DL_SEVERE_MD		2.12E-04	3.73E-04	0.569	0.569
DRILL_FRAC_INTERVAL		1.25E-02	1.05E-02	1.193	0.233
DL_FREQ_10PLUS		-8.45E-02	7.80E-02	-1.082	0.279

Risk Factor	Level	OR	LL	UL	Risk	Impact
DL_BEND_STRESS		0.999999	0.999966	1.000032	▼	1%
FRAC_SEASON	**Spring**	3.997192	0.564849	28.28638	▲	299%
	Summer	1.693661	0.206214	13.91026	▲	69%
	Winter	4.681156	0.445095	49.23273	▲	368%
MD		0.999872	0.999098	1.000647	▲	1%
ACID	**Yes**	1.292664	0.235832	7.085466	▲	29%
CEMENT	**Yes**	0.566000	0.111224	2.880286	▼	56%
BTH		0.967748	0.903955	1.036043	▼	4%
LATERAL_SHRINKAGE		1.163346	0.772904	1.751024	▲	16%
MAX_INCL		1.295689	0.881411	1.904687	▲	29%
DL_SEVERE_MD		1.000212	0.999482	1.000943	▲	0%
DRILL_FRAC_INTERVAL		1.012572	0.992008	1.033562	▲	0%
DL_FREQ_10PLUS		0.919001	0.788654	1.070892	▼	9%

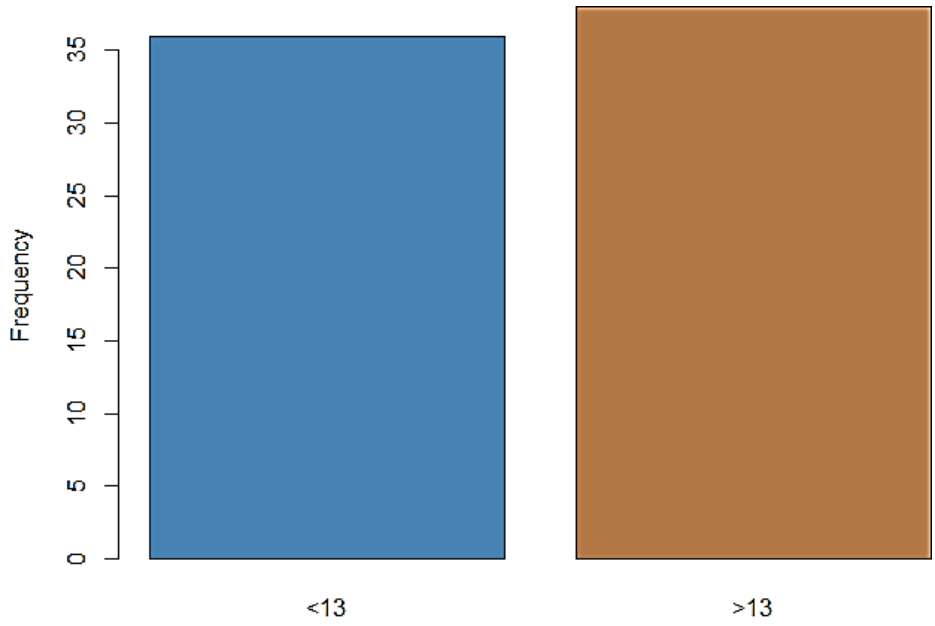
APPENDIX B



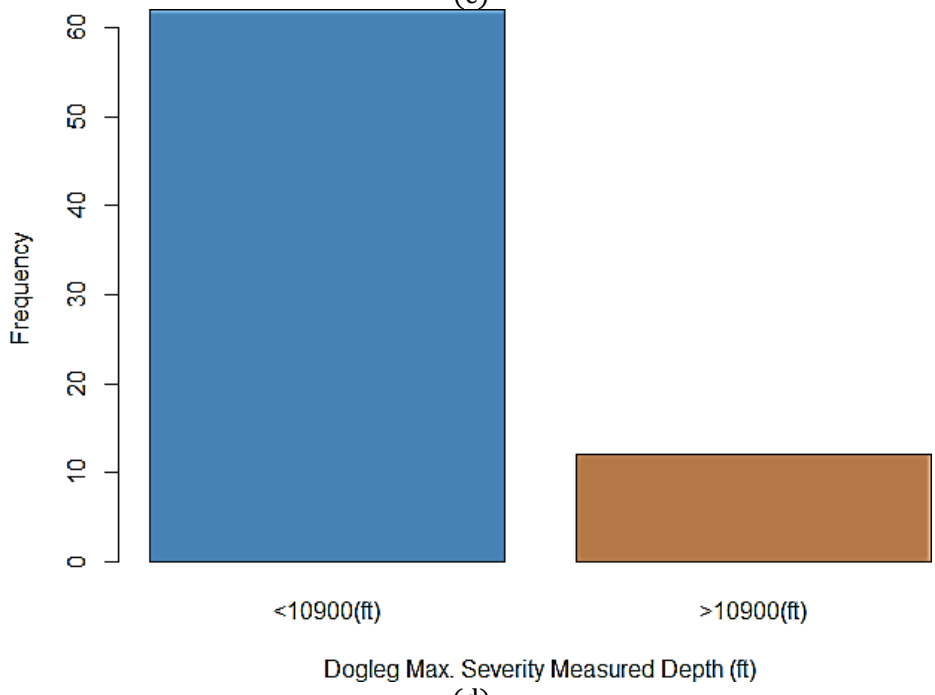
(a)



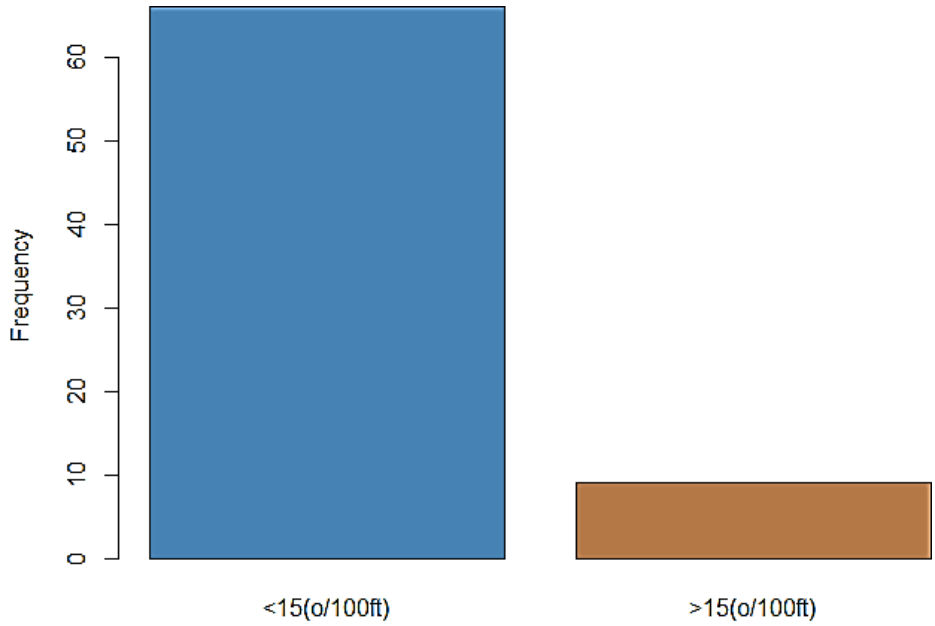
(b)


Frequency
0
5
10
15
20
25
30
35
<13
>13
Frequency of Doglegs > 10

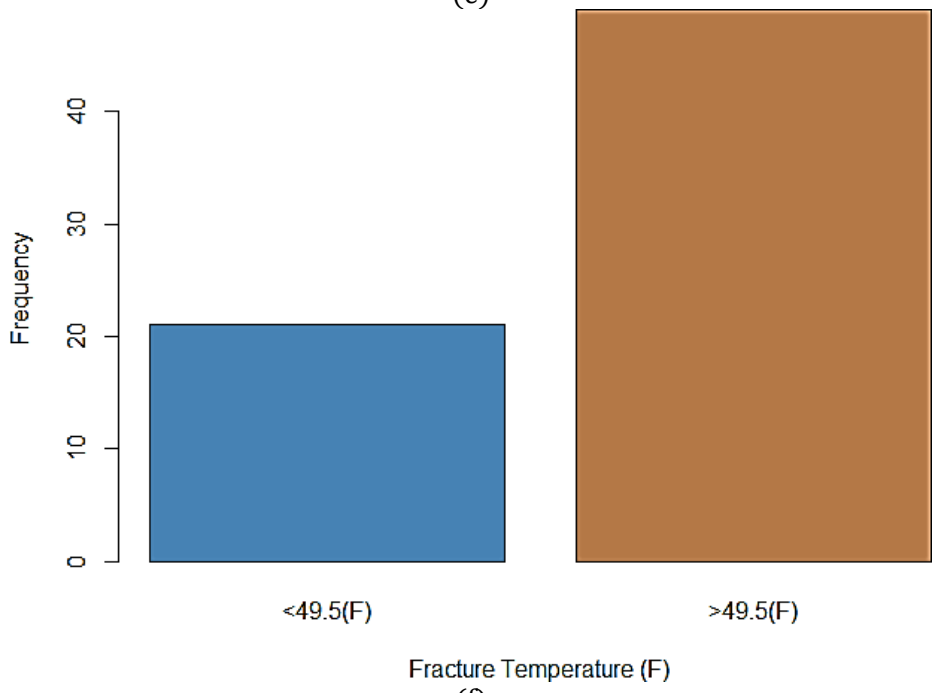
(c)


Frequency
0
10
20
30
40
50
60
<10900(ft)
>10900(ft)
Dogleg Max. Severity Measured Depth (ft)

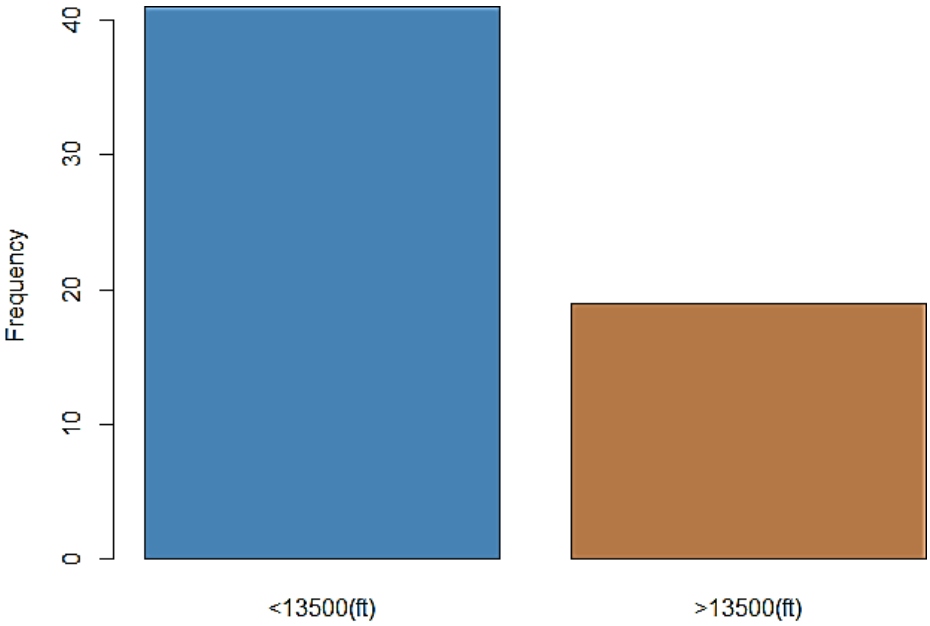
(d)


Frequency
0
10
20
30
40
50
60
<15(o/100ft)
>15(o/100ft)
Dogleg Severity (o/100ft)

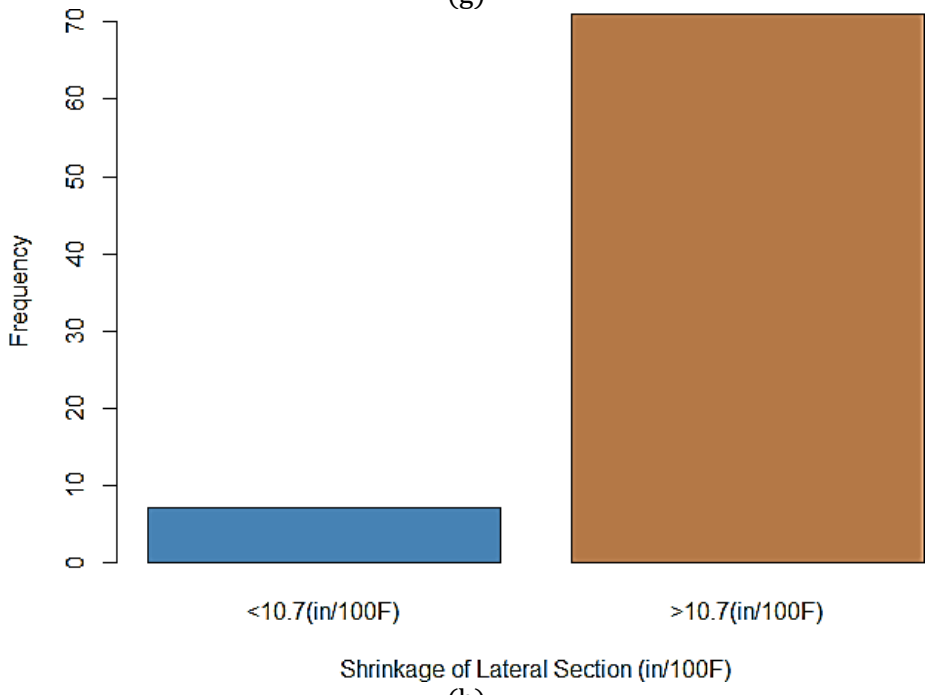
(e)


Frequency
0
10
20
30
40
<49.5(F)
>49.5(F)
Fracture Temperature (F)

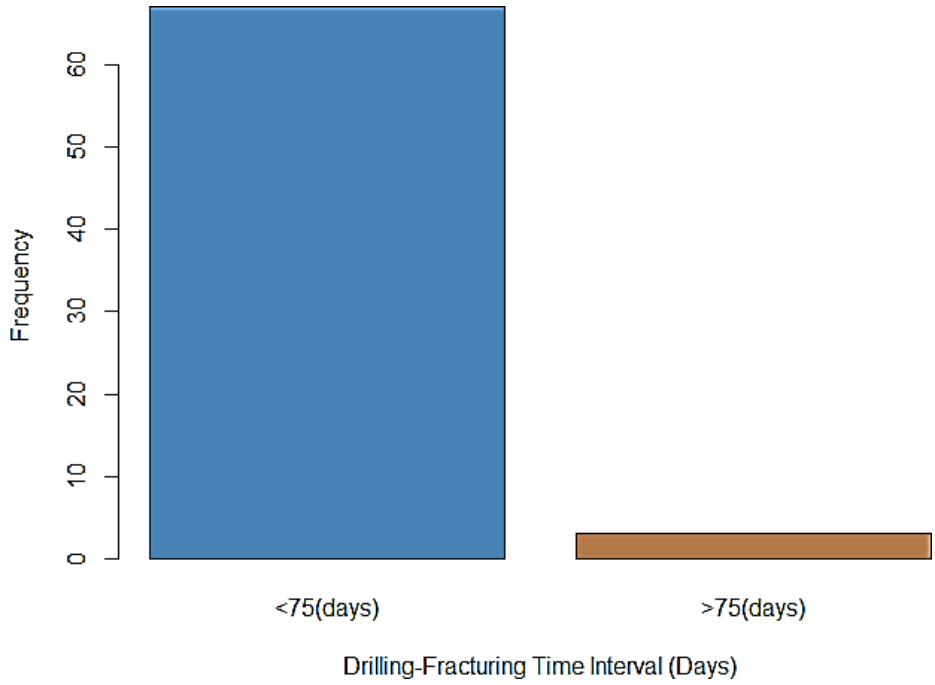
(f)


Frequency
0
10
20
30
40
<13500(ft)
>13500(ft)
Measured Depth (ft)

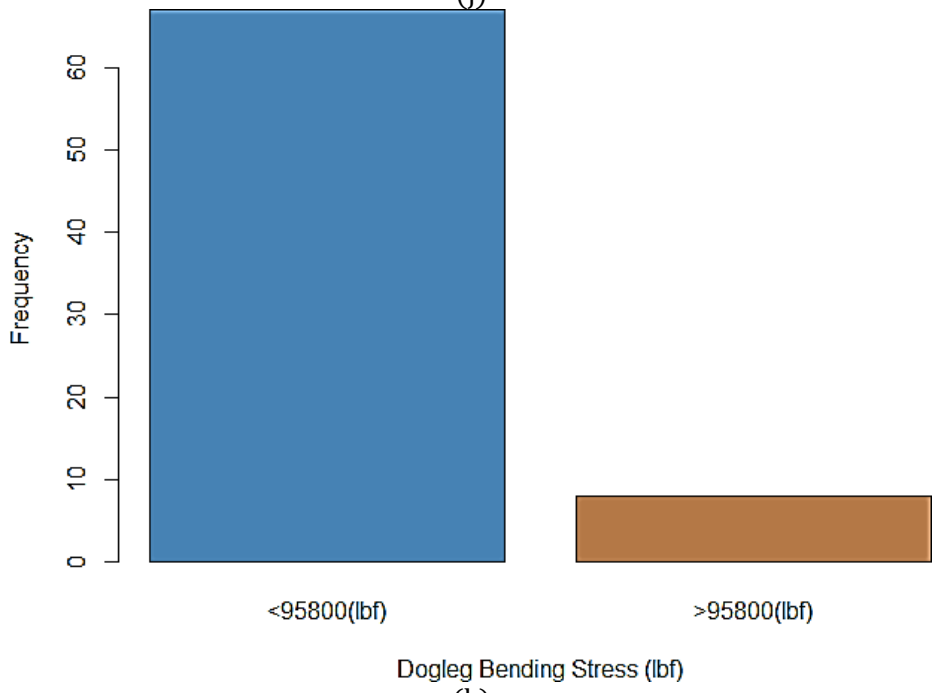
(g)


Frequency
0
10
20
30
40
50
60
70
<10.7(in/100F)
>10.7(in/100F)
Shrinkage of Lateral Section (in/100F)

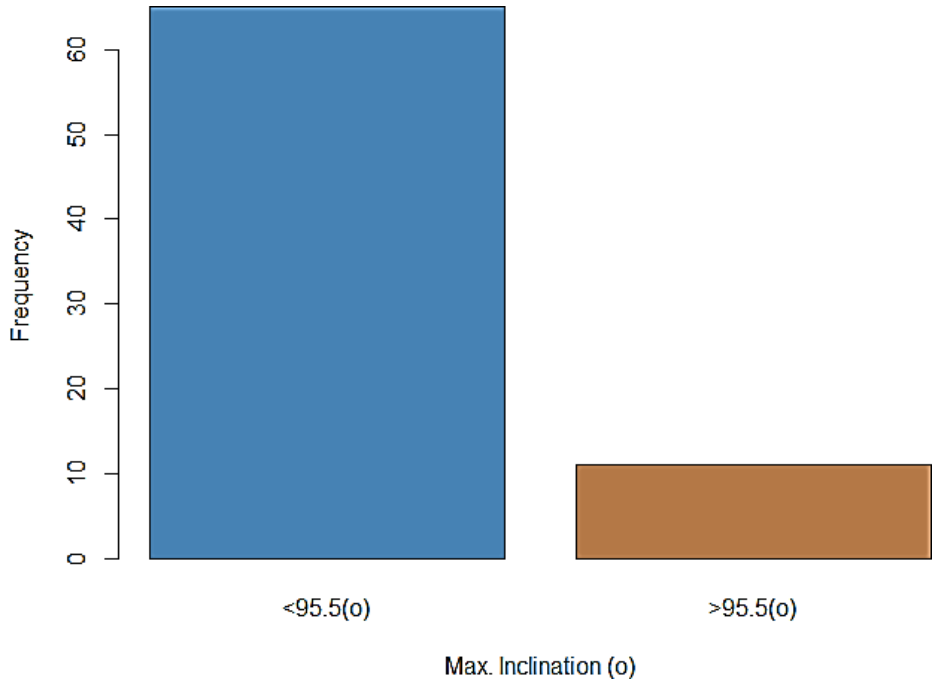
(h)


Frequency
0
10
20
30
40
50
60
<75(days)
>75(days)
Drilling-Fracturing Time Interval (Days)

(j)


Frequency
0
10
20
30
40
50
60
<95800(lbf)
>95800(lbf)
Dogleg Bending Stress (lbf)

(k)



(l)

Figure B-1. Graphical Representation of Exploratory Data Analytics (EDA) of Data Set.

APPENDIX C

In essence, parametric survival curve estimators assume that a survival function follows a parametric distribution. The distribution might be exponential; corresponding to constant survival/hazard, or what is known as Weibull distribution; corresponding to varying survival/hazard. Regardless of the distribution, it is essential that we calculate the unknown parameter(s). The theory of "*maximum likelihood estimation*" provides one way of doing that.

Maximum Likelihood Estimation

We start the discussion by considering a simple example; assuming that the survival function follows an exponential distribution. The likelihood is calculated, then, by taking a product of terms from the exponential distribution, one for each observation. In case there is no censoring, the likelihood function takes the general form

$$L(\lambda;\, d_1,\, d_2,\, \dots\,,\, d_n) = f(d_1,\, \lambda) \cdot f(d_2,\, \lambda) \cdot \cdots \cdot f(d_n,\, \lambda) = \prod f(d_i,\, \lambda) \qquad \textbf{(B - 1)}$$

If some observations are censored, we have to make an adjustment to this expression. For an observation of an observed failure, we put in the probability density function (PDF). as above. But for a right-censored observation, we put in the survival function, indicating that observation is known only to exceed a particular value. The likelihood in general then takes the form

$$L(\lambda;\, d_1,\, d_2,\, \dots\,,\, d_n) = \prod_{i=1}^{n} f(d_i,\, \lambda)^{\delta_i} \cdot S(d_i,\, \lambda)^{1-\delta_i} = \prod_{i=1}^{n} h(d_i,\, \lambda)^{\delta_i} \cdot S(d_i,\, \lambda) \qquad \textbf{(B -2)}$$

Censoring is basically a way of differentiating between individuals (drilling/producing wells) that experienced the outcome (casing failure) from those that did not. For an arbitrary well, d_i, that experienced a casing failure, the censoring indicator takes the value $\delta_i = 1$, In this case, we use a probability density function (PDF) for our calculations. When a well, d_i, is a censored observation, we have $\delta_i = 0$ we enter a survival factor. Alternatively, we may enter a hazard factor for each censored observation and a survival factor for every observation, censored or not.

For the *exponential* distribution, the general expression for the likelihood function can be simplified as follows:

$$L(\lambda; d_1, d_2, \dots, d_n) = \prod_{i=1}^{n} [\lambda e^{-d_i/\mu}]^{\delta_i} \cdot [e^{-\lambda d_i}]^{1-\delta_i} = \lambda^d e^{-\lambda V} \quad \textbf{(B - 3)}$$

where $d = \sum n\ \delta i$ is the total number of failures and $V = \sum n\ di$ is the total amount of depths i the data set.

After the definition of the likelihood function that constitutes the survival function, it is important that we find the value of λ that maximizes this function. It is for that reason it is also known as the *maximum likelihood estimate*.

Using the logarithmic transformation simplify the likelihood function; by converting it into a sum, a *log-likelihood* function can be defined as

$$l(\lambda) = d \log \lambda - \lambda V \quad \textbf{(B - 4)}$$

Since the log transformation is *monotonic*, it only has one minima and one maxima. This, in turn, means that the value of λ that maximizes the log-likelihood also maximizes the original likelihood function.

We use standard calculus to find the *first* derivative, also called the *score function*,

$$l`(\lambda) = \frac{d}{\lambda} - V \quad \textbf{(B - 5)}$$

which we set equal to zero to obtain the maximum likelihood estimate, $\hat{\lambda} = d/V$.

APPENDIX D

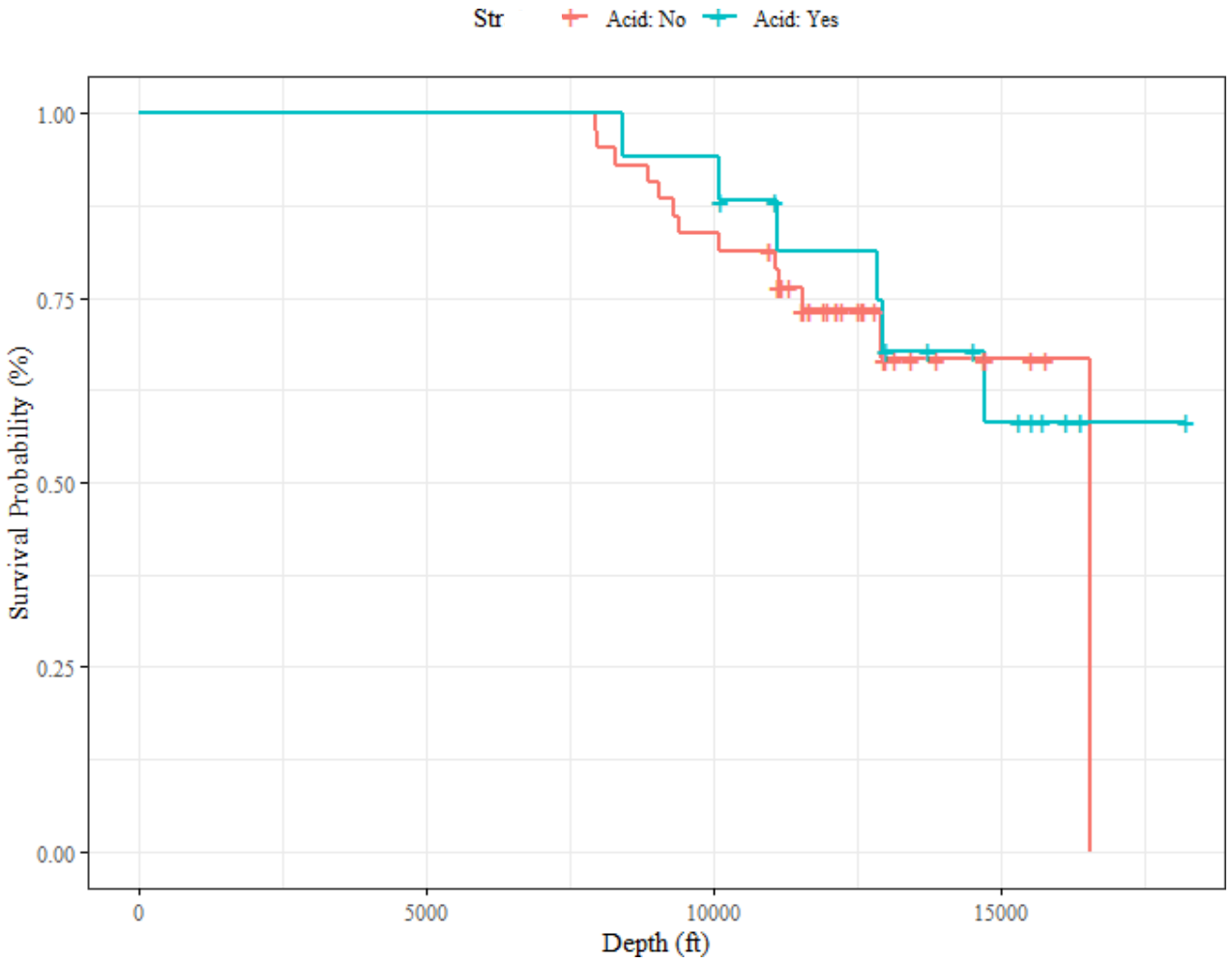


Figure D-1. Graphical representation of survival curves based on experience of acid treatment. Survival curve corresponding to "not having acid treatment" is highlighted in red. Survival curve corresponding to "having acid treatment" is highlighted in green.

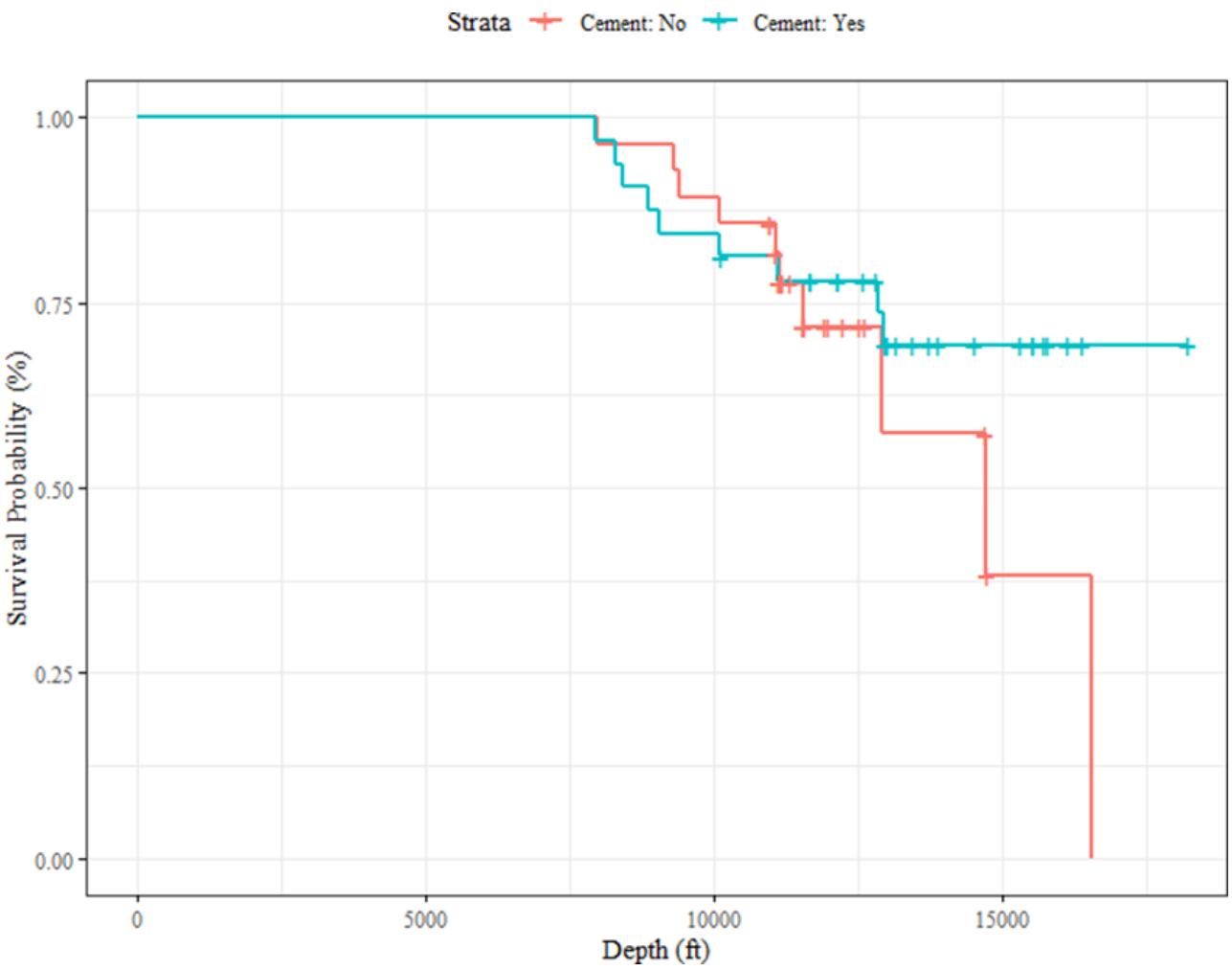


Figure D-2. Graphical representation of survival curves based on experience of cementing. Survival curve corresponding to "not having cement" is highlighted in red. Survival curve corresponding to "having cement" is highlighted in green.

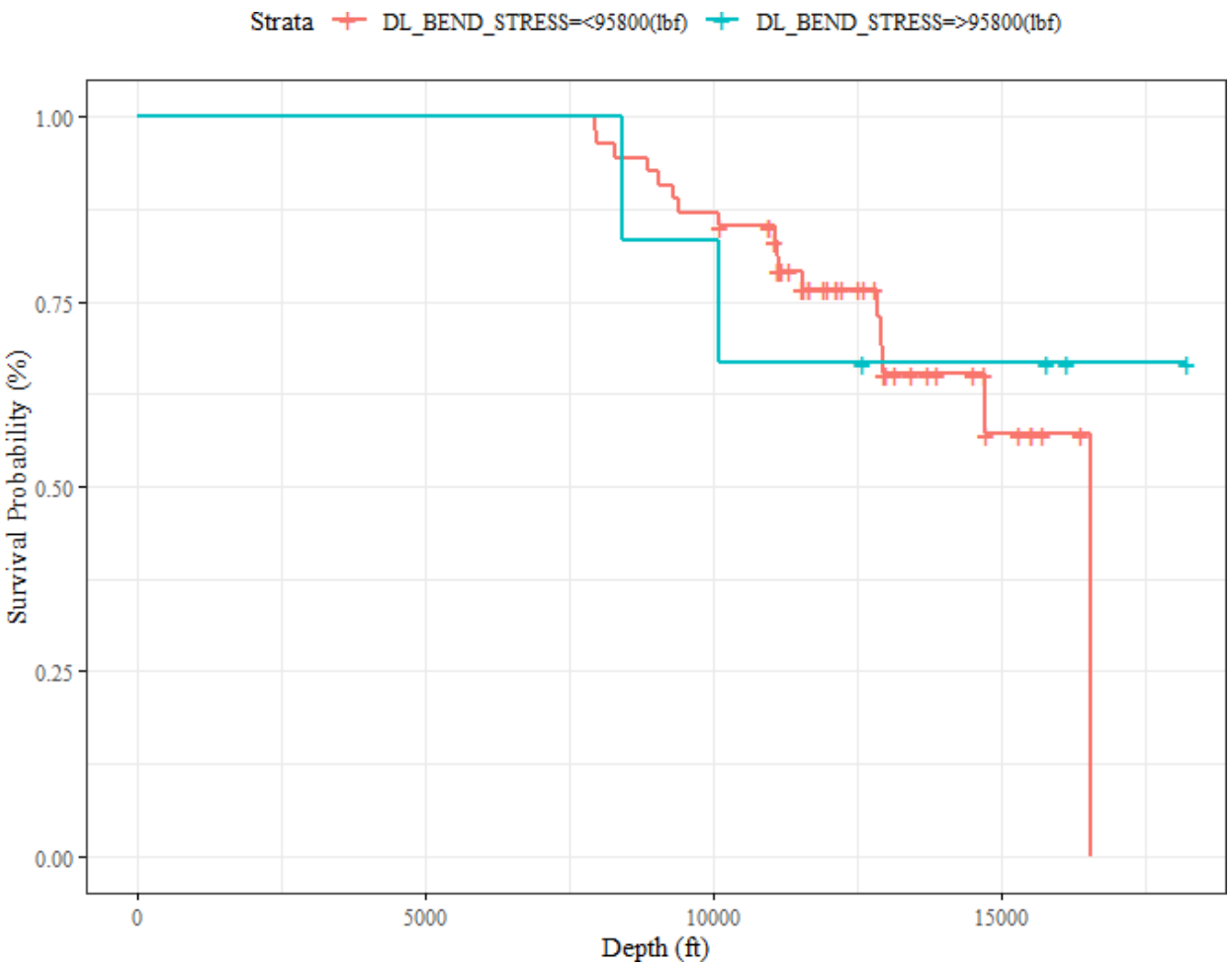


Figure D-3. Graphical representation of survival curves based on change of dogleg bending stress. Survival curve corresponding to "dogleg bending stress < 95.8k lb./ft^3" is highlighted in red. Survival curve corresponding to "dogleg bending stress > 95.8k lb./ft^3" is highlighted in green.

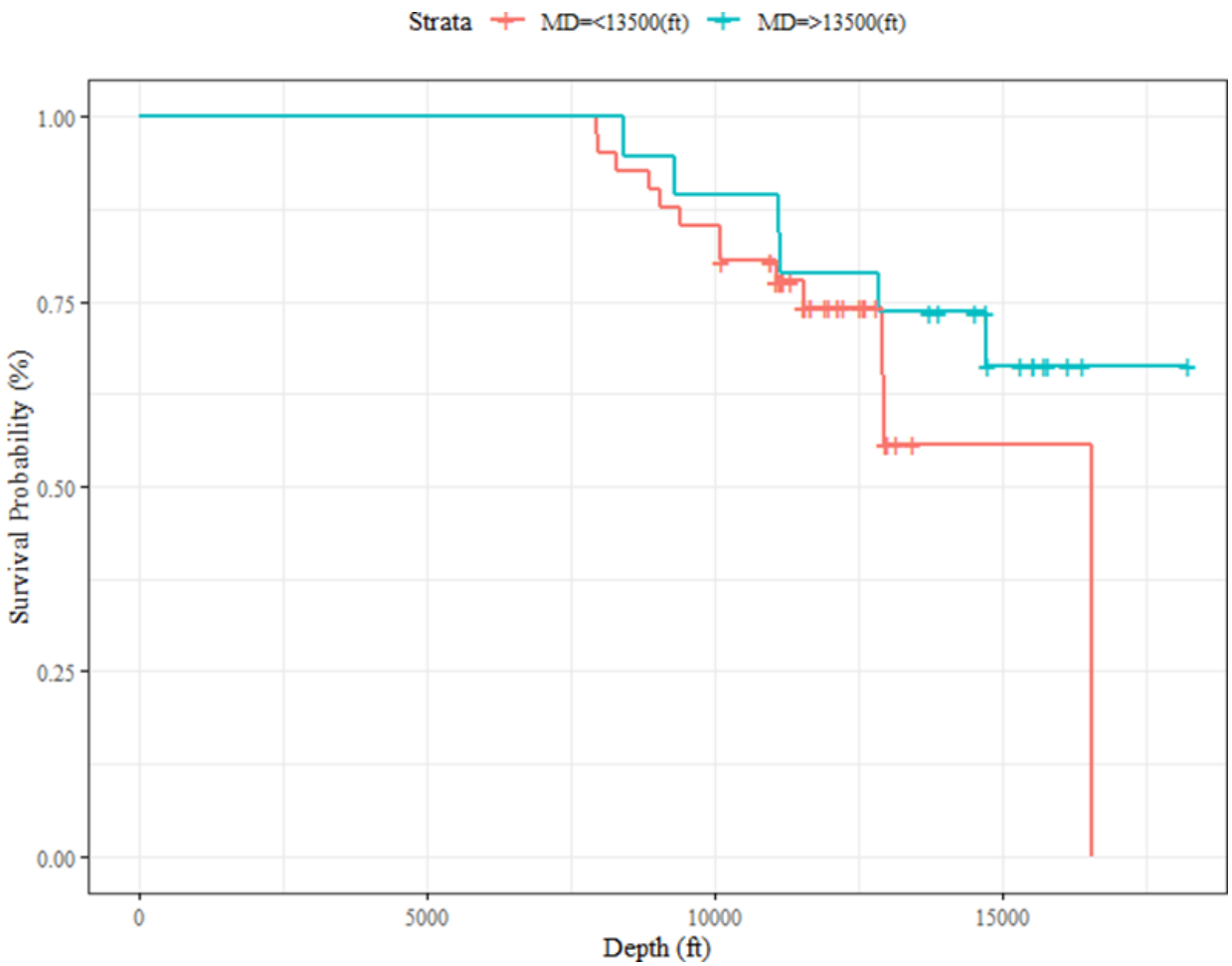


Figure D-4. Graphical representation of survival curves based on change of total measured depth. Survival curve corresponding to "measured depth < 13.5k ft" is highlighted in red. Survival curve corresponding to "measured depth > 13.5k ft" is highlighted in green.

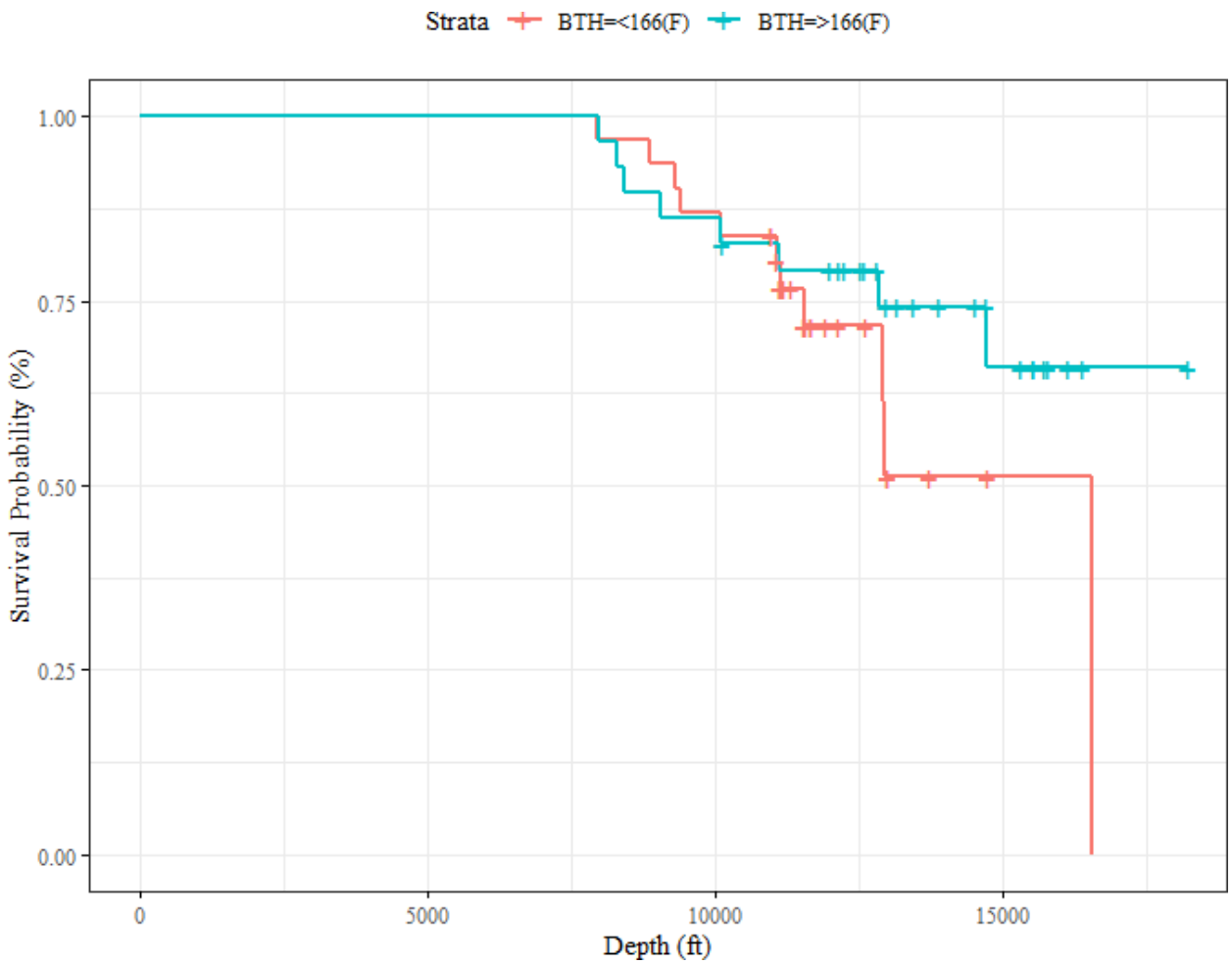


Figure D-5. Graphical representation of survival curves based on change of bottomhole temperature. Survival curve corresponding to "bottomhole temperature < 166 F" is highlighted in red. Survival curve corresponding to "bottomhole temperature > 166 F" is highlighted in green.

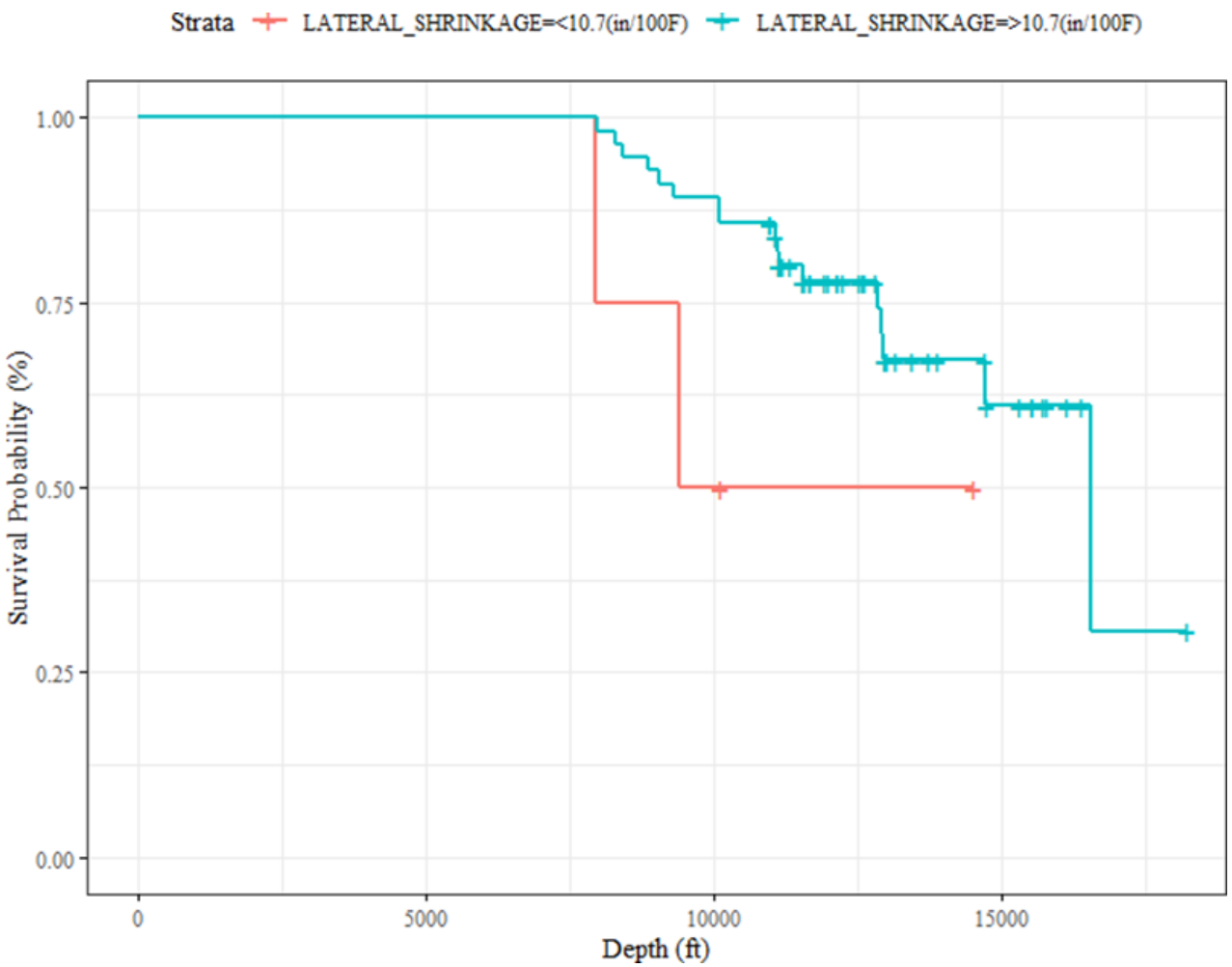


Figure D-6. Graphical representation of survival curves based on change of lateral section shrinkage. Survival curve corresponding to "lateral shrinkage < 10.7 in/100ft" is highlighted in red. Survival curve corresponding to "lateral shrinkage > 10.7 in/100ft" is highlighted in green.

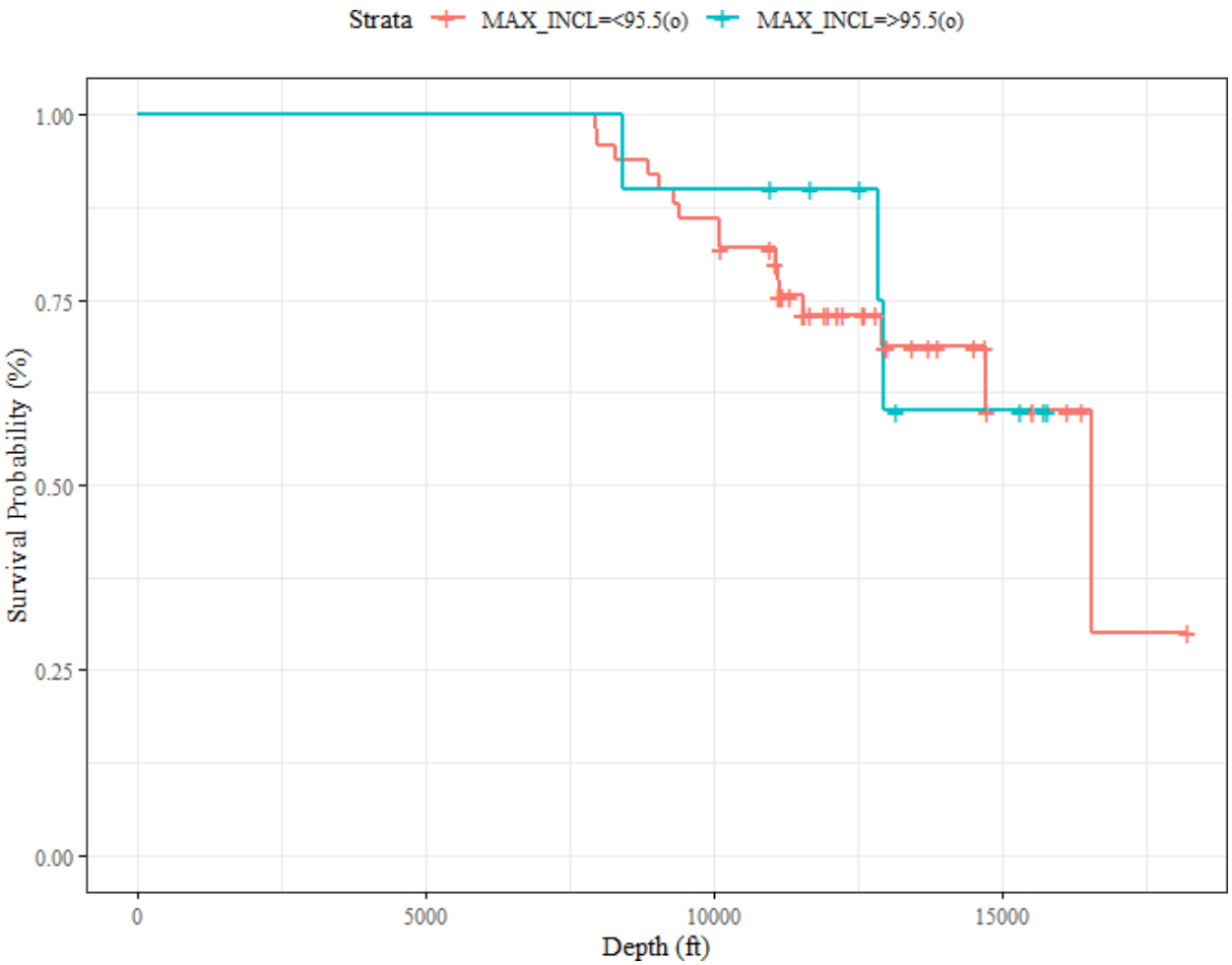


Figure D-7. Graphical representation of survival curves based on change of maximum inclination. Survival curve corresponding to "max. inclination < 95.5 º" is highlighted in red. Survival curve corresponding to "max. inclination > 95.5 º" is highlighted in green.

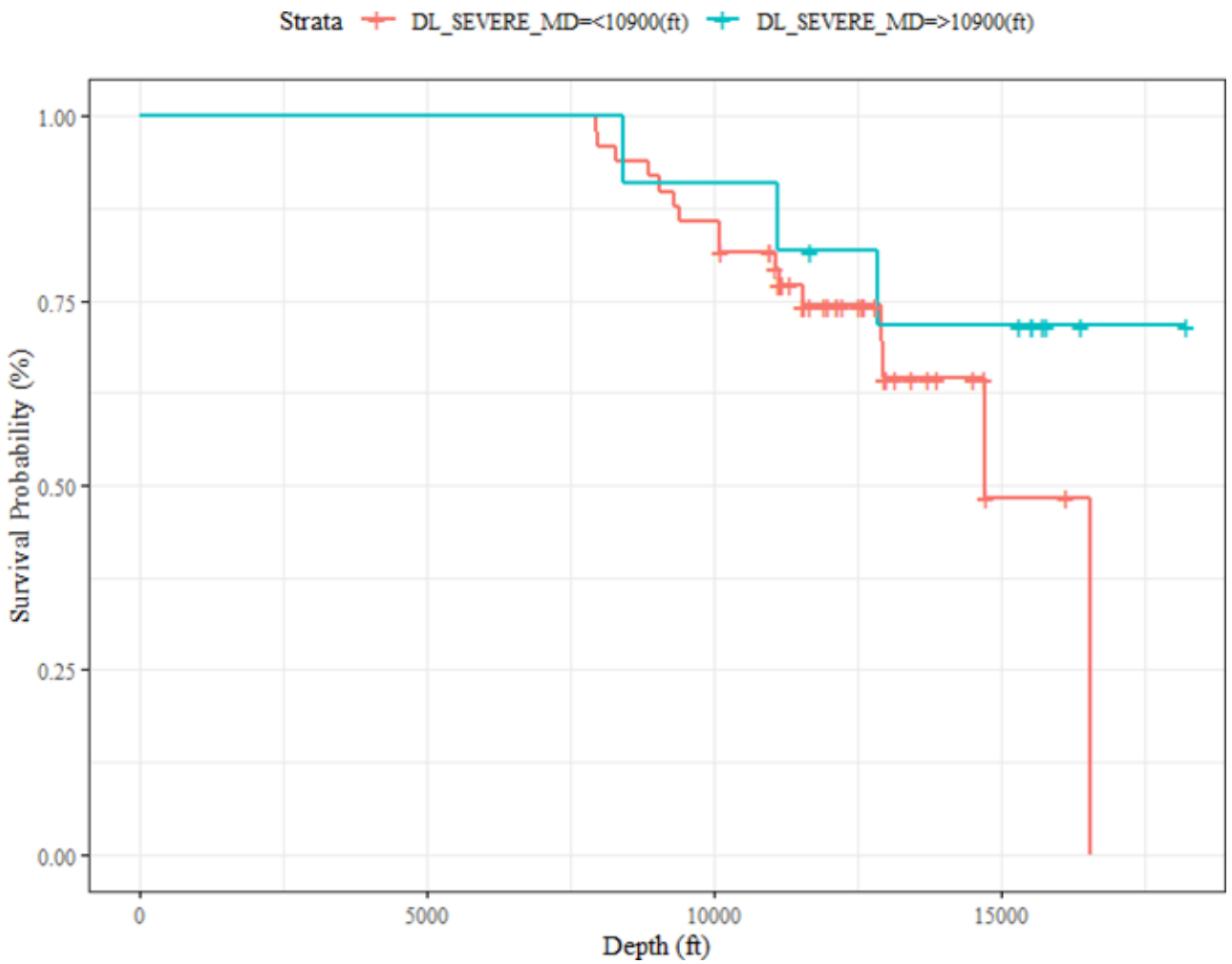


Figure D-8. Graphical representation of survival curves based on change of MD of max. DL severity. Survival curve corresponding to "MD of max. DL severity < 10.9k ft" is highlighted in red. Survival curve corresponding to "MD of max. DL severity > 10.9k ft" is highlighted in green.

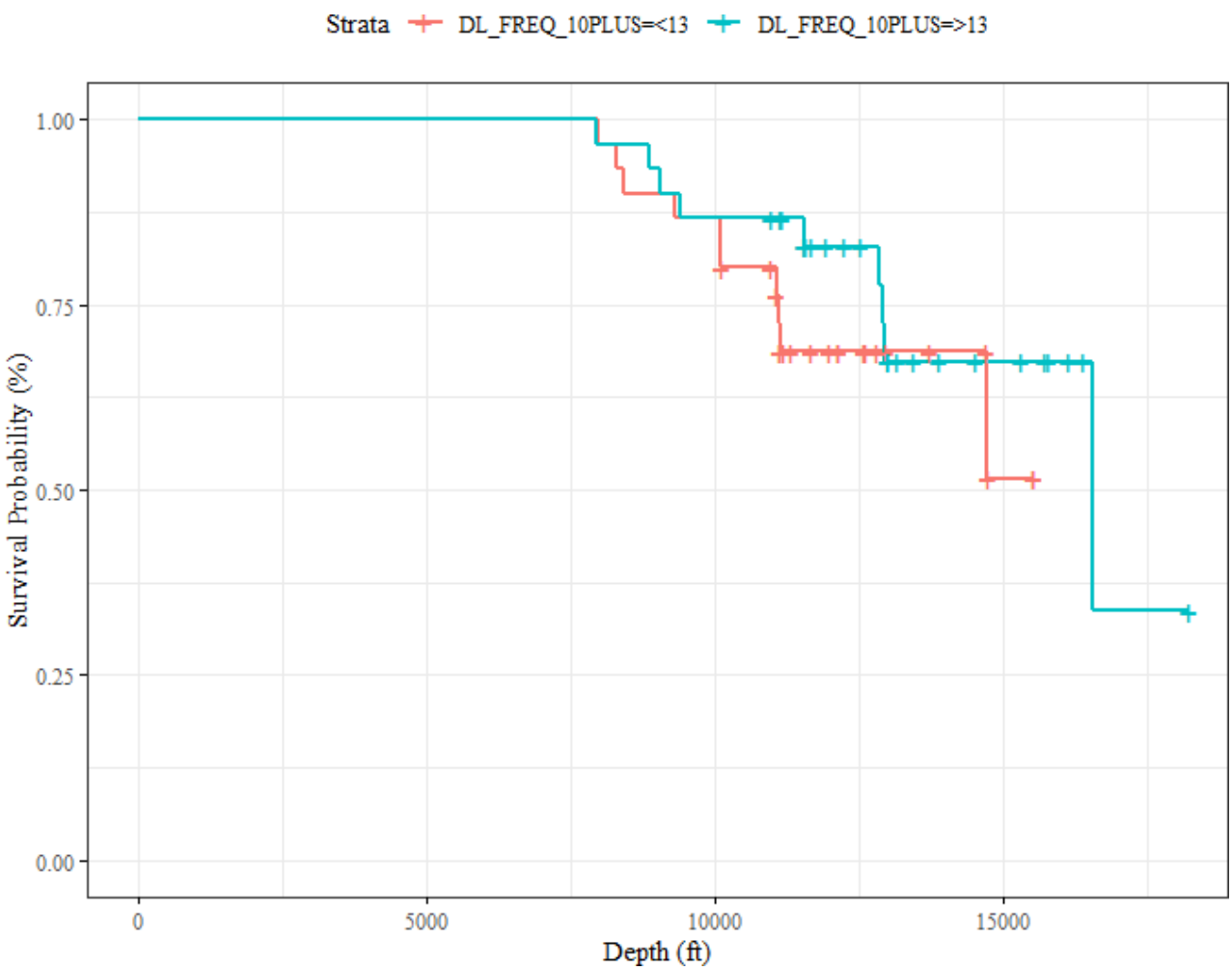


Figure D-9. Graphical representation of survival curves based on freq. of DL severity (>10 °/100ft). Survival curve corresponding to "freq. of DL severity (>10 °/100ft) < 13" is highlighted in red. Survival curve corresponding to "freq. of DL severity (>10 °/100ft) > 13" is highlighted in green.

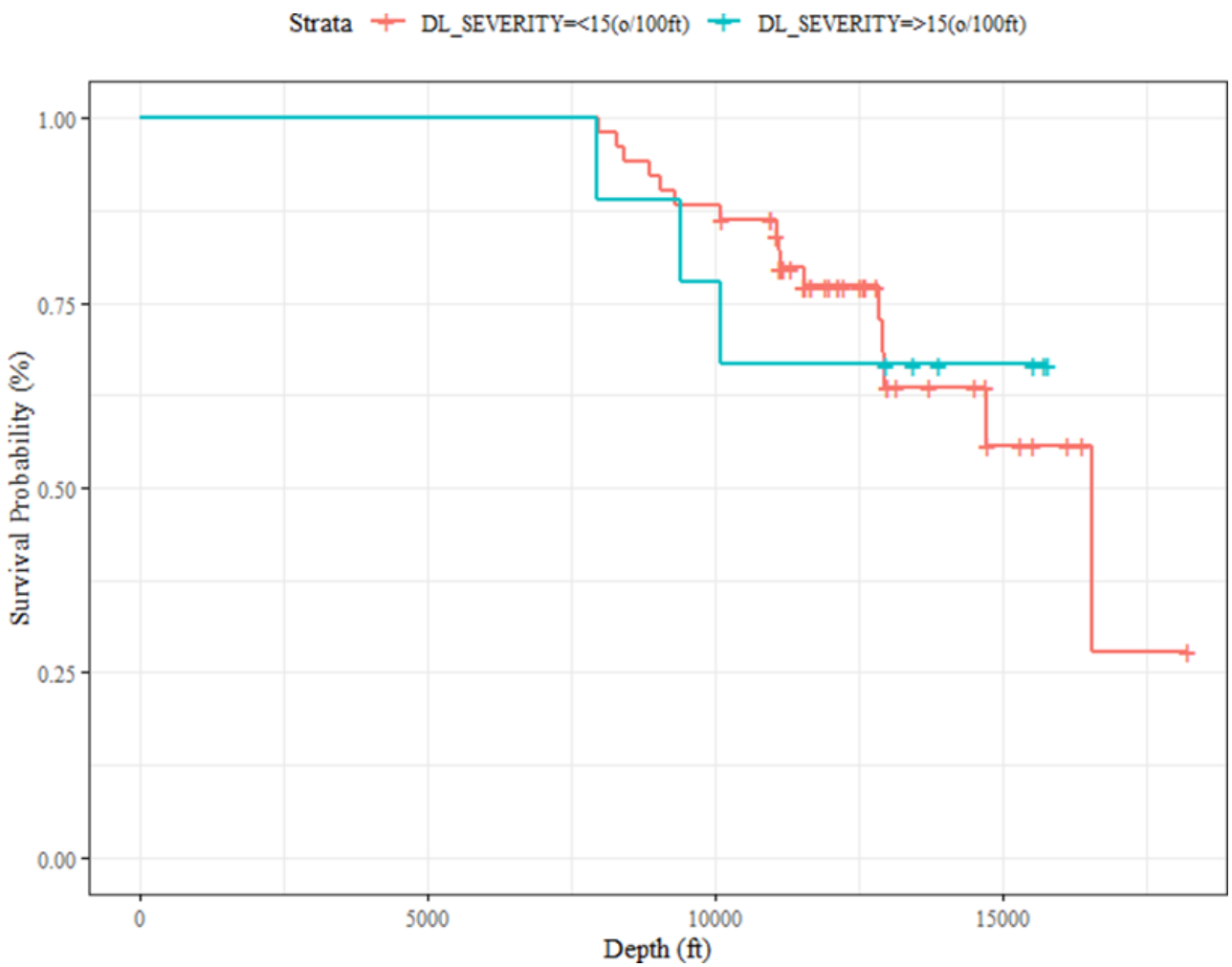


Figure D-10. Graphical representation of survival curves based on change of dogleg severity. Survival curve corresponding to "DL severity < 15 °/100ft" is highlighted in red. Survival curve corresponding to "DL severity > 15 °/100ft" is highlighted in green.

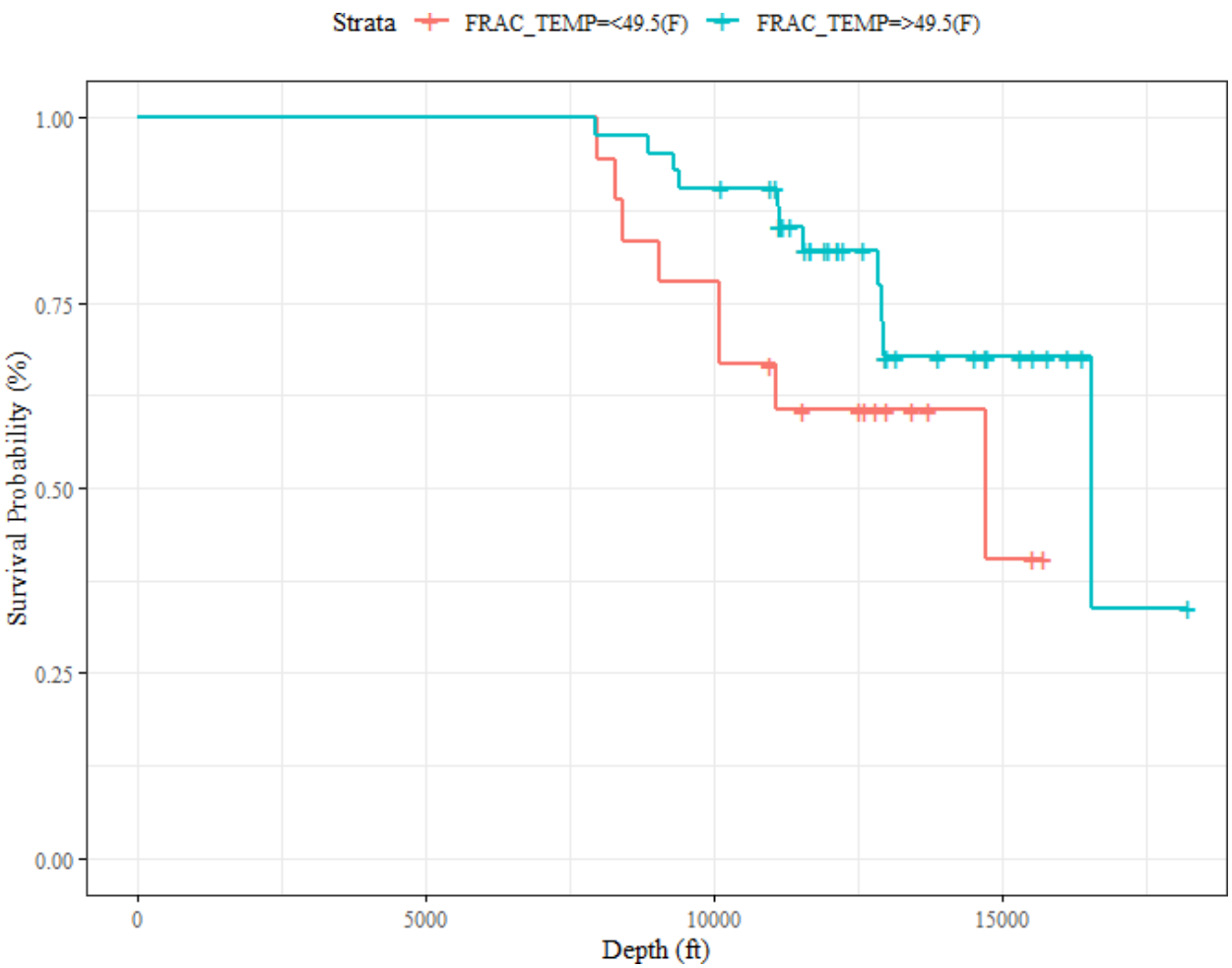


Figure D-11. Graphical representation of survival curves based on change of fracture temperature. Survival curve corresponding to "fracture temperature < 49.5 F" is highlighted in red. Survival curve corresponding to "fracture temperature > 49.5 F" is highlighted in green.

APPENDIX E

In this section, we present the full results obtained from the application of semi-quantitative risk assessment techniques (Chapter 4); namely, probability-impact risk assessment matrices (PI-RAMs) on the potential risk factors identified through the implementation of risk analysis association measurements.

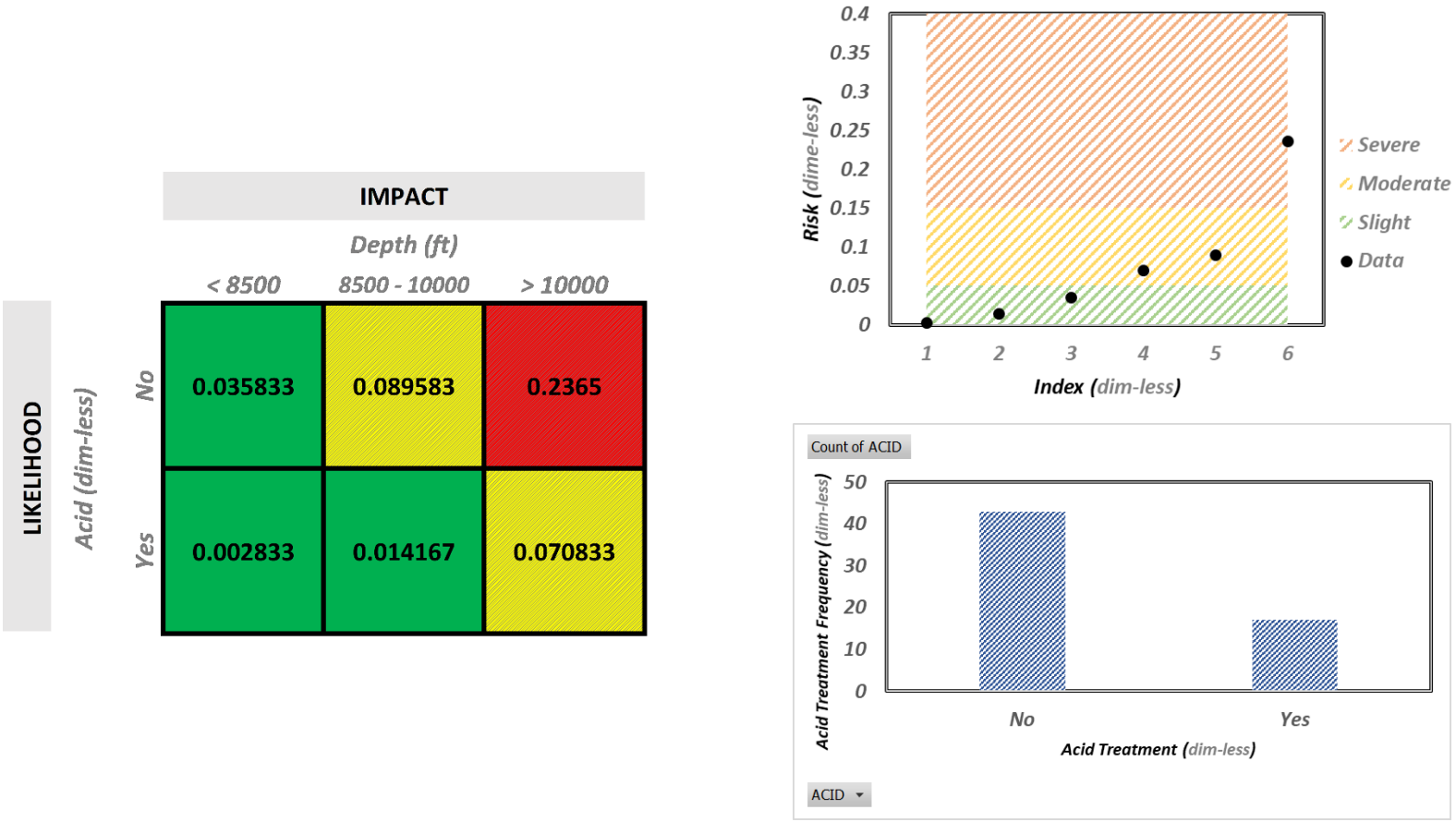


Figure E-1. (left) Risk Assessment Results for "Acid" Risk Factor. (Top-right) Calculated Risk Values for Different Variations of "Acid" Risk Factor. (Bottom-right) Summery of Frequency Analysis Used for Determining Likelihood of Various Combinations of "Acid" Risk Factor.

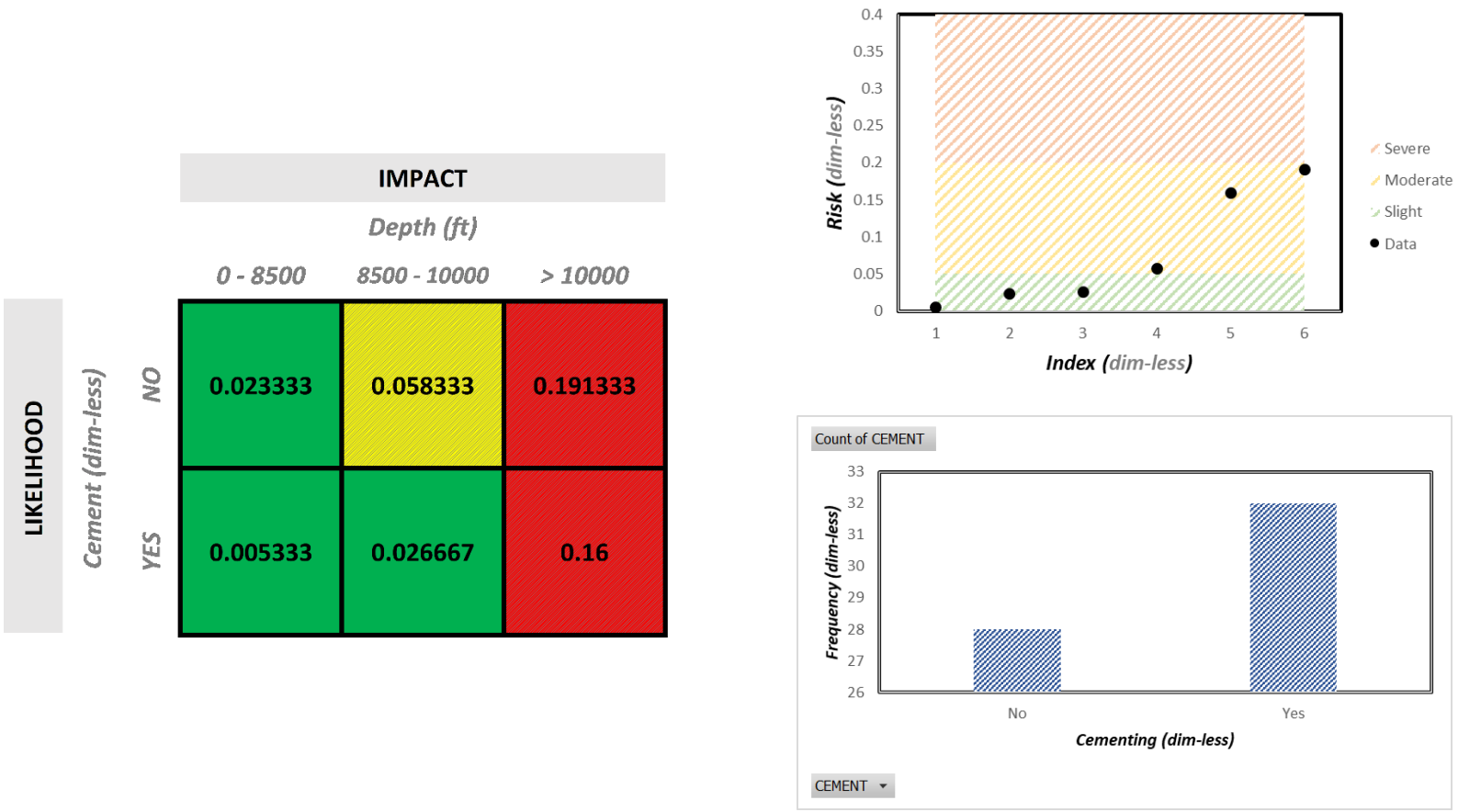


Figure E-2. (left) Risk Assessment Results developed in this study for "Cement" Risk Factor. (Top-right) Calculated Risk Values for Different Variations of "Cement" Risk Factor. (Bottom-right) Summery of Frequency Analysis Used for Determining Likelihood of Various Combinations of "Cement" Risk Factor.

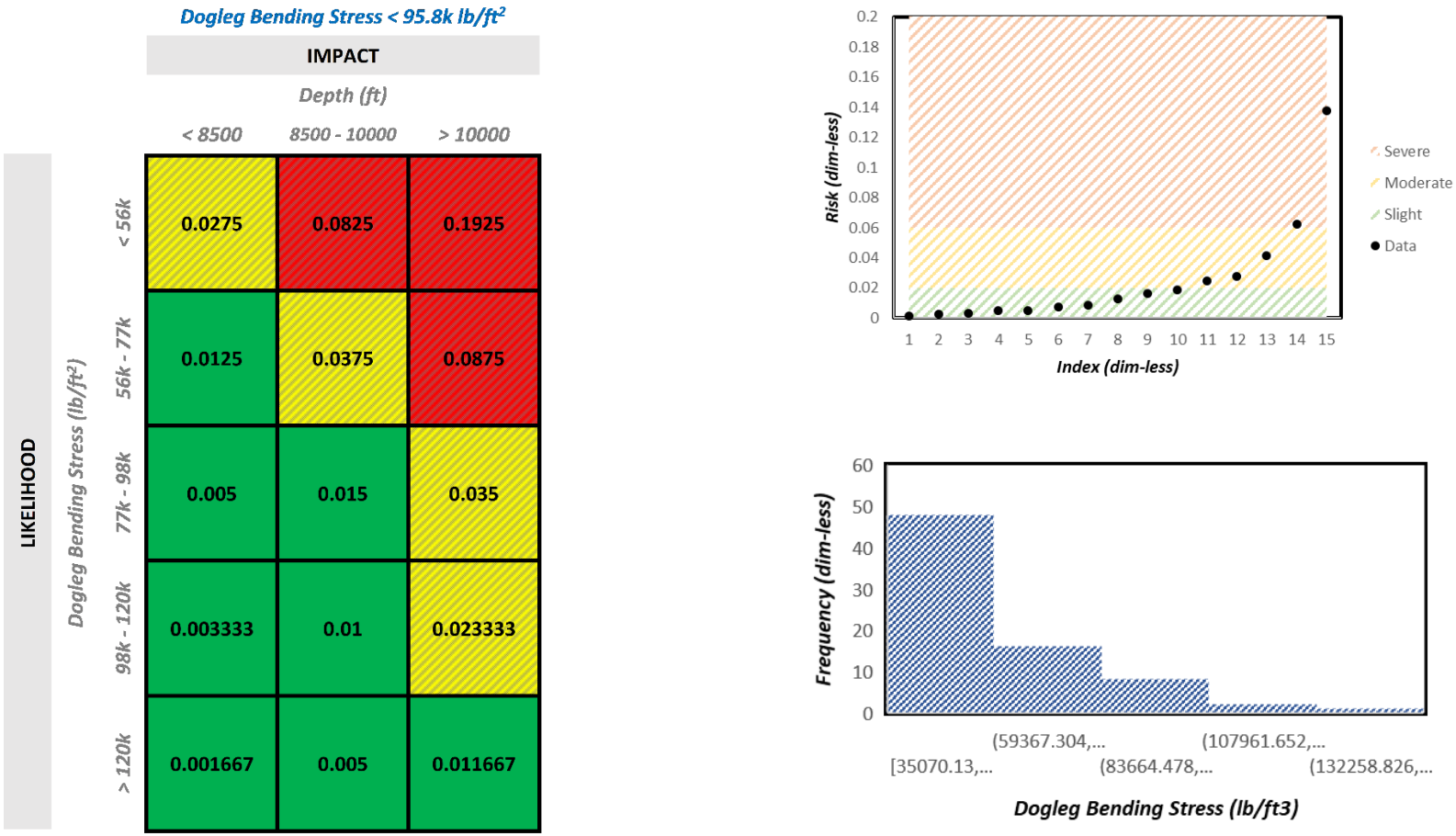


Figure E-3. (left) Risk Assessment Results for "Dogleg Bending Stress < 95.8" Risk Factor. (Top-right) Calculated Risk Values for Different Variations of "Dogleg Bending Stress < 95.8" Risk Factor. (Bottom-right) Summery of Frequency Analysis Used for Determining Likelihood of Various Combinations of "Dogleg Bending Stress < 95.8" Risk Factor.

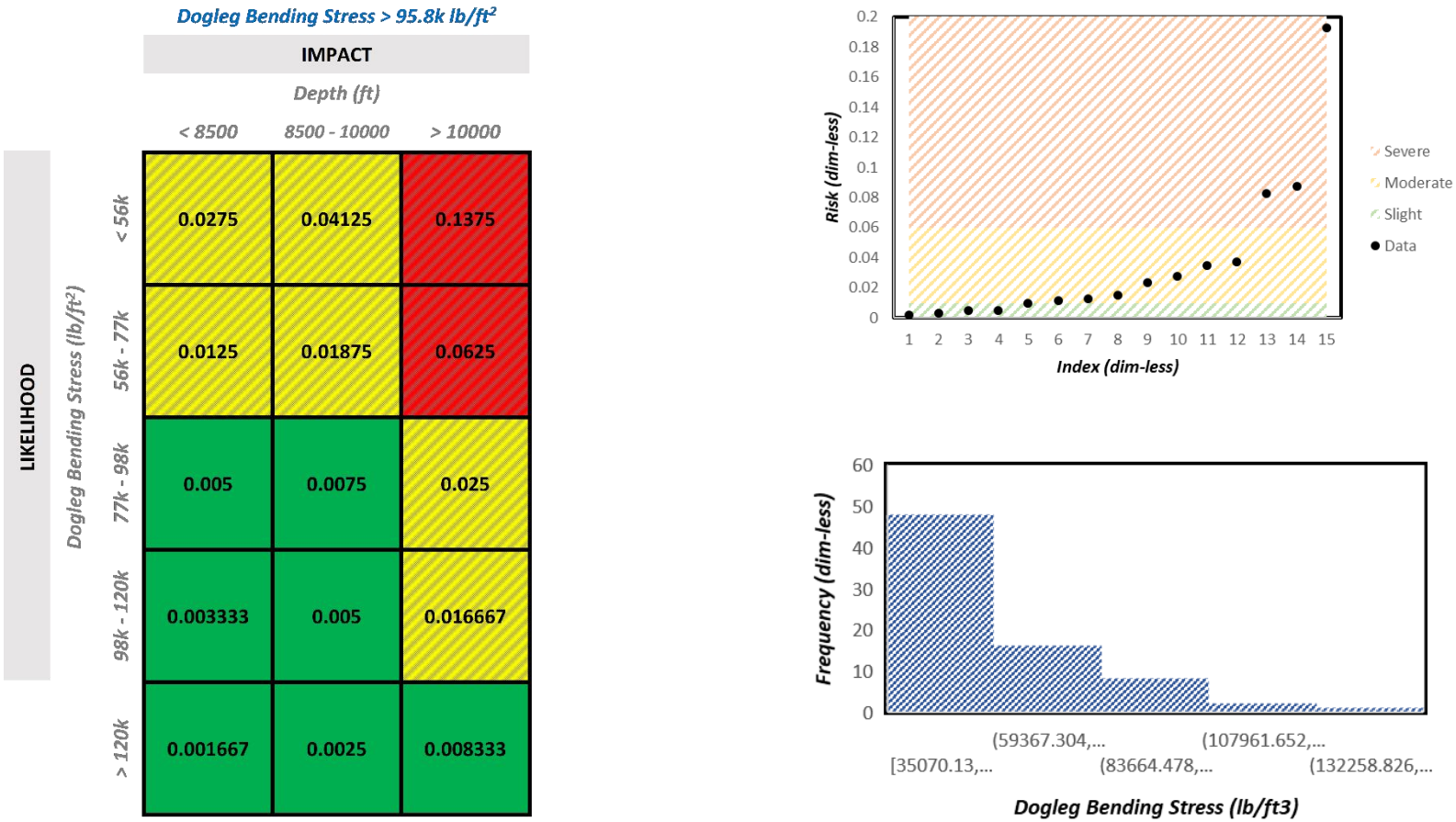


Figure E-4. (left) Risk Assessment Results developed in this study for "Dogleg Bending Stress > 95.8" Risk Factor. (Top-right) Calculated Risk Values for Different Variations of "Dogleg Bending Stress > 95.8" Risk Factor. (Bottom-right) Summery of Frequency Analysis Used for Determining Likelihood of Various Combinations of "Dogleg Bending Stress > 95.8" Risk Factor.

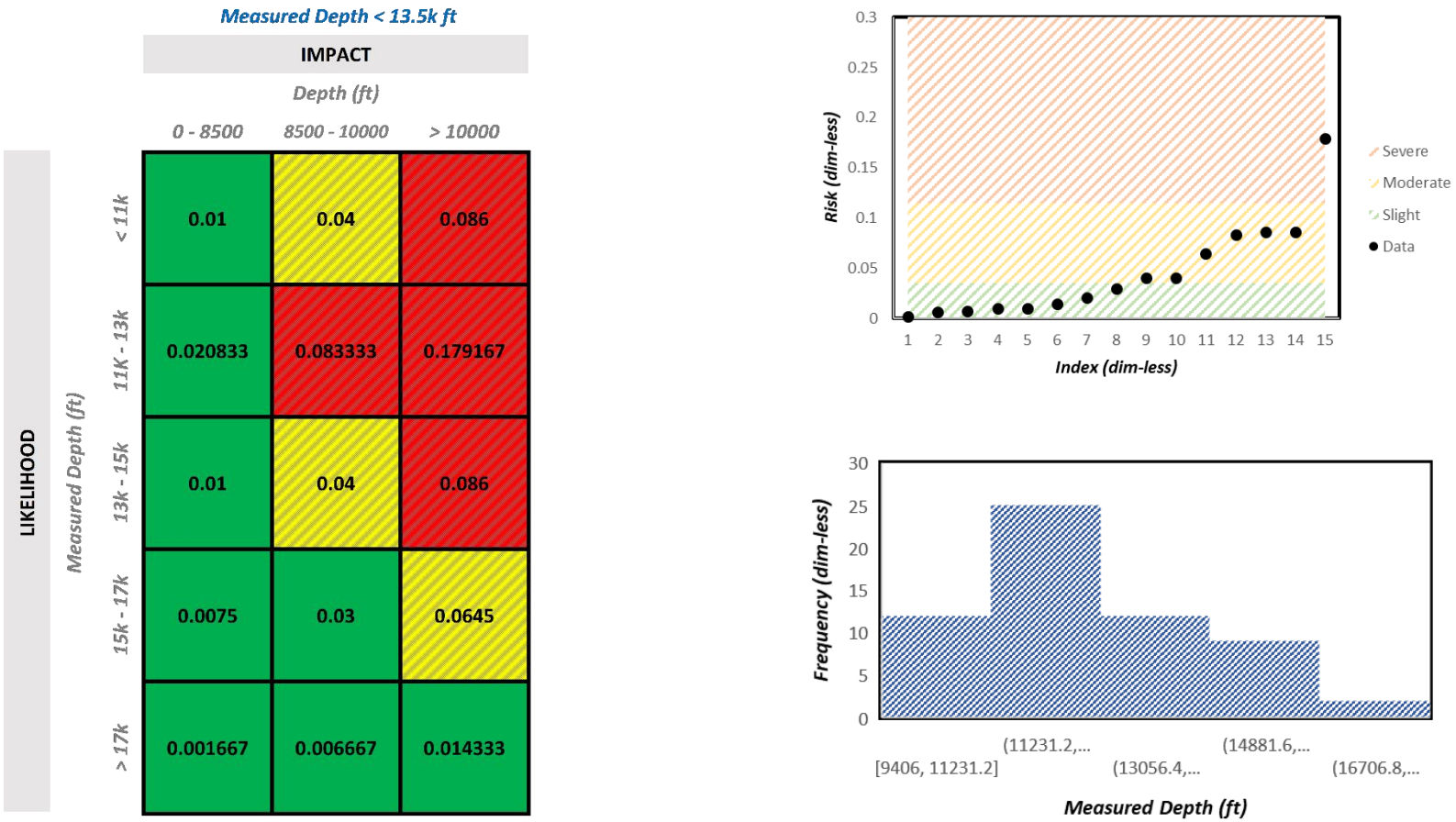


Figure E-5. (left) Risk Assessment Results for "Measured Depth < 13.5k" Risk Factor. (Top-right) Calculated Risk Values for Different Variations of "Measured Depth < 13.5k" Risk Factor. (Bottom-right) Summery of Frequency Analysis Used for Determining Likelihood of Various Combinations of "Measured Depth < 13.5k" Risk Factor.

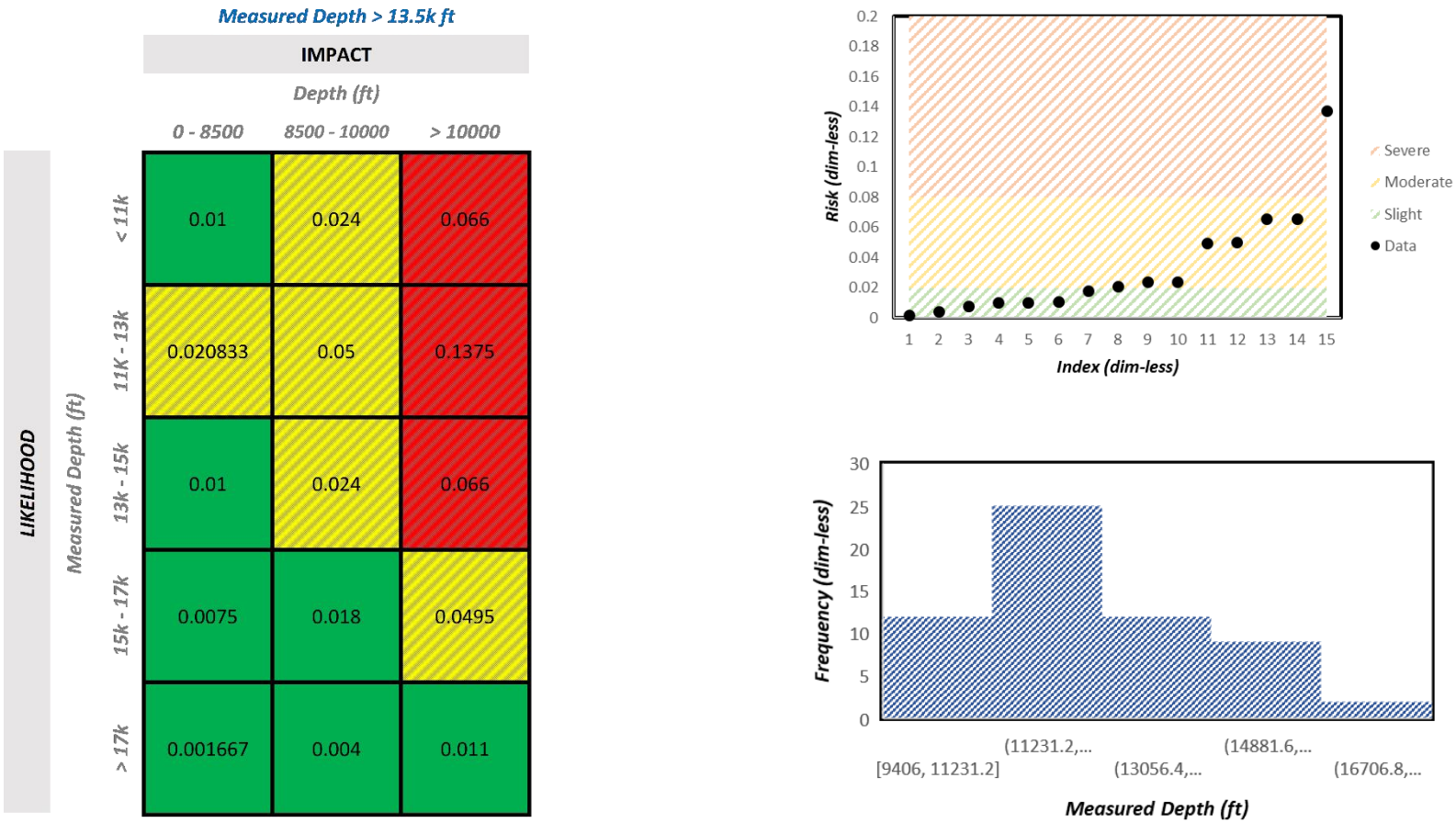


Figure E-6. (left) Risk Assessment Results for "Measured Depth > 13.5k" Risk Factor. (Top-right) Calculated Risk Values for Different Variations of "Measured Depth > 13.5k" Risk Factor. (Bottom-right) Summery of Frequency Analysis Used for Determining Likelihood of Various Combinations of "Measured Depth > 13.5k" Risk Factor.

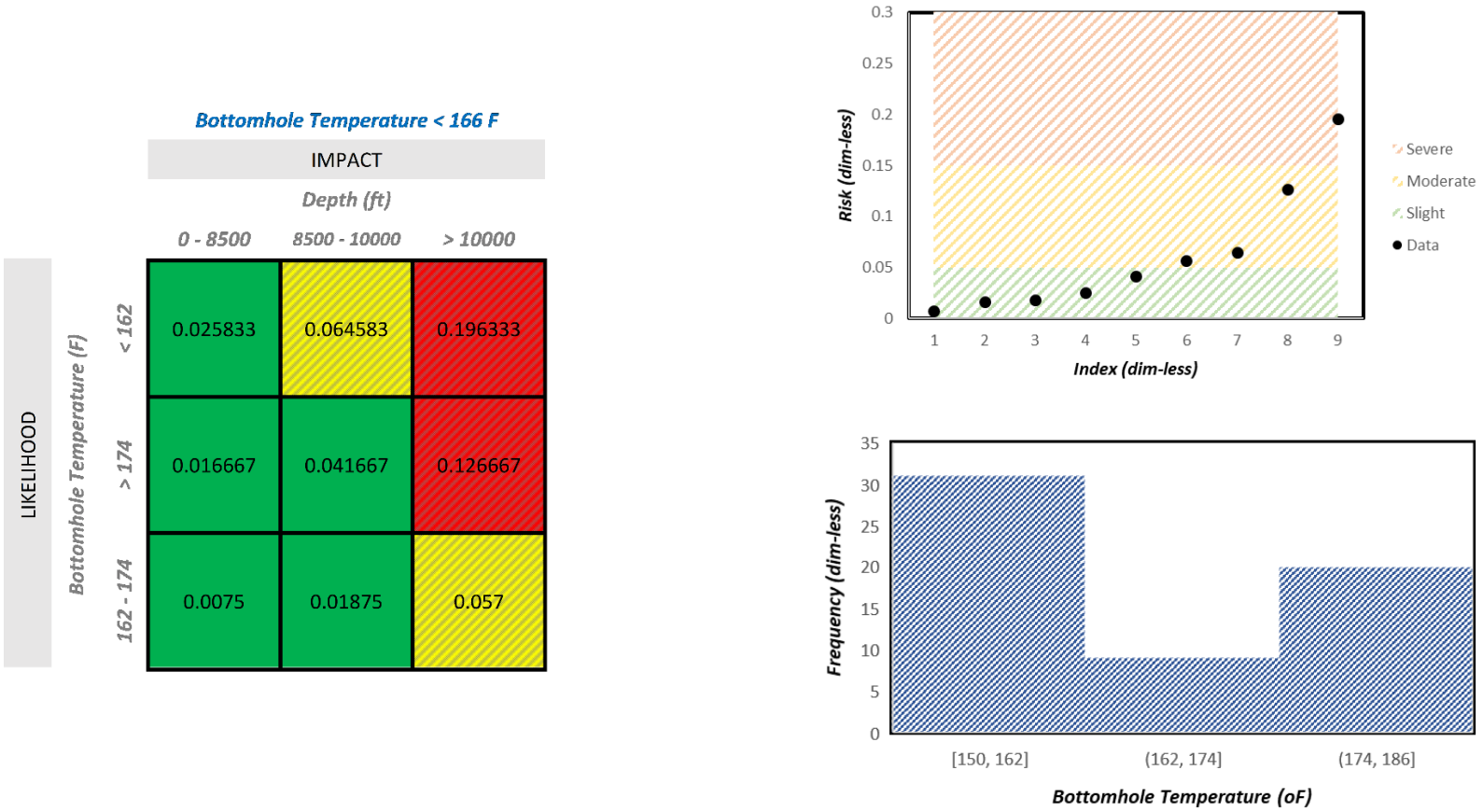


Figure E-7. (left) Risk Assessment Results for "Bottomhole Temperature < 166F" Risk Factor. (Top-right) Calculated Risk Values for Different Variations of "Bottomhole Temperature < 166F" Risk Factor. (Bottom-right) Summery of Frequency Analysis Used for Determining Likelihood of Various Combinations of "Bottomhole Temperature < 166F" Risk Factor.

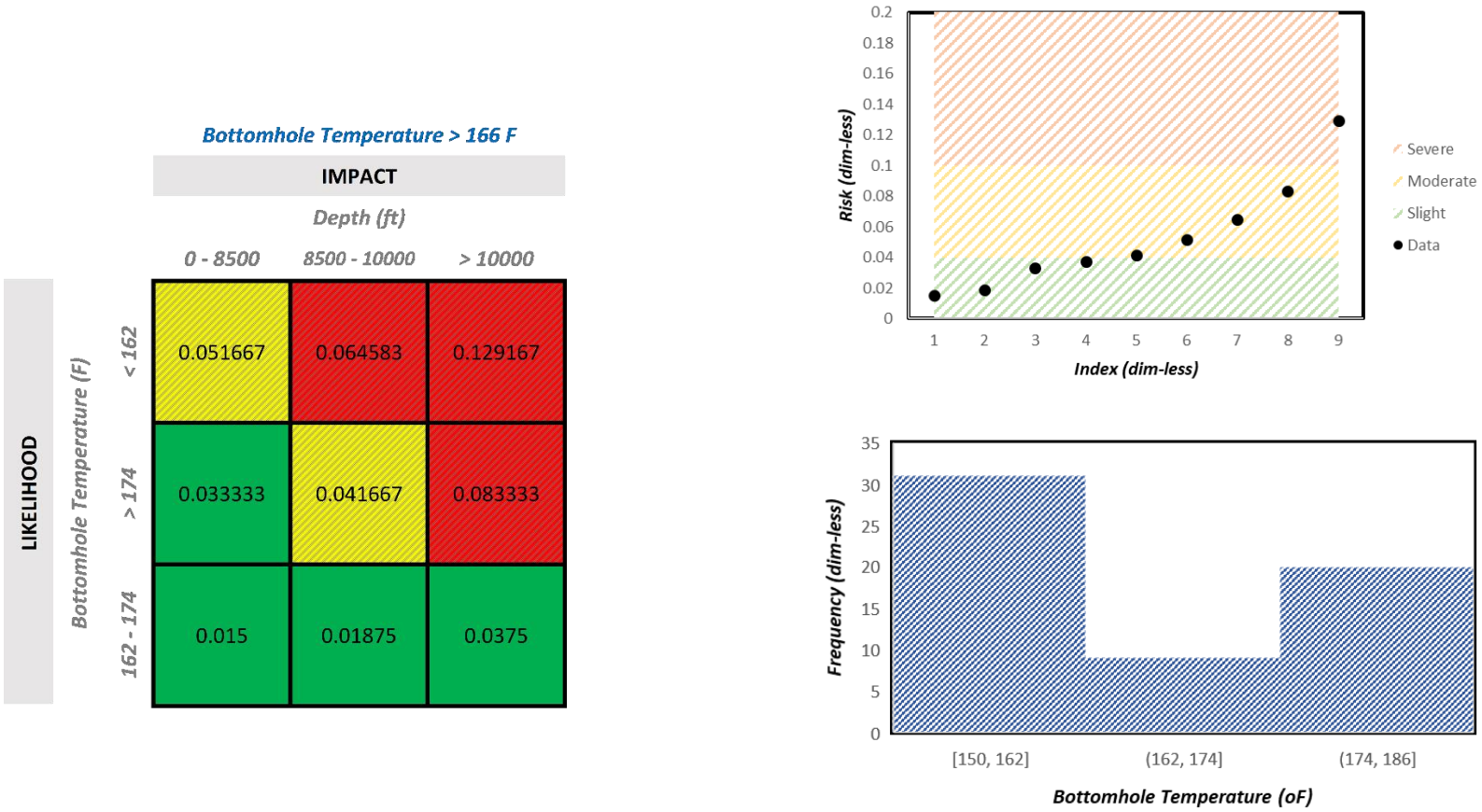


Figure E-8. (left) Risk Assessment Results for "Bottomhole Temperature > 166F" Risk Factor. (Top-right) Calculated Risk Values for Different Variations of "Bottomhole Temperature > 166F" Risk Factor. (Bottom-right) Summery of Frequency Analysis Used for Determining Likelihood of Various Combinations of "Bottomhole Temperature > 166F" Risk Factor.

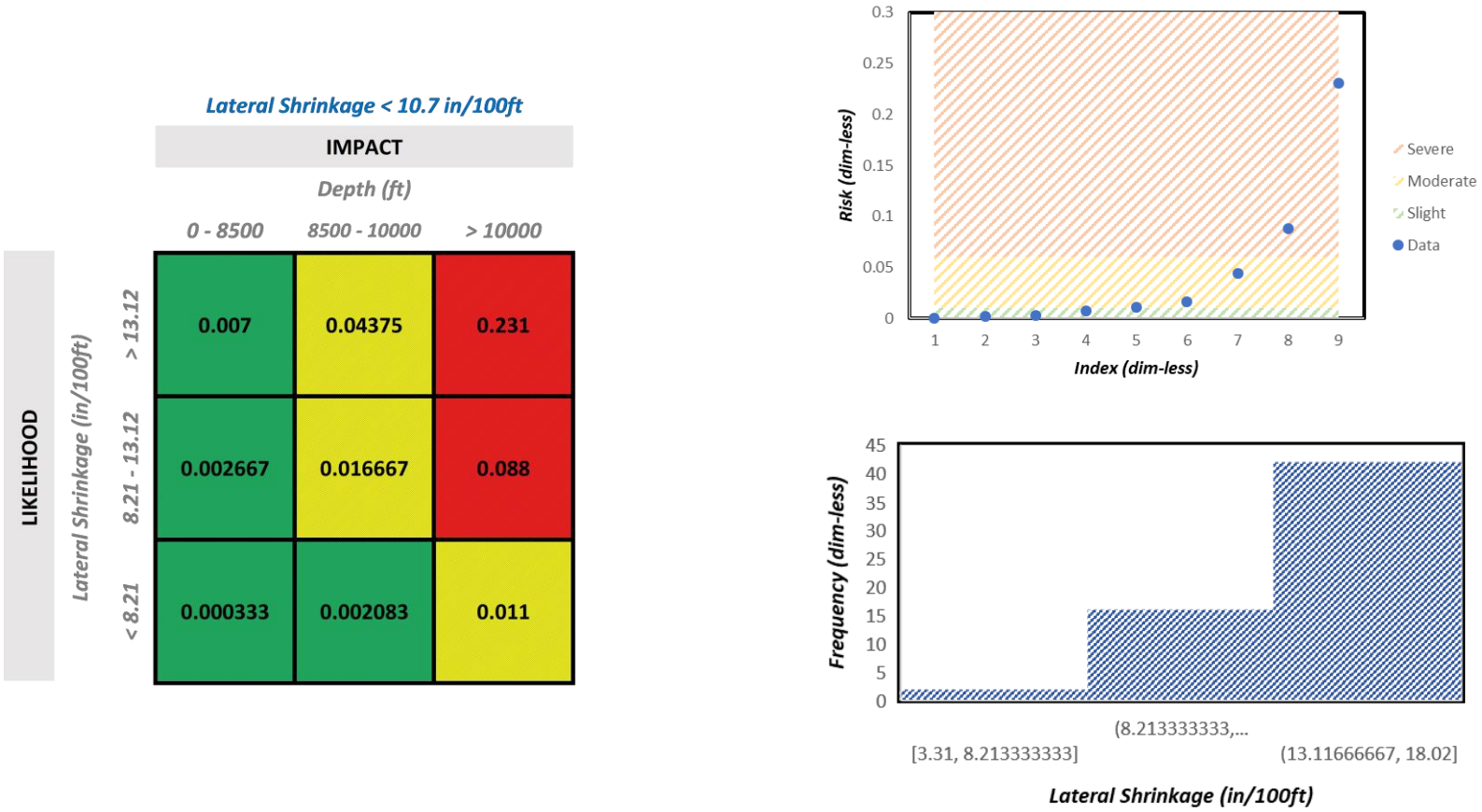


Figure E-9. (left) Risk Assessment Results for "Lateral Shrinkage < 10.7" Risk Factor. (Top-right) Calculated Risk Values for Different Variations of "Lateral Shrinkage < 10.7" Risk Factor. (Bottom-right) Summery of Frequency Analysis Used for Determining Likelihood of Various Combinations of "Lateral Shrinkage < 10.7" Risk Factor.

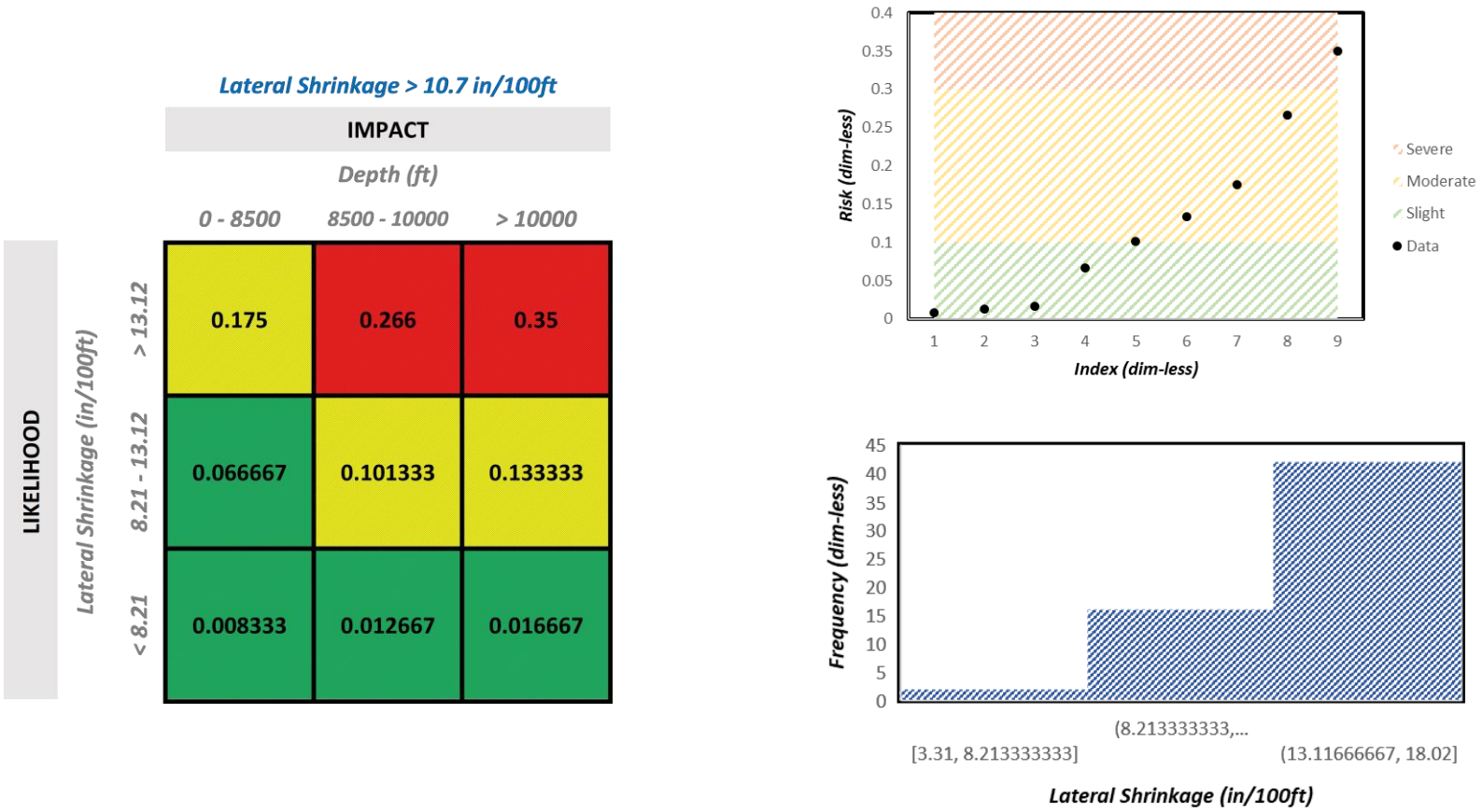


Figure E-10. (left) Risk Assessment Results for "Lateral Shrinkage > 10.7" Risk Factor. (Top-right) Calculated Risk Values for Different Variations of "Lateral Shrinkage > 10.7" Risk Factor. (Bottom-right) Summery of Frequency Analysis Used for Determining Likelihood of Various Combinations of "Lateral Shrinkage > 10.7" Risk Factor.

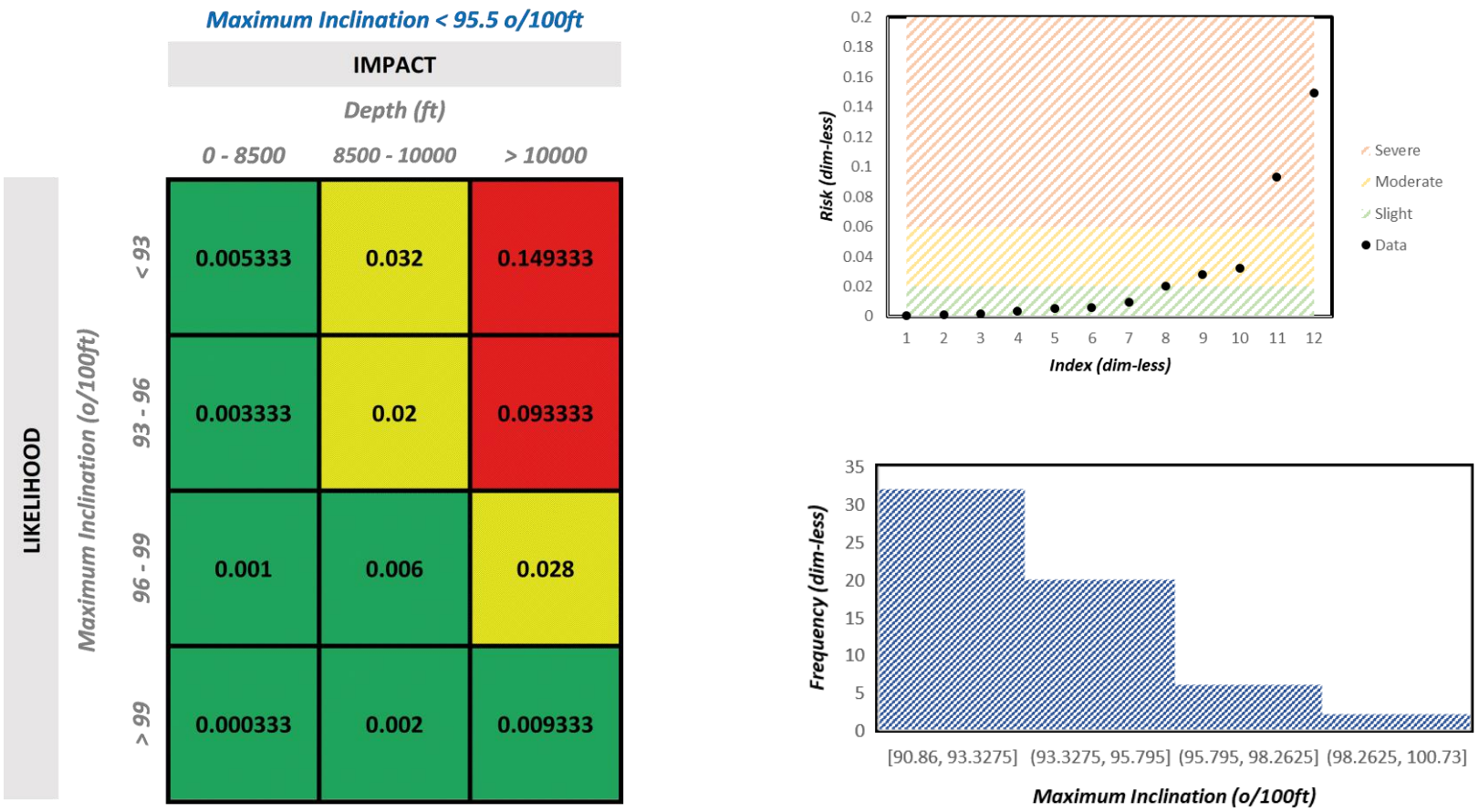


Figure E-11. (left) Risk Assessment Results for "Maximum Inclination < 95.5" Risk Factor. (Top-right) Calculated Risk Values for Different Variations of "Maximum Inclination < 95.5" Risk Factor. (Bottom-right) Summery of Frequency Analysis Used for Determining Likelihood of Various Combinations of "Maximum Inclination < 95.5" Risk Factor.

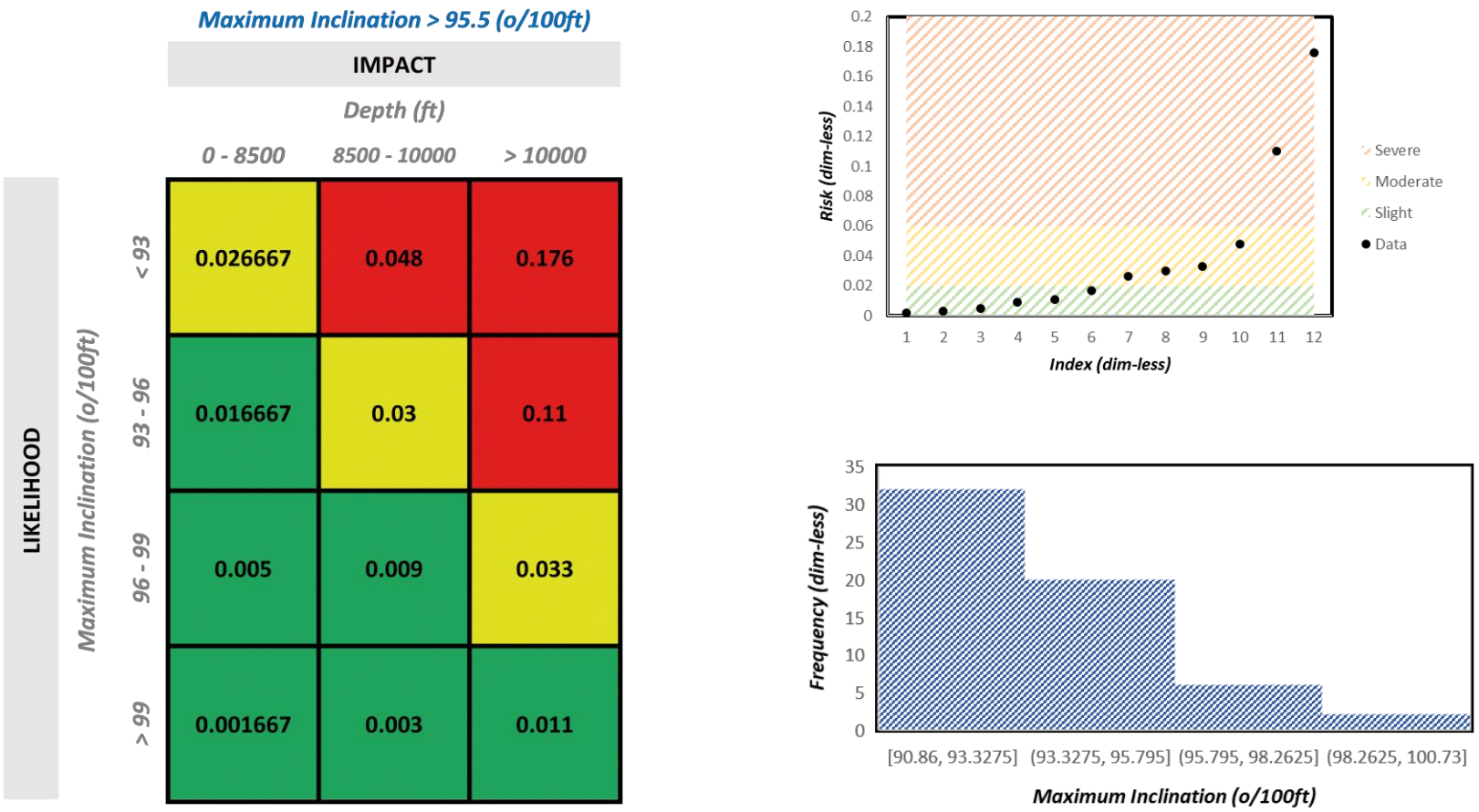


Figure E-12. (left) Risk Assessment Results for "Maximum Inclination > 95.5" Risk Factor. (Top-right) Calculated Risk Values for Different Variations of "Maximum Inclination > 95.5" Risk Factor. (Bottom-right) Summery of Frequency Analysis Used for Determining Likelihood of Various Combinations of "Maximum Inclination > 95.5" Risk Factor.

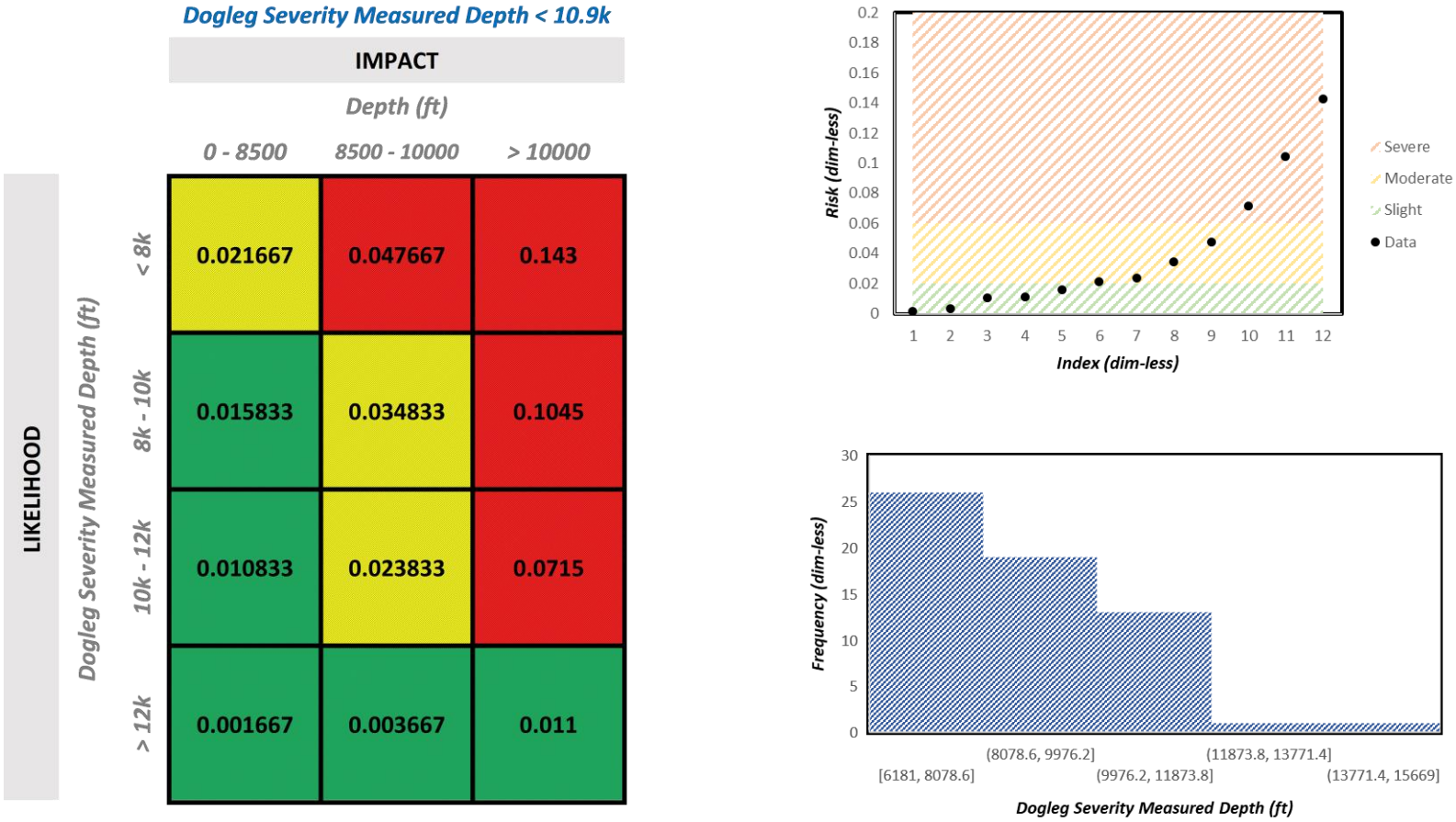


Figure E-13: (left) Risk Assessment Results for "Dogleg Severe Measured Depth < 10.9k" Risk Factor. (Top-right) Calculated Risk Values for Different Variations of "Dogleg Severe Measured Depth < 10.9k" Risk Factor. (Bottom-right) Summery of Frequency Analysis Used for Determining Likelihood of Various Combinations of "Dogleg Severe Measured Depth < 10.9k" Risk Factor.

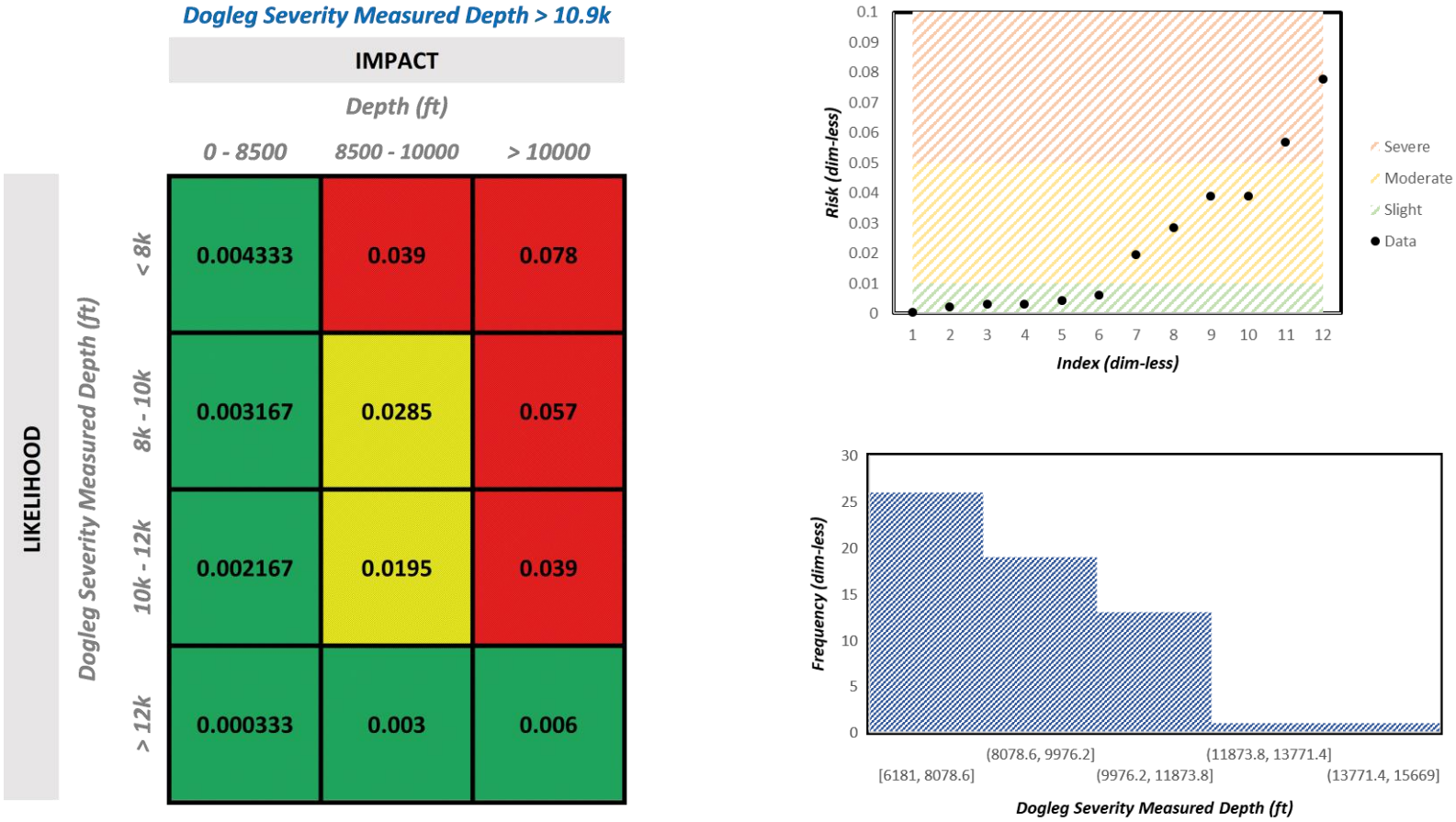


Figure E-14. (left) Risk Assessment Results for "Dogleg Severe Measured Depth > 10.9k" Risk Factor. (Top-right) Calculated Risk Values for Different Variations of "Dogleg Severe Measured Depth > 10.9k" Risk Factor. (Bottom-right) Summery of Frequency Analysis Used for Determining Likelihood of Various Combinations of "Dogleg Severe Measured Depth > 10.9k" Risk Factor.

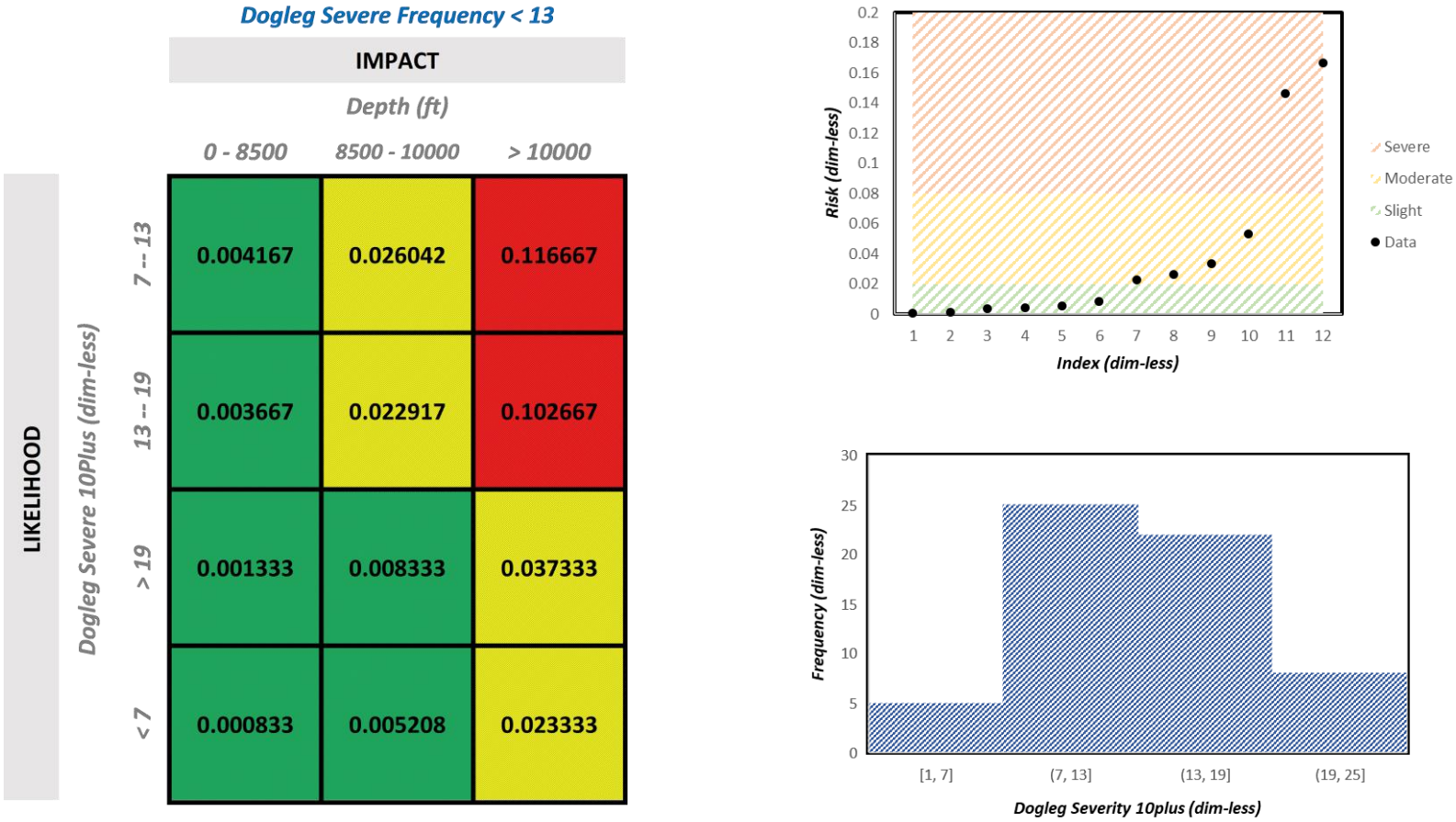


Figure E-15. (left) Risk Assessment Results for "Dogleg Severity Freq. < 13" Risk Factor. (Top-right) Calculated Risk Values for Different Variations of "Dogleg Severity Freq. < 13" Risk Factor. (Bottom-right) Summery of Frequency Analysis Used for Determining Likelihood of Various Combinations of "Dogleg Severity Freq. < 13" Risk Factor.

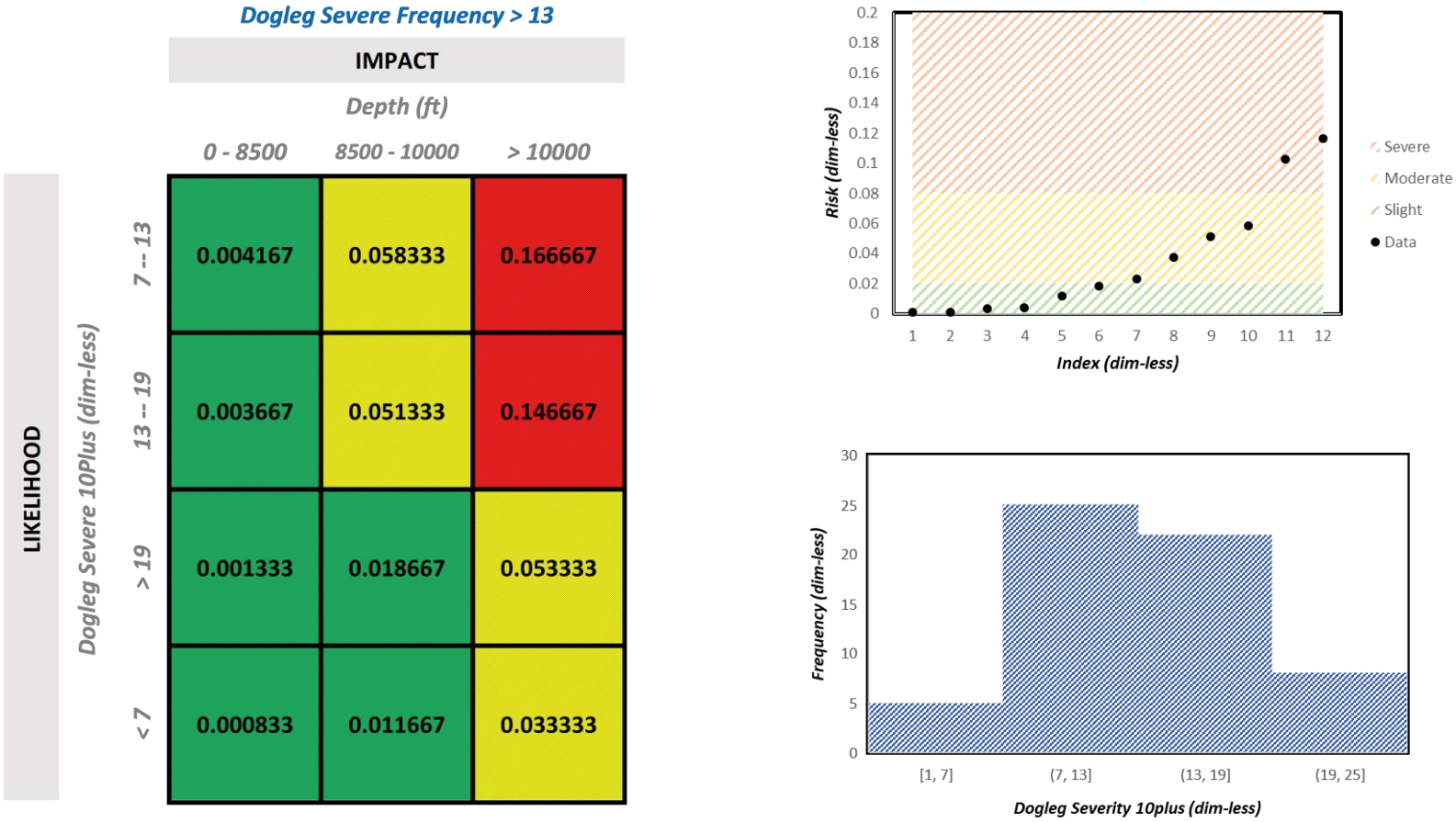


Figure E-16. (left) Risk Assessment Results for "Dogleg Severity Freq. > 13" Risk Factor. (Top-right) Calculated Risk Values for Different Variations of "Dogleg Severity Freq. > 13" Risk Factor. (Bottom-right) Summery of Frequency Analysis Used for Determining Likelihood of Various Combinations of "Dogleg Severity Freq. > 13" Risk Factor.

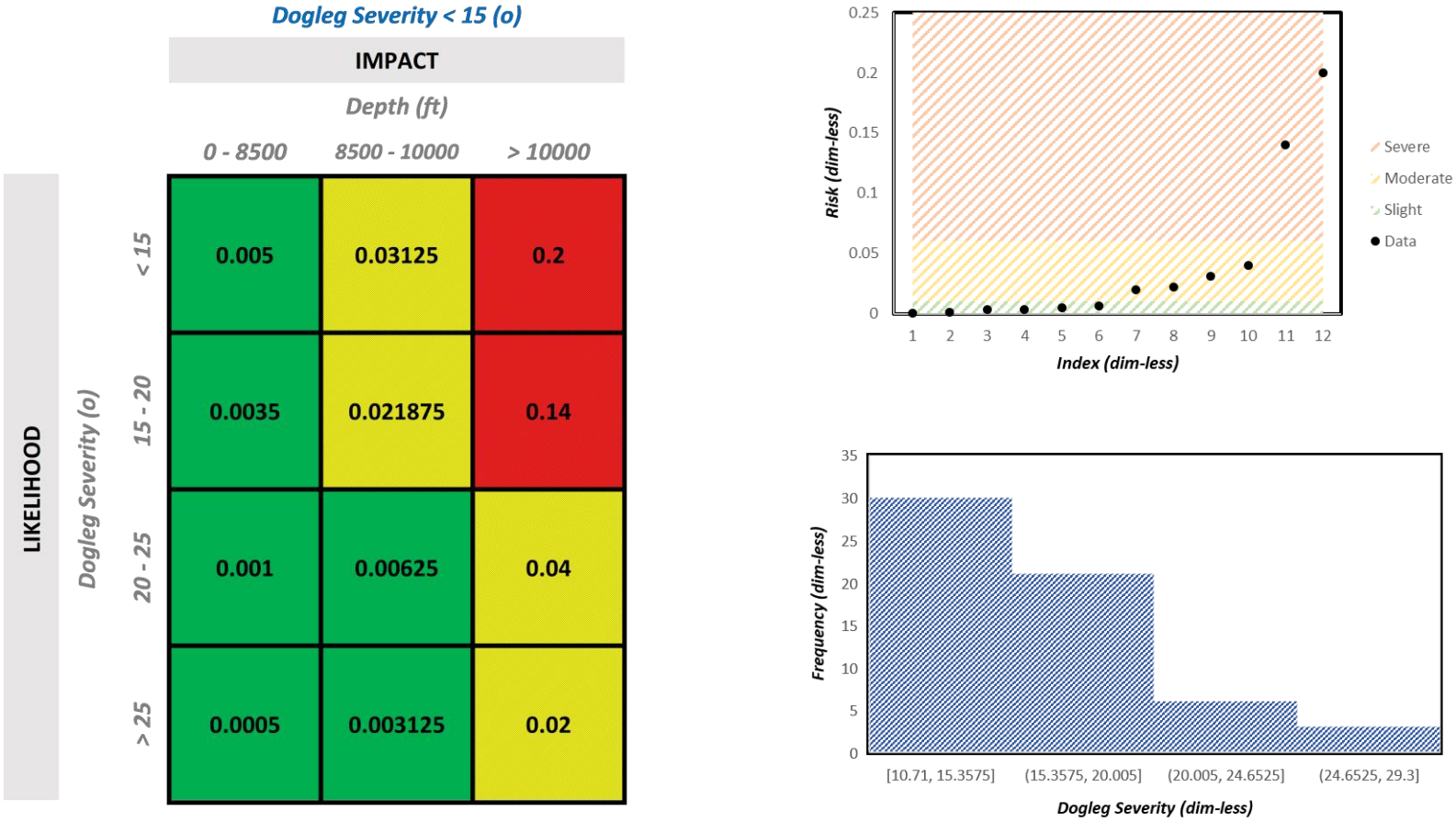


Figure E-17. (left) Risk Assessment Results for "Dogleg Severity < 15" Risk Factor. (Top-right) Calculated Risk Values for Different Variations of "Dogleg Severity < 15" Risk Factor. (Bottom-right) Summery of Frequency Analysis Used for Determining Likelihood of Various Combinations of "Dogleg Severity < 15" Risk Factor.

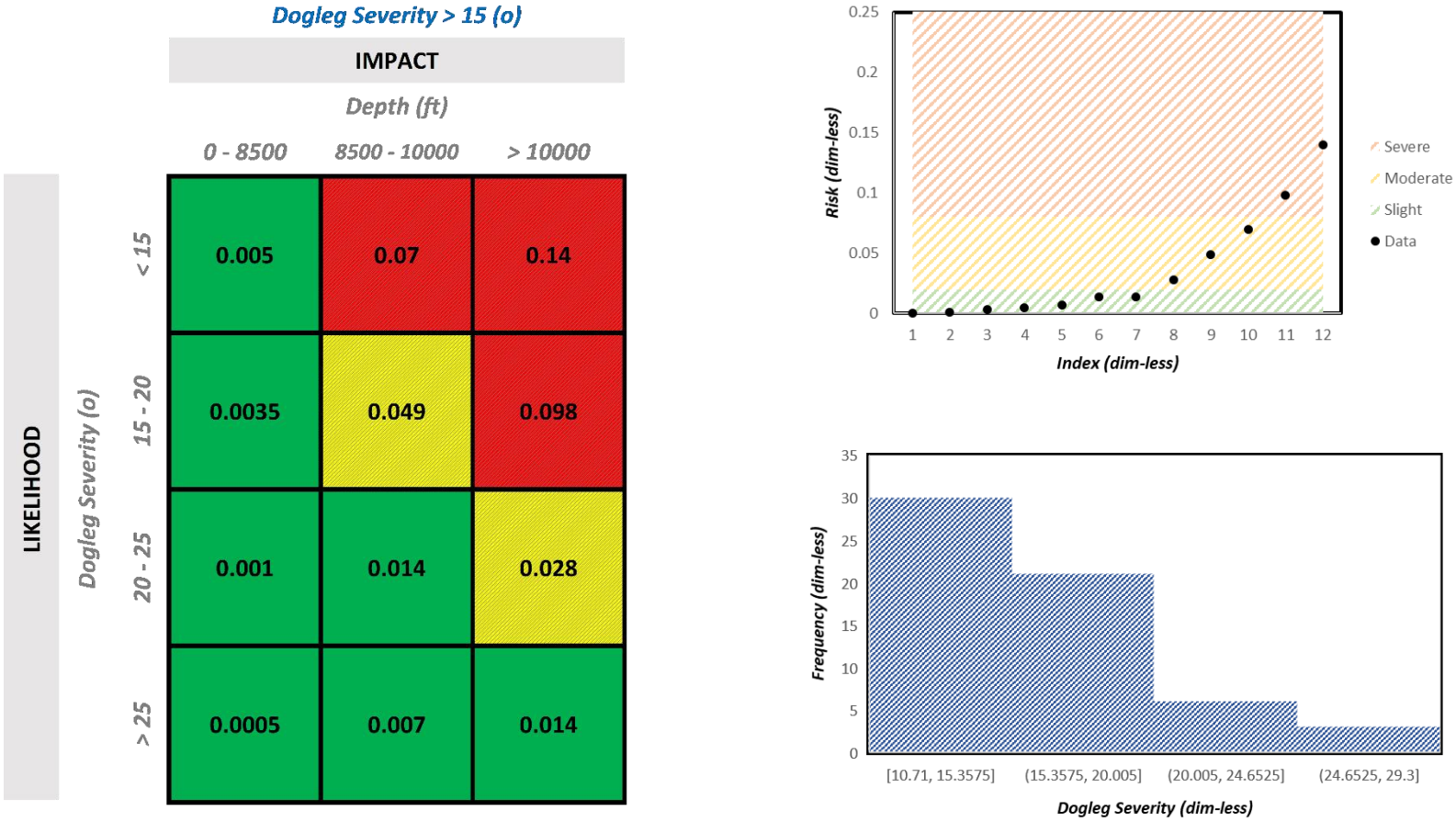


Figure E-18. (left) Risk Assessment Results for "Dogleg Severity > 15" Risk Factor. (Top-right) Calculated Risk Values for Different Variations of "Dogleg Severity > 15" Risk Factor. (Bottom-right) Summery of Frequency Analysis Used for Determining Likelihood of Various Combinations of "Dogleg Severity > 15" Risk Factor.

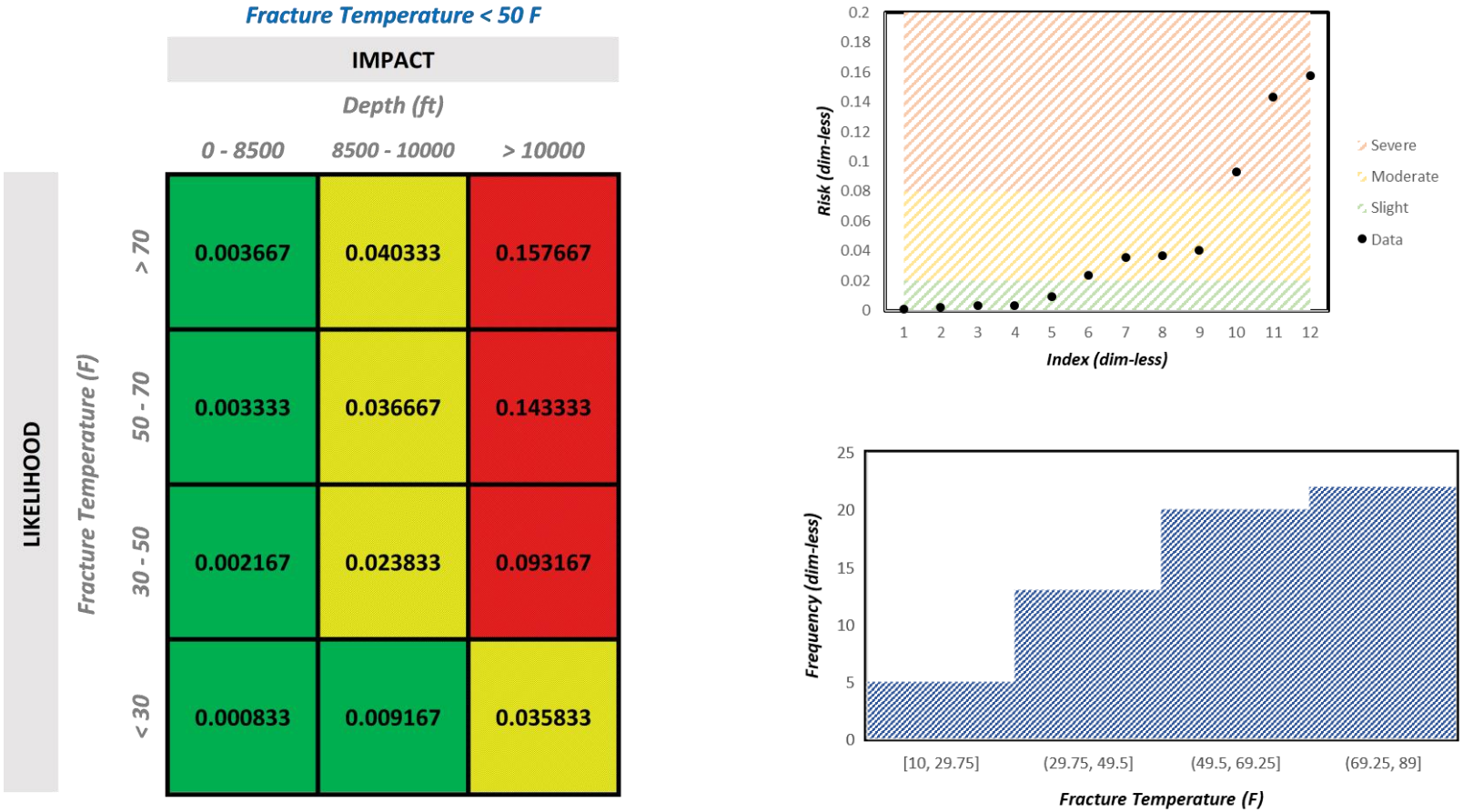


Figure E-19. (left) Risk Assessment Results for "Fracture Temperature < 50" Risk Factor. (Top-right) Calculated Risk Values for Different Variations of "Fracture Temperature < 50" Risk Factor. (Bottom-right) Summery of Frequency Analysis Used for Determining Likelihood of Various Combinations of "Fracture Temperature < 50" Risk Factor.

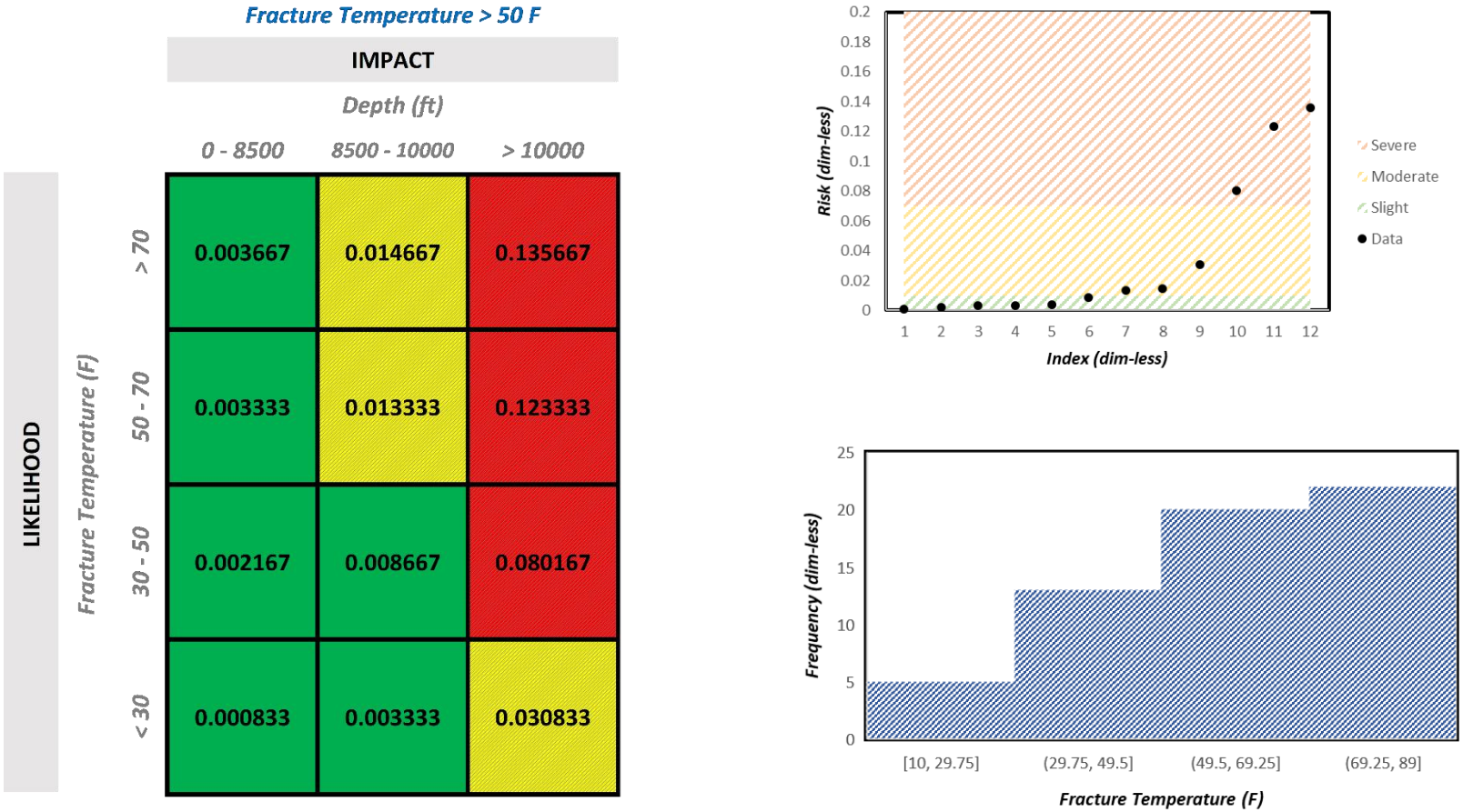


Figure E-20. (left) Risk Assessment Results for "Fracture Temperature > 50" Risk Factor. (Top-right) Calculated Risk Values for Different Variations of "Fracture Temperature > 50" Risk Factor. (Bottom-right) Summery of Frequency Analysis Used for Determining Likelihood of Various Combinations of "Fracture Temperature > 50" Risk Factor.

DEPTH	FEATURE												
	f_1	f_2	f_3	f_4	f_5	f_6	f_7	f_8	f_9	f_{10}	f_{11}	f_{12}	f_{13}
7927	Spring	No	Yes	69256.13	9406	150	7.19	94.2	7370	16	21.15	64	133
7954	Winter	No	No	62117.67	12588	170	15.2	95	10227	13	18.97	19	200
8280	Summer	No	Yes	37853.5	12343	175	13.45	93.4	7911	7	11.56	24	107
8400	Winter	Yes	Yes	101116.8	18532	174	11.26	95.74	15669	12	14.47	17	127
8846	Spring	No	Yes	42667	12114	150	13.47	93.9	7844	15	13.03	58	44
9052	Winter	No	Yes	84905.49	13260	180	14.31	93.2	8300	17	14.27	41	93
9297	Fall	No	No	56064.15	13746	150	15.94	92.9	8551	10	14.9	68	65
9403	Spring	No	No	83107.35	10472	150	10.04	94.4	7185	15	25.38	75	69
10074	Winter	Yes	Yes	126684.3	13525	170	12.94	92.33	9341	12	29.3	43	128
10100	Winter	No	No	42830.74	11950	150	13.57	93.4	7644	12	13.08	42	82
11065	Spring	No	No	37591.5	11350	150	13.93	94.4	6685	6	11.48	38	44
11096	Summer	Yes	Yes	48037.23	15960	180	14.7	93.15	11297	12	14.67	88	59
11129	Spring	No	No	46793	14720	150	17.85	92	8906	12	14.29	73	49
11537	Summer	No	No	49479.43	12246	150	13.31	93.1	8158	19	13.15	80	82
12830	Summer	Yes	Yes	48855.86	15689	180	13.69	95.89	11531	15	14.92	68	69
12920	Fall	No	No	53505.68	11120	150	13.49	94.5	6791	14	16.34	63	37
12938	Summer	Yes	Yes	56190.78	11673	150	14.91	96.1	9215	19	17.16	81	56
14705	Spring	Yes	No	43747.6	14705	182	18.01	95	8180	12	13.36	45	46
16558	Spring	No	No	42536.03	12937	150	13.95	93.1	8280	15	12.99	59	63

Figure E-21. Assessment of Already Identified Failure Cases in Granite Wash data set Using the Developed Risk matrices (RMs).

APPENDIX F

The purpose of this appendix is to highlight results of ANSYS simulation runs for various scenarios of casing support and induced thermal stresses induced in P-110 production casing.

Max. Strain Effect on Cement Support Percentage

Cement support volume has been changes by changing cement support thickness from (10-3/4" to 8-1/2") 100% to zero%. Change in casing temperature from 50^{o}C to 350^{o}C. Five different cases have been performed such as 0%, 20%, 40%, 60%, 80%, and 100% Cement support and change in temperature are 50^{o}C, 100^{o}C, 200^{o}C, 300^{o}C, and 350^{o}C.

Case 1

In case 1 at 50 ^{o}C temperature for all Cement Support Percentage

Table F-1 *Case 1 Max Strain values at 50^{o}C temperature for different cement percentage.*

Cement Support Percentage	Cement Thickness	Equivalent Strain (in/in)
100%	1.125	0.0013632
80%	0.9	0.0014394
60%	0.675	0.0015484
40%	0.45	0.0015807
20%	0.225	0.001669
0%	0	0.0017475

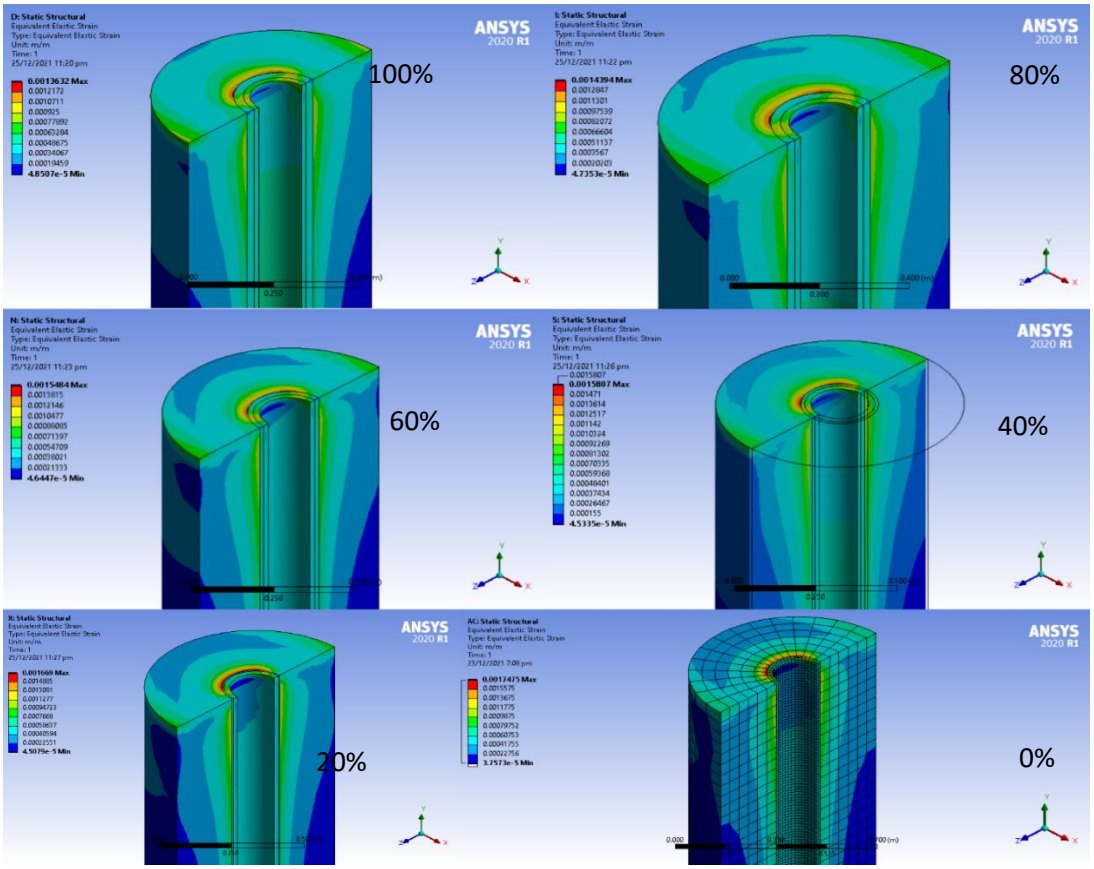


Figure F-1. Case 1 at 50ºC Temperature Differential and Different Cement support Percentage.

Case 2

In case 2 at 100oC temperature for all Cement Support Percentage

Table F-2 *Case 2 Max Strain values at 100°C temperature for different cement percentage.*

Cement Support Percentage	Cement Thickness	Equivalent Strain (in/in)
100%	1.125	0.0016955
80%	0.9	0.0017978
60%	0.675	0.0019452
40%	0.45	0.0019689
20%	0.225	0.0019882
0%	0	0.0020564

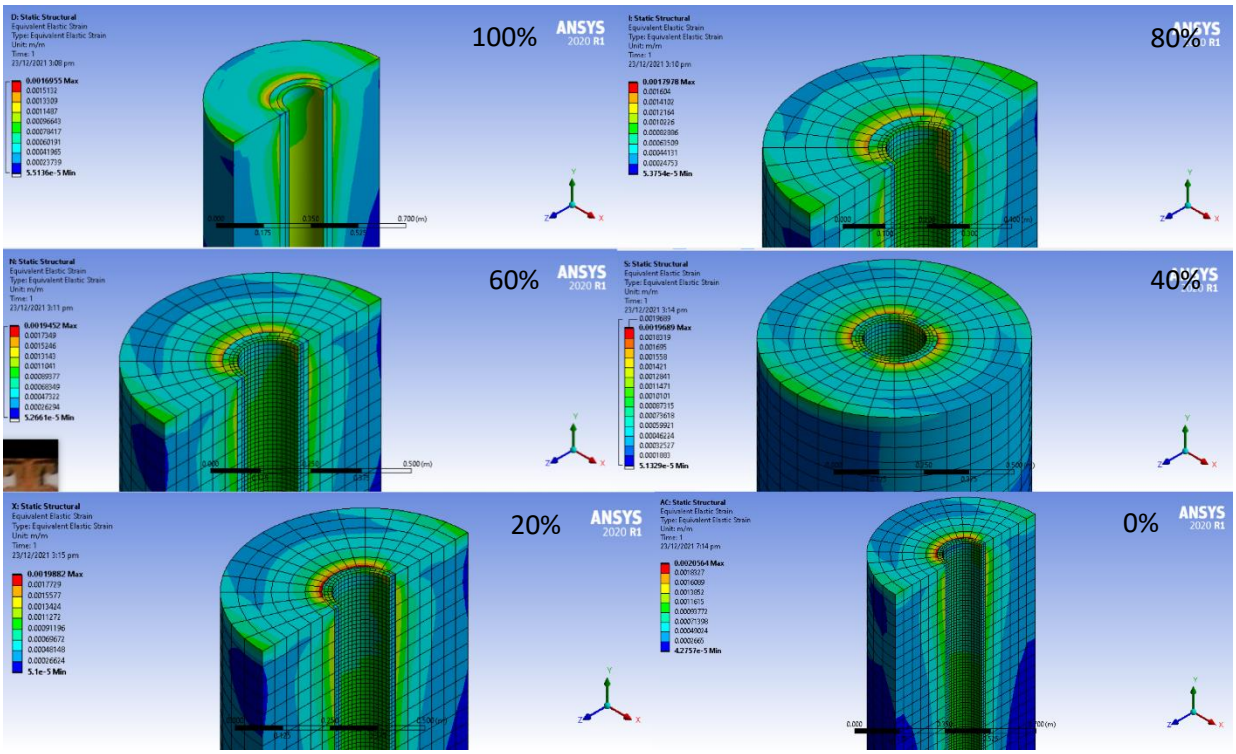


Figure F-2. Case 2 at 100°C Temperature Differential and Different Cement support Percentage.

Case 3

In case 3 at 200^{o}C temperature for all Cement Support Percentage

Table F-3 *Case 3 Max Strain values at 200oC temperature for different cement percentage.*

Cement Support Percentage	Cement Thickness	Equivalent Strain (in/in)
100%	1.125	0.0024107
80%	0.9	0.002569
60%	0.675	0.0027993
40%	0.45	0.0028386
20%	0.225	0.0028635
0%	0	0.002973

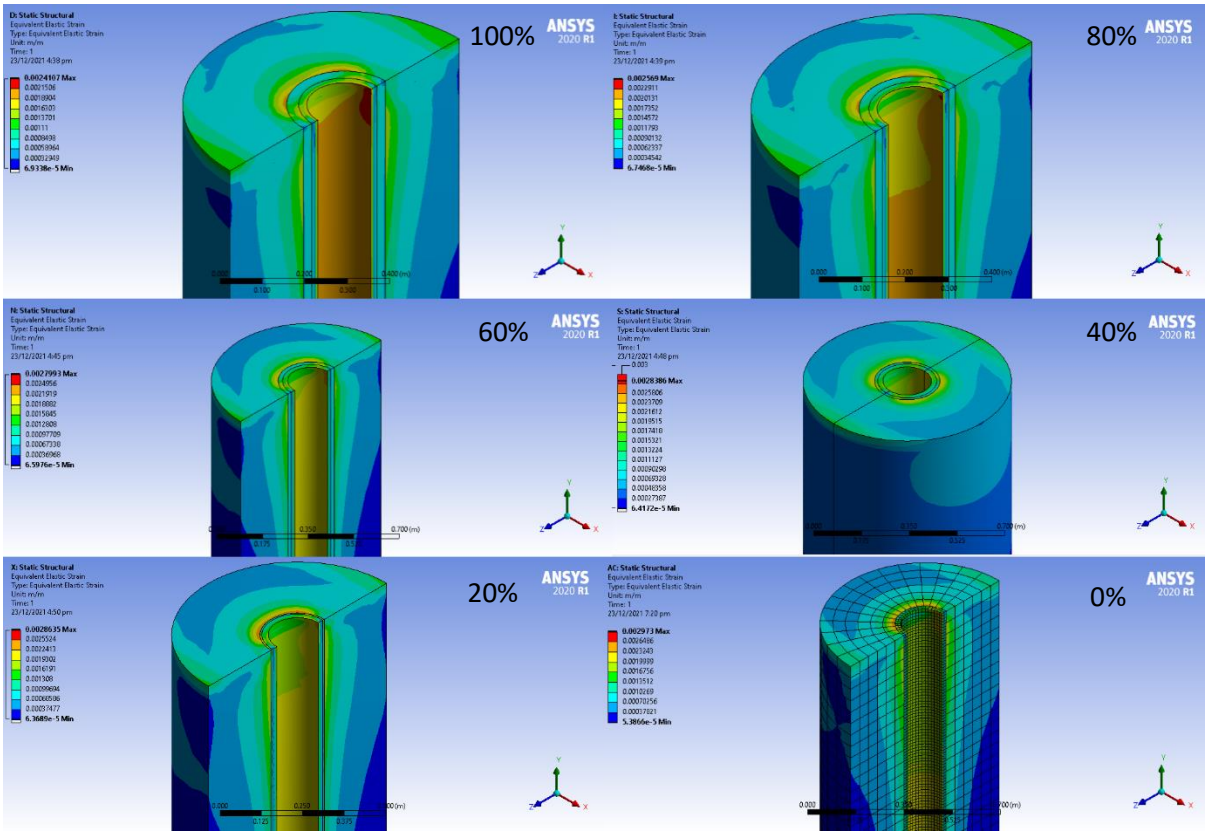


Figure F-3. Case 3 at 200^{o}C Temperature Differential and Different Cement support Percentage.

Case 4

In case 4 at 300^{o}C temperature for all Cement Support Percentage

Table F-4 *Case 4 Max Strain values at 300^{o}C temperature for different cement percentage.*

Cement Support Percentage	Cement Thickness	Equivalent Strain (in/in)
100%	1.125	0.0031632
80%	0.9	0.0034218
60%	0.675	0.0036414
40%	0.45	0.0036963
20%	0.225	0.003726
0%	0	0.0038771

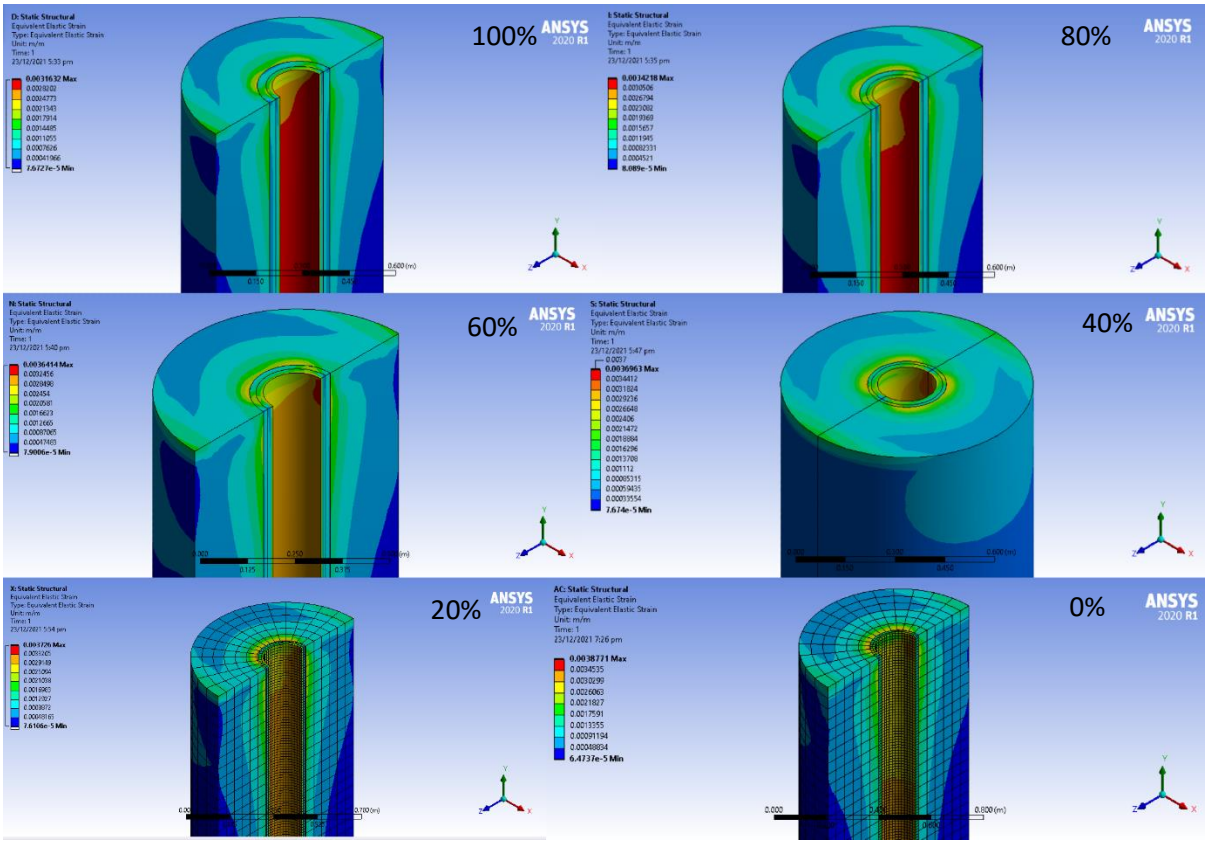


Figure F-4. Case 4 at 300^{o}C Temperature Differential and Different Cement support Percentage.

Case 5

In case 5 at 350^{o}C temperature for all Cement Support Percentage

Table F-5 *Case 5 Max Strain values at 350^{o}C temperature for different cement percentage.*

Cement Support Percentage	Cement Thickness	Equivalent Strain (in/in)
100%	1.125	0.0039393
80%	0.9	0.0039513
60%	0.675	0.0040633
40%	0.45	0.0041261
20%	0.225	0.0041581
0%	0	0.0043302

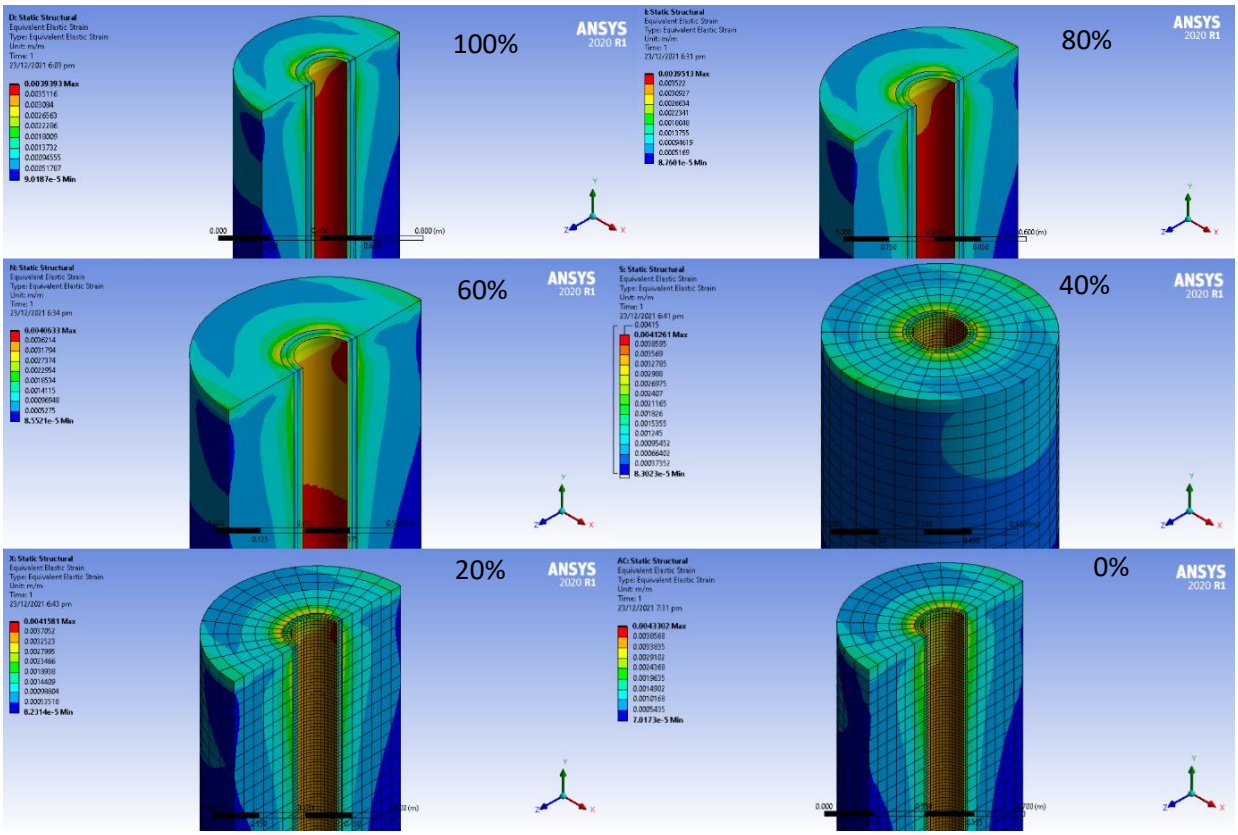


Figure F-5. Case 5 at 350^{o}C Temperature Differential and Different Cement support Percentage.