January 2023

A Data Driven Approach To Optimize Re-Fracturing Operations In The Williston Basin

Joshua Kroschel

A DATA DRIVEN APPROACH TO OPTIMIZE RE-FRACTURING OPERATIONS IN THE WILLISTON BASIN

by

Josh Kroschel

Bachelor of Science, Colorado School of Mines, 2011

Bachelor of Science, Colorado School of Mines, 2012

Master of Science, Colorado School of Mines, 2019

A Dissertation

Submitted to the Graduate Faculty

of the

University of North Dakota

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Grand Forks, North Dakota

May

2023

This thesis, submitted by Joshua Timothy Kroschel in partial fulfillment of the requirements for the Degree of Doctor of Philosophy in Petroleum Engineering from the University of North Dakota, has been read by the Faculty Advisory Committee under whom the work has been done and is hereby approved.

_____

Dr. Olusegun Tomomewo

_____

Dr. Sven Egenhoff

_____

Dr. Hu Pu

_____

Dr. Kegang Ling

_____

Dr. I-Hsuan Ho

This thesis is being submitted by the appointed advisory committee as having met all of the requirements of the School of Graduate Studies at the University of North Dakota and is hereby approved.

_____

Chris Nelson

Dean of the School of Graduate Studies

_____

Date

PERMISSION

Title A Data Driven Approach to Optimize Re-fracturing Operations in the Williston Basin

Department Petroleum Engineering

Degree Doctor of Philosophy

In presenting this thesis in partial fulfillment of the requirements for a graduate degree from the University of North Dakota, I agree that the library of this University shall make it freely available for inspection. I further agree that permission for extensive copying for scholarly purposes may be granted by the professor who supervised my thesis work or, in his absence, by the Chairperson of the department or the dean of the School of Graduate Studies. It is understood that any copying or publication or other use of this thesis or part thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of North Dakota in any scholarly use which may be made of any material in my thesis.

Josh Kroschel

5/3/2023

ACKNOWLEDGMENTS

I wish to express my sincere gratitude to my advisor, Dr. Olusegun Stanley Tomomewo for his knowledge, guidance, and support during my time in the PhD program at the University of North Dakota. I would like to extend my appreciation to the committee members, Dr. Sven Egenhoff, Dr. Hui Pu, Dr. Kegang Ling. And Dr. I-Hsuan Ho. I acknowledge the financial support of the North Dakota Industrial Commission (NDIC).

I also wish to acknowledge the help of multiple people throughout my career and in support of this dissertation. Chris Brown has been a great mentor and I enjoy our constant conversations about O&G operations. He has been a constant source of inspiration. Nathan Crawford has provided continual questioning that has greatly expanded my knowledge of the material and pushes me to better.

DEDICATION

I dedicate this dissertation to my mom, dad, sister, brother, brother-in-law, and most all, my fiancée Tami.

ABSTRACT

Because of the recent paradigm shift focusing heavily on cost minimization, many operators are now re-developing existing assets at much lower costs instead of developing newly drilled wells. Although it may seem that the hydraulic fracturing process on a well would be easier after initial stimulation, this is not usually the case and is often more difficult. Being able to identify high margin effects of treatment parameters will help engineers design hydraulic fracturing treatments to minimize average STP (STP) and minimize costs. This research develops a feature engineered multivariate regression model that identifies several high margin areas for STP reduction. These models also yield error around 2% when predicting average STP. Using the marginal effects estimated in this study, operators can start to consider minimizing STP as a design parameter that has implications for pump time, pump maintenance costs, fuel costs, and emissions.

Table of Contents

List of Tables

List of Figures

Chapter 1

1. Introduction

Portions of Chapter 1 were taken from Kroschel, J. Rabiei, M. Rasouli, V. Modeling Temporal Dependence of Average STP in the Williston Basin Using Dynamic Multivariate Regression. Energies 2022, 15, 2271 and Kroschel, J. Rabiei, M. Rasouli, V. Accounting for Fixed Effects in Re-Fracturing Using Dynamic Multivariate Regression. *Energies* 2022, 15, 5451. https://doi.org/10.3390/en15155451.

1.1 The Real "Energy Transition"

Recently in the U.S., an energy revolution has occurred. U.S. energy imports increased and peaked in 2005, coal production peaked in the year 2018, and coal consumption peaked the year prior (Energy Information Administration, 2023b). Figure 1.1 from the Energy Information Administration (EIA) shows that even with relatively steady energy consumption, the U.S. became a net exporter of energy around 2017.

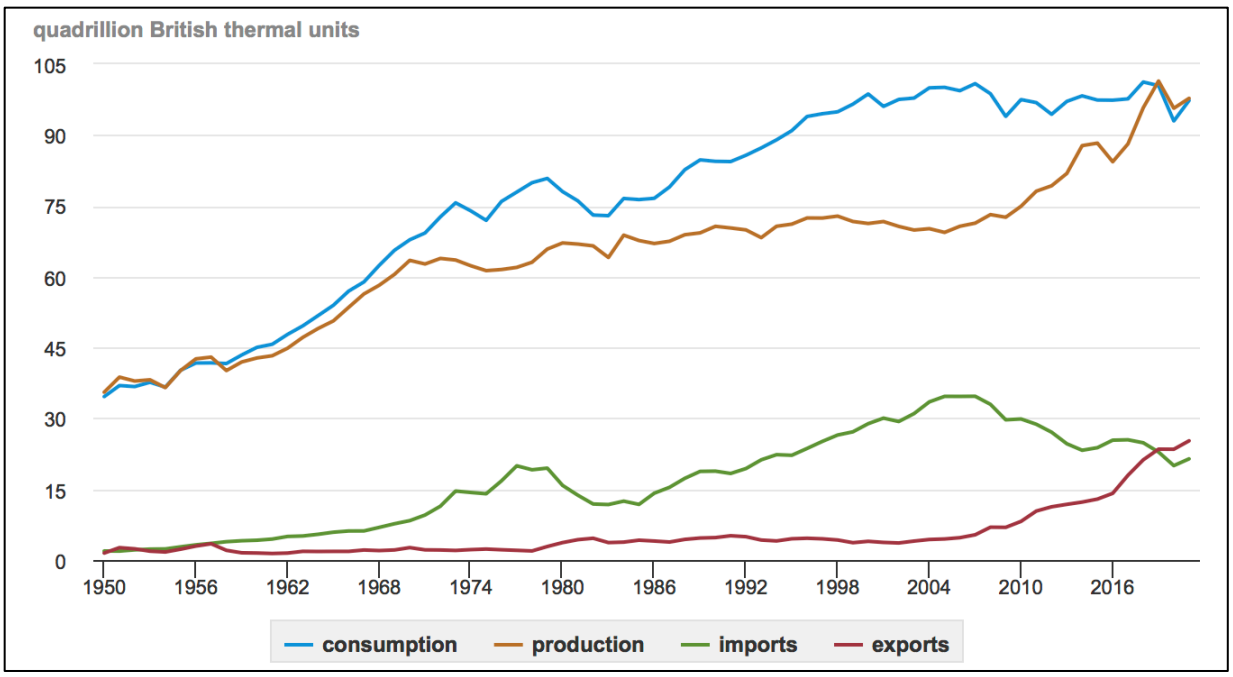

Figure 1. 1 - U.S. energy production, consumption, imports and exports from 1950 - 2021 (Energy Information Administration, 2023b)

Looking at figure 1.1, the year 2005 saw a sharp turnaround in U.S. energy imports and exports. This was due to the implementation of hydraulic fracturing (HF) in conjunction with recent advancements in horizontal drilling and the effects of this revolution have been extraordinary.

With this, it would be tempting to argue that this revolution and access to abundant oil and natural gas (O&G) would create more greenhouse gas emissions as these resources are used. However, this hasn't been the case. Figure 1.2 shows the trend in annual emissions with the associated percentage change. We see the general decreasing trend starting around 2006, just after the energy revolution in 2005. Even after the rebound from the COVID 19 pandemic, emissions in 2021 remained well below peak levels in 2007 (Energy Information Administration, 2022).

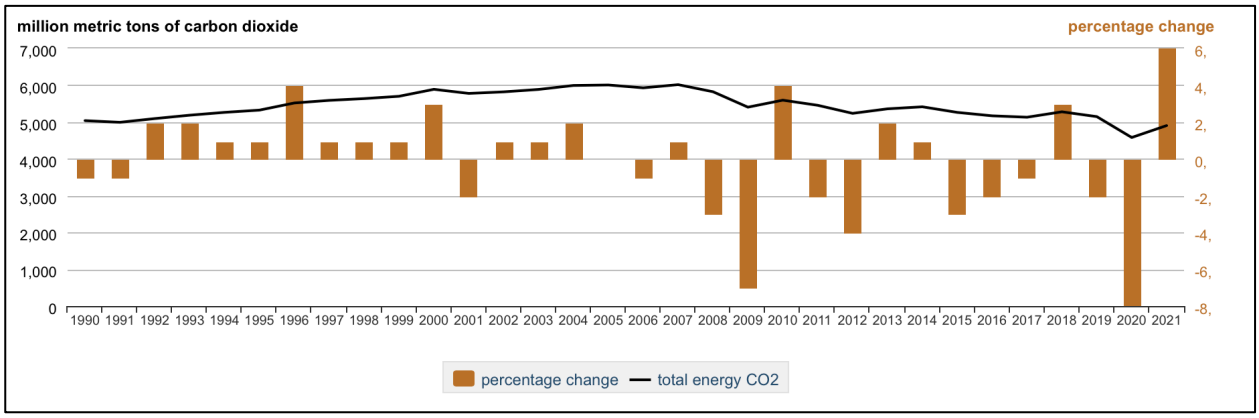

Figure 1. 2 - Data from the EIA showing trend in annual emissions and associated percentage changes (Energy Information Administration, 2022)

Figure 1.3 shows the CO_2 emissions reductions relative to 2005 levels specifically for electricity generation. Year over year, we see more emissions reductions from *more* consumption of natural gas, not less. According to figure 1.3, natural gas is responsible for

more CO_2 emissions reductions than non-carbon sources from electricity generation. This fact should not be taken lightly.

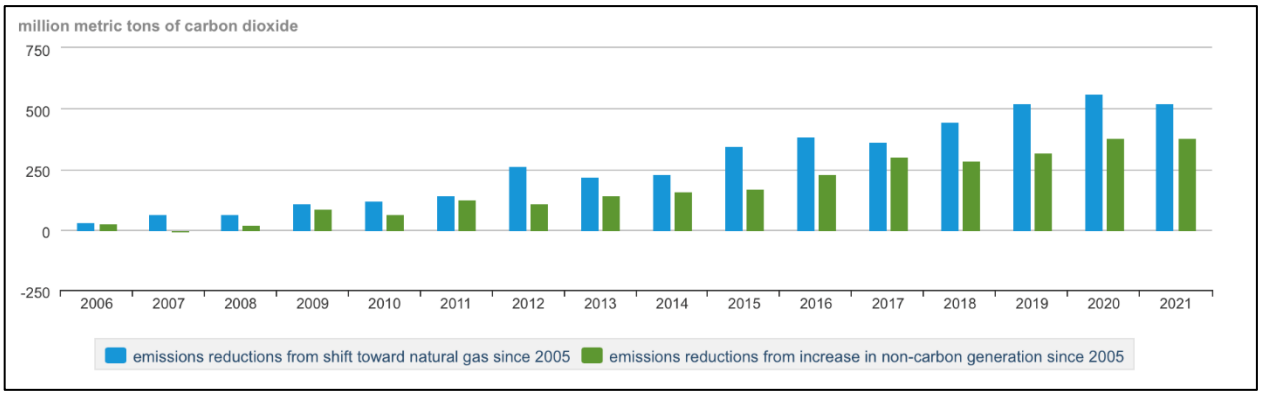

Figure 1. 3 - Data from the EIA showing emissions reductions realized by substituting natural gas (blue) and non-carbon sources (green) for electricity generation relative to 2005 emissions (Energy Information Administration, 2022)

Figures 1.2 and 1.3 show that if we are worried about emissions, and serious about reducing CO_2 emissions in particular, it is entirely possible that we should be focusing more heavily on natural gas production and consumption instead of limiting it. This is because natural gas consists mostly of CH_4 and has a high hydrogen content and combustion will produce far less CO_2 (Energy Information Administration, 2023a). Figure 1.4 shows the large decrease in coal consumption for energy generation. This is due to large scale substitution from coal to more cost efficient natural gas and other sources for electricity generation (Energy Information Administration, 2023b). This highlights the large scale substitution of coal for natural gas for electricity generation and the adoption of more natural gas for primary energy consumption as well.

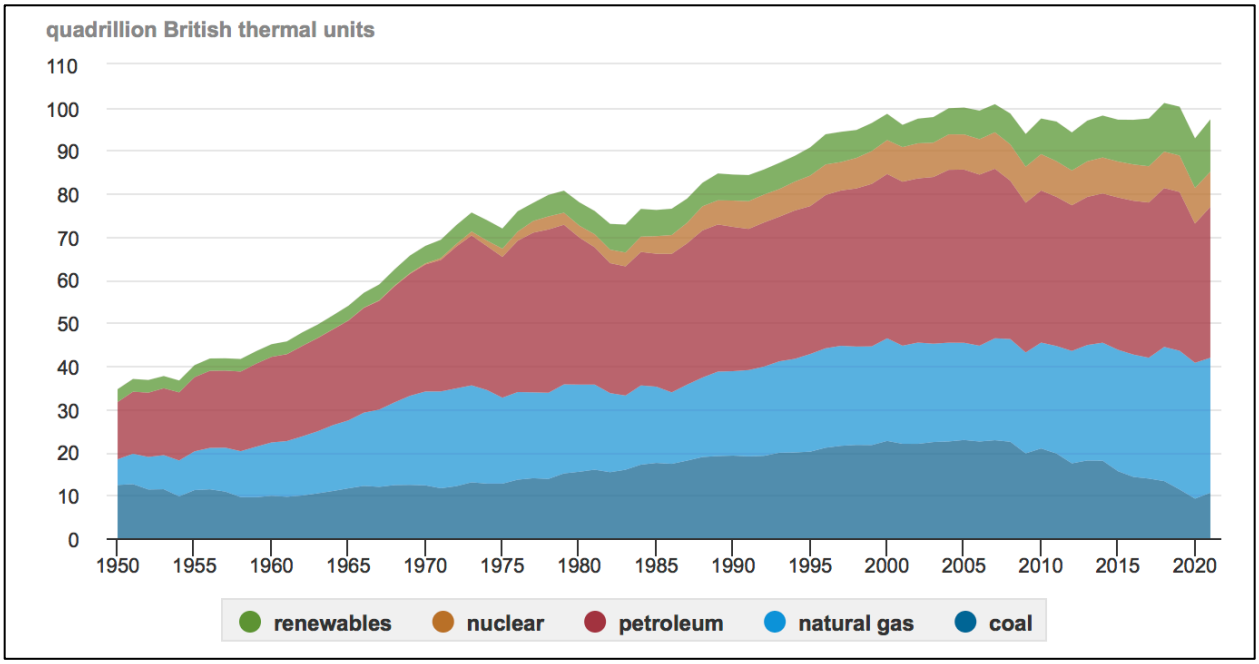

Figure 1. 4 - Primary energy consumption by source in the U.S. from 1950 – 2021 (Energy Information Administration, 2023b)

On December 08, 2021, President Biden signed an executive order to direct the federal government of the U.S. to achieve five goals. Two of the five goals is to 1) produce 100% carbon emission free electricity (CFE) by 2030, and 2) 100% zero emission vehicle (ZEV) acquisitions by 2035 which includes 100% zero emission light duty vehicle acquisitions by 2027 (White House, 2021). This shows that electricity is a primary target of energy policy. In line with this policy, research in renewable energy, energy storage, and battery and electric powered vehicles are becoming more common in our daily lives (Chen, Xiong, Li, Sun, & Yang, 2019). However, it's important to remember that electricity is an end-use product, not an energy source. This is highlighted in figure 1.5. The primary energy sources in the U.S include fossil fuels (petroleum, natural gas, and coal), nuclear energy sources, and renewable sources of energy (Energy Information Administration, 2023b). Energy consumption from each sector in the economy has a wide range of variability as shown in

figure 1.5 (Energy Information Administration, 2023b). We see that petroleum and natural gas are inputs into all sectors of the economy and comprise 68% of the total energy input.

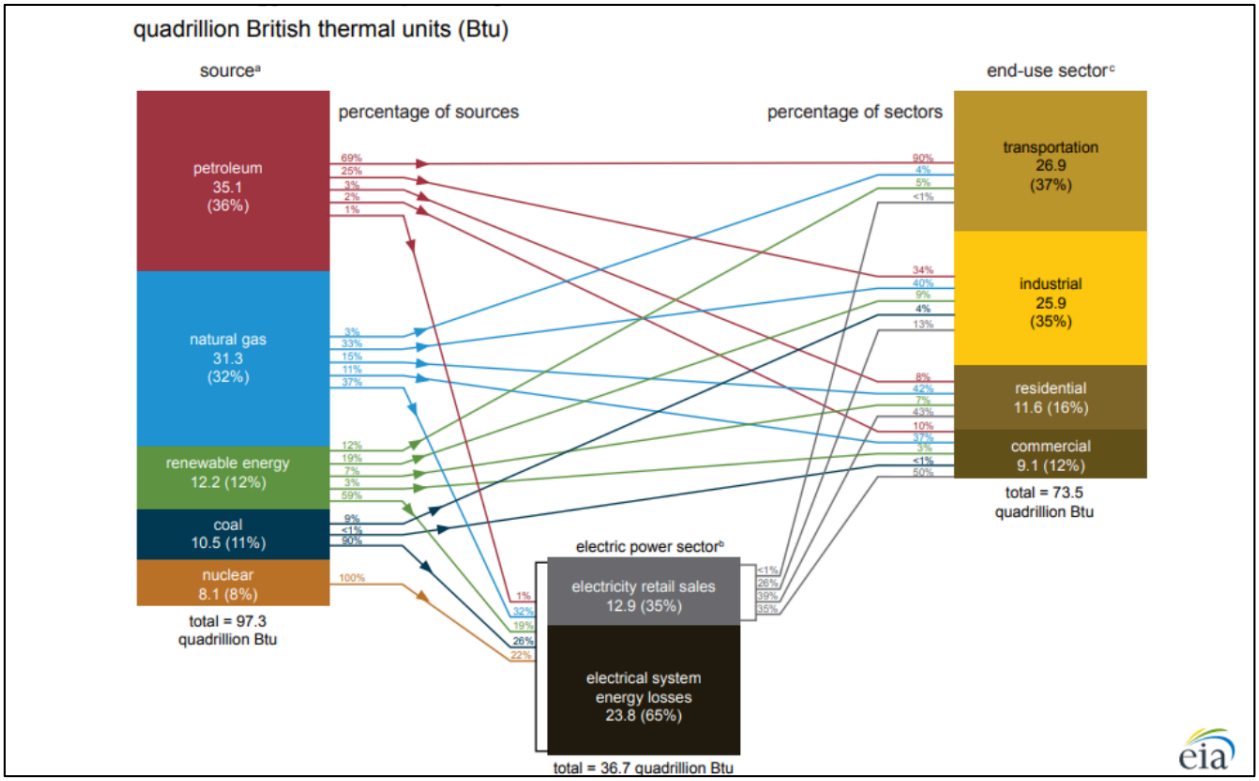

Figure 1. 5 - U.S. energy by individual sectors using different sources for year 2021 (Energy Information Administration, 2023b)

In contrast to U.S. energy consumption, households in developing countries depend on burning wood for energy (Dendup, 2022). Roughly 2.4 billion people around the globe rely on open fires or inefficient stoves for household use using fuel sources such as kerosene, biomass, and coal, which generates toxic air pollution as a result (World Health Organization, 2022). This indoor air pollution is responsible for around 3.2 million deaths every year and an estimated 86 million healthy life years were lost in

2019 as a results of indoor air pollution (Household Air Pollution, 2022). Given this, we shouldn't assume a priori that ceasing O&G use will result in benefits greater than costs. Given that fossil fuels are inexpensive, reliable, relatively easily transportable, and wind and solar are intermittent and expensive to store (Borenstein & Kellogg, 2021), we should also not assume a priori that tethering these developing countries to technologies like wind and solar instead of fossil fuels. Considering this, it is worthwhile to analyze the arguments for ceasing O&G use.

It is well documented that CO_2 causes global temperatures to rise through the greenhouse effect that traps heat from the sun (NASA, What is the greenhouse effect?, 2023). However, too much CO_2 in the atmosphere and the resulting greenhouse effect is an argument for reducing CO_2 in the atmosphere, not ceasing O&G extraction and consumption. Significant CH_4 emission in the natural gas supply chain could undermine the gains realized by natural gas combustion (Weller, Hamburg, & von Fischer, 2020). The fact that natural gas pipelines leak is an argument for actions like better policy and more robust construction, not getting rid of pipelines. Weller et al. (2020) conclude that targeting the oldest pipelines and identifying the largest leaks for repair can result in the largest emission reduction in the natural gas distribution system (Weller, Hamburg, & von Fischer, 2020). In addition, NASA's Earth Surface Mineral Dust Source Investigation was developed to map minerals in desert dust, but can also detect airborne methane (NASA, 2022).

Extending the same logic to intermittent sources of energy like wind and solar, intermittent supply is not an argument for not developing wind and solar, it is an argument for fixing the issue of intermittent supply. However, although the unit cost of

some low-emission technologies has decreased since 2010 (Shukla, 2022), the source energy intermittency problem with wind and solar generation still exists. To counteract this, storage technologies have traditionally been the solution. Although there has been substantial advancement in batteries, batteries are not anymore of an energy source than a pantry is a food source. If we also assume that technological advancements will reduce the cost of storage and increase efficiency enough to make it feasible to store large amounts of electricity, storage is still conditioned on intermittent supply, i.e. you can not store what isn't produced. This also requires assuming that there will be excess capacity to store.

Given the energy source problem and the remarkable increase in energy production realized in the U.S., along with large scale CO_2 reductions realized by utilizing natural gas, an argument can be made that we should not be limiting production of any energy source, especially oil and natural gas. We should also not assume that the "energy transition" the developed world is attempting to undertake necessarily applies to other areas of the globe.

The world needs energy as it is vital for economic opportunity (Song, et al., 2023) and policymakers are now focused on the relationship between energy supply and economic development (Oyekale & Molelekoa, 2023). Thus, limiting energy limits economic growth and opportunity. Also, when technology advances swiftly and governments try to decide who the winner is a priori, the results are often costly (Borenstein & Kellogg, 2021). So, it is not clear that we should decide a priori that electricity generated by intermittent supply sources is the way of the future. Also given that we know fossil fuels can produce many direct benefits and spillovers, it would be immoral to drag the rest of

the world along our energy transition without first letting them transition to industrialized societies and reap the benefits readily deployable, reliable, and energy dense fossil fuels.

1.2 Challenging the Unstated Assumption in the Energy Transition

In this section, I start by challenging the unstated assumption that we should not do things that increase global temperatures, which is required to logically conclude that we should cease O&G use.

The philosopher David Hume proposed that "ought" does not follow logically from "is". (Hume, 1985). More practically, there is no clear path way from facts that exist in the world (is) to actions that should be taken (ought). The current debate about what to do about climate change can be viewed through this lens. Essentially, CO_2 emissions increase global temperatures through the greenhouse effect (is). Therefore, we should cease emitting CO_2 (ought). However, this is not logically complete. To complete this logically, we need an assumption of another "ought" (Woodford, 2021). The previous "ought" follows logically from "is" if we assume that we should not do things that increase global temperatures. So, the complete logical statement looks like the following: "CO_2 emissions increase global temperature, and we should not do things that increase global temperature. Therefore, we should cease emitting CO_2.".

Burning fossil fuels creates CO_2 , H_2O, and heat, with the heat being utilized for energy (Energy Information Administration, 2023b). Figure 1.6 shows the primary energy consumption by sources specifically for the year 2021. We see that O&G made up about 68% of the total energy consumption for the year 2021. In contrast, energy from renewable sources made up only 12% of the overall energy consumption. Of that 12%, solar and wind

made up only 39% of the total renewable energy generation. Given this, it should come as no surprise that that majority of anthropogenic greenhouse gas emissions in the U.S. come from burning fossil fuels for energy (Energy Information Administration, 2023a). However, based on evidence, these emissions are simply a cost of our well-being and progress, as the rest of the section will argue.

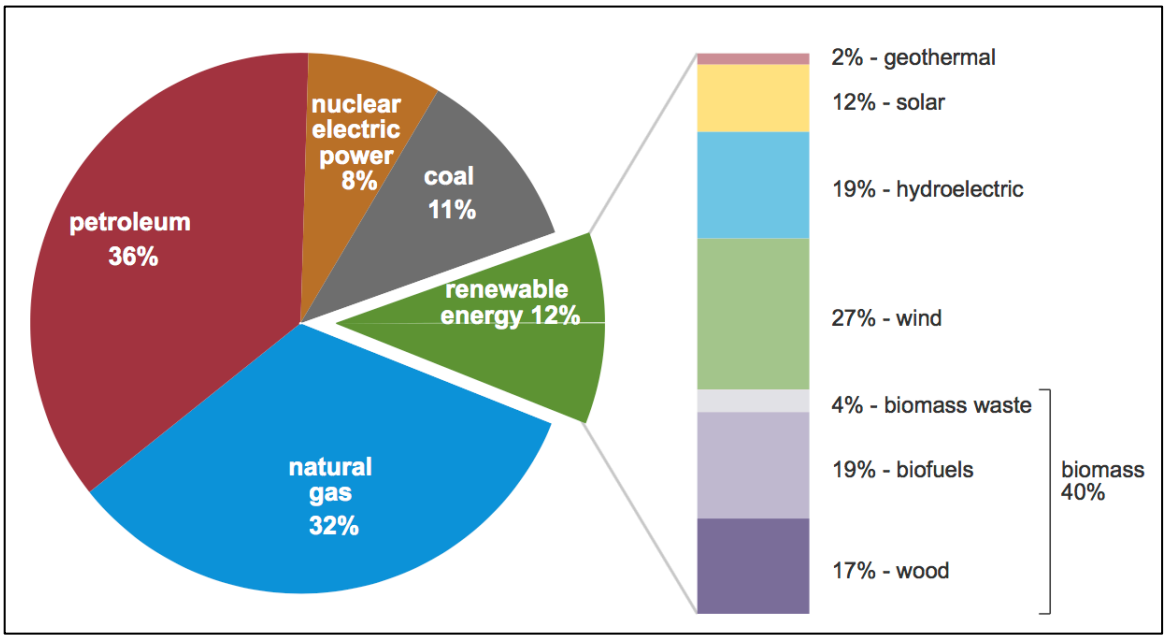

Figure 1. 6 - Primary energy consumption by source for year 2021 (Energy Information Administration, 2023b)

Christopher Hitchens proposed the following philosophical razor: "If it can be asserted without evidence, it can be dismissed without evidence" (Wikipedia, 2023). Therefore, the rest of this section will provide evidence that challenges the unstated assumption that we should not undertake any activity that warms the planet. This

discussion is important because of the debate about what to do about climate change and why re-developing O&G assets is justifiable, and moral, based on historical evidence.

1.3 Economics, Climate Change, and Human Well-Being

The economist Thomas Sowell has said, “There are no solutions, there are only trade-offs” and the Nobel prize-winning economist Milton Friedman expressed similar sentiments when he stated, “There is no such thing as a free lunch”. The larger point of these two quotations is that there is always a forgone cost, or opportunity cost, associated with an action. In doing A, you’re forgoing B, which also has some value and represents a forgone benefit, or, an incurred cost. For example, a college students economic cost of attending university is not only the cost of attendance, but also any forgone wages that they could’ve otherwise been earning in the marketplace over their time in college. One historical example was noted by Louis Stotz, in 1938; “the discovery of petroleum in Pennsylvania gave kerosene to the world, and life to the few remaining whales” (Stotz, 1938). This highlights two economic points: 1) Human ability to adapt and create substitutes, and 2) the opportunity cost of pursuing O&G production may have been the existence of whales on earth.

This brings us to the debate about what to do about climate change, which must consider the opportunity costs. There is a plethora of literature about the costs around climate change as a result of the greenhouse effect. It is well understood that economic outcomes are intertwined with climate and this relationship will have policy and market impacts in the future (Newell et al., 2018) and according to William Nordhaus, the foremost economist on climate economics and Nobel Prize winner, the Social Cost of Carbon (SCC)

is the most important economic metric when designing and implementing effective climate change policy (Nordhaus, 2017). William Nordhaus estimates the SCC is $31 per additional tone of CO_2 emitted in 2010 US$ (Nordhaus, 2017).

There are legitimate concerns around the changing climate. Local air quality from O&G operations during extraction may be of concern (Hausman & Kellogg, 2015). Drought may become a challenge as global temperatures rise which makes arid and semi-arid areas especially prone to impacts (Dhanya & Geethalakshmi, 2023). Hajdu and Hajdu (2023) analysed data from some 600,000 pregnancy losses between 1984 and 2018 in Hungary and estimate that higher temperatures tend to increase the risk of pregnancy (Hadjdu & Hajdu, 2022). Hajdu and Hajdu (2022) also estimate that the effect of increased exposure to heat is more pronounced during the first half of pregnancy *and* among women with less than high school education (Hadjdu & Hajdu, 2022). However, this is just as much evidence for increasing education levels as it is for reducing global temperatures. In addition, wealth is generally a metric for a willingness to pay for environmental quality (Timmins & Vissing, 2022). So, it seems that focusing on education and wealth generation might have more immediate impacts. In addition to the vast amount of academic literature, the Intergovernmental Panel on Climate Change (IPCC) recently published their sixth assessment report and found that cumulative net CO_2 emissions have increased due to anthropogenic activities since 1850 (Shukla, 2022). This seems to have been widespread knowledge in private, academic, and government circles for quite some time. The review of internal documents from Exxon scientists has led to the conclusion that the fossil fuel industry has known that the by-products of fossil fuel consumption could cause the earth to warm through the greenhouse effect (Supran, Rahmstorf, & Oreskes, 2023). However,

Supran, Rahmstorf, and Oreskes (2023) state that these models and knowledge were "consistent with, and at least as skilful as, those of independent academic and government models." (Supran et al., 2023). So, it seems most everyone knew about the potential costs of burning fossil fuels; however, no one was willing to forgo the benefits. This yields one of, if not the, biggest problem when dealing with climate change: intertemporal choice. Intertemporal choices are choices made over time, where costs and benefits are dispersed (Loewenstein & Thaler, 1989).

Even if we assume everything thus far regarding climate change is entirely accurate, we still need to assume that we ought not undertake activities that raise global temperature and may therefore impose a social cost from those activities. Based on historical evidence, I would argue this is not a valid assumption and that although the SCC may represent an actual cost, there is actually a *net benefit* in undertaking the anthropogenic activities that produce CO_2 emissions. O&G provide more than two-thirds of the energy Americans consume daily and are a large part of the standard of living, contributing in ways that are often not apparent (U.S. Department of Energy, 2020). Broadly speaking, Americans take for granted the energy infrastructure in place and the fact that it always works (Borenstein & Kellogg, 2021) and over 1/3 of the electricity generated in the U.S. is generated from burning natural gas (U.S. Department of Energy, 2020).

Figures 1.7 shows the world annual CO_2 emissions from 1960 to 2021 taken from Ritchie et al. (2021). The argument for continued O&G production will continue to refer back to this figure as this is usually seen as an existential threat. Although these emissions present problems, I would argue that, at least currently, it is simply a cost of improving the quality of life much like going to the dentist is a cost for a greatly improved oral health.

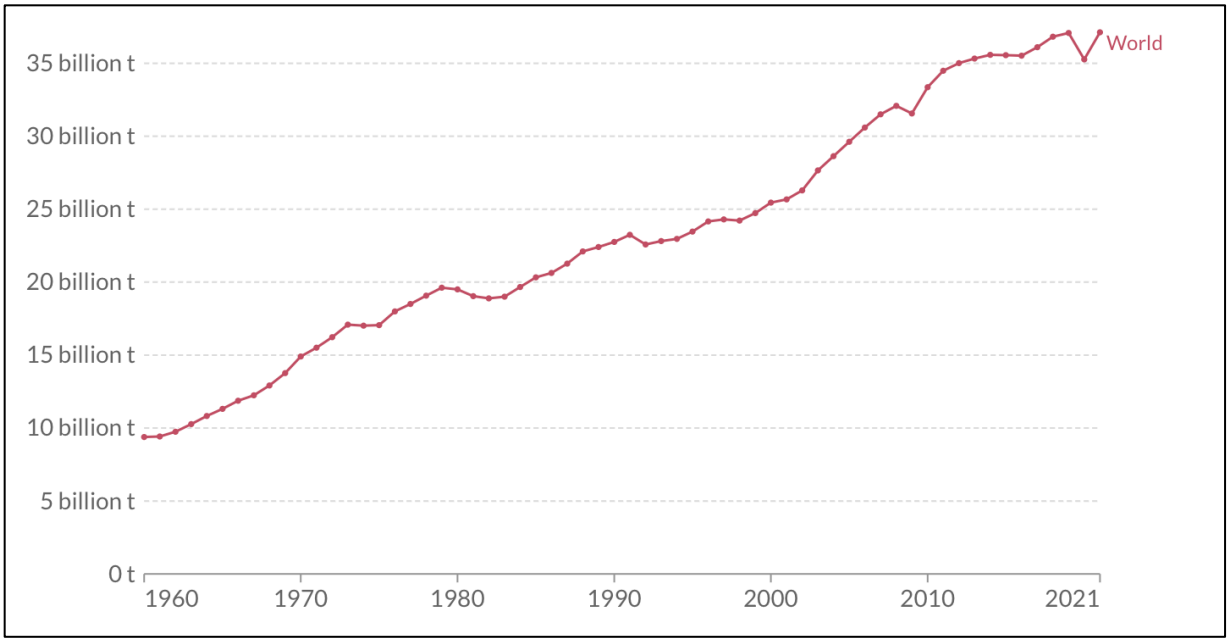

Figure 1. 7 - Annual global CO2 emissions since 1960 - 2021 (Ritchie et al., 2020)

Figure 1.8 shows world GPD (in current US$) from 1960 to 2021, respectively. This tracks very closely with world CO_2 emissions, and the relationships aren't only specific to the world in aggregation. Figures 1.9 shows that GDP has seen large increases since the 1960s for low and middle-income countries. However, when comparing these results with figure 1.10, we see that the upper and lower middle income countries are responsible for nearly all of the increase in GDP, measured in current U.S. $.

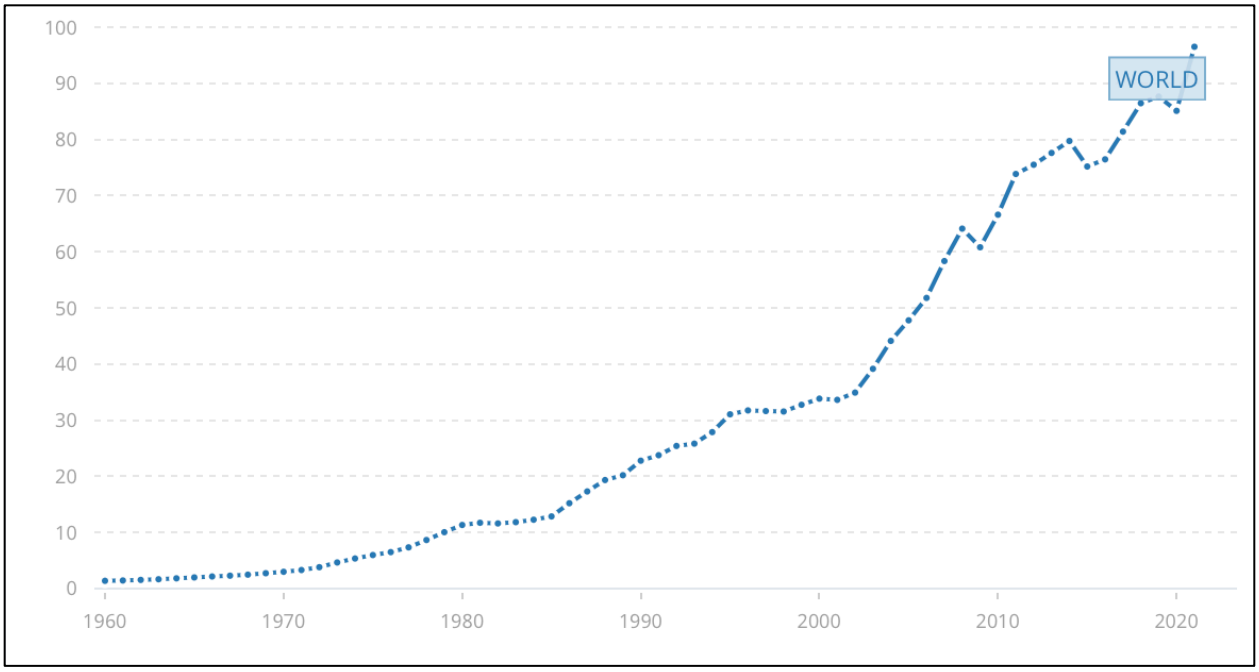

Figure 1. 8 - World GDP in trillions of current U.S. $ from 1960 – 2021 (World Bank, 2023)

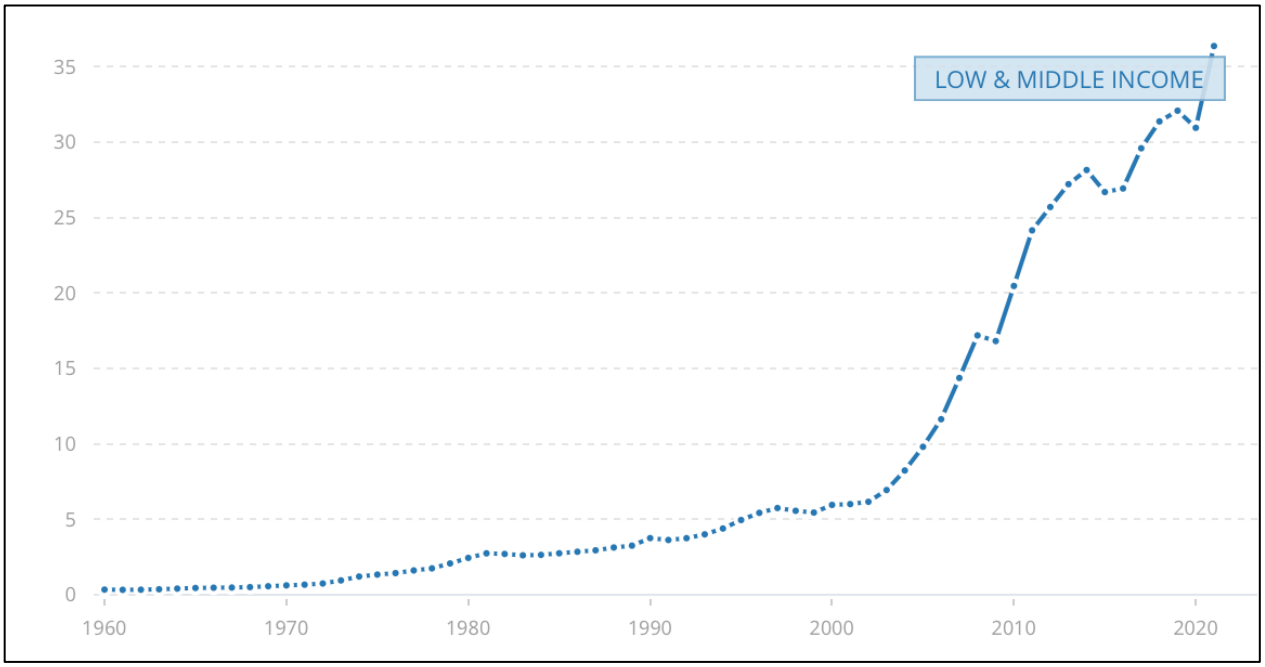

Figure 1. 9 – GDP in trillions of current U.S. $ from 1960 – 2021 for low and middle-income countries (World Bank, 2023).

Now, it should come as no surprise that the countries with highest GDP tend to be wealthier. However, when comparing figure 1.10 with figure 1.11, which shows the emissions for middle and low income countries, we see the same stratification with upper-middle, lower-middle, and low income countries in terms of emissions. The argument here is not that increasing CO_2 emissions increases wealth, but taking advantage of opportunity costs does. When societies are able to utilize energy dense and reliable resources like fossil fuels, they are able to redistribute their time to other activities such as healthcare, medicine, engineering, education, etc.

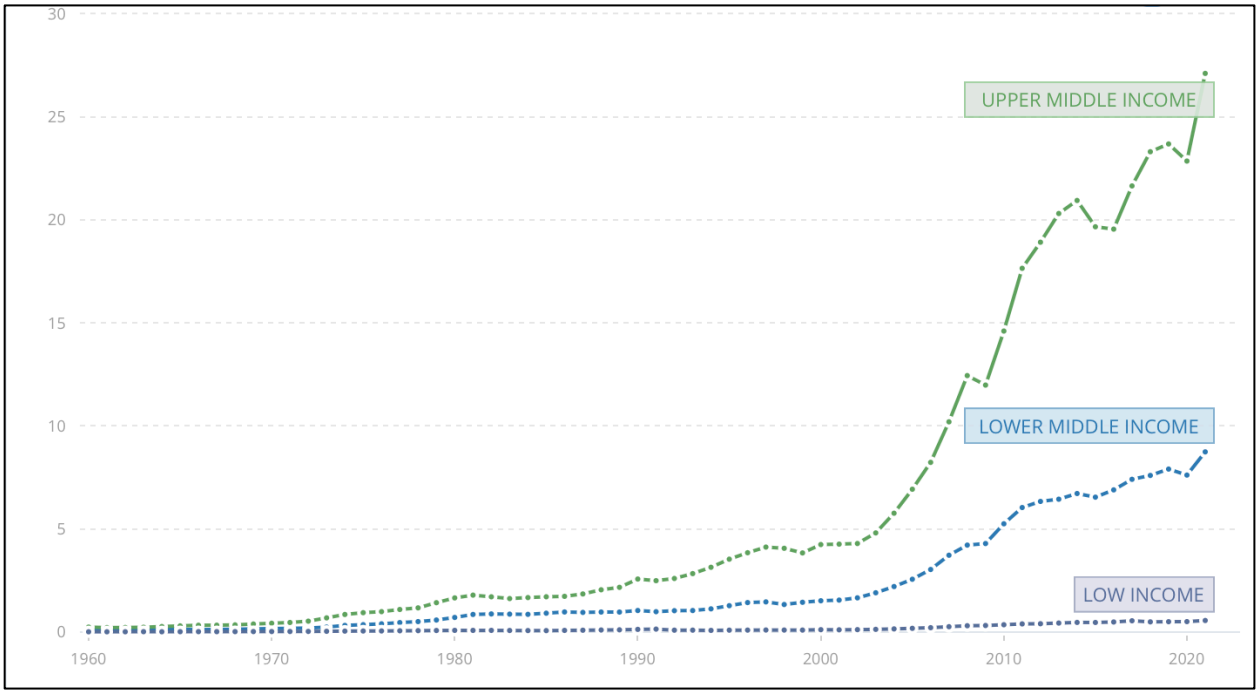

Figure 1. 10 – GDP in trillions of current U.S. $ from 1960 – 2021 for low and middle-income countries (World Bank, 2023)

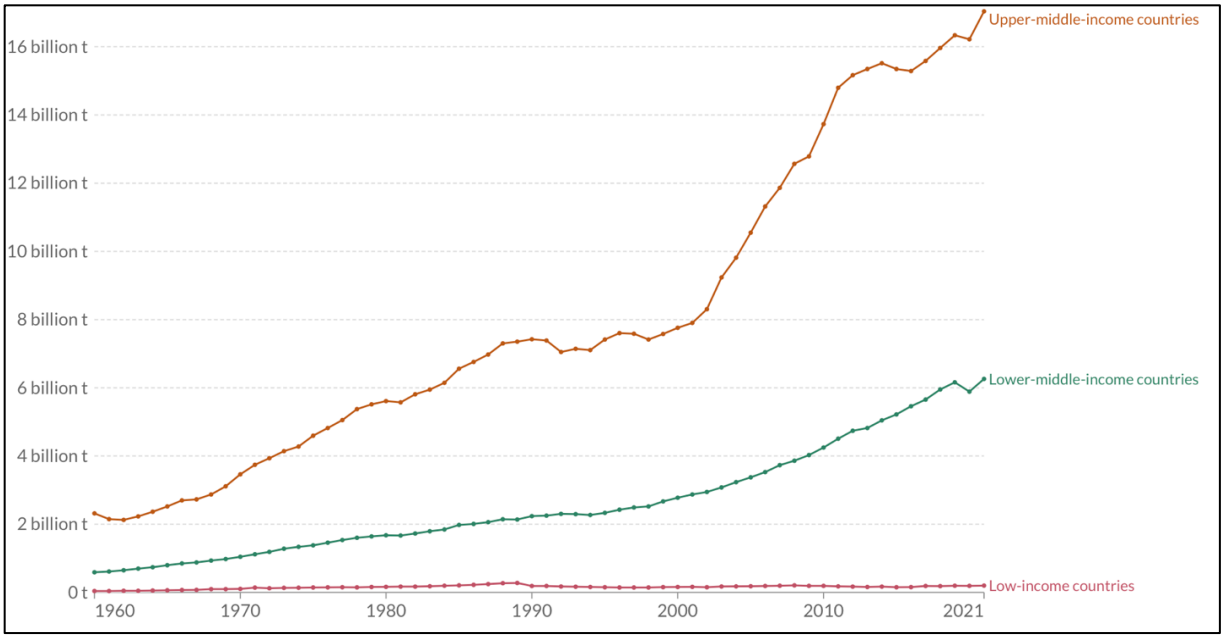

Figure 1. 11 - CO_2 emissions from middle and low income countries from 1960 - 2021 (Ritchie et al., 2020)

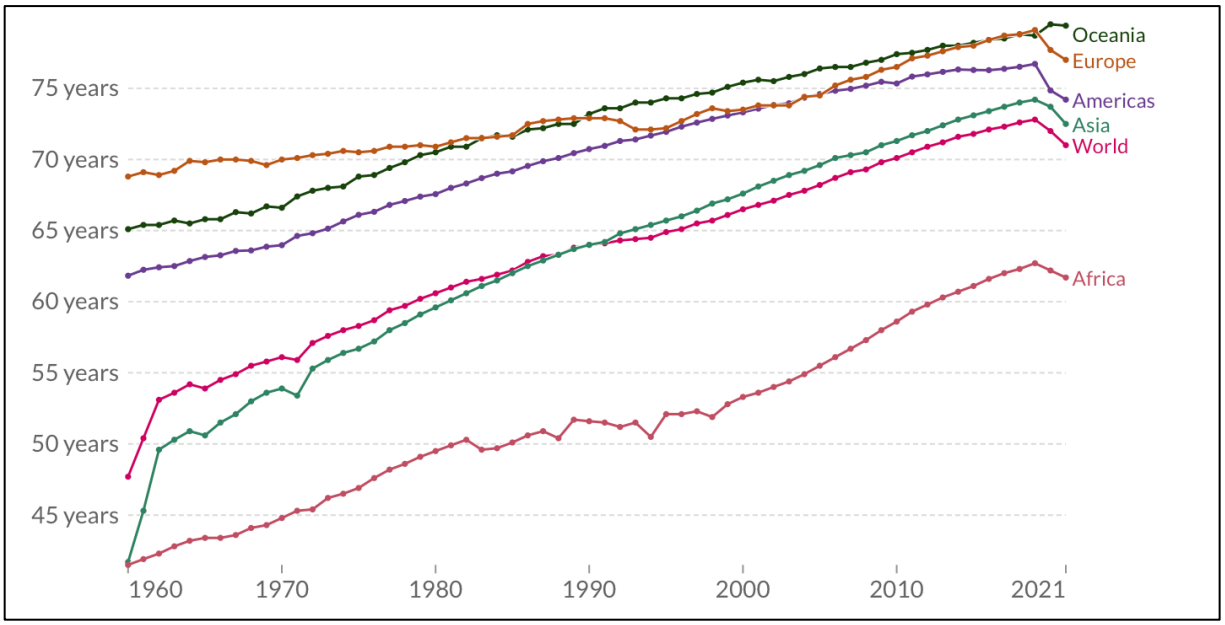

Figure 1. 12 - Life expectancy from 1960 to 2021 (Roser et al., 2013)

Figures 1.12 (and D.5 in Appendix D) show that life expectancy has also has largely increased worldwide, certainly since the 1950s. Now, referring back to CO_2 emissions in figure 1.7, it is not that CO_2 emissions are directly responsible for the improvement in life expectancies and GDP, the exogenous variation is the opportunity cost people are able to take advantage of when provided a cost effective and energy dense resource like O&G. Cheung et al. (2020) note that Hong Kong has the highest life expectancies in the world despite cross-boundary air pollution effects from the industrial center Pearl River Delta Economic Zone (PRDEZ) (Cheung, He, & Pan, 2020). The Cheung et al. (2020) study was unique in that it estimates the effects of high pollution in a high income setting (Cheung, He, & Pan, 2020). Their study provides evidence that the negative health effects from air pollution can be mitigated through more robust health care and high quality medical institutions (Cheung, He, & Pan, 2020). I would argue that this is evidence of the opportunity cost we are able to take advantage of when we're not spending time worrying about and procuring energy. Not only was Hong Kong able to provide quality health care and health institutions (which require considerable and sustained energy), but was also able to provide academics the opportunity to study these effects.

Figure 1.13 shows the average worldwide deaths from climate events over each decade from the 1960s through the 2010s and current estimates through the 2020s thus far. It's pretty clear that people are much safer today from climate events than in decades prior. This is even more impressive given that population has increased over the same time period.

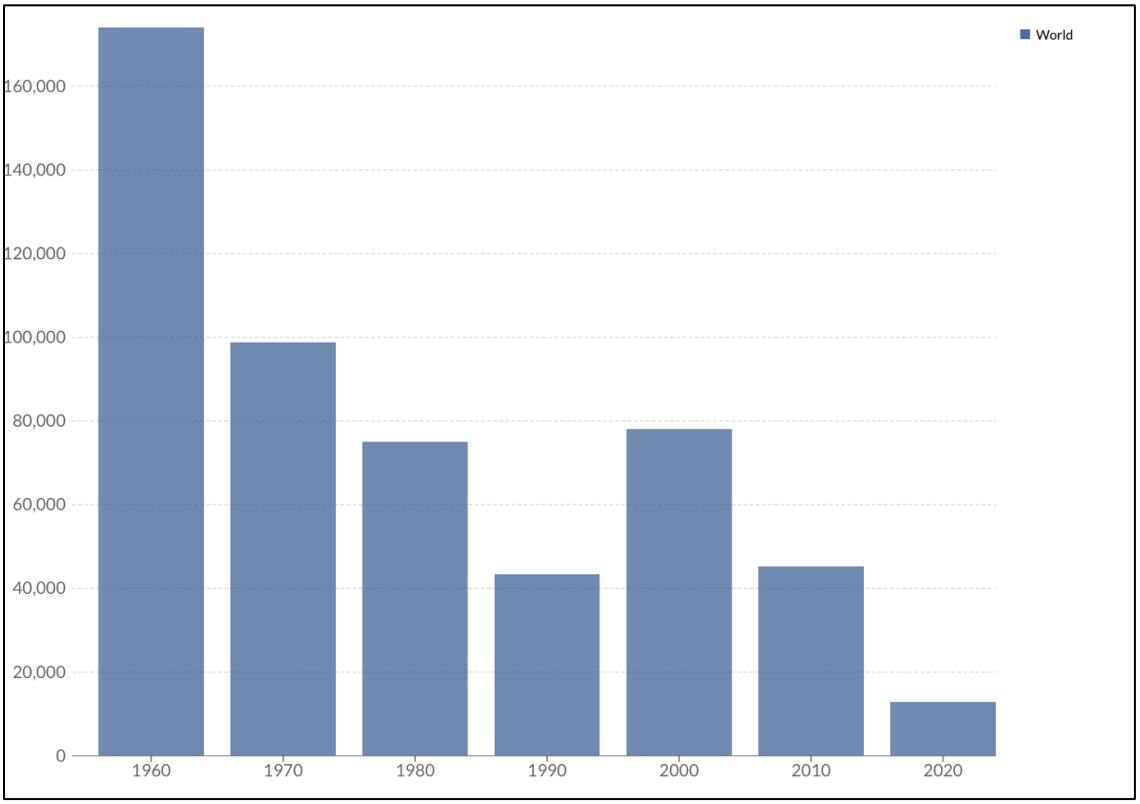

Figure 1. 13 - Decade average worldwide deaths due to geophysical, metoerological, and climate events including earthquakes, volcanoes, landslides, drought, wildfire, storms, and flooding from 1960s to 2010s and current through the 2020s (Ritchie et al., 2022)

Examining figures 1.7-1.13, it seems life has improved for a large majority of people on the planet despite the imposed SCC. These improvements are also in spite of any temporal effects of incurred SCC from CO_2 emitted in the past. It stands to reason that we should be seeing the SCC for CO_2 emitted in the middle of the century and paying for it daily. However, it seems the SCC is outweighed by the benefits of our anthropogenic activities that create a net benefit from the opportunity cost.

Fossil fuels have been the worlds primary energy source because they differ from renewables in that they can be readily supplied and at low cost (Borenstein & Kellogg, 2021). This has produced societies where seeking out a certain degree of discomfort is a

past time (e.g. camping) and it may even be argued that these are luxury goods given the time and resources needed to undertake these activities. This is because we're largely disassociated from the actual harshness of the environment and are no longer at its' mercy. This shift from surviving in the environment to actively seeking discomfort in the environment is a product of the activities we undertake that produce emissions.

This chapter has provided strong evidence, from increased GDP and life expectancy, to challenge the assumption that we should undertake activities that produce CO_2 emissions. Now, after having established a foundation and justification for re-development of O&G assets, the rest of the chapter will focus on economics of O&G production and re-fracturing.

1.4 Counteracting Peak Oil

As production naturally declines from O&G production, new technologies and processes are needed to counteract the natural reservoir pressure declines and replace extracted reserves. A prevailing ecological assumption is that resource scarcity is an inevitable consequence of sustained human growth on a finite planet (Seibert & Rees, 2021). More narrowly, in the O&G sector, Marion King Hubbert proposed an extraction profile in the 1950s resembling a "bell shaped" curve for the lower 48 states (Hubbert, 1962). Hubbert predicted peak U.S. oil production in the lower 48 would occur around the year 1970. Figure 1.14 shows the trend in U.S. oil production from 1920 to the late 2000s. We can see a peak around 1970 with the production following a bell-shaped curve just as Hubbert predicted. The nature of this bell-shaped production profile indicates there should be much debate on how resource extraction should be planned (Bardi & Lavacchi, 2009)

and those in charge of making decisions are either unaware of "peak production" or see it as unjustified (Bardi & Lavacchi, 2009). Figure 1.14 shows the bell shaped curve Hubbert predicted and his prediction of the peak year was extremely accurate.

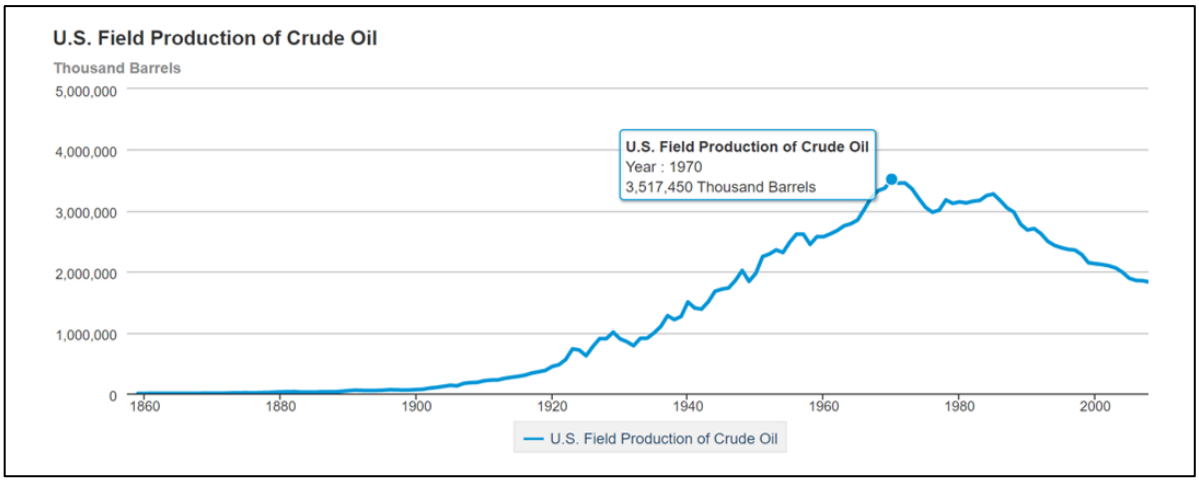

Figure 1. 14 – U.S. field production before the revolution highlighting (Energy Information Administration, 2023c)

The peak oil production described by Hubbert was the prevailing sentiment until roughly 2005. However, Hubbert neglected any advancement in technology, largely discounting the ingenuity and adaptability of humans. Widespread implementation of the advanced technologies of horizontal drilling and HF adopted 2005 reversed the course of O&G production in the U.S. (Barati & Alhubail, 2021). Figure 1.15 shows the most current data on U.S. crude oil production. We see that recent production has surpassed that of the peak predicted by Hubbert. In fact, Worldwide O&G reserves increased by more than 50% in 2009 (U.S. Energy Information Administration , February 2011).

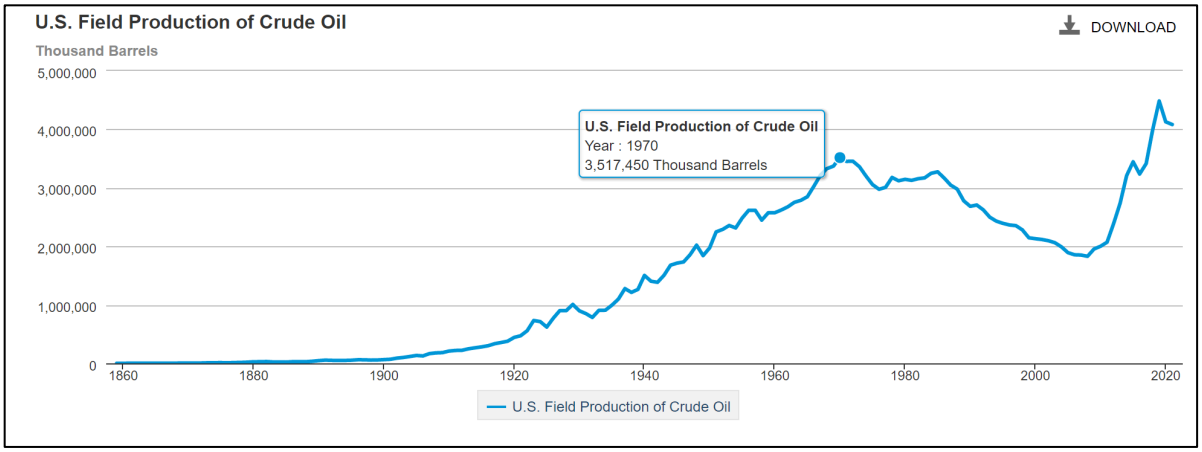

Figure 1. 15 - U.S. field production of O&G highlighting decline counteraction from (Energy Information Administration, 2023c)

In addition to added reserves, average finding costs for Financial Reporting System (FRS) companies worldwide decreased to $18.31 per barrel of oil equivalent (BOE) of oil reserves added in the 2007-2009 period. This represents a decline of $5.79 per BOE from 2008 (U.S. Energy Information Administration , February 2011). These figures are in spite of the fact that investment in production and development have been much higher than exploration since the early 1980s, as shown in figure 1.16. There is a market incentive to disseminate proprietary technology from those who developed it and rent it to others, which is often most profitable (Baumol, 2002). This indicates that we do not necessarily need to find more oil, we just need to get better at extracting and producing it through technology.

Futures prices and contracts are one of the best indications of resource scarcity available. Figure 1.16 indicates that with an increase in investment in production and development, production can be substantially increased, even with a relatively constant

rate of investment in exploration. Figure 1.16 is an example of what the economist William J. Baumol would call "routinized" innovation. Essentially, firms will not risk falling behind competitors and will make research and development an internal and routine procedure (Baumol, 2002).

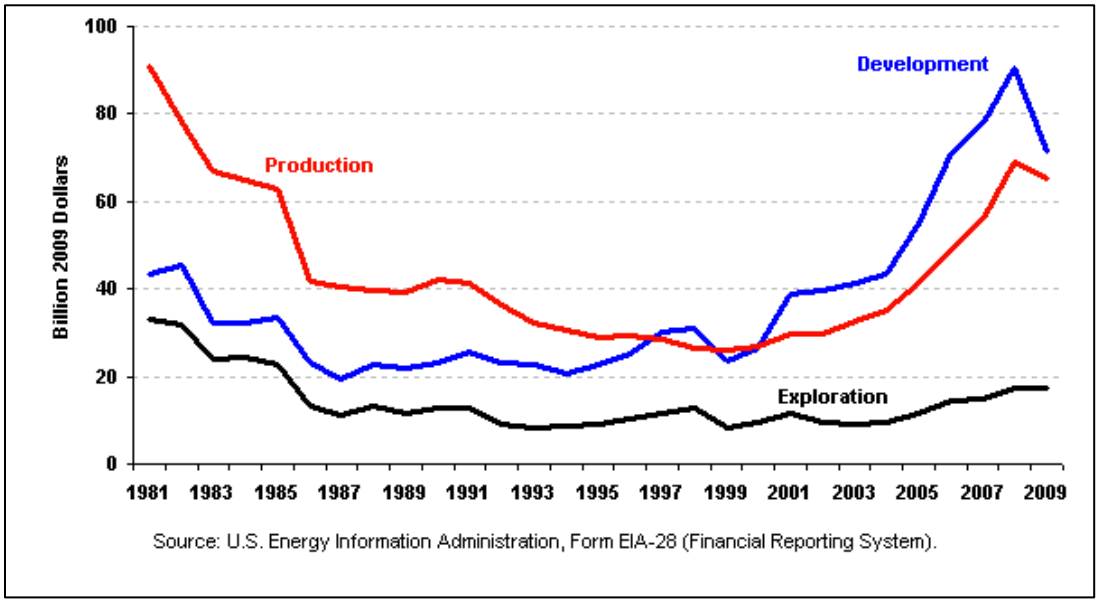

Figure 1. 16 - Worldwide expenditure on Production, Exploration, and Development (Energy Information Administration, n.d.)

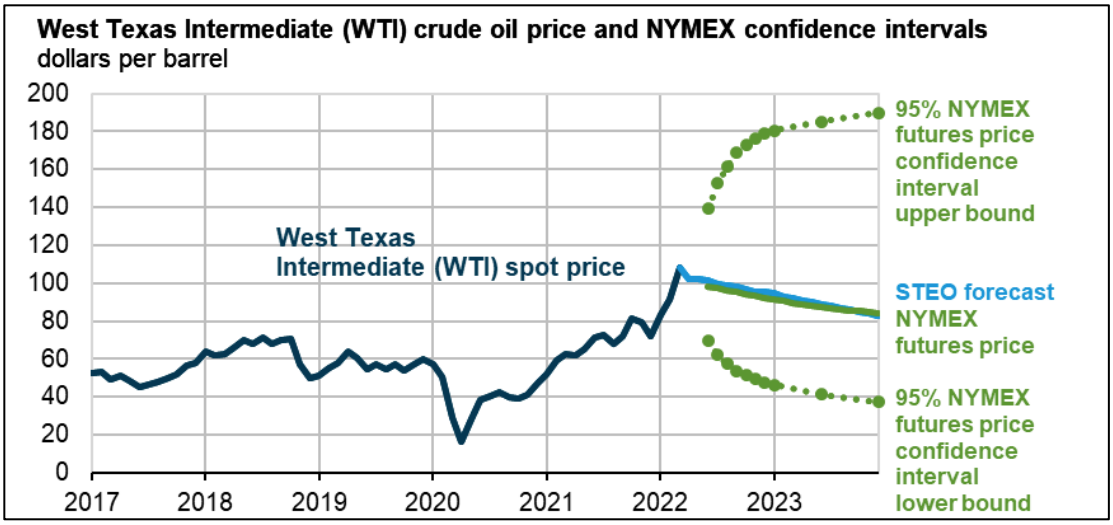

Figure 1. 17 – West Texas Intermediate and NYMEX crude oil futures prices with confidence intervals (Energy Information Administration, 2023d)

1.5 Re-fracturing: Potential to Recover Stranded Resources in the Williston Basin

Given the need for continued O&G production and further development of the resource given in the previous sections, redevelopment of O&G assets is essential. There is also sufficient evidence that the world is not running out of O&G, we just need to get better at extracting it. Additionally, the largest portion of the emissions reductions in the U.S. since 2005 have been realized due to substitution to of coal for natural gas for electricity generation, as shown in figure 1.4. This indicates that continued development of O&G and continued substitution of natural gas for electricity generation will not only have positive impacts on human well-being, but also environmental benefits in terms of emissions reductions.

Combining the technological developments of horizontal drilling and HF has not only made the U.S. energy independent in terms of O&G but has also made the country an

exporter of these resources (Rignol & Bui, 2020). Likely due to economic conditions, stimulation until around 2015 largely consisted of sliding-sleeve/packer completion types. These consisted of a series of sleeves that would activate a shear pin given a large enough pressure differential after a particular-size composite ball would seat and seal off zones below it, and thus exposing new perforations in the liner to allows entry into formation in a new zone. With such high oil prices from 2010-2015, this technology made it possible to extract large quantities oil as quickly as possible.

There were almost certainly information asymmetries in the early days of HF due to its' novelty. Combine this with an incentive to extract as much oil as quickly as possible and there exists a recipe for stranded resources (Barba, Allison, & Villarreal, 2022). This presents an opportunity for operators to redevelop assets through re-fracturing (Li, Han, LaFollette, & Kotov, 2016). Re-fractured wells in the Bakken have seen an increase in production rate and the estimated recoverable reserves (Wan, Rasouli, Damjanac, Torres, & Qiu, 2019). Rignol and Bui (2020) estimate production from re-fractured wells to be as high as 92% of initial production in the Bakken which is most likely due to insufficient initial stimulation (Rignol & Bui, 2020). Figure 1.18 shows the production gains specific to the Bakken after analyzing re-fracturing data captured from FracFocus (Shammam et al., 2021). Shammam et al. (2021) estimate production gains as high as 160% of original production in the Bakken, conditioned on proper well selection. This provides evidence that there is opportunity in the Bakken to a make economic wells by redeveloping current assets and forgoing the cost of new drills.

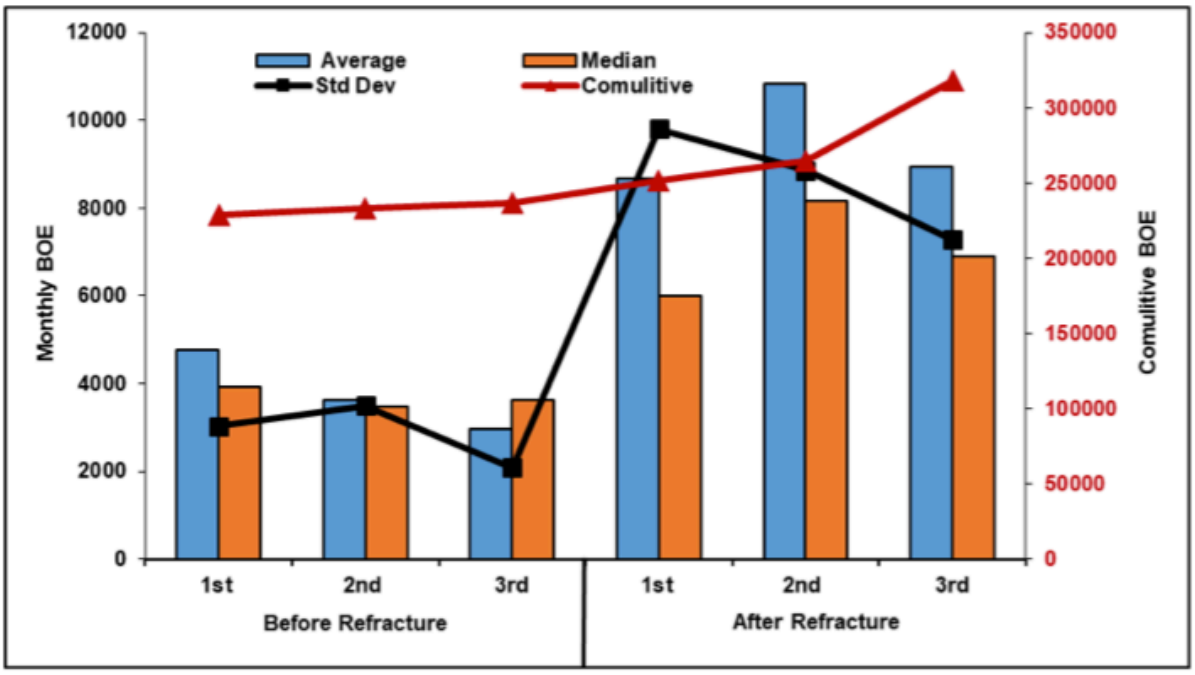

Figure 1. 18 - Expected production gains in Bakken after re-fracturing treatment (Shammam et al., 2021).

However, given the potential of re-fractured wells, the re-fracturing process is even more complex than initial stimulation operations. Factors such as wellbore degradation, depletion, possible stress reversal due to depletion, etc. all play a role in treatment implementation. Although work has been done to investigate fracture models based on physics to optimize production, there has been little done with a focus on treatment implementation in the field to reduce costs. Also, there is a gap in literature for providing marginal estimates for the effects of treatment parameters in the Williston basin for re-fractured wells. With this, there exists a need to develop data driven models, guided by field experience and actual treatment data, to provide insights into what knobs to turn to make treatment easier to implement in the field and reduce costs. With many information asymmetries working against petroleum engineers, and often complex relationships that are often times unobservable, this study will use economic models in an attempt to

describe these relationships. These models identify the high margin areas where the most efficient reductions can be made. These economic models are also appropriate since economists often deal with the same issues of lack of knowledge about entire populations and their actions, but still require data driven decisions.

1.6 Hypothesis and Research Questions

The continued need for O&G extraction combined with market forces has prompted operators to focus on minimizing costs while producing O&G. This is because technological advancements have been made and the economics of the industry has changed. However, the process of re-fracturing is complicated and requires a different approach from initial stimulation since expected pressures will be different and pressure limitations have changed due to wellbore degradation.

1.6.1 Hypothesis

So, this study will state the following hypothesis and set of accompanying research questions. The stated hypothesis is as follows:

Application of multivariate regression models can identify high margin areas for surface treatment pressure (STP) reduction and lead to a more cost-efficient hydraulic fracture implementation.

In order to answer this, the following research questions that will be investigated. These are important as they will guide the research and provide other insights into the re-fracturing process.

1.6.2 Research Question 1

Can multivariate regression models yield inferences about treatment behaviour?

It is important to identify marginal effects that will help us to better understand how changes in completion parameters effect treatment implementation in terms of average STP. Factors such as stress shadow effect, perforations, proppant, etc. may drive up average STP.

1.6.3 Research Question 2

Can more complex models resolve endogeneity and increase predictive power of average STP?

This will go a step further in identifying the causal relationship amongst distance between stages and previous stage average STP.

1.6.4 Research Question 3

Can we use the lessons learned from the inferential models to accurately predict average STP?

This is important because it will be the culmination of the previous models and, ultimately, test whether thy hypothesis can be answered. How can these predictive models be used to change parameters that affect treatment implementation and cost?

1.7 Methodology

This project will utilize the following approaches guided by the hypothesis and research questions to investigate if average STP can be characterized by general completion parameters and can be predicted from these parameters:

1. Obtain field treatment data
2. Use domain knowledge about HF treatments, wellbore mechanics, field operations, and literature and incorporate it into model specification. Domain knowledge will be helpful when selecting independent variables for the regressions as well as building other models to answer more complicated questions.
3. Construct pooled dynamic multivariate regression model as an initial inferential analysis. The useful inferences from this model are marginal effects of completion parameters on average STP. This will allow operators and field personnel to identify those parameters that are most likely to affect average STP during re-fracturing implementation. This pertains to research question 1 and will be the first step in the process of understanding marginal effects of treatment parameters and temporal stress shadow effects.
4. Construct dynamic panel model accounting for within-well differences. This model will also serve as inferential and build upon the pooled model. This model should serve as a guide to encroach on high margin areas and retreat from low margin areas when optimizing and predicting average STP. This model is related to research question 2 and will help solidify our understanding of stress shadow effects.
5. Perform feature engineering using guidance from the pooled and fixed-effects models to improve predictive potential of multivariate models. This will help

answer the overall hypothesis of the study, incorporating the knowledge gained from the previous models.

1.8 Significance

The significance of this research is to construct models that estimate the effects of completion parameters on treatment implementation. This will be useful as a cost minimization tool for operators. This will include the following objectives:

1. Introduction of simple and useful multivariate regression models and approaches and much simpler statistical models to use in the highly stochastic systems of the wellbore and reservoir;

2. Provide a repeatable framework for operators in the Williston basin using R (R Core Team, 2022. https://www.R-project.org/) to provide causal analysis in determining the completion parameters that affect average STP. This approach will hopefully scale to other basins as basic HF treatment data is available;

3. Present an approach to identify causal elements that affect average STP. This is important because it gives engineers a tool to control that parameters that can be controlled in a treatment and increases the probability of treatment optimization from a cost perspective;

4. Provide a feature engineered model that can use the inputs that have a causal effect on average STP to predict on other data sets. This will be useful when designing new treatments;

5. Provide a reproducible workflow that operators can apply on smaller scale that is cost effective and only requires prior or offset treatment data;

6. Provide insight into why average STP is crucial to operations and how costs can be minimized by focusing on average STP under operational conditions;

1.9 Dissertation Structure

This dissertation consists of six chapters.

Chapter 1 will start with a brief economic and philosophical analysis to support the a priori assumption that continued and optimal O&G extraction is and will continue to be beneficial to human progression, therefore creating a need to re-fracture wells. Chapter 1 also will provide a brief introduction to the project, the hypothesis, research questions, methodology, and glossary for this study.

Chapter 2 will provide a thorough investigation of the current literature on operations, machine learning methods for optimizing O&G extraction, and a thorough introduction to the multivariate regression models used in this study.

Chapter 3 will introduce a pooled multivariate regression model for estimating the marginal effects of completions parameters on average STP. This model also provides insight into the temporal dependence of STP from previous stages on subsequent stages.

Chapter 4 will build on the model constructed in Chapter 3 by accounting for unobservable, within-well fixed effects through the use of panel data and fixed effects multivariate regression models.

Chapter 5 will use the results from the multivariate regressions from chapters 3 and 4 as a base to perform feature engineering on the models to improve predictive performance of average STP.

In Chapter 6, a list of conclusions will be provided as well as discussion about field applications for cost minimization and areas for improvement in the process.

1.10 Glossary

Multivariate Regression – Ordinary least squares that includes multiple predictor variables (Bailey, 2017)

Added Variable Plot – Plot highlighting individual effects, all others held constant

Opportunity Cost – The cost incurred from an opportunity forgone

Endogenous – Changes in a predictor variable are related to changes in the error term (Bailey, 2017)

Predictor (Independent) Variable – A variable we have reason to believe influences the dependent variable in a regression (Bailey, 2017)

Dependent Variable – An outcome variable we are interested in modelling (Bailey, 2017)

Fixed Effects Model – Model utilizing panel data to account for unit and/or time effects unique to each unit and/or period in the dependent variable (Bailey, 2017)

Margin (Marginal) – The next or last unit of analysis

Residual – The difference between an observed value and a predicted value (Bailey, 2017)

Statistical Significance – We will say a coefficient is statistically significant if we can reject the null hypothesis that the coefficient estimate is zero (Bailey, 2017)

Standard Error (coefficient) – The standard error of a coefficient estimate refers to how wide the distribution of the parameter estimate is (Bailey, 2017)

Standard Error (regression) – The standard error of the regression is defined as the square root of the variance and is therefore measured in the units of the dependent variable. Essentially, it is the average distance between the fitted values and the observed values.

Causality – If changes in an independent variable, X, increase the probability that the dependent variable, Y, will change as a result, we assume there is a causal relationship.

Inferential Models – Models used to gain insight into relationships as opposed to solely making predictions

Chapter 2

2. Literature Review

2.1 The Bakken Petroleum System

Chapter 1 outlined an argument for continued O&G extraction and use. This argument provides a foundation for the entire study since it is focused on re-developing existing O&G assets through re-fracturing. This section will start with a review of the Bakken petroleum system. This is followed by a review of the HF process including discussion about operations and equipment. Next, an analysis of literature pertaining to data mining and domain knowledge. Lastly, a review of multivariate regression models and some guiding principles in the development of the models conducted for this study will be discussed. The objective for this chapter is to provide a broad and sufficient background discussion for the reader about the information pertaining to the objectives of this dissertation.

The Williston basin is an unconventional shale play that spans several states and reaches up into Canada. Specifically, it is located in northwestern SD, eastern Montana, most of North Dakota, Manitoba and Saskatchewan (Gerhard & Anderson, 1988). Figure 2.2 shows the extent the Bakken reaches in each direction. The Williston basin was known to be oil bearing since the 1950s when Amerada struck economically viable quantities of

oil south of Tioga, ND in the Silurian Interlake Formation with following development targeted in the Mississippian Madison Group (Ling, et al., 2013). Recently, the widescale implementation horizontal drilling augmented with HF in the 2000s has provided a measurable increase in production and development of the Williston basin (Fry & Paterniti, 2014; Ling, et al., 2013). By 2018, the Bakken formation was producing about 1.2 MM barrels per day (bpd) (Wan et al., 2018). According to the EIA, the Bakken is currently producing around 1,200 thousand bpd and has seen a substantial increase in daily production since 2013 (EIA, 2022). These results are shown in figure 2.1.

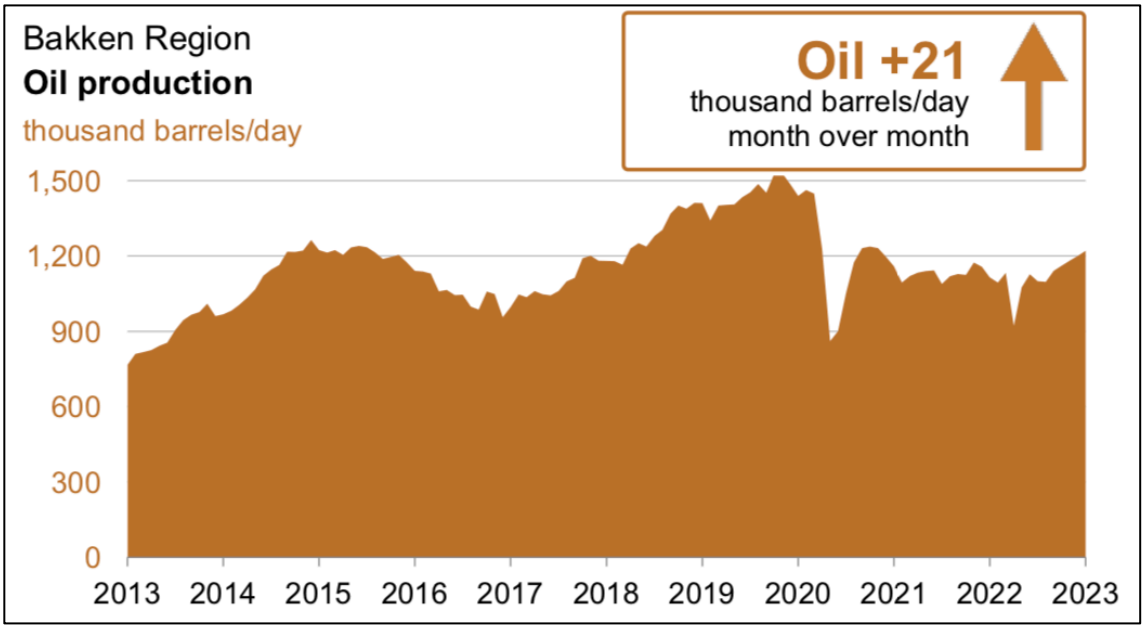

Figure 2. 1 - Recent oil production trends for the Bakken region (EIA, 2022)

Figure 2. 2 - Map of the Williston Basin (Gerhard & Anderson, 1988).

The Bakken reaches its' maximum thickness near the center of the basin and pinches out towards the edges of the basin to near zero (Smith & Bustin, 1995). The Bakken is thickest around Watford City at around 145 feet (Fry & Paterniti, 2014). At the deeper, center part of the basin, the Bakken overlies the Three Forks formation conformably and becomes unconformable towards the margins of the basin as the formations pinch out (Fry & Paterniti, 2014). Because of its' structure, there are very few outcrops of the Bakken that are available for study so most of the subsurface knowledge comes from petrophysical logs, well logging data, core samples, and seismic surveys (Gerhard & Anderson, 1988). Figure 2.3 shows the shape of the Bakken petroleum system from a profile view. We can see from figure 2.3 that the Bakken pinches out towards the east and the west and the maximum thickness mentioned earlier is near the Nesson Anticline (center of the basin).

Figure 2.4 shows the stratigraphic column for the Williston basin with oil bearing formations in blue and gas bearing formations in red (Ling et al., 2013). Figure 2.5 highlights the Bakken system specifically in the stratigraphic column.

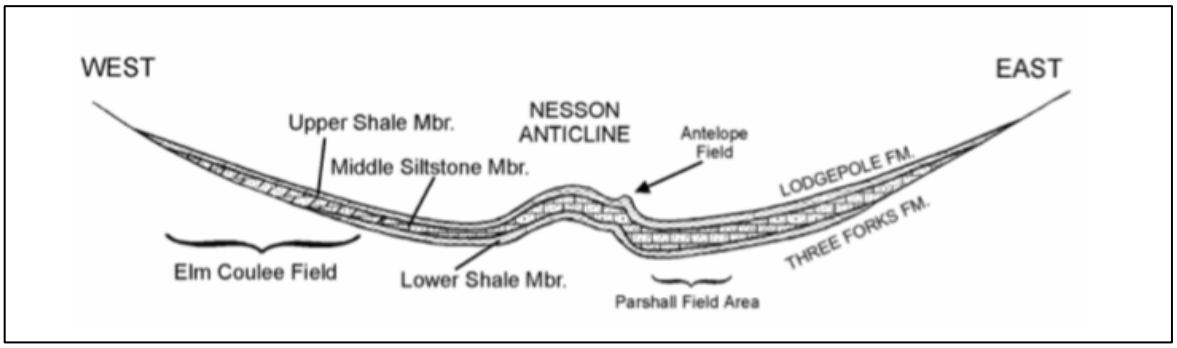

Figure 2. 3 - Cross sectional view of Bakken petroleum system (Meissner, 1978).

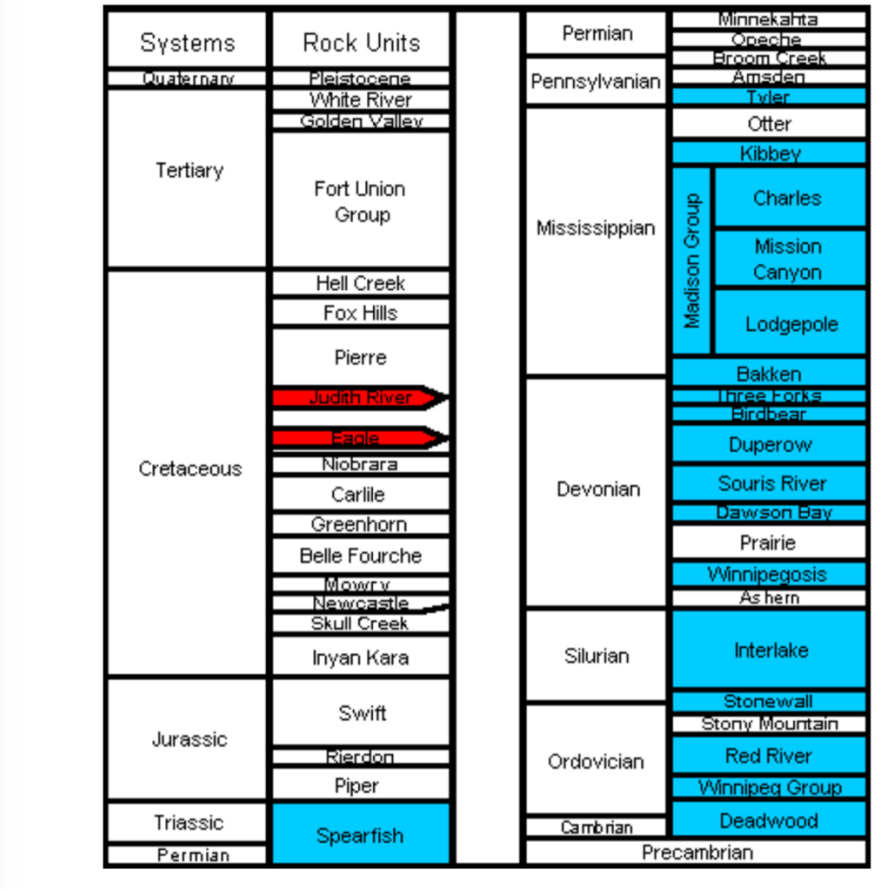

Figure 2. 4 - Stratigraphic column of the Williston basin (Heck et al., 2002)

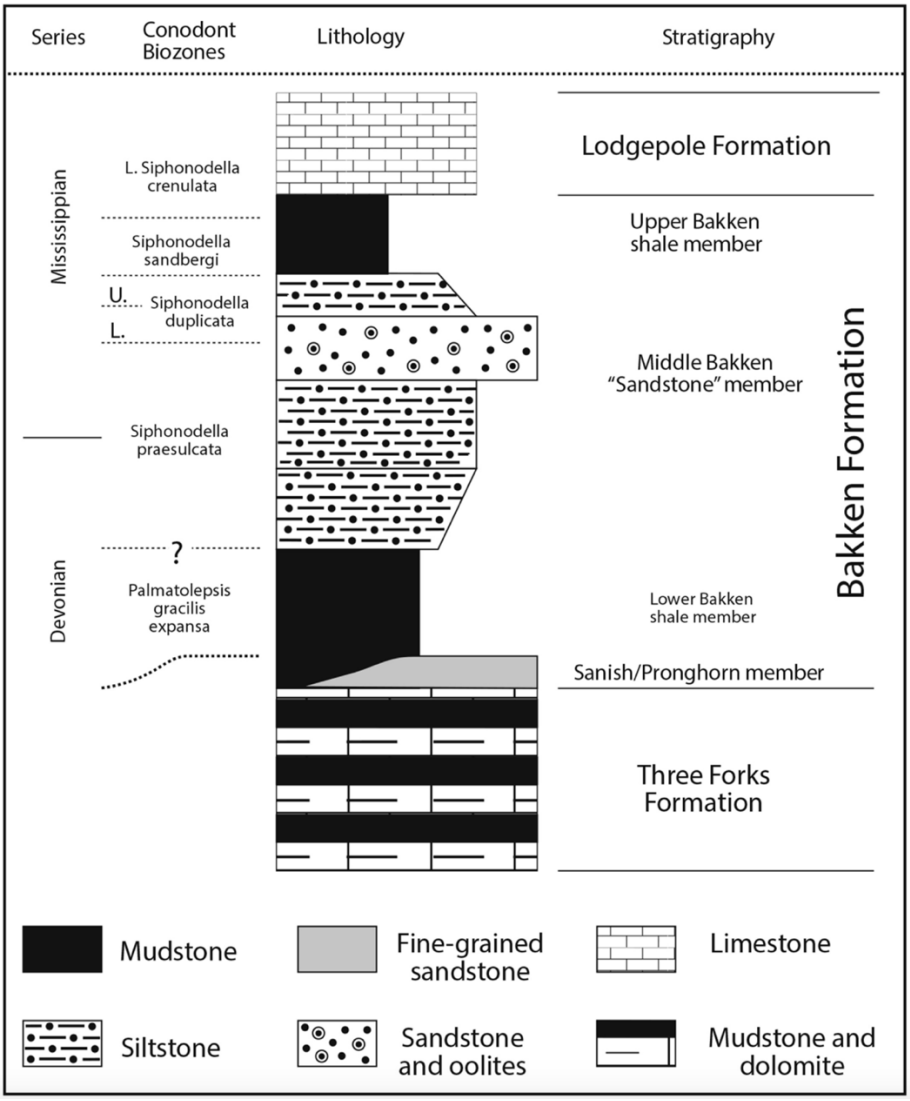

Figure 2. 5 - Stratigraphic column focusing on the four Bakken members including the Sanish/Pronghorn, Lower Bakken, Middle Bakken, and Upper Bakken (Egenhoff, 2017)

The Bakken system consists of four members: Upper Bakken, Middle Bakken, Lower Bakken, and the Sanish/Pronghorn member (Egenhoff, 2017). The upper and lower members are dark, organic rich source rocks for the middle member, which is a mixture of sandstone and mudstone (Ling et al., 2013; Wan et al., 2019). The upper and lower memebrs also serve as a seal for the middle member (Fry & Paterniti, 2014; Ling et al.,

2013) with the upper member being deposited under dynamic and occasional dynamic conditions (Egenhoff & Fishman, 2013). The lower and middle member contains a thin carbonate transition unit as opposed to a grain coarsening gradation (Egenhoff, 2017). These processes have had causal impacts on the rock characteristics that determine how they store fluid and how they may react to HF (Fishman, Egenhoff, Boehlke, & Lowers, 2015). This suggests that the Bakken system has undergone two encompassing sea-level-oscillations with the first represented by the carbonate boundary present in the lower to middle Bakken transition and the second marked by the transition from the middle Bakken to the upper Bakken shale (Egenhoff, 2017).

2.2 Hydraulic Fracturing Operations and Treatment

Unconventional reservoirs such as the Bakken require HF to realize sufficient production to make them economical (Manchanda & Sharma, 2012). Horizontal wells are now used in conjunction with multi-stage, multi-fracture treatments to successfully produce from low permeability formations (Fry & Paterniti, 2014; Havens & Batzle, 2011; Ling, et al., 2013; Roussel & Sharma, 2011; Wan et al, 2018). Successfully generating multiple fractures in each stage of a horizontal well is necessary for sufficient production (Abobobaker& Olson, 2015; Havens & Batzle, 2011). To achieve this, special equipment is required.

Equipment can be divided into four categories which are transport, servicing equipment, iron and pipe equipment, and pressure pumping equipment (Josifovic et al., 2016). As the targeted formations have become deeper and higher pressure, positive displacement pump design has travelled together with pressure and rate requirements to accommodate these formations (Josifovic et al., 2016). Positive displacement pumps used

for HF are able to operate over a broad range of pump rates at high pressure over long periods of time (Josifovic et al., 2016). The physical process of HF is carried out operationally by moving fluid from a static source, as shown in figure D.1 (Appendix D), through equipment that increases in pressure throughout the process to eventually move the fluid into the wellbore at stimulation pressure. Figure 2.6 shows the general layout for a HF operation. Operationally, there are two different circuits for low and high pressure fluid between the pumps and the missile or monoline (Josifovic et al., 2016). A blender is required to pull fluid from holding tanks (D.1, Appendix D) via a centrifugal pump, where it is mixed with chemicals and sand to create a slurry. The slurry is then discharged via the blender to a monoline or missile that all of the HF pumps are connected to. Figure D.2 and D.2 in Appendix D show pictures of a monoline with connected HF pumps. Therefore, the pressure at the discharge side of the blender becomes the low pressure side of the positive displacement pumps used for HF. The pressure at this point ranges from roughly 60 psi to 100 psi depending on the designed slurry rate and is called the blender discharge pressure. Here, the positive displacement HF pumps will convert the fluid at blender discharge pressure to a stimulation pressure known as surface treating pressure (STP). This is the pressure monitored at the wellhead during treatment.

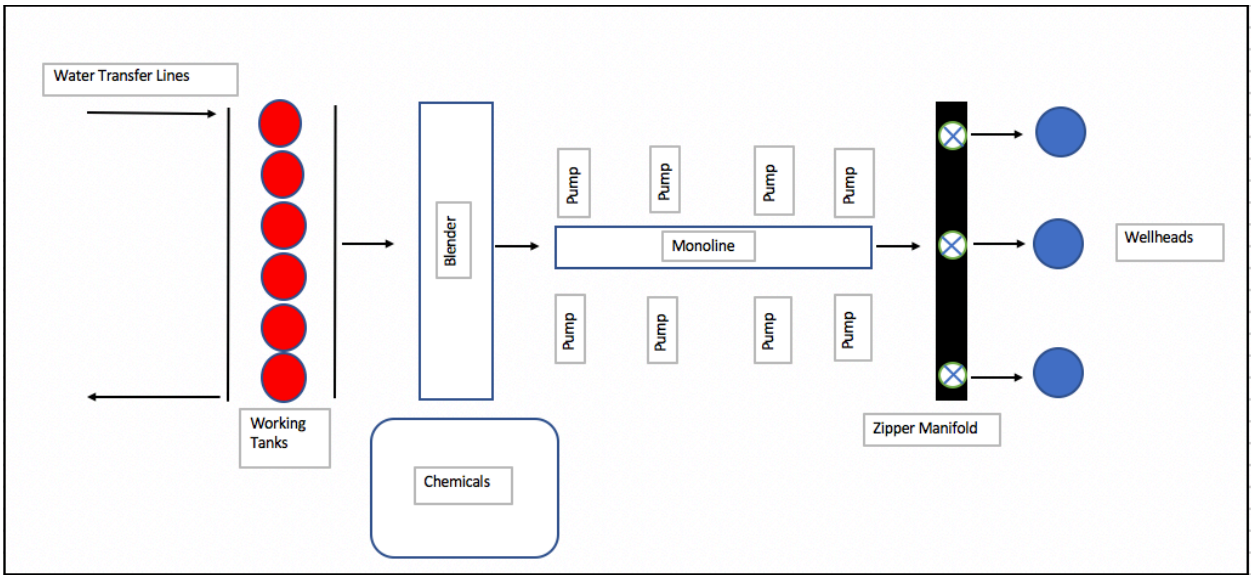

Figure 2. 6 - General layout for operations

Although there are numerous suppliers of positive displacement HF pumps, there is surprising homogeneity amongst pump design (Josifovic et al., 2016). Each positive displacement pump is powered by a diesel engine connected to a transmission that drives the crankshaft of the fluid end of the pump (Josifovic et al., 2016). A fluid end shown in figure D.4 (Appendix D). All of these pieces are mounted on a single trailer to make the transportation process easier (Josifovic et al., 2016). All of this culminates in delivering a treatment to formation to hydraulically stimulate the reservoir.

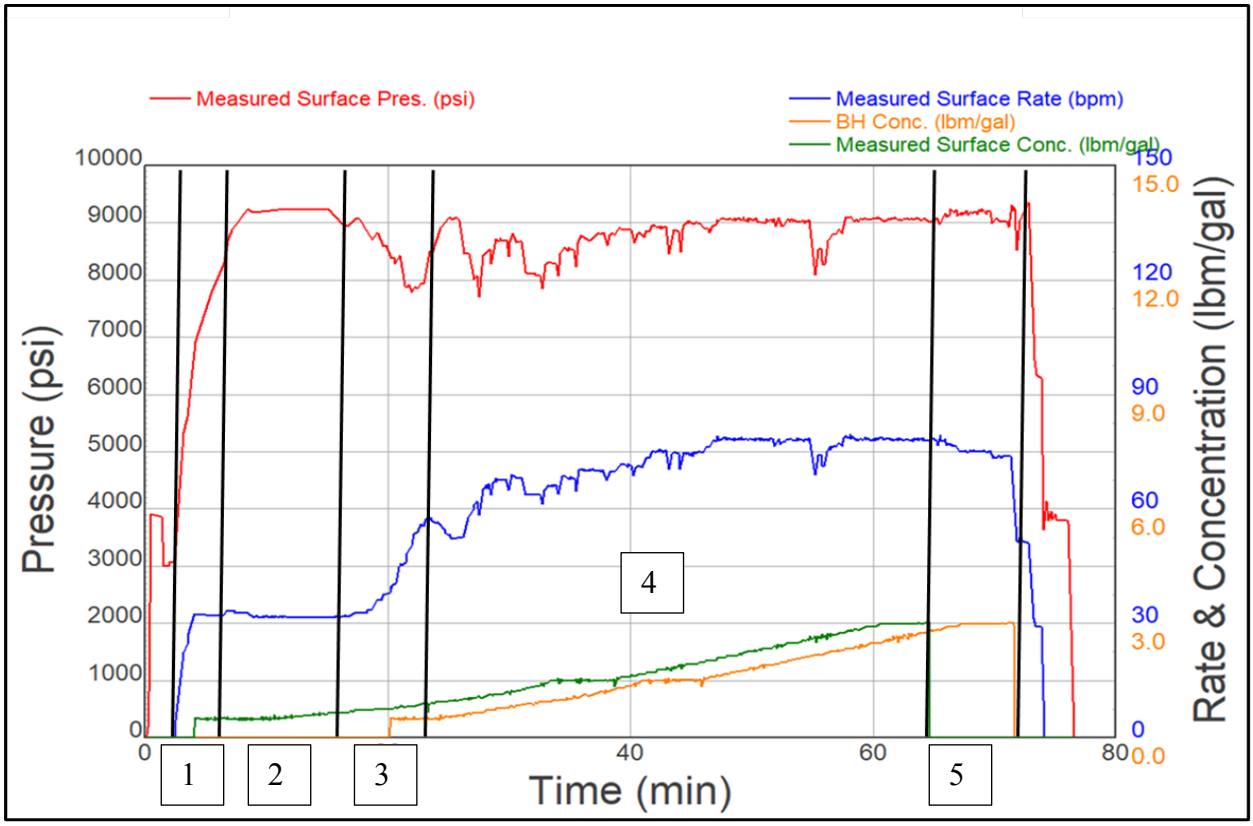

Figure 2. 7 – Treatment plot from actual HF treatment

Figure 2.7 is an actual treatment plot which shows the culmination of operations and theory. Section 1 is known as the breakdown phase. Lower rates are generally required during formation breakdown, before rate is established and proppant is started (Josifovic et al., 2016). During this phase, hydraulic fractures initiate from coalescing tensile failures when fracturing fluid pressure is increased above the local least principle stress (Zoback, 2010).

Section 2 is a wellbore displacement of acid to clean up perforations and near wellbore damage. Numerous skin factors from operations culminate in a total skin factor which has

implications for well productivity (Abobaker et al., 2021). Skin factor is a dimensionless production efficiency estimate where positive skin values indicate resistance to flow and negative skin values indicate flow enhancement (Schlumberger, 2023). Section 3 shows the acid, in combination with sand, cleaning up any near wellbore damage as indicated by the decrease in STP.

Section 4 in figure 2.7 shows the fracture propagation phase of the treatment. The process of hydraulically fracturing formations extends tensile fractures when fracturing fluids are pumped at pressures above the least principle stress and are then filled with proppant in an attempt to maintain the fracture dimensions (Zoback, 2010). This process leverages the fracture networks that are created to increase the stimulated reservoir volume (SRV) through micro and macro channels (Barati & Alhubail, 2021). Section 5 in figure 2.7 is known as a wellbore flush and is used to clean up the wellbore after treatment and before any further operations.

2.3 Wellbore Dynamics to Derive Average STP

Figure 2.8 shows a general wellbore schematic with the associated pressures during treatment. From this, we can derive the equation for STP that will serve as the basis for independent variable selection for the models constructed in this study.

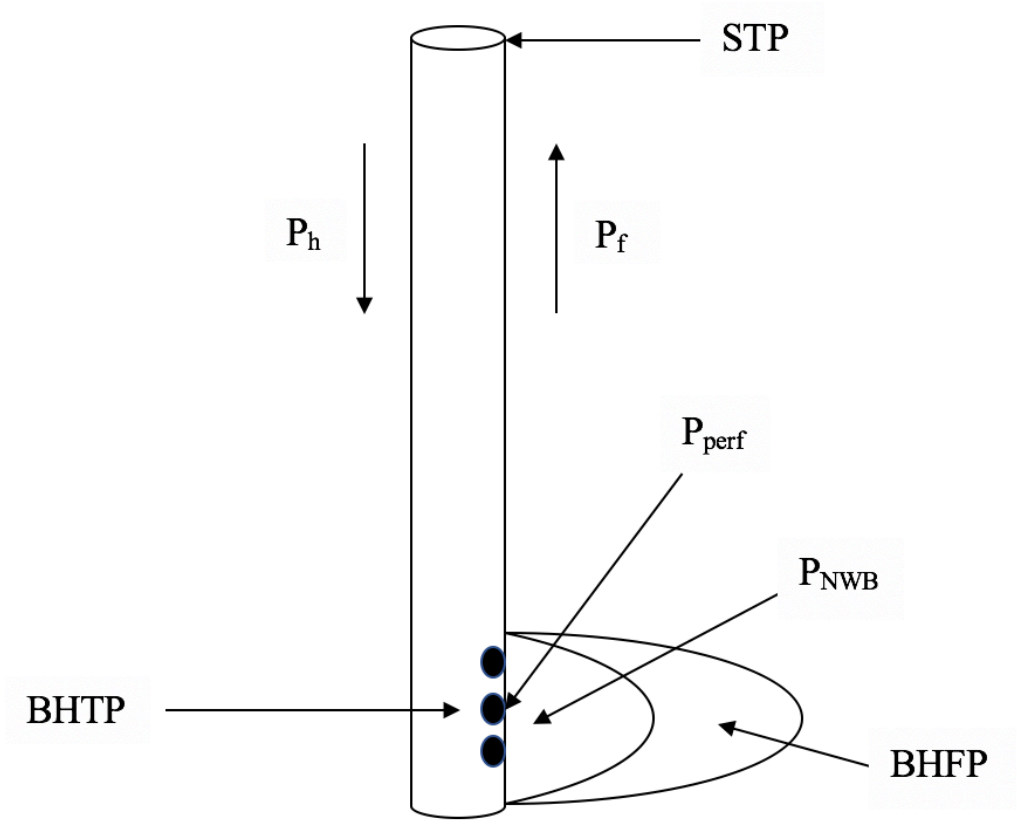

Figure 2. 8 - Schematic showing the relationship between STP and other wellbore pressures

The variables in figure 2.8 are defined as the following with the associated units:

STP = Surface Treating Pressure (psi)

BHTP = Bottom Hole Treating Pressure (psi)

BHFP = Bottom Hole Fracturing Pressure (psi)

P_h = Hydrostatic Pressure (psi)

P_f = Pipe Friction Pressure, (psi)

P_{perf} = Perforation Friction Pressure (psi)

P_{NWB} = Near Wellbore Friction Pressure, (psi)

Now, we can derive an equation for BHFP, BHTP, and ultimately STP which will be the primary dependent variable for the multivariate regression models. Equation 1 shows the equation for BHFP accounting for perforation friction and any NWB friction effects.

$$BHFP = BHTP - P_{perf} - P_{NWB} \quad (1)$$

Next, we can define the BHTP as follows with down as positive effects ad up as negative effects:

$$BHTP = STP + P_h - P_f \quad (2)$$

So, substituting equation 2 into 1, we obtain the following for BHFP:

$$BHFP = STP + P_h - P_f - P_{perf} - P_{NWB} \quad (3)$$

Rearranging equation 3 for STP, we eventually find the following relationship for STP to other wellbore factors occurring during treatment:

$$STP = BHFP - P_h + P_f + P_{perf} + P_{NWB} \quad (4)$$

Equation 5 shows the standard equation in the industry for estimating perforation friction (Gustavo et al., 2016):

$$P_{perf} = \frac{0.2369Q^2\rho}{N_p^2 D_p^4 C_d^2} \quad (5)$$

Where,

Q = Flow rate, volume/time, (bbl/min)

ρ = Fluid density, mass/volume (lb/gal)

N_p = Number of perforations, count

D_p = Diameter of perforations, L, (inches)

C_d = Coefficient of discharge, unitless

Using equations 4 and 5, hypothesis will be formed about how certain completion parameters will affect STP for the re-fracturing process. This domain knowledge about wellbore dynamics will serve as the basis for independent variable selection in chapter 3.

2.4. Data Mining in the O&G Industry

Data mining techniques are good at finding relationships and making predictions under stable conditions based off these relationships, but poorly equipped to explain why the relationship exists (Huntington-Klein, 2022). Here it is useful to determine whether a model is inferential or will be used for predictive purposes. In model construction, parsimony is a primary goal, especially for models that will be used for inference (Kuhn & Johnson, 2020). Although complexity can be a solution to inadequate accuracy (Kuhn & Johnson, 2020), more complex models are not necessarily likely to be more accurate, but they are more likely to contain mistakes (Bailey, 2017).

There are many variables that we are uncertain about in O&G reservoirs and the HF process. However, this does not mean these variables or events are random. We are

simply uncertain about them which is not the same as randomness. Randomness deals with probability and uncertainty stems from information asymmetry (Çambel, 1993). If a process were entirely random or consisted of truly random events, there would no meaningful time series (Çambel, 1993). However, decision making is all but moot in a completely deterministic system because the conclusion is forgone (Çambel, 1993). So, it is obvious that models in the O&G space operate somewhere between these two extremes. The fact that O&G reservoirs are non-deterministic and suffer from degrees of randomness and uncertainty need not prove itself. This can be thought of as a first principle: O&G reservoirs are non-deterministic systems and need not be characterized by deterministic models.

Data mining techniques are seeing widespread implementation throughout the O&G industry. Data mining is the large sphere that encompasses machine learning, data science, and artificial intelligence. Data science projects formalize data from a given domain into a mathematical summary or model and make data driven decisions (Bangert, 2021; Ramirez & Iriarte, 2019). These ML models provide powerful tools for understanding complex, often non-linear relationships (Cross et al., 2021). Due to the large amount of data and a general lack of understanding of the underlying physics of unconventional reservoirs, data-driven techniques may be better suited to deal with these complexities and information assymetries (Darabi et al., 2020). It's important to remember that "non-linear" has multiple meanings and does not necessarily mean exponential. Petroleum engineers deal with non-linear problems regularly for which there is no explicit mathematical general solution (Çambel, 1993).

Fry and Paterniti (2014) peformed a study to analyze the effect of a new, presumabley cleaner type of fluid system than older guar based systems. They wanted to see if the fluid quality had affect on clean-up in the fracture and wellbore and wanted to see if this helped production. With so many factors that may affect production, the uncertainty was reducd by using wells within a one mile radius of the wells pumped with the cleaner fluid system as offsets to compare production to.

Maldonado and Aoun (2019) successfully created a re-fracturing well selection mechanism using data from 50 producing wells in the Messaoud Field in Algeria (Maldonado & Aoun, 2019). Peirce and Bunger (2014) predicted fracture growth and propagation by deriving a mathematical model accounting for flow of fluid, rock breakage, and pressure loss. The authors argued that the goal of the simple model was to: (1) create more general models that can make broad predictions by individuals without narrow expertise, and (2) Provide a guide for more complex models (Peirce & Bunger, 2015).

Mohaghegh et al. (2017) developed "Shale Analytics" as a self-contained workflow using ML to maximize hydrocarbon production from shale resources in the Utica, Marcellus, Niobrara, and Eagle Ford plays (Mohaghegh et al., 2017). Shale Analytics provides multiple services, most importantly, optimizing well spacing, prediciton of the best candidates for re-fracturing, and Decline Curve Analysis (DCA).

Complex fracture interactions occur in zones of depletions (Brady et al., 2022). Mohaghegh (2016) developed a re-fracturing candidate selection algorithm. (Mohaghegh, 2016). Mohaghegh (2016) found that the optimal re-fracturing job was defined by creating the largest fracture network possible that maximizes hydrocarbon production. While this

was certainly a primary goal, the authors of this study added this goal has tradeoffs with cost. Correctly analyzing the tradeoffs between production and costs is vital to the economcis of re-fracturing.

Cross et al. (2020) used ML models to generate Shapley values for scaling factors in completions designs. Producing the Shapley values allows for the isolation of each parameter to estimate its' effects through time on production (Cross et al., 2020). Cross et al. (2021) used ML models to predict O&G production in multiple unconventional basins in the U.S. These models incorporated survey data to account for well-well interaction as well as key completion parameters normalized over lateral length. They then generated Shapley values which estimate the contribution of each completion feature in the model (Cross et al., 2021). Sakhardande and Devegowda (2021) used data driven causal analysis to optimize well spacing between parent (control group) and child (treatment group) wells. They approach the problem by looking at how well performance (normalized over lateral length) is affected by a set of independent variables, including well spacing. They claim this to be the first study in the O&G industry focusing on causal analysis (Sakhardande & Devegowda, 2021).

Figure 2.9 highlights the problem of distinguishing correlation from causality. This was demonstrated by Mohaghegh (2019) in trying to draw conclusions from correlations between well productivity and field treatment data (Mohaghegh, 2019) .

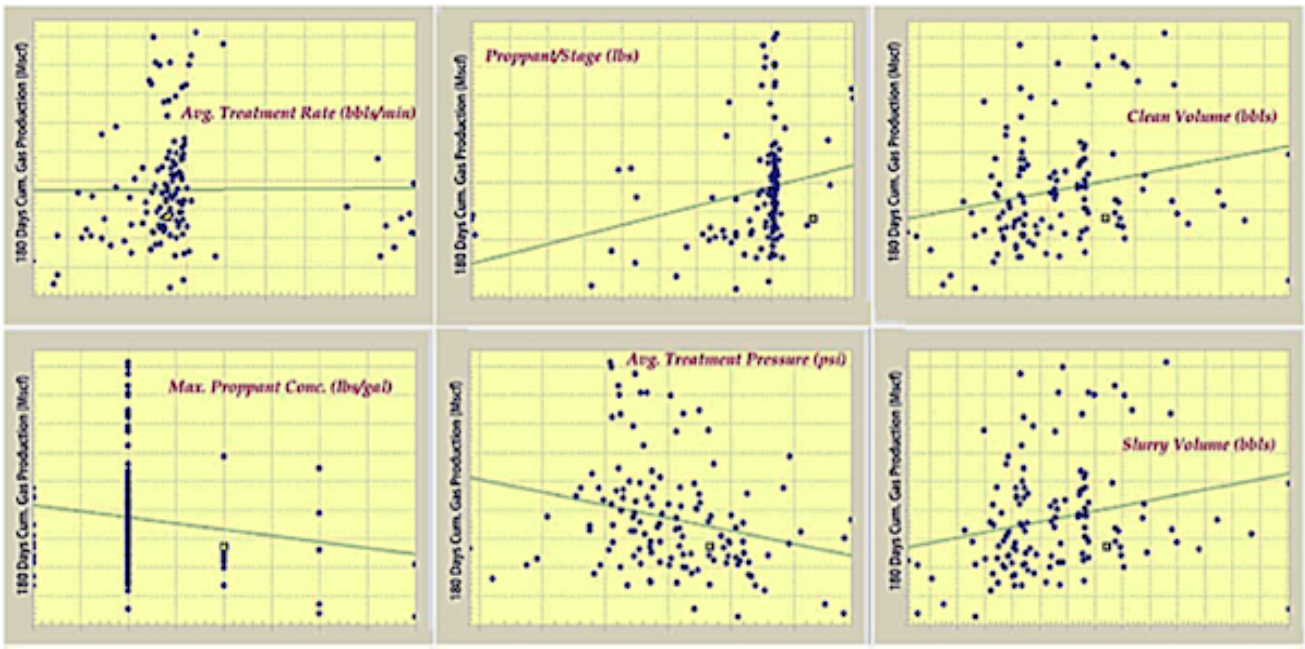

Figure 2. 9 - Correlations between well productivity and field treatment data from Mohaghegh (2019).

Mohaghegh (2019) attempts to distinguish between correlation and causation by using the state-of-the-art soft cluster analysis Shale Analytics workflow to estimate how these parameters will affect well productivity and predictions (Mohaghegh, 2019).

2.5 Domain Knowledge: A Guide for Independent Variable Selection

The domain knowledge contained in human capital has been steadily decreasing in the O&G industry. Between January 2015 and November 2021, the number of employees in the O&G extraction sector decreased from around 200,000 to about 140,000 (U.S. Bureau of Labor Statistics, 2021). This decrease represents a major loss of human capital and anecdotal and domain knowledge. Having a proper amount of domain knowledge will help guide research questions and lead to better results. This lost expertise is not usually possessed by the data scientist and can therefore not be incorporated into the ML models

(Bangert, 2021). Domain knowledge will make data science projects faster, more cost effective, and increases the probability of producing a practical result (Bangert, 2021). Domain knowledge may then be thought of as a sort of wisdom passed on to the algorithms, bridging the gap between smart and wise (Peterson, 2017). One practical example is the R801 drilling rig. Although this rig is impressive, it's far from an end point of fully autonomous drilling operations (Rassenfoss, 2021). John Willis, VP of Drilling and Completions for Onshore and Carbon Sequestration for Occidental Petroleum, says the company's goal for utilizing such automation is to "capture knowledge". This refers to capturing the most efficient way to perform tasks in control systems so lesser skilled workers may be at least as productive as more skilled workers (Rassenfoss, 2021). Admittedly, Willis says, they have not found a system that accomplishes this goal (Rassenfoss, 2021). One reason may be due to the anecdotal and domain knowledge contained within the human capital that is being replaced by the automation.

Therefore, it must be captured and managed to successfully develop practical results from ML and automation from the data. Symptoms of knowledge mismanagement include good ideas and best practices that are not effectively dispersed leading to repetition of past events and operations (Van der Spek, 2017). These symptoms can increase overall costs and may be quantified as the "cost of ignorance" (Van der Spek, 2017). This is increasingly important as the industry shifts paradigms to cost minimization (Barree, 2020). Darabi et al. (2020) constructed a novel "Augmented AI" framework to leverage domain knowledge and incorporate it into an AI workflow and showed positive results in the training, testing, and validation data (Darabi, et al., 2020). This augmented approach

was used to identify opportunities in completions and locations for new drills to create value in an unconventional reservoir (Darabi et al., 2020).

Data mining techniques are good at finding relationships and making predictions under stable conditions based off these relationships, but poorly equipped to explain why the relationship exists (Huntington-Klein, 2022). In short, data mining focuses on what's in the data but not why it's in the data (Huntington-Klein, 2022). Therefore, data mining techniques must be augmented not only with good research questions (Huntington-Klein, 2022), but with domain knowledge and a focus on causality.

2.6. Multivariate Regression Models

Clearly, there is a need to develop and implement models in O&G applications that are flexible enough account for uncertainty and randomness while also deriving useful insights. Although anecdotal observations are important and may create new ideas, but they are insufficient to draw any conclusions. We must present some sort of evidence for the conclusions drawn. However, it's also evident that these models must include domain knowledge to be self-contained. Multivariate regression models are one ML tool that allow for this combination. The multivariate regression method is based on the general bivariate regression model (Niu et al., 2021). Multivariate regression allows for identification of statistically significant factors that influence a dependent variable (Niu et al., 2021).

Equation 1 shows the bivariate regression model that is the basis for the models used in this study (Bailey, 2017). Here, Y_i is the dependent variable of interest, X_i represents the independent variable that is thought to affect the dependent variable, β_0 is the expected value when the independent variable is 0, and β_1 represents the marginal effect that the

independent variable will have on the dependent variable. Therefore, if the independent variable X_i is increased by one unit, the dependent variable is expected to change by the coefficient β_1, all else being equal. Equation 6 represents the true model. This means that the coefficient estimates β_0 and β_1, are the true values and the error term ϵ_i captures all of the variation not explained by β_0 and β_1.

$$Y_i = \beta_0 + \beta_1 X_i + \epsilon_i \tag{6}$$

It's important to note that when using data sets, the true values are not known because we almost certainly will not have population level data. Therefore, whenever a regression is constructed from a sample of the population, the β_0 and β_1 coefficients are estimates. These will be denoted with a hat and follow the notation in Bailey (2017).

$$\widehat{Y_i} = \widehat{\beta_0} + \widehat{\beta_1} X_i \tag{7}$$

Equation 7 represents a fitted value for the regression and is in fact a regression line (Bailey, 2017). Here, as opposed to the error term ϵ_i, we are left with the residual value. The residual value is the difference between a predicted value and the actual observation and is shown in equation 8 (Bailey, 2017). This can be thought of as the amount of variation not explained by the estimated coefficients $\widehat{\beta_0}$ and $\widehat{\beta_1}$. Although we will never know the true model, we can pull variables out of the error term and estimate a fitted value. It is also assumed this error is identical and independently distributed (iid). This means that "the theoretical distribution of the error term, ϵ_i, is unrelated to error terms of other observations and the other variables for the same observation (independent) as well as the same for each observation (identically distributed)." (Huntington-Klein, 2022).

$$\widehat{\epsilon_i} = Y_i - \widehat{Y_i} \tag{8}$$

The $\hat{\beta}$ coefficient estimates are then calculated by minimizing the residual sum of squares. Since $\widehat{Y_i} = \widehat{\beta_0} + \widehat{\beta_1} X_i$, we can equivalently write equation 9 (Bailey, 2017).

$$\widehat{\epsilon_i} = Y_i - \widehat{\beta_0} - \widehat{\beta_1} X_i \tag{9}$$

The algorithm will then choose the $\hat{\beta}$ coefficients that minimize the residual sum of squares shown in equation 10 (Bailey, 2017).

$$\sum_{i=1}^{N} \widehat{\epsilon_i^2} = \sum_{i=1}^{N} (Y_i - \widehat{\beta_0} - \widehat{\beta_1} X_i)^2 \tag{10}$$

This minimization algorithm provides specific coefficient estimates that are unique to the given sample data set. The hope is that the data is representative of the population and thus the coefficient estimates are representative of the true values. The estimate of $\widehat{\beta_1}$ is shown in equation 11 (Bailey, 2017).

$$\widehat{\beta_1} = \frac{\sum_{i=1}^{N} (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^{N} (X_i - \bar{X})^2} \tag{11}$$

Equation 11 describes how X and Y move together. The numerator shows the product of how far X deviates from its mean, $\bar{X}$, for the *i*th observation and how far Y deviates from its mean, $\bar{Y}$, for the *i*th observation. So if X tends to be above its mean when Y tends to be above its mean, the numerator will tend to be positive as well, creating a plethora of positive values. We'll also get positive values if X and Y tend to be below their means through the multiplication of tow negative numbers. However, if X and Y tend to be opposite in direction, if X or Y tends to be below their mean while the other is above their mean, then

we'll tend to get negative values for the summation (Bailey, 2017). Then, $\widehat{\beta_0} = \bar{Y} - \widehat{\beta_1}\bar{X}$ (Bailey, 2017).

In HF, because of factors like reservoir heterogeneity and wellbore dynamics, there is a need to account for multiple independent variables in treatment analysis. This study will utilize multivariate regression models to not only construct descriptive models, but to estimate the marginal effects of multiple independent variables of interest. Just as the coefficient estimate for β_1 in the bivariate model gives an estimate of the marginal effects X_i will have on Y_i, so too do the β_n coefficient estimates provide marginal effect estimates on the dependent variable. Multivariate regression models are an extension of bivariate regression models. Equation 12 shows the basic construction of multivariate regression models (Bailey, 2017).

$$Y_i = \widehat{\beta_0} + \widehat{\beta_1}X_1 + \cdots + \widehat{\beta_n}X_n + \epsilon_i \quad (12)$$

One important addition to the multivariate model is the error term ϵ_i. The error term is meant to capture all other factors not included in the regression model that may have a causal relationship with the dependent variable. Leaving variables in the error term that have a causal relationship with the dependent variable will lead to what is known as endogeneity. Broadly speaking, endogeneity is defined as explanatory variables being correlated with the error term (Bailey, 2017). This also leads to problems known with collinearity. If two independent variables are correlated with each other, it makes it hard to decipher the causal affect that either has on the dependent variables. Endogeneity leads to the deduction of causal relationships that may be spurious. One example may be a potential relationship between ice cream sales and an increase in drowning rates in swimming pools

(Wikipedia, 2021). To regress one variable on the other and then conclude that an increase in ice cream sales causes people to drown would be spurious. In fact, it would most likely be the case that heat was driving people to buy more ice cream and drive more people to pools and thus increasing the likelihood of drownings (Wikipedia, 2021).

The $\widehat{\beta_n}$ are coefficient estimates given the data sample and if X_n is uncorrelated with the error term ϵ_i, then $\widehat{\beta_n} = \beta_n$ in the limit (Bailey, 2017). In other words, the estimate equals the true underlying value in the limit if all exogenous variation is accounted for. The primary precision estimate for $\widehat{\beta_n}$ is the variance which is a measure of how wide the distribution around $\widehat{\beta_n}$ is and is shown in equation 13 (Bailey, 2017).

$$var(\widehat{\beta_n}) = \frac{\widehat{\sigma^2}}{N \; x \; var(X)} \tag{13}$$

Where X is the independent variable, N is the number of observations, $\widehat{\sigma^2}$ is the variance of the regression defined by equation 14 (Bailey, 2017) and explains how well the models explains variation in Y (Bailey, 2017). In equation 14, k is the number of predictors in the regression and (N-k) is known as the degrees of freedom, Y_i is each observation, and $\widehat{Y_i}$ is the fitted value.

$$\sigma^2 = \frac{\sum_{i=1}^{N}(Y_i - \widehat{Y_i})^2}{N-k} \tag{14}$$

For practical purposes, R software (R Core Team, 2022. https://www.R-project.org/) will export standard errors (se) for each coefficient estimate and the regression. The se for each estimate is the square root of the variance of the estimate and the se for the regression, $\hat{\sigma}$, is the is the square root of the variance of the regression. So, the se of the

regression will be in the same units as the dependent variable and is roughly equivalent to the average distance between the actual and fitted values (Bailey, 2017). The se for the coefficient estimates is then essentially a measure of how wide the distribution is around each coefficient estimate, $\widehat{\beta_n}$ (Bailey, 2017). The larger the se compared to the coefficient estimate, the less confident we are the estimate. Likewise, the larger the se of the regression is relative to the dependent variable, the less confident we are in the model.

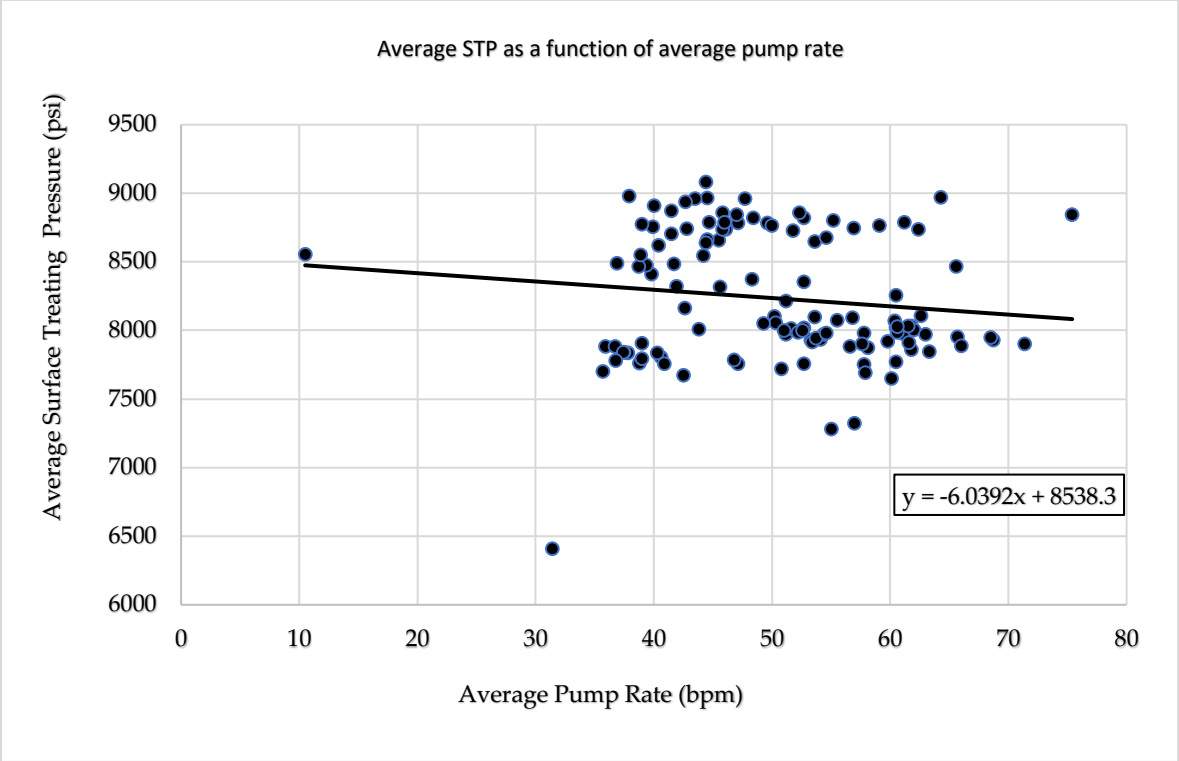

Figure 2. 10 - Bivariate regression model regressing average STP on average pump rate.

Figure 2.10 shows a bivariate regression showing the relationship between average pump rate and average STP using data from re-fracture jobs used in this study. Only accounting for average pump rate would lead one to conclude that increasing the average pump rate will decrease the average STP. However, this is not congruent with domain

knowledge. This is an example of an endogenous model and highlights the necessity to account for more variables when trying to draw conclusions about HF treatment effects and implementation, i.e. exogenous variation.

Huchton et al. (2020) performed a multivariate analysis of completion parameters on the STACK acreage in Oklahoma. Using six month oil production to evaluate impacts at the well level and six month oil production per acre to evaluate area effects, they concluded that fluid volume has a strong and significant relationship with production at the well and area level (Huchton et al., 2020). Niu et al. (2021) constructed multivariate regression models to estimate the ultimate recovery of gas wells in the Weiyuan block. This study considered 172 shale gas wells and constructed predictive multivariate models with reasonable accuracy (Niu, Lu, & Sun, 2021). However, there are additional benefits to using multivariate regression models for analysis. Wang et al. (2022) used multivariate regression to estimate sensitivity of a range of geologic and reservoir properties on the recovery factor in the Daqing Oilfield to estimate suitability for CO_2 flooding (Wang et al., 2022).

Recall that data mining focuses on what's in the data but not why it's in the data (Huntington-Klein, 2022). Therefore, data mining techniques must be augmented not only with good research questions (Huntington-Klein, 2022), but with domain knowledge as well and a focus on causality. Combining data mining, domain knowledge, and causal analysis will lead to a better understanding of the complexities of O&G systems, reduce the costs of knowledge mismanagement, reduce uncertainty, and increase the effectiveness of ML and AI implementations.

After constructing multivariate regression models and estimating the marginal effects from an independent variable(s) on the dependent variable, we can start to develop a mindset focusing on margins and engineers can then make decisions at the margin.

2.7. Marginal Mindset and Marginal Analysis

In economics and economic models, decisions and discussion almost always revolve around the margins. Marginal thinking requires making decisions at the margin. I would argue that most discussions in the O&G are actually questions about margins or questions that are best answered at the margin. Questions about how much sand, how many perforations, how many clusters, fluid volume, etc. to add to a treatment are questions about margins and require marginal estimates. Figure 2.11 shows a simplified schematic on how to think about margins.

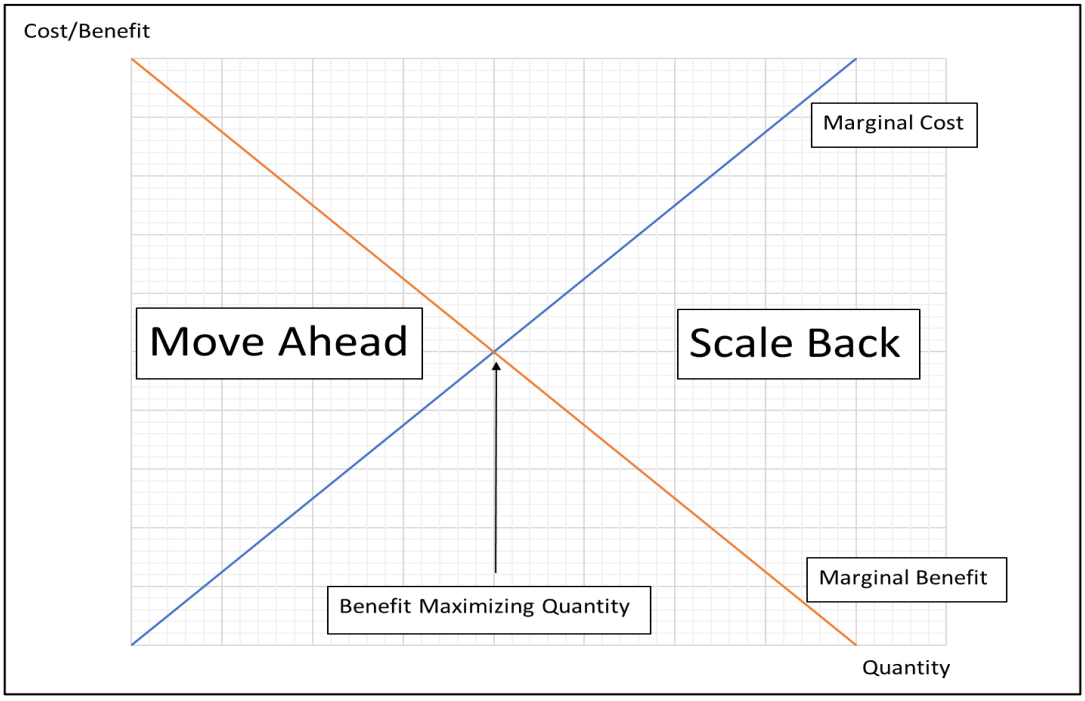

Figure 2. 11 – Plot highlighting marginal thinking

The basic idea behind figure 2.11 is straightforward. The red line is the marginal benefit (MB) from producing the next unit and the blue line represents the marginal cost (MC) of producing the next unit. If the MB>MC, it makes economic sense to move ahead in production. Tethering this to a treatment, if the marginal benefit of pumping an extra 100 bbl of slurry yields production greater than the cos of the 100 bbl, then it makes economic sense to include the extra 100 bbl of slurry into the treatment design since there would benefits left on the table. On the other hand, if the MC>MB, it makes economic sense to scale back production. Therefore, the benefit maximizing point is where MC=MB. Here, you are not incurring unnecessary costs nor leaving benefits on the table. This analysis is straightforward, but finding estimates on marginal effects in the O&G industry sparse.

This study focuses on average STP as the outcome and is what the marginal benefits/costs will be measured against. For example, what if exceedingly high STPs are experienced creating abnormally long pump times which has a direct effect on costs. What can be done operationally to reduce the pressure? Practically speaking, there are only a few knobs to turn in order to reduce the STP. One knob is the addition of perforation to provide more entry points into the reservoir. What is the desired STP? Based off of this, how many perforation should be added to the tool string? Given the tradeoff with potential decreases in cluster efficiency, is adding perforation even the best option for reducing pressure? Moving forward with design, what other factors determine STP and how can more effective wellbore designs change the expected STP. Will it be beneficial to install larger liners in the lateral section of a wellbore?

These types of questions are difficult to answer without thinking about margins and having marginal estimates of costs and benefits The cost of adding an additional perforation may be straightforward, but what are the expected benefits in terms of STP and will these offset the costs? The models in this study can help answer these types of questions and allow engineers to consider margins when designing treatments and incorporate marginal thinking into a cost minimization framework and augment these models with physics based fracture models used for production maximization.

The large decrease of roughly 60,000 employees between 2015 and 2021 (U.S. Bureau of Labor Statistics, 2021) and the new focus on cost minimization (Barree, 2020) has created a more capital-intensive industry that is just focused on minimizing costs as much as maximizing production. This highlights a need for models that can identify where treatments may be improved and field implementation made more efficient. This study will construct multivariate regression models to identify high margin areas for improvements in implementation with regards to average STP. Average STP is often a binding constraint in the field, especially in re-fractured wells with wellbore degradation, and often limits parameters such as pump rate. However, limiting pump rate also increase stage pump time which has direct impact on costs. Higher STPs also create a need for more hydraulic horsepower (HHP) which increases equipment costs and fuel costs. The high margin areas identified in this study highlight areas where improvements can be made to minimize these costs.

2.8. Methodology

Re-fracturing success in terms of production depends largely on the well selection (Rignol & Bui, 2020). There is no universal method for candidate screening, but it should be based on well potential, production performance, and the success of the initial treatment (Rignol & Bui, 2020). Candidate selection is beyond the scope of this study and, here, it is assumed that the best candidates were chosen for re-fracturing treatment. Some of the wells selected for re-fracturing treatment were also completed during zipper frac operations on pads with new drills. Any effects of these operations are assumed to be negligible compared to the effect from treatment of previous stages. With the increased complexity and smaller margins of re-fractured wells, minimizing costs is a crucial component that should be considered in combination with proper well selection.

The process of model construction to investigate how completion parameters will effect treatment implementation is described in figure 2.12. First, domain knowledge will guide independent variable selection. Using domain knowledge to select variables will allow for investigation into parameters that can actually be altered in the field and account for underlying physical relationships. Next, an initial pooled regression model will be built as a first attempt to identify causal relationships between completion parameters and average STP. Next, a fixed effect models will be constructed to account for any within well fixed effects. Finally, a more complex feature engineered model will be constructed, using understanding from the first two models, attempting to identify more complex relationships without becoming more complex than necessary.

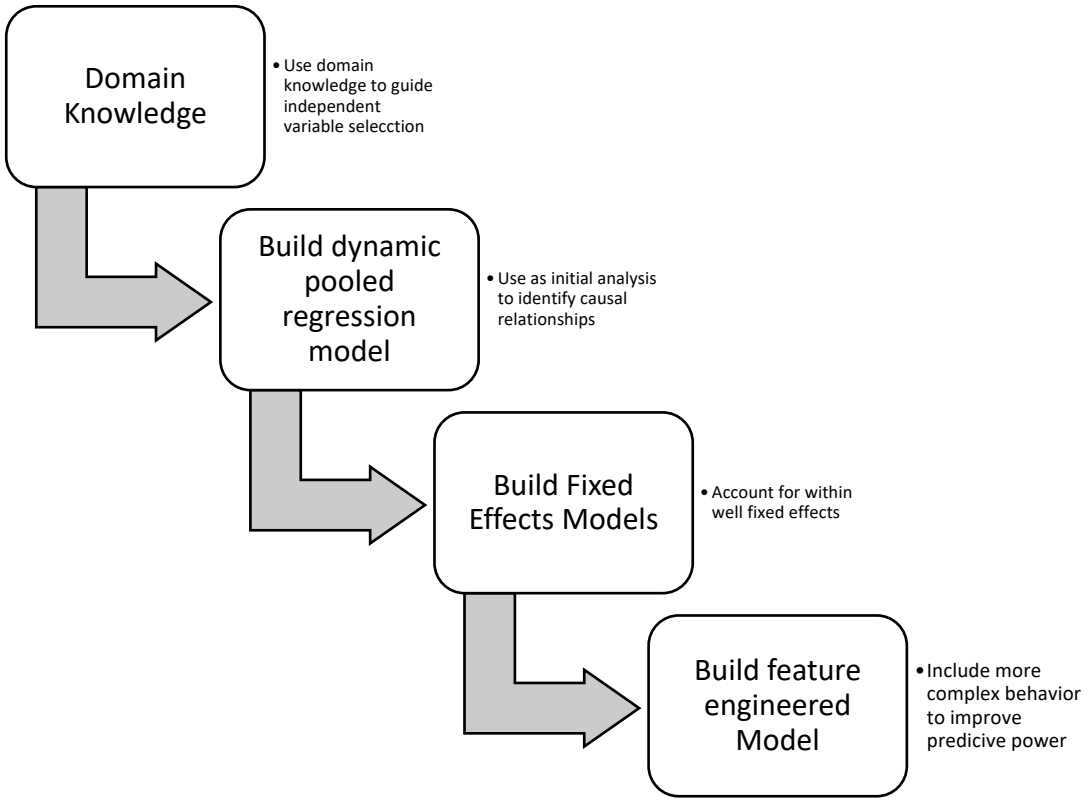

Figure 2. 12 – Flowchart for process of constructing multivariate regression models for study

Chapter 3

3. Modeling Temporal Dependence of Average STP

This chapter was published in Energies 2022 15(16). Full citation: Kroschel, J. Rabiei, M. Rasouli, V. Modeling Temporal Dependence of Average STP in the Williston Basin Using Dynamic Multivariate Regression. Energies 2022, 15, 2271. **https://doi.org/10.3390/en15062271**. As the first author, I have contributed more than 75% of the effort to this work including initial draft preparation, experiment design, and model construction.

3.1. Re-fracturing

Given the complexity of re-fracturing and the market push for cost minimization, there is a need to identify places where improvements can be made and cost can be minimized. This chapter will construct a pooled multivariate regression model that will serve as an initial inferential model to being the analysis conducted in this study. **This is**

the first step described in figure 2.12 and will attempt to answer research question 1 by identifying statistically significant predictors of average STP, their directional impact on average STP, and provide an initial estimate of stress shadow effects by accounting for temporal average STP using a dynamic model.

Using domain knowledge, initial hypotheses will be stated with justification as to why these variables should be accounted for in the multivariate models. Including these in the model will provide initial estimates of their marginal effects on average STP and identify areas where improvements can be made to improve treatment implementation. This will also be a first attempt to identify causal relationships between independent variables and average STP that we believe exist based off of our physical understanding of wellbore dynamics derived using the wellbore relationships described in section 2.3.

The initial inferential model accounts for the temporal nature of stress shadow effects. By adding a lagged dependent variable, this relationship can be captured to account for temporal effects. The general dynamic equation is shown in equation 15.

$$Y_{it} = \gamma Y_{i,t-1} + \beta_0 + \beta_1 X_{1it} + \cdots + \epsilon_{it} \quad (15)$$

3.1.1 Perforation Standoff

Net pressure and stress contrast both increase substantially with the number of sequential fractures and decreased fracture spacing (Roussel & Sharma, 2011). Induced stress shadow from previously fractured stages also has effects on the subsequent zone as the effects extend beyond the top perforation and may thus change the fracture initiation at the next bottom perforation for the next stage (Barree, 2020). Therefore, the standoff

between the top perforation for one stage and the bottom perforation for the subsequent stage may affect treatment through stress shadow effects and will be included in the regression models. Figure 3.1 is a schematic of the perforation standoff in the wellbore. The hypothesis is that increasing the perforation standoff will decrease stress shadow effects from previous stages and thus decrease average STP.

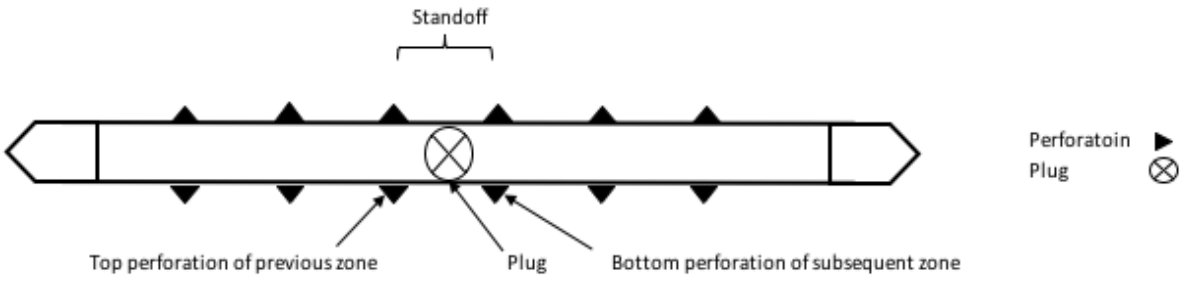

Figure 3. 1 – Flowchart for process of constructing multivariate regression models for study (Kroschel et al., 2023a)

3.1.2. Stage Proppant Weight

The amount of proppant placed during a stage affects bottom hole treating pressure by increasing the hydrostatic pressure. This in turn affects the STP. The amount of proppant pumped may also affect STP through perforation erosion as treatment progresses through hole erosion and rounding (Behrmann & Nolte, 1998). The hypothesis is that the effects of increasing stage proppant weight will be negative as more sand increases in hydrostatic pressure and increases perforation erosion, thus decreasing the average STP.

3.1.3. Total Clean Volume

It is reasonable to assume that treatment size will affect the STP. Total clean volume pumped should be a good indication of stage size. The hypothesis is that an increase in total clean volume will increase average STP.

3.1.4. Previous Stage Average STP

This dynamic variable will attempt to capture the stress shadow effects from the previous stage treatment. HF treatment with multiple stages and clusters leads to complex interactions resulting from different propagating fractures (Damjanac et al., 2018). When multiple fractures are growing near each other, the stress fields created around each fracture must be considered, as these will affect fracture growth and treatment pressures (Barree, 2020).

This may be even more significant in re-fracturing treatments, since decreases in pore pressure can lead to redistributing stresses or even reversal of principle stress in regions around initial fractures (He et al., 2021). Stress interference from reorientation increases with the number of fractures created and depends on the sequence of fracturing (Roussel & Sharma, 2011).

Fractures on each end of the stage will behave as individual fractures, regardless of the number of fractures between the two (Barree, 2020). These end fractures will tend to dominate and become the primary fractures (Barree, 2020). Poroelastic effects change the net vertical stress along the wellbore (Barree, 2020). Although it is not possible to know the rock properties and poroelastic effects for each stage, average STP may capture

these effects during and after treatment. This is also related to the perforation stand-off. Considering the dominance of the end fractures and the associated stress and poroelastic effects that fractures create, the hypothesis is that a higher STP from previous stages will increase the average STP for the subsequent stage. Referring to equation 15, this will be accounted for by $\gamma Y_{i,t-1}$ where γ will estimate the escalation factor of average STP due to stress shadow effects.

3.1.5. Number of Perforations

The number of perforations is important to any treatment, as these are the conduit from the wellbore that will ultimately deliver the treatment slurry into formation. Although only a limited number of perforations may take fluid and preferentially propagate major fractures, usually the toe and heel perforations (Roussel & Sharma, 2011), it is still necessary to have a certain number of perforations above the number of fractures for pressure relief during treatment. The hypothesis is that an increase in the number of perforations will decrease average STP.

3.1.6. Presence of a 3.5" Liner

The presence of a 3.5" liner will increase friction due to the smaller inner diameter relative to treatments pumped through a 4.5" liner. Installation of a new liner in the horizontal section is usually necessary. If the well was initially completed with sliding sleeves and packers, these will need to be removed or milled and a new liner installed to provide a means for zonal isolation. If the well was initially completed using a cemented liner in the horizontal, a new liner may be required to avoid the issues that will arise with wireline running into previous perforations and possible casing issues from previous

treatment and production. Previous open hole completions also require the installation of a new 4.5” liner. The hypothesis is that the presence of a 3.5” liner will increase the average STP. This is a binary variable, with “1” indicating the presence of a 3.5” liner and “0” indicating a 4.5” liner. The presence of a 3.5” liner may also affect average pump rate by increasing friction pressure and thereby decreasing the maximum rate achievable.

3.1.7. Average Pump Rate

Increased pump rate creates more friction along the wellbore as well as the fracture face. The average pump rate will most likely have an effect on treatment costs as service companies generally charge for pump time. Therefore, longer pump times will increase treatment costs. The assumption is that an increase in average pump rate will increase the average STP.

3.1.8. Acid Volume

Acid is important in designs to clean up any near wellbore damage caused by the perforation and cement. It can also have drastic effects on treating pressure if the formation is carbonate. The assumption is that an increase in acid volume will decrease the average STP. It’s also reasonable to think that acid will affect average pump rate. If acid provides pressure relief, it’s reasonable to think maximum pump rate can be achieved more quickly and therefore increase average pump rate.

3.1.9. Formation Type

In an attempt to capture any geologic differences between the formations that may affect treatment, a binary variable will be used in Model 2 to see if there are any

statistical differences between the models and if accounting for geological differences has any statistical impact on the models and coefficient estimates for the other independent variables. These estimates will be relative to the Three Forks formation.

3.1.10. Causal Diagram of Relationships

Figure 3.2 describes the hypotheses in graphical form. These are simply representations of a data generating process (DGP) (Huntington-Klein, 2022). These diagrams have two components: 1) each variable in the DGP, represented by a node, and 2) an arrow from the cause to the effect representing a causal relationship (Huntington-Klein, 2022). It's important to note that these arrows do not indicate anything about positive or negative relationships (Huntington-Klein, 2022). The direction of the relationship will be estimated by the multivariate regression models.

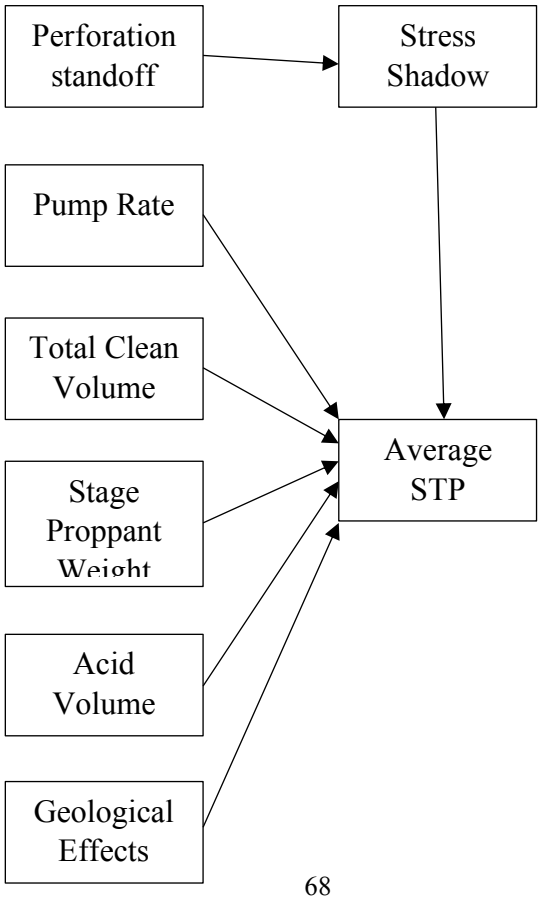

Figure 3. 2 – Flowchart for process of constructing multivariate regression models for study

3.2 Re-fracturing Data, Well Location, and Summary Statistics

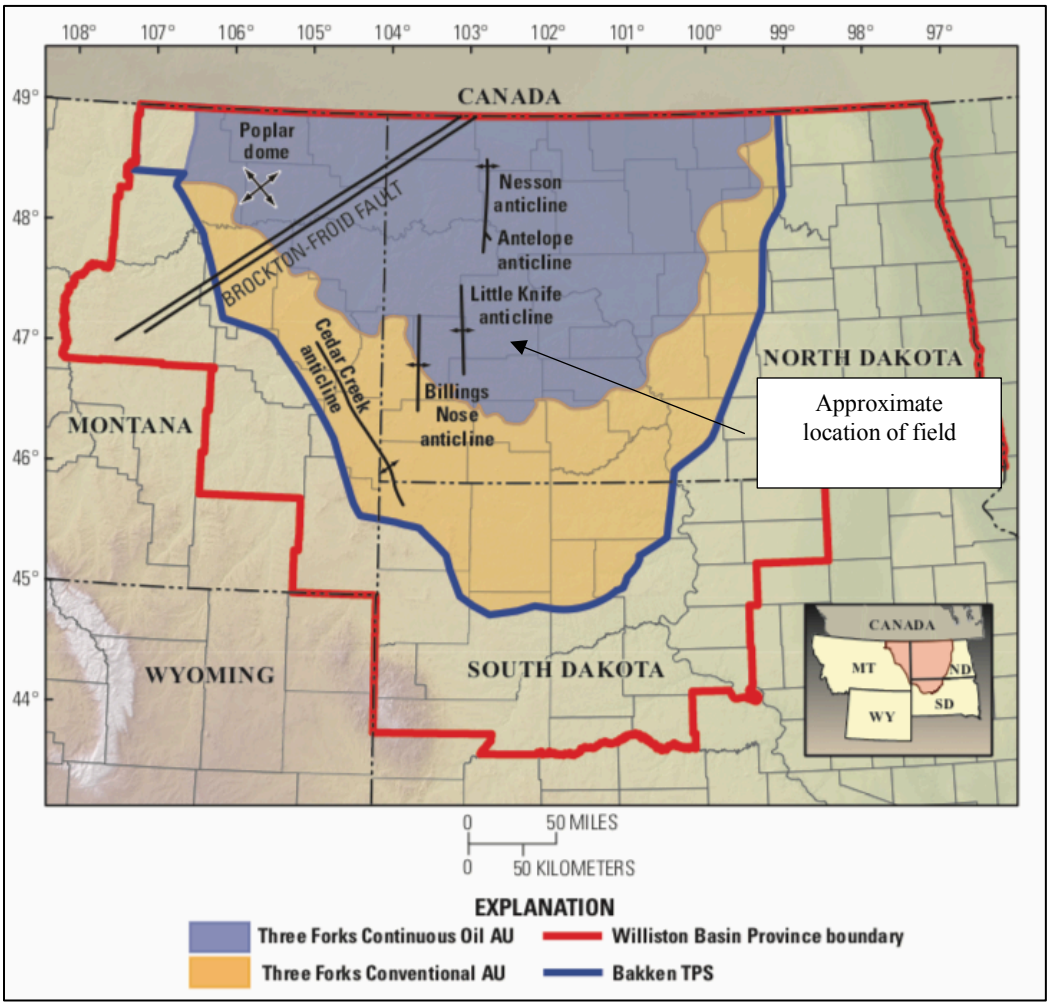

Figure 3. 3 - Map showing location of re-fractured well in dataset (Gaswirth et al., 2013)

Figure 3.3 shows the approximate location of the wells where data was collected for this study. Data for this study were not normalized as it is not unnecessary because of the uniqueness of multivariate regression models. These models attempt to measure marginal changes in the independent variable and how these will affect the dependent variable. Therefore, since the units of each of the independent variables used in the regression are

used in the field, they will be kept in these units to make coefficient interpretation easier. If the models attempted to account for something like Poisson's ratio, which has a range from 0 to 0.5, then a marginal unit increase does not make sense in this context. In other words, the units used for the variables in the regression models are in context with the units used in the field so there is no need for standardization. In fact, statistical significance and model fit would be the same for standardized and unstandardized results (Bailey, 2017). Table 3.1 shows the summary statistics for the data set and table 3.2 provides variable definitions and associated units.

Table 3. 1 - Summary statistics for re-fracturing datasets used in this study.

Variable	N	Mean	Std. Dev.	Min	Pctl. 25	Pctl. 75	Max
avg_pump_rate	121	50.331	9.966	10.5	42.6	57.8	75.4
avg_stp	121	8234.355	465.448	6411	7903	8726	9083
acid_volume	121	20.711	30.92	0	11	23	227
total_clean_volume	121	3950.039	839.746	2920.091	3606.364	4024.941	8940.621
stage_prop_weight	121	220,207.884	37,671.035	25,649	212,565	218,373	405,515
perfs	121	24.165	1.562	22	24	24	36
liner_3.5	121	0.686	0.466	0	0	1	1
formation	121	0.752	0.434	0	1	1	1
perf_standoff	117	25.573	4.415	18	22	26	49
avg_stp_prev	117	8235.778	467.02	6411	7907	8726	9083

Table 3. 2 - Variable definitions for data set used in this study.

Variable	Variable Definition
avg_pump_rate	Average pump rate, bpm
avg_stp	Average STP, psi
acid_volume	Acid volume, bbls
total_clean_volume	Total clean volume for stage, bbls
stage_prop_weight	Proppant weight pumped for stage, lbs.
perfs	Number of perforations over treatment interval
liner_3.5	Presence of 3.5" lateral liner (1 = yes, 0 = no)
formation	Binary variable indicating formation (1 = middle Bakken, 0 = Three Forks)
perf_standoff	Distance between top perforation of one stage and bottom perforation of subsequent stage, ft
avg_stp_prev	Average STP for previous stage, psi

3.3 Regression Models

Table 3.3 summarizes the regression results from R, including the independent variables discussed above. The model investigates whether these have a statistically significant effect on the dependent variable (average STP) from the given data set. Two models were run, with model 2 including a binary variable for which formation the well was drilled (1 for middle Bakken and 0 for Three Forks). The coefficient estimate for formation is then relative to the Three Forks formation. This was done to see if there were any statistical difference between models accounting for geologic properties and those that do not. It is important to note that whenever coefficient estimates and results are discussed, they are only expectations as the true coefficient estimates, and therefore marginal effects, are not known (Bailey, 2017). This extends from the fact that statistical models do not yield facts, only expectations. To test the marginal effects, experiments would have to be

conducted in the field by altering only one parameter by one unit while holding all others constant.

Table 3. 3 - Regression results showing two different models using stargazer package in R (Hlavac, 2022).

	Dependent Variable:	
	Average Surface Treating Pressure (psi)	
	Model (1)	**Model (2)**
Perforation Standoff (ft)	14.752 *	14.615 *
	(7.723)	(7.796)
Stage Proppant Weight (lb)	−0.001	−0.001
	(0.002)	(0.002)
Total Clean Volume (bbl)	−0.043	−0.044
	(0.090)	(0.090)
Number of Perforations	−22.859	−23.170
	(15.884)	(16.050)
3.5-inch Liner	228.412 **	221.913 **
	(91.925)	(99.258)
Average Pump Rate (bpm)	9.561 **	9.471 **
	(3.692)	(3.742)
Acid Volume Pumped (bbl)	8.075 ***	7.785 **
	(2.628)	(3.098)
Previous Stage Average Surface Treating Pressure (psi)	0.714 ***	0.713 ***
	(0.074)	(0.075)
Formation (1 = middle Bakken, 0 = Three Forks)		−14.937
		(83.683)
Constant	2217.077 ***	2260.782 ***
	(821.258)	(860.533)
Observations	117	117

R^2	0.712	0.712
Adjusted R^2	0.691	0.688
Residual Std. Error	259.541 (df = 108)	260.712 (df = 107)
F Statistic	33.423 *** (df = 8; 108)	29.446 *** (df = 9; 107)

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Model 1 and model 2 yielded similar results, with statistical significance for acid volume being the only change and formation not showing statistical significance. Coefficient estimates were also largely unchanged after accounting for the different formations. The coefficient estimate for the formation type was also not statistically significant, indicating that any heterogeneity between the formations had no statistical impact on the average STP.

The significance codes' corresponding significance levels are shown in the results as well. The marginal estimate for each variable are shown in table 3.3 with corresponding standard errors for each coefficient estimate in parenthesis. For example, in model 1, the coefficient estimate for perforation standoff is 14.752 with a corresponding standard error of 7.723. Although not all independent variables were statistically significant, all of the coefficient estimates for model 2 will be discussed.

3.3.1. Perforation Standoff

Perforation standoff was found to have a statistically significant effect at the 90% level. This means that the null hypothesis that the coefficient estimate is 0 can be rejected with 90% confidence. However, the coefficient estimate is (+), which runs counter to the initial hypothesis. The positive coefficient estimate indicates that, in expectation, a 1 ft.

increase in perforation standoff between the top perforation from one zone and the bottom perforation for the subsequent zone would yield a 14.615 psi increase in average STP.

3.3.2. Stage Proppant Weight

The initial hypothesis was that increasing stage proppant weight would decrease average STP due to increased hydrostatic pressure and perforation erosion effects. The coefficient estimate of −0.001 indicates the initial hypothesis was correct based on the data set. The coefficient estimate means that a 1 lb. increase in stage proppant weight yields a decrease of 0.001 psi in average STP. However, stage proppant weight was not a statistically significant predictor of average STP.

3.3.3. Total Clean Volume

The initial hypothesis was that an increase in total clean volume would tend to increase the average STP simply based on the fact that there is only a finite fracture volume. The coefficient estimate of −0.044 counters the initial hypothesis and estimates that a 1 bbl. increase in total clean volume will decrease the average STP by 0.044 psi. Total clean volume was not a statistically significant predictor of average STP.

3.3.4. Number of Perforations

The initial hypothesis that an increase in the number of perforations will tend to decrease average STP was backed by the model. The coefficient estimate of −23.17 indicates that the addition of one perforation will decrease the average STP by 23.17 psi, although it was not statistically significant.

3.3.5. Presence of a 3.5” Liner

The initial hypothesis that a 3.5” liner will increase average STP was backed by the model. The coefficient estimate of 221.913 indicates that a 3.5” liner will increase the average STP by 221.913 psi, which agrees with the initial hypothesis. The presence of a 3.5” liner is also a statistically significant predictor of average STP at the 95% level, meaning the null hypothesis that the coefficient estimate is 0 can be rejected with 95% confidence.

3.3.6. Average Pump Rate

The initial hypothesis that an increase in average pump rate will tend to increase the average STP is backed by the model. The coefficient estimate of 9.471 indicates that a 1 bpm increase in average pump rate will increase the average STP by 9.471 psi. The coefficient is statistically significant at the 95% level, indicating that the null hypothesis that the coefficient estimate is 0 can be rejected with 95% confidence.

3.3.7. Acid Volume

The initial hypothesis was that an increase in acid volume would decrease the average STP. This hypothesis was not backed by the model. It is also interesting to note that the coefficient estimate decreased from model 1 to model 2 and significance level dropped from 99% to 95%. This may be because of endogeneity issues and may be resolved by including omitted variables or gathering more data. It is also possible that the variable is endogenous because the pressure relief seen from acid allows for greater treatment rate

which will also increase the STP. The coefficient estimate of 7.785 indicates that a 1 bbl. increase in acid volume will increase average STP by 7.785 psi. The estimate is statistically significant at the 95% level.

3.3.8. Previous Stage Average STP

This variable was included to try and capture stress shadow effects on average STP from one stage to the next. The (+) coefficient agrees with the hypothesis that a higher average STP from the previous stage will tend to increase the STP for the subsequent stage. The coefficient estimate indicates that a 1 psi increase in average STP from the previous stage will increase the average STP of the next stage by 0.713 psi in expectation.

Previous stage average STP was statistically significant at the 99% level, meaning the null hypothesis that the coefficient estimate is 0 can be rejected with 99% confidence. This provides statistical evidence that stress shadow effects may tend to dominate treatment from an operational standpoint.

3.3.9. Formation Type

Accounting for the two different formations (middle Bakken and Three Forks) from model 1 to model 2, they did not have a significant effect overall on the model. Although the coefficient estimate for acid decreased in magnitude and significance, there was no clear evidence that the different formations, and therefore differing geological and mechanical properties between the two, had any statistically significant effect on average STP. The coefficient estimate of −14.937 indicates that if the well were drilled in the

middle Bakken, the average STP is expected to be 14.937 psi lower relative to the Three Forks formation.

3.3.10. Model Visual and Verification

Figure 3.4 shows the added variable (AV) plots for model 2 from R (R Core Team, 2022. https://www.R-project.org/). Since creating 2-D visuals for multiple regression is impossible, one effective way to view individual effects is through AV plots. The plots are constructed so that the dependent variable (average STP) is on the *y*-axis and an independent variable is on the *x*-axis. These plots differ from the bivariate regression constructed in Mohaghegh (2019) in that these are accounting for other independent variables that are being held constant. Therefore, each plot shows the individual effects of one independent variable, with all other independent variables being held constant on the dependent variable.

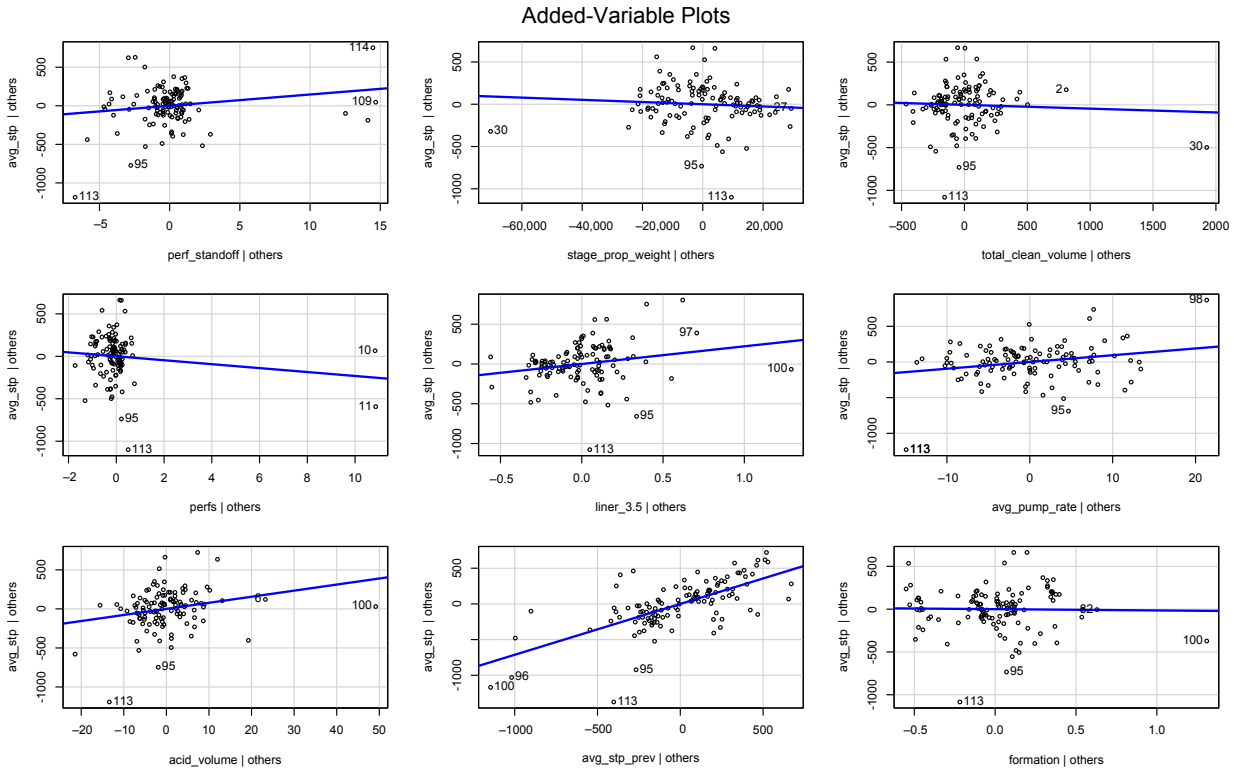

Figure 3. 4 - AV plots from R (R Core Team, 2022. https://www.R-project.org/) for model 2 showing the isolated relationship for each independent variable. This approach holds all other variables constant.

The trendline shows the direction of effect with a positive slope indicating a positive marginal effect and a negative slope indicating negative marginal effect. For a visual comparison between AV plots and simple bivariate regression plots, refer back to figure 2.10 and notice the direction of the estimated effect. Notice the positive relationship between average pump rate and average STP in the AV plot (which we would expect due to friction) and the negative relationship in the bivariate regression plot from figure 2.10. In the bivariate model, we would conclude that the average STP will decrease as the average pump rate increases which does not make sense intuitively. Therefore, without accounting for other factors, we might draw an incorrect causal inference about the

relationship between pump rate and STP. This further highlights the usefulness of multivariate regression models as opposed to bivariate models.

Figure 3.5 is a plot of the residuals vs. the fitted values for model 2. A residual is what is not explained by the model and is simply the fitted value minus the actual value at each data point and is calculated using equation 8. The most important thing to note is that there is not a clear trend in the residual values, which would indicate a different regression model may work better. With no clear and present trend in the residuals, this is evidence that the linear model constructed in this study may be appropriate.

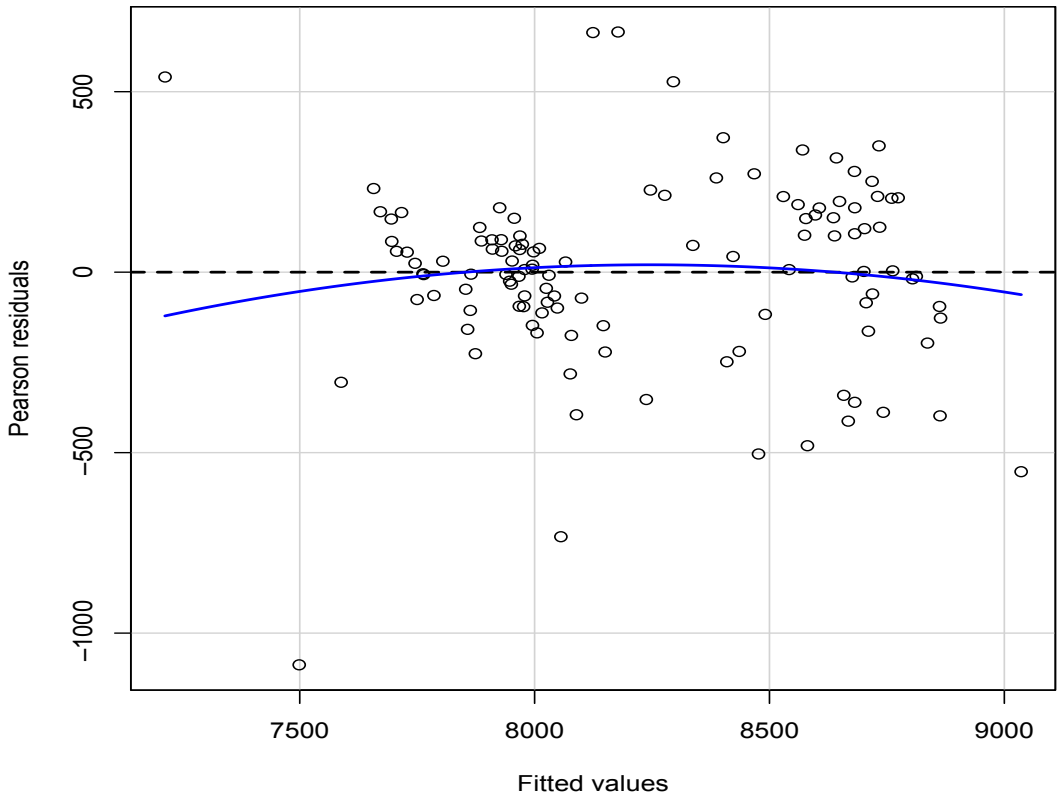

Figure 3. 5 - Residual plot using R (R Core Team, 2022. https://www.R-project.org/)of fitted values for model 2.

One model concern is the possibility of induced bias in the model by including a lagged dependent variable as an independent variable. This makes sense intuitively because average STP is essentially being used as a control variable for its' future self. This may create a positive feedback loop with the other independent variables since they will affect average STP in a given period, which will then be used as an independent variable in the next period, thus propagating the effects. This feedback would create patterns or serial correlation in the errors of the model and induce bias (Bailey, 2017). Therefore, we need to make sure that the errors are not serially correlated using the following model shown in equation 16 (Bailey, 2017):

$$\epsilon_t = \rho\epsilon_{t-1} + \upsilon_t \tag{16}$$

where ϵ_t and ϵ_{t-1} are the residuals and lagged residuals, respectively, from the model in question. Here, we are simply looking for a statistically significant value for ρ, which would indicate a correlation between the errors for each period. Table 3.4 shows the results from the regression.

Table 3. 4 - Regressions results using stargazer package (Hlavac, 2022) from error model.

	Errors
	Err1
Lagged Errors	0.067
	(0.093)
Constant	−1.836
	(23.327)
Observations	116
R^2	0.005
Adjusted R^2	−0.004

Residual Std. Error	251.236 (df = 114)
F Statistic	0.523 (df = 1; 114)

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

From the model, we see a coefficient estimate of 0.067 for the lagged error that is not statistically significant, indicating that there is no correlation of the errors in the model from one period to the next. This indicates that the pooled model is appropriate for characterizing the factors that affect the average STP.

3.4 Chapter Summary

There are additional considerations regarding the models constructed in this study. First, although there were numerous independent variables included in the regression and there is no clear trend in the residuals and the lagged errors are not statistically significant, there still may be variables in the error term that are causing endogeneity. This may be why the acid volume seems to increase average STP and the perforation stand-off coefficient does not match the intuitive understanding of stress shadow effects from previous zones. One way to combat this is to simply add more data from future re-fracturing treatments. The assumption of negligible interaction between wells on the same pad may also need to be addressed.

The proppant concentration may also be considered, although this data was not available for this study. Since the proppant travels at a different speed than the slurry and falls out due to gravity (Gorucu et al., 2021), this may influence the STP near wellbore and in the fracture. Cleary et al. (1993) implemented a system using proppant slugs to estimate

pressure responses and clean up near wellbore issues that affect treating pressure (Cleary et al., 1993). This indicates that sand concentration may be significant predictor of STP.

Further analysis will need to be conducted to attempt to isolate individual well characteristics. Things such as percentage of wellbore drilled out of zone, natural fracture networks, variations in geomechanical properties along the wellbore, perforation characteristics, etc. may also have a causal role in determining average STP. **This will be the basis for constructing fixed effects models in the next chapter.**

Chapter 4

4. Accounting for Unobservable, Within-Well Fixed Effects

This chapter was published in *Energies* 2022 15(16). Full citation: Kroschel, J. Rabiei, M. Rasouli, V. Accounting for Fixed Effects in Re-Fracturing Using Dynamic Multivariate Regression. *Energies* 2022, 15, 5451. https://doi.org/10.3390/en15155451. As the first author, I have contributed more than 75% of the effort to this work including initial draft preparation, experiment design, and model construction.

4.1 Introduction

In chapter 3, multivariate regression models were constructed to estimate the marginal effects of completion parameters on average STP. As mentioned in chapter 3, there is a need to account for more factors that potentially drive average STP. Things such as wellbore effects, natural fracture networks, and geologic factors may also affect the STP. This chapter will build off the previous models by introducing the idea of panel data and fixed effects in an attempt to account for important, albeit unobservable, parameters. Kroschel, Rabiei, & Rasouli (2022a) created a dynamic multivariate regression model to investigate the temporal dependence of average STP on subsequent

stages for re-fractured wells since reservoir depletion has a strong influence on reservoir stresses (Wan et al., 2019).

The model from chapter 3 found that the distance between stages, presence of a 3.5" liner, average pump rate, and acid volume were all statistically predictors of average STP at the 95% level. The model also found that the average STP from the previous stage was a statistically significant predictor of average STP at the 99% level. The estimated marginal effects of an increase in the previous stage average STP was 0.713 (Kroschel et al., 2022a). This pressure translation from one stage to the next may be caused by significant stress shadow effects due to depletion and possible stress reversal. However, these may also be due to things like faulty packers and cement due to aging wellbores. Therefore, it is necessary to try and control for these unobservable factors using more advanced multivariate regression models. **This chapter will construct dynamic fixed effects models treating the data as panel data in an attempt to answer research question 2.**

4.2 Data

Panel data is data that consists of multiple observations for an individual over time (Bailey, 2017). For instance, data collected for an individual person over years or data collected on a single well over years. For this study, the data consists of multiple HF treatments per well combined into a single data set. So, the data for the models in this chapter are treated as cross-sectional as they consists of multiple observations for the same well over time.

Figure 4.1 shows the data set as cross-sectional, separating out the groups by well with associated trendlines. Although all trend in the same direction, they do not do so at the same rate and well 3 seems to treat drastically different than the other wells in the data set. We presume these can be at least partially explained fixed differences within-well. Variables such as wellbore trajectory, percent of wellbore out of zone, formation, etc. are expected to have different effects on treatment for each well. This presumption and the graphical evidence are why it is necessary to treat panel data appropriately and highlights the fixed effects inherent within each well. Data used for this model is the same data set used in chapter 3. Figures 4.2 - 4.5 show each well individually, further highlighting the differences and trends in average STP.

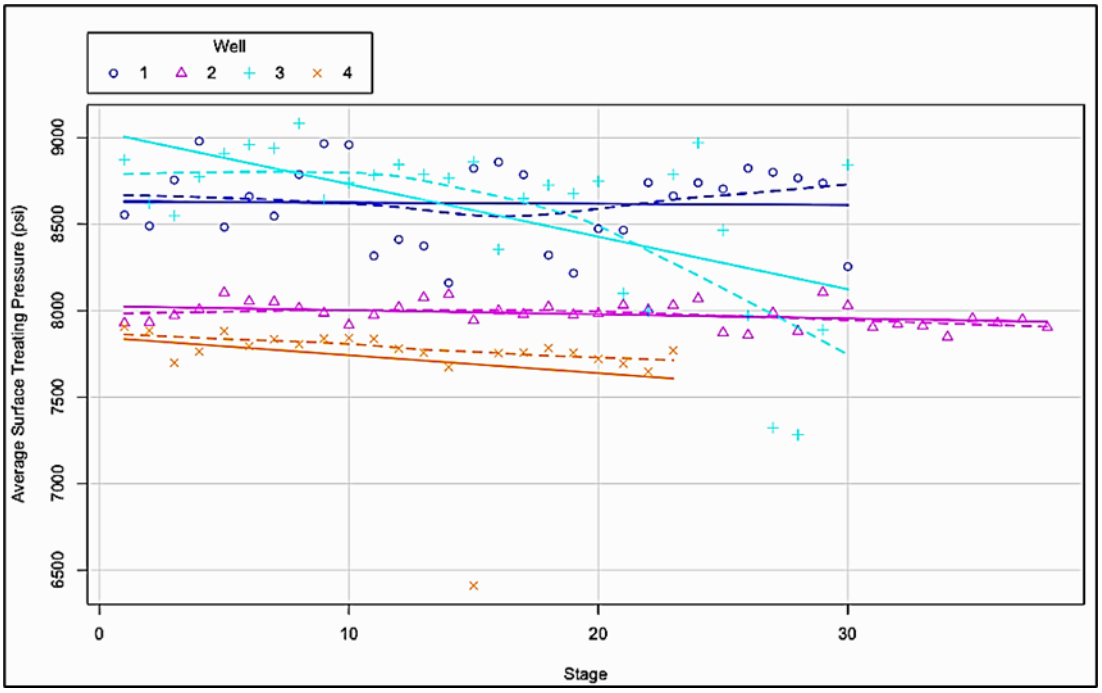

Figure 4. 1 - Cross-sectional data set from re-fractured wells with associated trendlines highlighting the effect of different within-well characteristics. Plot created in R (R Core Team, 2022. https://www.R-project.org/)

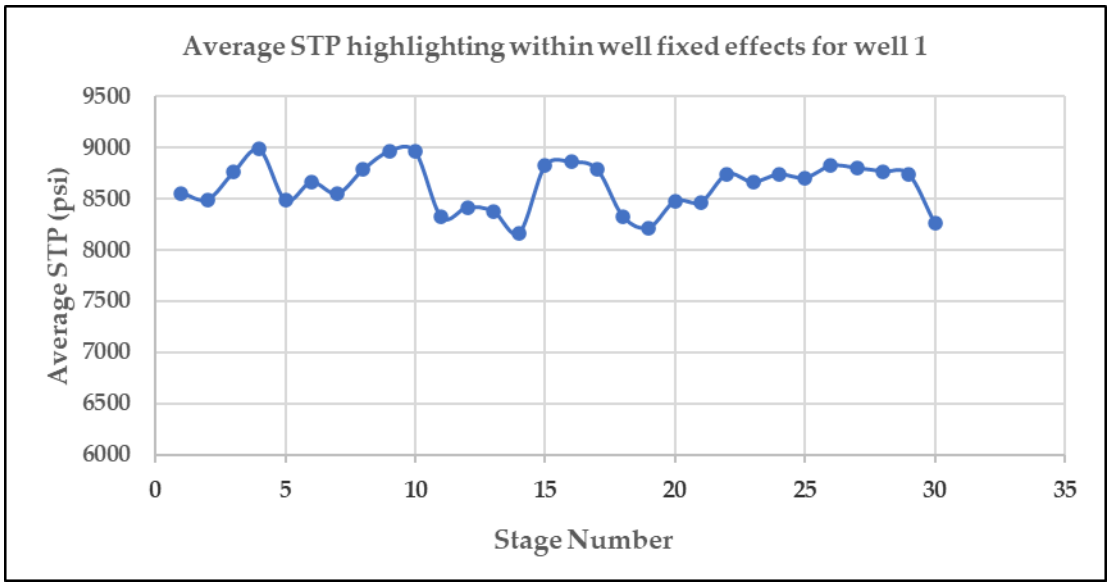

Figure 4. 2 - Plot showing trend in average STP for well 1

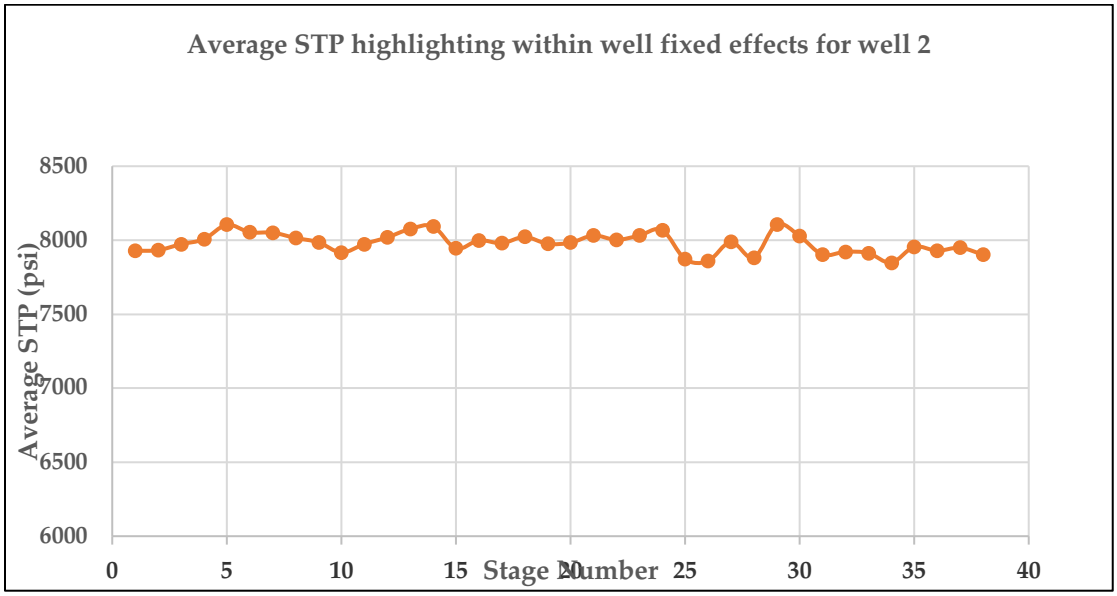

Figure 4. 3 - Plot showing trend in average STP for well 2

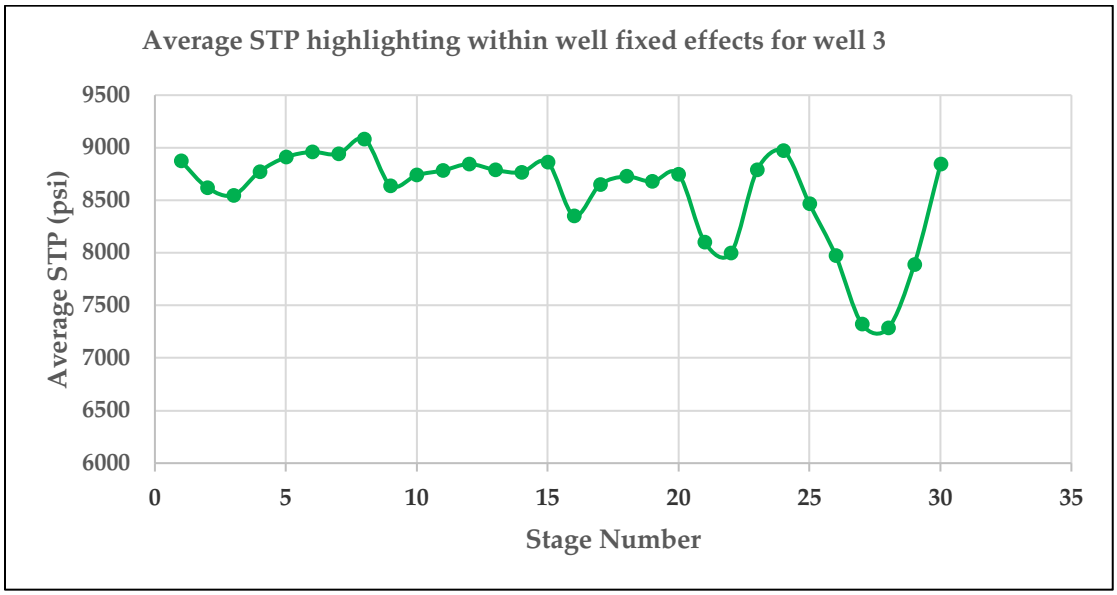

Figure 4. 4 - Plot showing trend in average STP for well 3

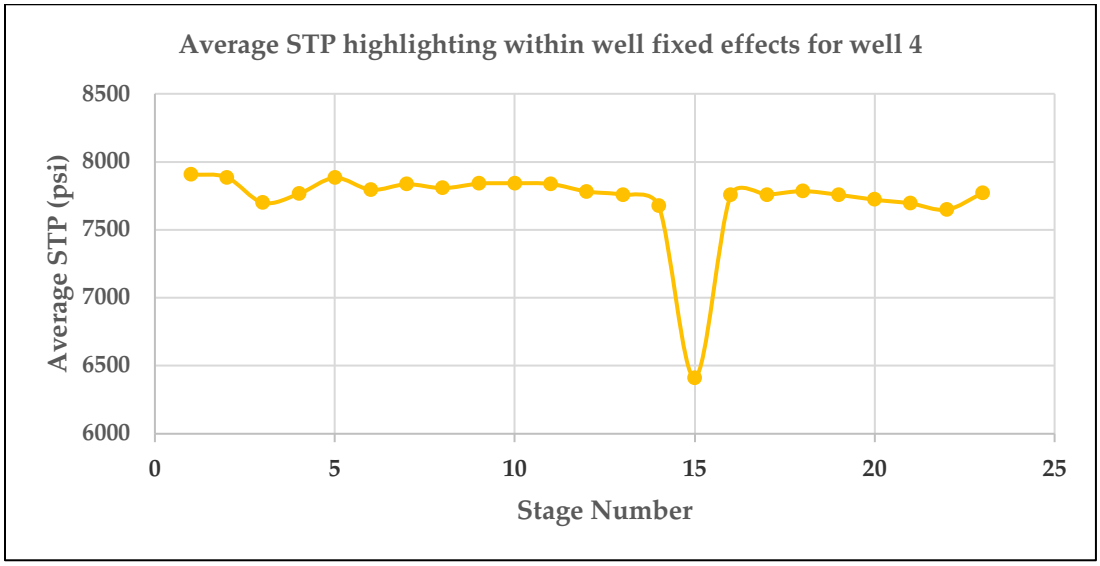

Figure 4. 5 - Plot showing trend in average STP for well 4

4.3 Panel Data Model

A pooled data model will treat all observations as independent (Bailey, 2017). So, using a pooled data model, as in chapter 3, for panel data will ignore the panel nature of

the data and any within well variation (Bailey, 2017). Here we will introduce the panel model. Equation 17 shows the generic form of a panel data model accounting for well *i* over time period *t* (Bailey, 2017).

$$Y_{it} = \beta_0 + \beta_1 X_{1it} + \ldots + \beta_n X_{nit} + \epsilon_{it} \quad (17)$$

Although this model accounts for the panel nature of the data by specifying individual observations over time, it will be helpful to account for effects inherent within individual wells. This may include data that is expensive to obtain or is outright unobservable. Often, these are the parameters we're most interested in. For instance, it is reasonable to assume that a natural fracture network that exists for a given well is unobservable in part or in entirety. Recall here the difference between uncertainty and randomness. We know the fractures are not spread randomly and could very well be mapped given enough resources and data. Although there may be some indications from data sources such as logs, tracers, offset well data such as fiber optic, microseismic, etc., knowing the extent of the entire natural fracture network will almost certainly not be observable for every well. Therefore, we must make attempts to account for parameters such as these given the inherent uncertainty. However, it is reasonable to assume that this natural fracture network will be constant prior to stimulation and constant from stage to stage during stimulation. This is where we introduce the concept of fixed effects (FE).

4.4 Fixed Effects Models

Equation 18 shows the equation of a dynamic FE model (Bailey, 2017). Again, the dynamic term, $\gamma Y_{i,t-1}$, accounts for the dynamic stress shadow effects between stages. FE

models attempt to deal with endogeneity by dividing the error term into two parts. The first part is a random error term, υ_{it}, and the second part is a FE term, α_i. The FE term α_i essentially accounts for every unobserved factor that is constant over time for individual *i*.

$$Y_{it} = \gamma Y_{i,t-1} + \beta_0 + \beta_1 X_{1it} + \ldots + \beta_n X_{nit} + \alpha_i + \upsilon_{it} \quad (18)$$

Pulling the FE term out of the error term allows for control of anything that may be correlated with the error term and/or an independent variable which causes endogeneity and bias (Bailey, 2017). FE may be thought of as taking a long list of variables, condensing them into one variable, and then controlling for that variable (Huntington-Klein, 2022). Things like perforation erosion, geologic properties and structure, reservoir heterogeneities, natural fracture networks, etc. will play a critical role in determining average STP. It is presumed that these characteristics that are inherent within a well will affect treatment and may be driving the differences in trends seen in figures 4.2 - 4.5. However, FE models allow us to control for the variation that can be expected within each well.

Complex fracture geometry can not be measured by any fracture diagnostic technique (Wu & Olson, 2015). If many of these variables are presumed to be constant, they can be accounted for using a FE model without costly data collection and computationally expensive models. For example, fracture initiation depends on stress around the wellbore caused by drilling, cementing, perforating, etc. This is unobservable. However, there is no reason to assume it will change prior to treatment and could possibly

remain constant over the course of a wellbore. In this regard, FE models have certain advantages over deterministic physics-based models.

From literature, there seems to be several trends that do not change significantly throughout HF treatments. Roberts et al. (2020) conducted a study on perforation erosion that included more than 6,000 clusters over 600 stages and noted the tendency for non-uniform proppant placement across a stage and a strong preference for heel-side perforation proppant placement and erosion (Roberts et al., 2020). It is also reasonable to assume that large geologic structures and wellbore characteristics will also remain constant through HF stimulation. For existing structures such as a natural fracture network, HF treatment will neither seal a pre-existing fracture nor, by definition, can it create a pre-existing fracture. Therefore, these may be thought of as constant between each well prior to and during treatment. Evidence from literature also suggests that other geologic and geomechanical properties may be treated as constant in the middle Bakken as well. Vertical and horizontal moduli show similarity in the middle Bakken and thus indicate elastic isotropy (Havens, 2012); Ostadhassan et al., 2012; Ostadhassan, 2013). The assumption is that any of these properties that do not vary wildly from their mean can be considered constant and captured by the FE term and thus differenced out and accounted for in the regression models using a de-meaned approach.

Production logs indicate that as many as 30% of perforations do not contribute to production due to under-stimulation from stress shadow effects (Huang & Datta-Gupta, 2017). Understanding and characterizing these effects is important to increasing production and implementing effective treatments in the field. Kroschel et al. (2022a) have already

modeled the temporal dependence of average STP from one stage to the next. However, the FE model in this chapter should give a better estimation of the temporal dependence of stress shadow effects and previous stage treatment after accounting for fixed wellbore and geologic effects, while accounting for other factors that affect treatment. Properly characterizing the temporal dependence of these effects from one stage to the next may aid engineers in effective perforation and treatment design and thus help treatment implementation in the field. The FE model is also an attempt to resolve what may endogeneity as evidenced by the coefficient estimates for perforation standoff and acid volume in chapter 3.

4.5 Fixed Effects Model Results

Two FE models were initially constructed: one with dynamic average STP and one without. The results from a FE (model 1) and dynamic FE (model 2) models are shown in table 4.1. It is important to note that any constant term was differenced out of the regression. So, the constant β_0, the presence of a 3.5" liner, and the formation are no longer in the regression results. These, along with any other unobservable fixed effect captured by the variable α_i, are now accounted for in the regression results. We see that there is still a statistically significant temporal dependence on average STP from stage to stage, although the coefficient estimate of 0.537 is smaller than that estimated from Kroschel et al. (2022a).

Table 4. 1 - Results from FE and Dynamic FE models using stargazer package from R (Hlavac, 2022)

	Dependent variable:	
	Average STP (psi)	
	(1)	(2)
Previous Stage Average STP (psi)		0.537***
		(0.089)
Perforation Standoff (ft)	10.243	17.112**
	(8.539)	(7.482)
Stage Proppant Weight (lb)	0.0003	0.001
	(0.002)	(0.002)
Total Clean Volume (bbl)	-0.091	-0.102
	(0.101)	(0.088)
Number of Perforations	0.569	-16.432
	(17.552)	(15.459)
Average Pump Rate (bpm)	-0.003	5.418
	(4.230)	(3.772)
Acid Volume Pumped (bbl)	8.786**	7.457**
	(3.409)	(2.961)
Observations	117	117
R^2	0.080	0.317
Adjusted R^2	0.003	0.252
F Statistic	1.556 (df = 6; 107)	7.014*** (df = 7; 106)
Note:		*p<0.1; **p<0.05; ***p<0.01

The results from the models closely match those in Kroschel et al. (2022a). However, the coefficient estimate for previous average STP is significantly lower. Model

2 estimates a temporal effect of 0.537 that is statistically significant. Perforation standoff is also positive and statistically significant which matches model estimates from Kroschel et al. (2022a) but is counterintuitive.

One interesting note is that the marginal effects from volume of acid pumped is still positive and statistically significant which also matches Kroschel et al. (2022a). Reasons for this will be explored in the next section.

We can also perform an F-test using under the null hypothesis H0: $\alpha 1 = \alpha 2 = \ldots = \alpha i = 0$ to see if at least one of the fixed effects are non-zero (Bailey, 2017). This will compare the models in this study and the pooled model from Kroschel et al. (2022a). Table 4 shows the model results from this study as well as the results from the Pooled model in Kroschel et al. (2022a).

Table 4. 2 - Model summaries from Panel Linear (FE) models and Pooled models from chapter 3 (Kroschel, Rabiei, & Rasouli (2022a)) using stargazer package from R (Hlavac, 2022)

	Dependent variable:	
	Average STP (psi)	
	Pooled	Panel Linear
	(1)	(2)
Previous Stage Average STP (psi)	0.713***	0.537***
	(0.075)	(0.089)
Perforation Standoff (ft)	14.615*	17.112**
	(7.796)	(7.482)
Stage Proppant Weight (lb)	-0.001	0.001

	(0.002)	(0.002)
Total Clean Volume (bbl)	-0.044	-0.102
	(0.090)	(0.088)
Number of Perforations	-23.170	-16.432
	(16.050)	(15.459)
3.5 Liner	221.913**	
	(99.258)	
Average Pump Rate (bpm)	9.471**	5.418
	(3.742)	(3.772)
Acid Volume Pumped (bbl)	7.785**	7.457**
	(3.098)	(2.961)
Formation	-14.937	
	(83.683)	
Constant	2,260.782***	
	(860.533)	
Observations	117	117
R^2	0.712	0.317
Adjusted R^2	0.688	0.252
Residual Std. Error	260.712 (df = 107)	
F Statistic	29.446*** (df = 9; 107)	7.014*** (df = 7; 106)
Note:		*p<0.1; **p<0.05; ***p<0.01

Using R (R Core Team, 2022. https://www.R-project.org/) and the pFtest() package, and using the above Pooled and FE models, the F-statistic is 11.318 with a p-value of 0.00107 under the null hypothesis that fixed effects are zero and the alternative hypothesis that there

are significant effects. Using the qf() function in R, the critical F value is 3.931. Since The F-statistic is greater than the critical value (11.318 > 3.931), we can reject the null hypothesis that at least one of the FE is zero.

4.6 Further Model Investigation

The standard error (se) refers to accuracy of the parameter estimate which is determined the by distribution of the parameter estimate (Bailey, 2017). This in turn affects possible values for the estimate and statistical significance (Huntington-Klein, 2022). Figures 4.2 – 4.5 show the individual plots for average STP vs. stage for each well. Separating the data out by well highlights an apparent cyclical trend in average STP that is different in each well. Given the nature of panel data, and the apparent trends shown in figures 4.2 – 4.5, it may necessary to report clustered standard errors.

Clustered standard errors are used when we think there are correlations between the error terms within the group. This is common in data with multiple time periods and known as temporal autocorrelation (Huntington-Klein, 2022). This violates the assumption that the error term is independent and identically distributed. This assumption requires that the error term be uncorrelated with error terms for each observation (independent) and have the same distribution for each observation (independently distributed) (Huntington-Klein, 2022). Therefore, if the error term for one observation can predict another (as is in the case of cycles or seasonal trends), the error terms are correlated. Figures 4.2 - 4.5 indicate there be some autocorrelation between stages.

Table 4. 3 - Results from Dynamic FE model with clustered se using stargazer package from R (Hlavac, 2022)

	Dependent variable:
	Average STP (psi)
Previous Stage Average STP (psi)	0.537***
	(0.097)
Perforation Standoff (ft)	17.112***
	(3.391)
Stage Proppant Weight (lb)	0.001
	(0.001)
Total Clean Volume (bbl)	-0.102
	(0.093)
Number of Perforations	-16.432***
	(2.437)
Average Pump Rate (bpm)	5.418**
	(2.305)
Acid Volume Pumped (bbl)	7.457***
	(1.388)
Note:	*p<0.1; **p<0.05; ***p<0.01

4.6.1. Previous Stage Average STP

The coefficient estimate for the previous stage average STP indicates that a 1 psi change in average STP for a stage is linearly associated with a temporal effect of a 0.537 psi increase in average STP for the subsequent stage. This estimate is lower than the dynamic pooled models constructed in chapter 3.

4.6.2. Perforation Standoff

The coefficient estimate for perforation standoff indicates that a 1 psi change in perforation standoff between stages is linearly associated with a 17.112 psi increase in average STP for the subsequent stage. This results still does not match what we would expect from physics and our domain knowledge. Further investigation will be done in the next chapter in an attempt to account for any endogeneity caused by exogenous not accounted for in the model.

4.6.3. Number of Perforations

The coefficient estimate for number of perforations indicates that adding 1 perforation is linearly associated with a 16.432 psi decrease in average STP for the stage. This makes sense intuitively as adding entry points and conduits from well-bore to formation should yield pressure relief. This is also observed when cleaning up a wellbore and retrieving the frac ball for the previous stage after flowing back a well. Well cleanup is much easier when pumping across two zones as opposed to one.

4.6.4. Average Pump Rate

The coefficient estimate for average pump rate indicates that a 1 bpm change in average pump rate is linearly associated with a 5.418 psi increase in average STP for the stage. This also makes sense intuitively since increasing rate increases friction along the wellbore and fracture

4.6.5. Acid Volume Pumped

The coefficient estimate for acid volume pumped indicates that a 1 bbl change in acid placed in formation is linearly associated with a 7.457 psi increase in average STP for the subsequent stage. This result is still counterintuitive. Possible reasons are examined in the next section.

4.7 Chapter Summary

The FE model in this study can not be used for prediction since all of the fixed well-bore effects and intercept were differenced from the data. Prediction would not be clear since the models are not accounting for other wells which have their own within-well variation that is not captured by the model. This model only provides an estimation of effect sizes and would need to be augmented with a physics based model or direct observation from fiber optic. Experiments would need to be set up in the field by design. This could be tested anecdotally, for example, by increasing the perforation standoff between zones. The hypothesis could also be tested by changing the perforation standoff halfway through well completions and constructing a difference in difference model and investigating the effects of the change in design. This will hopefully be investigated in a future study.

More data for re-fracturing treatments is necessary to expand on this idea and provide estimates of temporal dependence that may differ by company, location, formation, etc.

It is interesting to note that the coefficient estimate for acid volume from both the pooled model (chapter 3), the FE model, and the dynamic FE model are all positive. This

yields evidence that increasing acid volume tends to increase average STP in re-fractured wells. One possibility may be that the acid is opening more interior perforations thus increasing overall fracture width, but decreasing individual fracture width as individual fractures compete at a constant rate and pressure (Cleary, et al., 1993), and increasing total friction along the fracture face. There is also the possibility of faulty acid or insufficient mixing on location as the acid is delivered to location raw and cut with water on-the-fly. One area of increasing interest is that of supramolecular complexes fracture fluids. Bhat, et al. (2021) introduced a novel supramolecular structure-based fluid that is used in alkaline conditions (Bhat, et al., 2021). Assuming acid is opening interior perforations, which will increase overall width while decreasing individual width, these new fluid structures may help alleviate individual fracture width restrictions. One advantage of this fluid is that it's properties are amplified in the presence of salt (Bhat et al., 2021). This may prove even more useful as produced water reuse is becoming more common in HF designs.

It is also possible that model still suffers from endogeneity, which is why the coefficient estimates for perforation standoff and acid volume are positive and statistically significant. One possible source of endogeneity is that independent variables like perforations and perforation standoff vary only slightly within the dataset. More data may be needed with more diverse perforation stand-off, clusters, and number of perforations per cluster.

It's important to reiterate that the coefficient estimates will be bias due to the structure of the model. One way to combat this is to use what is known as an instrumental variable (IV) but these can be complicated and lead to imprecise results (Bailey, 2017).

However, it is worth noting that the bias becomes less severe as the number of observations for each panel (in this case each well) increases. Observations over ~20, the problem of bias becomes less serious (Bailey, 2017). Thus, since there are more than 20 observations for each well in the data set, the bias will be accepted as tolerable.

This FE model does not account for well-well interactions as this would be a different type of model, for example - parent/child or re-frac/new-drill well interactions. Interactions between wells that are on a single pad may also exist and were not accounted for in this study. However, further model specification needs to be done to resolve the issue of endogeneity.

In the next section, a feature engineered model will be constructed utilizing more complex relationships called interaction effects. The models will also add additional variables in an attempt to resolve suspected endogeneity in the pooled and FE models.

Chapter 5

5. Feature Engineered Model for Improved Performance

5.1 Introduction

Chapter 3 constructed pooled multivariate regression models to estimate the marginal effects of completions parameters on treatments. Chapter 4 built upon this model by accounting for within-well fixed effects to better understand these marginal effects. The results from these models show there is significant temporal stress shadow effects from stage to stage. However, some concern still remains about potential endogeneity since the results for perforation standoff show positive marginal effects which doesn't match what we know physically. The coefficient estimates for acid volume also hint there may be endogeneity. The inferential models in chapter 3 and chapter 4 are what are known as high bias-low variance models, meaning the coefficient estimates may deviate from the true value, but they are not very sensitive to small changes in the underlying data set (Kuhn & Johnson, 2020). However, they may benefit from feature engineering which is a process of manipulating the model to better represent the data set (Kuhn & Johnson, 2020). In particular, there is reason to think that some predictor variables in the model may exhibit interaction effects which will relate to the outcome (Kuhn & Johnson, 2020).

Haustveit et al. (2022) measured poroelastic responses from seven subsequent frac stages using a bottom hole pressure gauge (BHPG) installed in the first stage of a well. The authors state the treatment designs were held relatively constant throughout this process. The authors measure a non-linear, attenuating poroelastic responses the further the stage is from the BHPG (Haustveit, Elliot, & Roberts, 2022). The tools developed by the authors

provide effective models for estimating a pororelastic response that occurs away from a fracture face (Haustveit et al., 2022). This insight will be useful to account for non-linear features into the previous multivariate models from chapter 3 and chapter 4 as well as testing the hypothesis that there are interaction effects present between perforation standoff and temporal stress shadow effects. There exists a trade-off between bias and variance in statistical modelling and is commonly known as the "bias-variance tradeoff" (Kuhn & Johnson, 2020). Variance refers to the extent to which data can differ and bias refers to the degree to which the data or predictions are different from the true value (Kuhn & Johnson, 2020).

Supplementing low variance models with engineered representations of data that decrease bias may also diminish the bias-variance trade-off mentioned (Kuhn & Johnson, 2020). Ponomareva et al. (2021) concluded the bottom hole flowing pressure (BHFP) of a well is governed by complex relationships (Ponomareva et al., 2021). The authors approached the problem of complexity by constructing a feature engineered multivariate regression models for the Solikamsk depression in the Perm Krai by taking a multilevel approach. Using a multilevel approach, the authors were able to eliminate irrelevant factors and account for different domes within the Un'vinskoye deposit (Ponomareva et al., 2021). Darabi et al. (2020) used feature engineering to improve the predictive power of a machine learning model. The authors use raw data to derive a set of important wellbore, operational, and reservoir properties to incorporate into their model (Darabi et al., 2020). Kroschel et al. (2022b) performed some feature engineering on previous work by accounting for within-well fixed effects, treating the data as panel data, to enhance the inferential dynamic model in Kroschel et al. (2022a)

Extracting knowledge from data and establishing a fast and reliable model to predict and carry out optimization is of interest for engineers (Luo, et al., 2022). According to economist John A. List, however, if the data are not causal, decisions made are like "pushing on a string" (List, 2022). The basic idea behind causality is that causal inference necessitates prediction; but one can not infer causality from the ability to predict (Sakhardande & Devegowda, 2021). Therefore, variables should be selected that are thought to have a causal relationship between predictor and the outcome variable. In this chapter, we will focus on the supervised prediction of average STP using feature engineering. Supervision refers to predicting known outcomes using given predictors (Kuhn & Johnson, 2020). This chapter will build off of the inferential work from chapter 3 (Kroschel et al., 2022a) and chapter 4 (Kroschel et al., 2022b) by performing feature engineering to enhance the prediction performance and resolve endogeneity issues from previous models using interaction effects identified in domain knowledge. **This feature engineered model will attempt to answer research question 3 and increase predictive capabilities of the multivariate regression models while also attempting to resolve the issue of endogeneity.**

By increasing the predictive power of the models, we will have a better estimate of the average STP for re-fractured wells. This is important because STP is often a binding constraint on re-fractured wells due to wellbore degradation and field personnel must have a way to estimate the effects of altering completion parameters on STP if they are to make effective design changes in real time. In addition, by resolving the endogeneity issues, informed decisions can be made based off causal evidence.

5.2 Interaction Effects

For predictive purposes, the majority of the variation in the data may be explained by the cumulative effect of variables in a regression (Kuhn & Johnson, 2020). Additional variation may be explained by interaction between variables (Kuhn & Johnson, 2020). However, inclusion of erroneous predictors can reduce model performance and feature engineering can help alleviate this problem by getting rid of those predictors and reducing the extra noise they cause (Kuhn & Johnson, 2020). Kuhn and Johnson (2020) give an example of how interactions between variables might work by describing farming with no water and fertilizer, farming with no fertilizer and water, and then farming with both fertilizer and water. The interaction between fertilizer and water will almost certainly produce more crops than either of the previous scenarios (Kuhn & Johnson, 2020). Equation 19 shows the dynamic form for the interaction model constructed in this study (Kuhn & Johnson, 2020).

$$Y_i = Y_{i,t-1} + \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_{\mathrm{N}} X_1 X_2 + \varepsilon \tag{19}$$

The interaction term may characterize the interaction between continuous or continuous and dummy variables. Looking at the basic model, we see the coefficient estimate β_{N} estimates the response for the interaction. Here, if the coefficient is positive and greater than one, then a one unit increase in X_2 will yield a more than unit increase in X_1 as a response. In addition, if the response in positive and less than one, then a one unit increase in X_2 will yield a less than one unit decrease in X_1as a response. The opposite interpretation will be true if the coefficient estimate is negative an greater than one.

As a fracture is initiated and dilated, a resulting force that is perpendicular to the fracture face creates a resulting stress on the formation that attenuates with distance from the fracture (Haustveit et al., 2022). Net pressure and stress differences both tend to increase, maybe substantially, with increasing number of sequential fractures combined with decreased fracture spacing (Roussel & Sharma, 2011). This may affect subsequent HF treatment implementation. These effects extend beyond the top perforation and may therefore affect the fracture initiation at the bottom perforation of the subsequent stage (Barree, 2020). Therefore, it is reasonable to assume that the interaction between stress shadows from HF treatment and perforation standoff will affect subsequent HF treatments and may be included in the feature engineered. The updated causal diagram is shown in figure 5.1, with an interactive relationship between perforation standoff and stress shadow effects. The cross-over in figure 5.2 as well the statistically significant results from table 5.1 identify interaction effects between perforation standoff and the square root of the previous stage avg. STP. Another addition to the feature engineered model is the stage depth which may more accurately account for pipe friction.

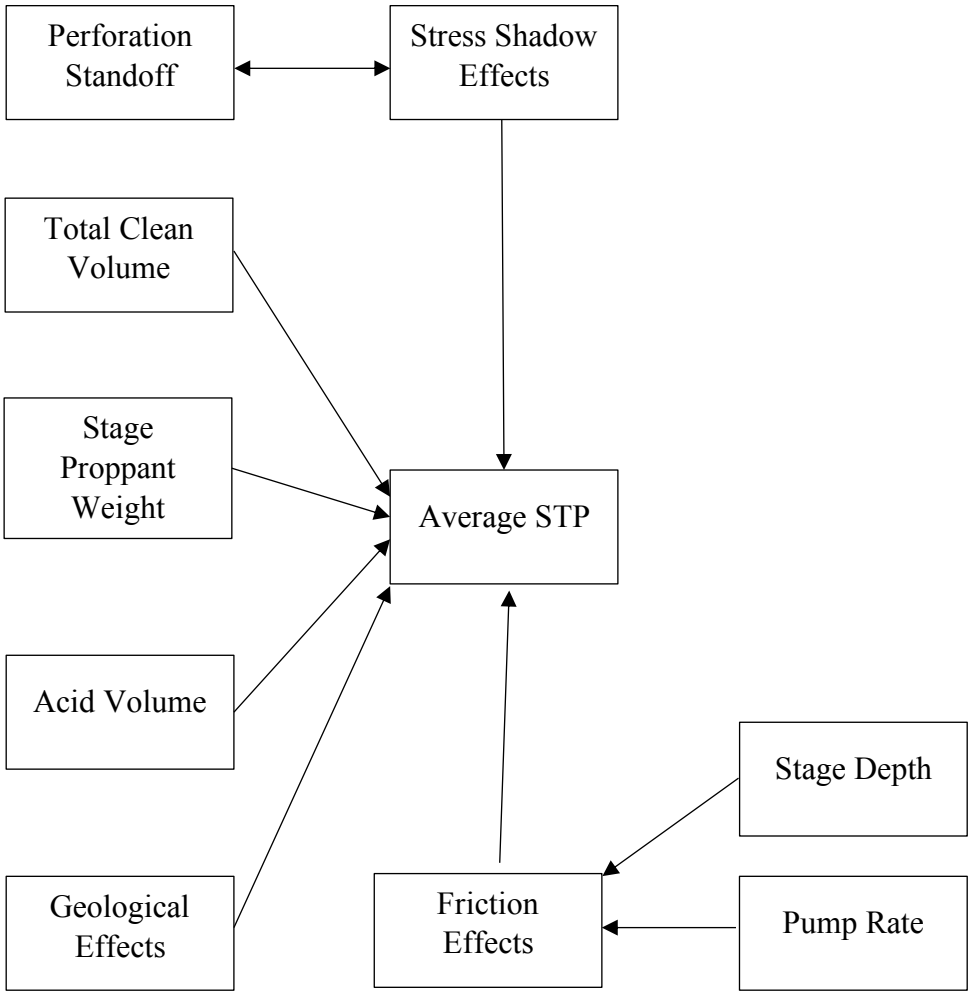

Figure 5. 1 - Updated causal diagram assuming an interactive relationship between perforation standoff and stress shadow effects and stage depth to more accurately account for pipe friction effects on average STP

Table 5. 1 – Summary table for interaction of perforation standoff and non-linear average STP using stargazer package in R (Hlavac, 2022. https://CRAN.R-project.org/packages=stargazer)

	Dependent variable:
	Average Surface Treating Pressure (psi)
Perforation Standoff (ft)	-8.370**
	(2.785)
Square Root of Previous Stage Average Surface Treating Pressure (psi)	-426.283***
	(36.122)
Interaction	5.007***
	(0.433)
Constant	8,286.808***
	(158.490)
Observations	121
R^2	0.547
Adjusted R^2	0.536
Residual Std. Error	317.133 (df = 117)
F Statistic	47.163*** (df = 3; 117)
Note:	*p<0.1; **p<0.05; ***p<0.01

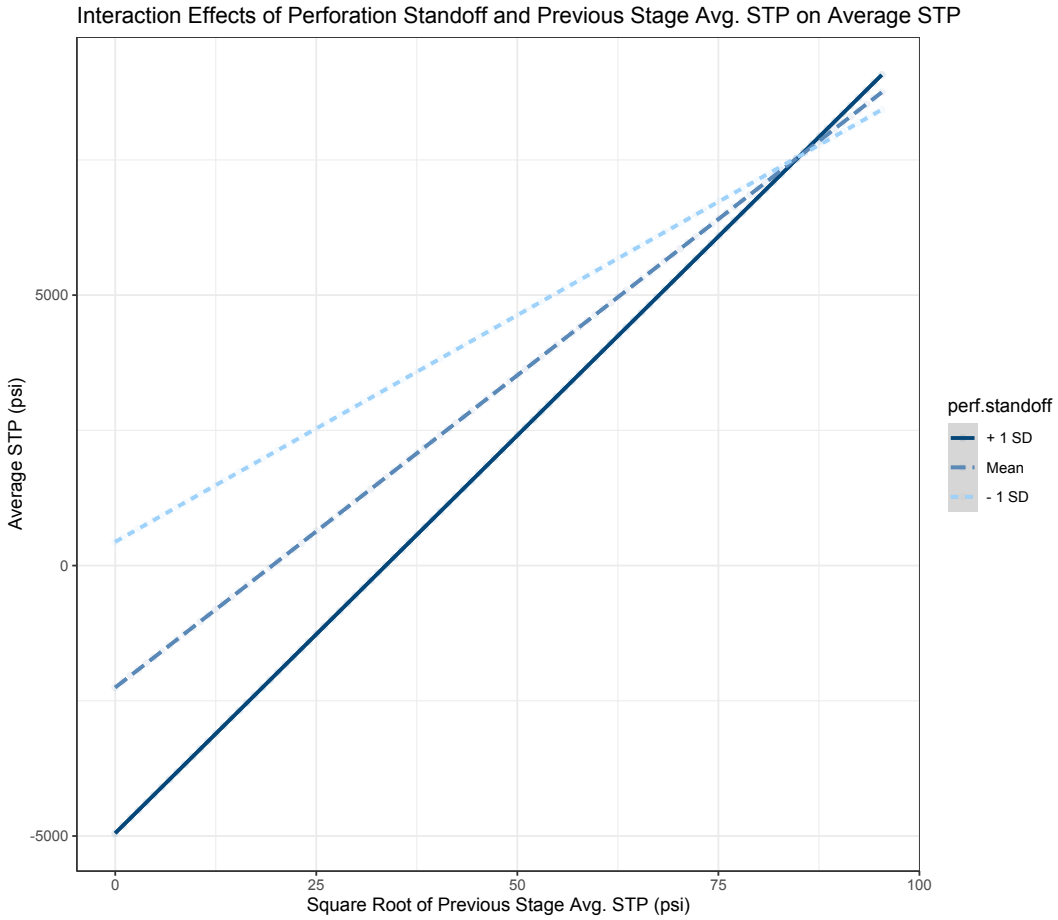

Figure 5. 2 - Interaction effects of temporal avg. STP measurements and perforation standoff using R (R Core Team, 2022. https://www.R-project.org/)

The interaction coefficient shown in table 5.1 is positive and statistically significant and both of the other terms are statistically significant, which indicates the relationship may be synergistic (Kuhn & Johnson, 2020).

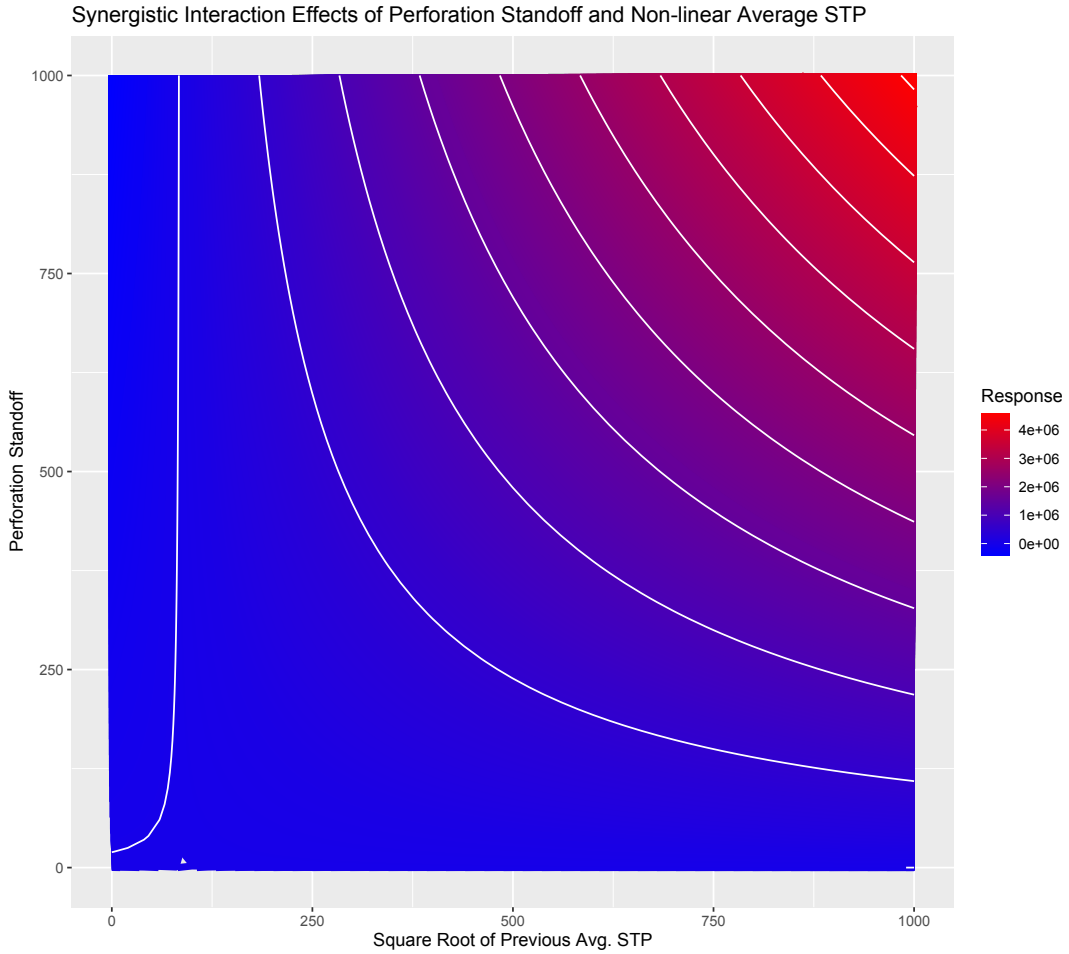

Figure 5. 3 - Evidence of possible synergistic interaction effects between perforation standoff and square root of previous stage avg. STP using R (R Core Team, 2022. https://www.R-project.org/)

Figure 5.3 shows the predicted response between perforation standoff and the non-linear temporal dependence of average STP, the increasing response as each predictor increases indicates there may be synergistic effects (Kuhn & Johnson, 2020).

5.3 Chapter Summary

Table 5.2 shows the model results for the full feature engineered model. The left column provides coefficient estimates using standard OLS and the right column shows the results after clustering standard errors. Recall from chapter 4 that clustered standard errors are used when we think there are correlations between the error terms within the group. Recall that clustering se is common in data with multiple time periods and known as temporal autocorrelation (Huntington-Klein, 2022) which violates the assumption that the error term is independent and identically distributed. Clustering the se simply provides more confidence in the coefficient estimates.

Table 5. 2 - Summary of feature engineered model results using stargazer package from R (Hlavac, 2022)

	Dependent variable:	
	Average STP (psi)	
	OLS	*Clustered Std. Errors*
	(1)	(2)
Square Root of Previous Stage Avg. STP (psi)	-3.457	-3.457
	(3.890)	(3.270)
Perforation Standoff (ft)	-234.463***	-234.463**
	(40.481)	(91.575)
Total Proppant (lbs)	-0.004***	-0.004***
	(0.001)	(0.002)
Number of Perforations	-25.354	-25.354***
	(15.372)	(2.036)
Acid Volume Pumped (bbls)	-1.570	-1.570
	(1.785)	(1.136)
3.5 Liner (binary)	495.091***	495.091***
	(95.032)	(164.191)
Total Clean Volume (bbl)	0.041	0.041
	(0.052)	(0.046)
Average Rate (bpm)	26.280***	26.280***

	(4.075)	(9.750)
Formation (Relative to Middle Bakken)(binary)	238.993***	238.993***
	(74.108)	(84.917)
End Depth of Stage (ft)	0.105***	0.105**
	(0.016)	(0.045)
Interaction Term	2.780***	2.780**
	(0.500)	(1.072)
Constant	6,131.671***	6,131.671***
	(650.116)	(1,154.717)
Observations	121	
R^2	0.739	
Adjusted R^2	0.712	
Residual Std. Error	249.656 (df = 109)	
F Statistic	28.009*** (df = 11; 109)	
Note:		*p<0.1; **p<0.05; ***p<0.01

Looking at table 5.2, the coefficient estimate for the interaction term is positive and statistically significant at the 95% level after clustering se and accounting for other variables that we presume affect average STP based off domain knowledge. However, the coefficient for the square root of previous stage average STP is no longer significant while the coefficient estimates for perforation standoff and the interaction are. This is evidence that the interaction may in fact be atypical instead of synergistic as presumed (Kuhn & Johnson, 2020).

As is the case for the fixed models in chapter 4, clustering the se at the well level may yield more accurate results for the coefficient se since it may be necessary to account for spatial autocorrelation (well location and trajectory) and temporal autocorrelation (stress shadow effects) when estimating average STP (Huntington-Klein, 2022). This will only change the se of the coefficient estimates. After clustering, the se for number of perforations drastically decreased and is statistically significant at the 99% level. The

standard error for the interaction term and perforation standoff increased and significant at the 95% level. The following subsections will discuss the results of the model with clustered se and discuss the implications each estimated coefficient estimate and the associated marginal effects.

The marginal effects estimated from the feature engineered model are discussed in the following subsections. Notice that the coefficient estimates are the same for both models in table 5.2 with un-clustered and clustered se. The statistical significance for each coefficient estimate in the following sections will refer to significance levels after clustering.

5.3.1. Square Root of Previous Stage Average STP

The coefficient estimate for non-linear temporal average STP is -3.457 with a se of 3.27 after clustering at the level of the well. So, The se is almost as large as the estimate providing evidence of little confidence in the coefficient estimate. This estimate is not statistically significant.

5.3.2. Perforation Standoff

The coefficient estimate for perforation standoff is -234.463 and is statistically significant at the 95% level. This estimate indicates that increasing the distance between the top perforation of one stage and the bottom perforation of the next stage will yield a decrease of roughly 235 psi in average STP. It is important to note that the limits of this estimate are probably within the range of the data. For instance, it is not reasonable to assume that increasing the distance between stages by 20 ft will yield 4700 psi decreases

in average STP. However, we interested in making decisions at the margin. The implications of this finding will be discussed in chapter 6.

5.3.3. Total Proppant

The coefficient estimate for the total proppant pumped for a stage is -0.004 and is significant at the 99% level. This indicates that increasing the stage proppant by one lbm yields an expected decrease of 0.004 psi in average STP. The direction of this relationship is consistent with domain knowledge and the relationship derived in chapter 2 (figure 2.8 and equation 4).

5.3.4. Number of Perforations

The coefficient estimate for the number of perforations is -25.354 and is significant at the 99% level. This indicates that adding one perforation per stage yields expected decreases in average STP by roughly 25 psi.

5.3.5. Acid

The coefficient estimate for acid is -1.57 and is not statistically significant. This yields evidence of what is observed qualitatively in the field where this data was collected. From field observation, there does not seem to be any difference in STP or pumping duration for a stage when acid is pumped, indicating no significant pressure relief that allows for more rate (C. Brown, personal communication, April 18, 2023).

5.3.6. 3.5” Liner

The coefficient estimate for the presence of a 3.5 inch liner is 495.091and is significant at the 99% level. This indicates that a 3.5 inch liner yields average STPs roughly 495 psi higher than average STPs on wells with a 4.5” liner.

5.3.7. Total Clean Volume

The coefficient estimate for total clean volume is 0.041 and is not statistically significant. This indicates that stage size is not a statistically significant predictor of average STP.

5.3.8. Average Rate

The coefficient estimate for average rate is 26.280 and is significant at the 99% level. This indicates that a one bbl/minute increase in average rate will increase average STP by an expected 26.28 psi

5.3.9. Formation

The coefficient for formation is 238.993 relative to the Middle Bakken indicating we expect roughly 239 psi higher average STP in the Three Forks formation. This is often observed in the field and matches what we expect since the Three Forks formation has a higher TVD and lies below the Middle Bakken (as shown in figure 2.5).

5.3.10. End of Stage Depth

The coefficient estimate for stage depth is 0.105 and is significant at the 95% level. This addition into the feature engineered model captures the effects of pipe friction and differences in pipe friction due to stage depth.

5.3.11. Interaction Term

The coefficient estimate for the interaction between non-linear temporal average STP and perforation standoff is 2.78 and is significant at the 95% level.

Figure 5.4 - figure 5.6 shows the coefficient plots for the statistically significant individual coefficient estimates. These plots provide a visual representation of the coefficient estimates and the associated 95% confidence interval. Meaning, the confidence band around the coefficient estimate shows with 95% confidence where the true value is expected to lie and we can be 95% confident in the direction of the estimates.

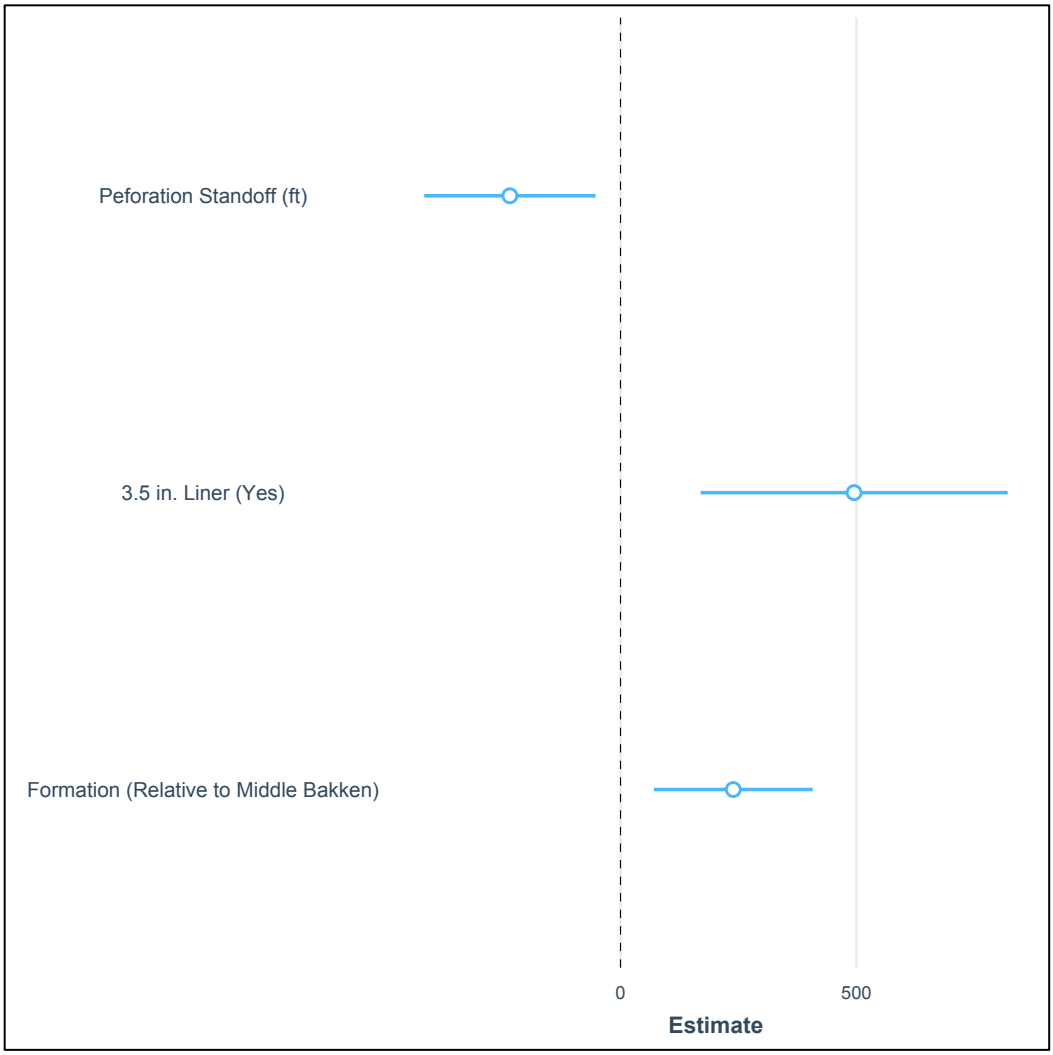

Figure 5. 4 - Coefficient plot showing the effects of perforation standoff, presence of a 3.5 inch liner, and formation with 95% confidence interval

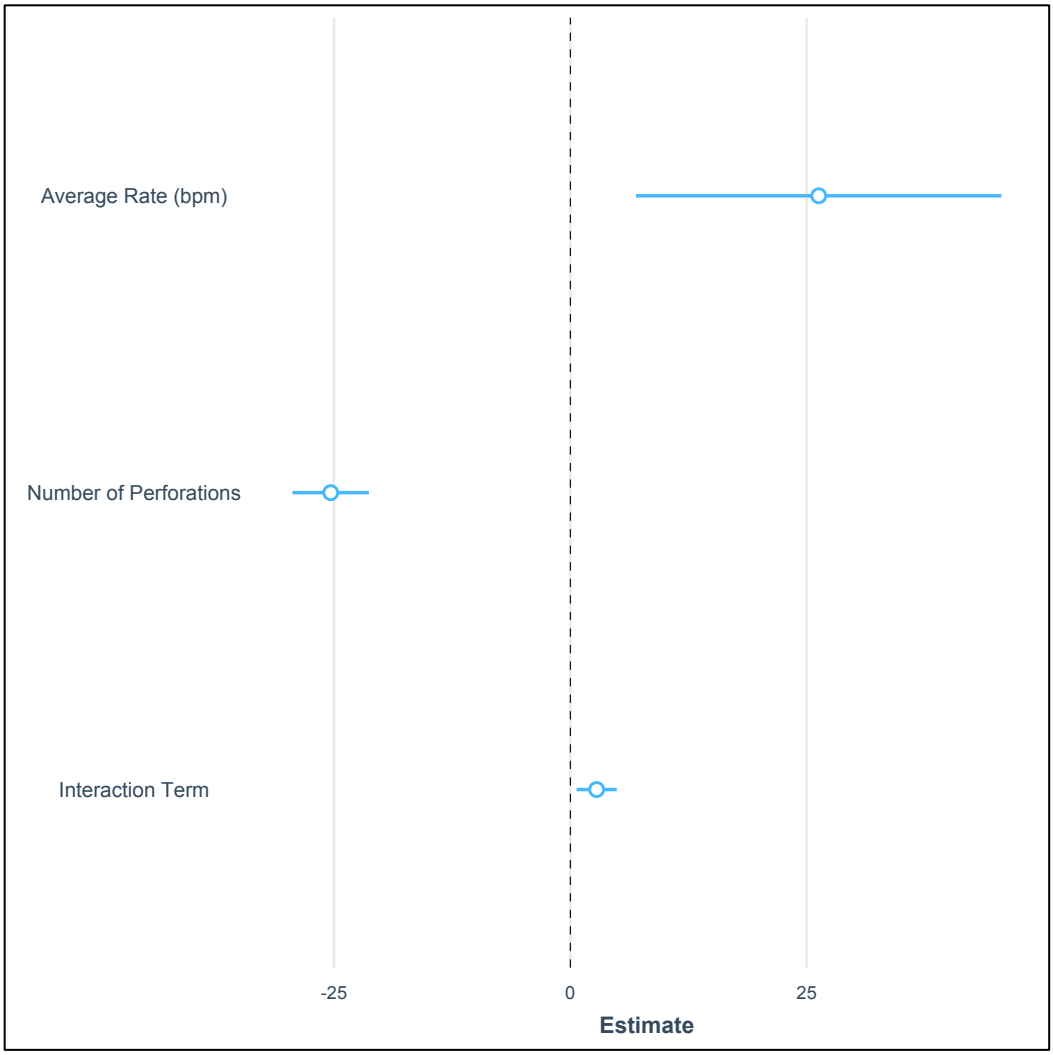

Figure 5. 5 - Coefficient plot showing effects of average rate, number of perforations, and the interaction of non-linear, temporal average STP and perforation standoff with 95% confidence interval.

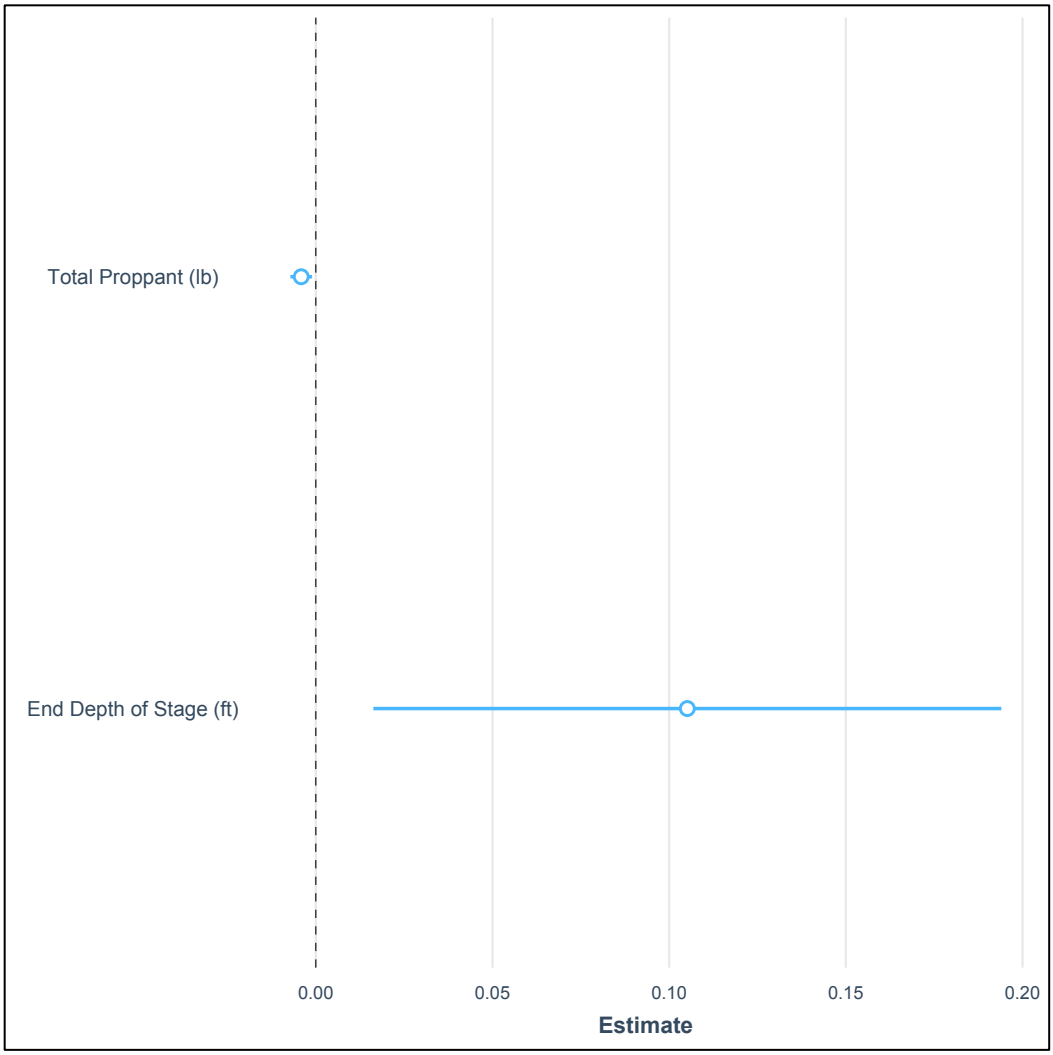

Figure 5. 6 - Coefficient plot showing the coefficient estimates for total proppant and end depth for the stage with 95% confidence interval.

Figure 5.7 shows the variable importance of each predictor and ranks them in descending order. Building off of the analysis of the coefficient estimates, the square root of the previous stage average STP is the least impactful variable in the model, while perforation standoff and the interaction are the 3rd and 4th most impactful variables, respectfully. Refer to table 3.2 for variable definitions with respective units.

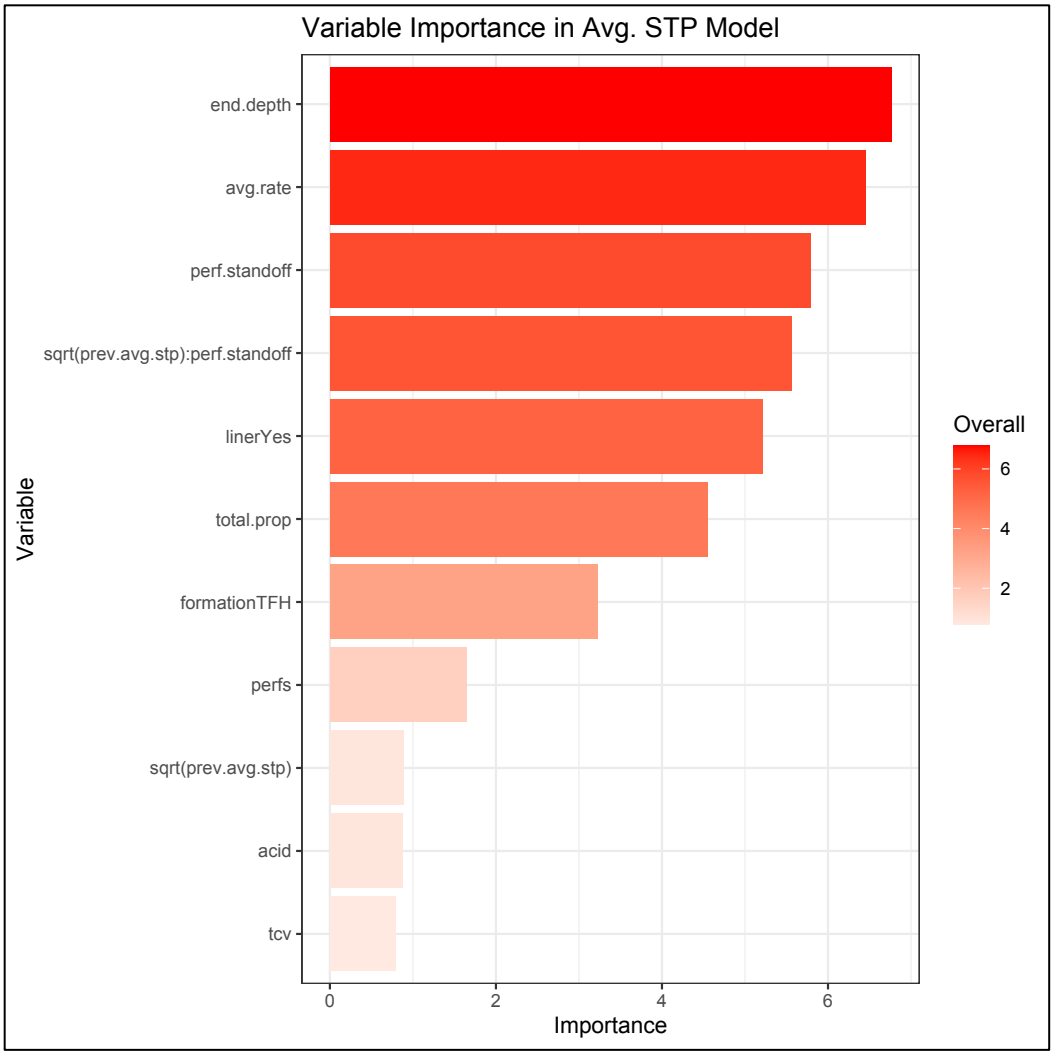

Figure 5. 7 - Variable importance for all independent variables in final feature engineered model

The addition of these variables changed the model significantly and greatly improved the interpretability of the coefficients. It also appears the model suffered from what is known as omitted variable bias. Omitted variable bias is when an independent variable is correlated with another variable in the error term, thus omitting an important predictor variable from the regression model and thus violates the exogeneity assumption (Huntington-Klein, 2022). In this case, the bias refers to spurious coefficient estimates that

will, on average, yield incorrect insights from coefficient estimates (Huntington-Klein, 2022). The coefficient estimates are now in line with what our domain knowledge hypothesized. Acid volume was no longer statistically significant, indicating it provides no pressure relief over the stage. This is observed anecdotally in the field and acid is often cut out of the design early in well treatment. Perforation standoff now has a negative and statistically significant effect on average STP indicating that increasing the distance between stages decreases the effects of stress shadows. This indicates that there was endogeneity present on both the pooled and fixed effects models and there was exogenous variation from pipe friction and interaction between perforation standoff and stress shadows that was unaccounted for. By accounting for this exogenous variation, not only can we draw conclusions about causality of the predictors that affect average STP, but the predictive power of the model was improved.

Chapter 6

6. Discussion, Applications, and Conclusions

Portions of Chapter 1 were taken from Kroschel, J. Rabiei, M. Rasouli, V. Modeling Temporal Dependence of Average STP in the Williston Basin Using Dynamic Multivariate Regression. Energies 2022, 15, 2271 and Kroschel, J. Rabiei, M. Rasouli, V. Accounting for Fixed Effects in Re-Fracturing Using Dynamic Multivariate Regression. *Energies* 2022, 15, 5451. https://doi.org/10.3390/en15155451.

6.1 Discussion and Application of Models

The next section will address the conclusions of the study and tie them back to the hypothesis and research question using the results from the feature engineered model to answer each question. Table 6.1 are the results from the feature engineered model constructed in chapter 5 with the identified high margin predictors highlighted in red. The coefficient estimates for each independent variable represent the expected marginal effects of each variable on average STP.

Table 6. 1 - Summary of feature engineered model results including clustered standard errors using stargazer package from R (Hlavac, 2022)

	Dependent variable:	
	Average STP (psi)	
	OLS	*Clustered Std. Errors*
	(1)	(2)
Square Root of Previous Stage Avg. STP (psi)	-3.457	-3.457
	(3.890)	(3.270)
Perforation Standoff (ft)	-234.463***	-234.463***
	(40.481)	(91.575)
Total Proppant (lbs)	-0.004***	-0.004***
	(0.001)	(0.002)
Number of Perforations	-25.354	-25.354***
	(15.372)	(2.036)
Acid (Binary)	-1.570	-1.570
	(1.785)	(1.136)

3.5 Liner (binary)	495.091***	495.091***
	(95.032)	(164.191)
Total Clean Volume (bbl)	0.041	0.041
	(0.052)	(0.046)
Average Rate (bpm)	26.280***	26.280***
	(4.075)	(9.750)
Formation (Relative to Three Forks)(binary)	238.993***	238.993***
	(74.108)	(84.917)
End Depth of Stage (ft)	0.105***	0.105**
	(0.016)	(0.045)
Interaction Term	2.780***	2.780*
	(0.500)	(1.072)
Constant	6,131.671***	6,131.671***
	(650.116)	(1,154.717)
Observations	121	
R^2	0.739	
Adjusted R^2	0.712	
Residual Std. Error	249.656 (df = 109)	
F Statistic	28.009*** (df = 11; 109)	
Note:		*p<0.1; **p<0.05; ***p<0.01

These marginal effects are important for re-fracturing design. Using these estimates, engineers can alter completion designs and minimize average STP to minimize cost. For example, the coefficient for perforation standoff indicates that increasing the length between the top perforation from one stage to the bottom perforation of the next stage will decrease expected average surface STP by roughly 230 psi, thus mitigating stress shadow effects. Using this data, engineers can consider tradeoffs between increasing perforation standoff, expected pressure decrease, and lateral length loss. The same logic applies for running a 3.5” liner with a coefficient estimate of roughly 485 psi, engineers can consider the trade-offs between the expense of running a 4.5” liner and the expected pressure increases in running a 3.5” liner. Given their unique cost structure for

inputs such as casing costs, rig-time, and fuel, engineers can now compare the cost of running a 4.5” liner with a 3.5” liner and the resulting pressure. Liner size will be an important factor in reusing produced water as fields mature and are re-fractured. Appendix A constructs a difference-in-difference model that estimates positive and statistically significant effects from using produced water as a base fluid for stimulation.

Negative and statistically significant marginal effects on average STP from using a 4.5” liner as opposed to 3.5” liner for re-fracturing has temporal considerations. For instance, it may be beneficial to install 5.5” liners instead of the 4.5” liners in the lateral section as discussed in Appendix C. The installation of larger liners will at least minimize the effects of reusing produced water for stimulation as well as more productive wireline practices as discussed in Appendix C. Although there are additional considerations with fluid volumes required for wireline pump down operations due to minimal drag in larger ID casings, this can be mitigated through string design with parachutes or larger OD gun strings.

The model utilizing feature engineering increases the predictive power substantially. The average % error was reduced from 5.8% in the initial regression models to 2.1% per stage with the current feature engineered, predictive model. Table 6.2 shows the predictive results using the model constructed in this study and includes the results from Kroschel et al. (2022). The offset well used for prediction was in the Three Forks formation and contained a 3.5” liner. So, these binary values are “1” in the models. This form of model validation is known as external validation and using the feature engineered model to predict

average STP on a different well in the same geographic is the second most stringent form of external validation (Harrell, 2015)

Table 6. 2 – Average STP prediction results comparing Pooled models from chapter 4 and feature engineered model from chapter 6.

Stage	Perforation Standoff	End Depth	Total Proppant	Total Clean Volume	Avg. Rate	Previous Stage Avg. STP	Avg. STP	Predicted STP From First Paper	% Error from Pooled Model	Predicted from Feature Engineering	% Error Feature Engineering
1	0	19974	61748	4230	21.6	NA	8719	NA	NA	NA	NA
2	26	19675	225181	8317	29.3	8719	8903	7990	10.3	8798	1.2
3	26	19377	218337	6065	28	8903	8749	8285	5.3	8713	0.4
4	26	19078	224202	5689	29.7	8749	8968	8142	9.2	8626	3.8
5	26	18779	227257	5225	29.9	8968	8799	8313	5.5	8647	1.7
6	26	18480	226392	5007	28.2	8799	8731	8191	6.2	8503	2.6
7	26	18182	226608	4960	29.1	8731	8848	8153	7.9	8469	4.3
8	26	17883	226521	5077	29.9	8848	8730	8140	6.8	8510	2.5
9	26	17584	227421	4666	33.6	8730	8626	8108	6	8507	1.4
10	26	17285	226288	4396	35.3	8626	8859	8065	9	8474	4.4
11	26	16987	227810	4300	37.6	8859	8806	8255	6.3	8580	2.6
12	26	16688	228025	4185	39.1	8806	8732	8235	5.7	8563	1.9
13	26	16389	228048	4365	42.6	8732	8582	8207	4.4	8606	0.3
14	26	16091	228459	4172	46.5	8582	8780	8145	7.2	8609	1.9
15	26	15792	230832	4187	44.6	8780	8881	8266	6.9	8597	3.2
16	26	15493	213499	3933	43.9	8881	8608	8359	2.9	8644	0.4
17	26	15194	227737	3903	42	8608	8631	8134	5.8	8404	2.6
18	26	14896	217414	3782	44.6	8631	8650	8191	5.3	8487	1.9
19	26	14597	217405	3771	47.8	8650	8728	8235	5.6	8547	2.1
20	26	14298	217358	4018	41.2	8728	8258	8131	1.5	8390	1.6
21	26	14000	217387	3742	49.1	8258	8813	7969	9.6	8372	5
22	26	13701	220905	4441	48.3	8813	8301	8323	0.3	8555	3.1
23	26	13402	216973	3813	51.5	8301	8339	8020	3.8	8397	0.7
24	26	13103	217031	3796	49	8339	8347	8024	3.9	8317	0.4
25	26	12805	217689	3877	50.6	8347	8277	8040	2.9	8334	0.7
								Average	5.8	**Average**	2.1

6.1.1 Average STP Effects on Fuel Consumption

Wang and Chen (2018) highlight the difficulty in broad, physics-based models by identifying latitude and longitude as significant factors in predicting production for wells in the Montney Shale (Wang & Chen, 2018). The information asymmetry is simply too large to create accurate models rooted in physics. However, what operators almost certainly have is a wealth of treatment data. This abundance of treatment data is a direct measurement of treatment behavior for all of the operators acreage.

One implication for designing hydraulic treatments around average STP is fuel consumption and fuel efficiency. Fuel consumption is defined as the amount of fuel consumed over a given period of time (Grisso, Perumpral, Vaughan, Roberson, & Pitman, 2010). This fuel consumption may be thought of as the specific volume fuel consumption and is measured in units of gallons per horsepower-hour (gal/hp-h) (Grisso et al., 2010). The specific volume fuel consumption is not affected by the size of the engine and is used to compare efficiency of diesel engines under varying operational situations (Grisso et al., 2010). This has a range from 0.0476 – 0.1110 gal/hp-h (Grisso et al., 2010). The inverse of this yields the fuel efficiency which ranges from 12-21 hp-h/gal for diesel engines (Grisso et al., 2010).

It is well known that the hydraulic horsepower (HHP) needed operationally for stimulation is calculated using equation 20.

$$HHP = \frac{P\ x\ Q}{40.8} \tag{20}$$

Where P is the expected pressure measured in psi, Q is the designed flow rate in bpm. So, higher pressures require more HHP and thus increases diesel consumption. This can be a major cost in implementation. Therefore, optimizing the treatment around expected average STP, or estimating the marginal effects of completion parameters, can help minimize expected diesel costs.

6.1.2 Average STP Considerations for Operations

The idea of circular economies (CE) has been gaining momentum in academic literature. Although there are many definitions, the most frequently used definition is a combination of reduce, reuse, and recycle activities (Kirchherr et al., 2017). CE is a unique economic concept using restorative designs to help sustainable economic development and address environmental challenges with waste (Aloini et al., 2020). The idea has also touched the O&G industry as more rigorous practices in resource management to minimize environmental impacts from O&G production has increased (Castilla et al., 2021). One major concern in resource management is the reuse and recycling of produced water (PW). PW has generally been disposed of through using disposal wells that are drilled into deep basement formations (Tomomewo et al., 2021). As environmental concerns continue to grow about fresh water (FW) consumption, there is now a desire to effectively reuse or recycle PW by incorporating it into operations. If PW can be effectively recycled, it may provide an opportunity to reduce environmental problems with disposal as well as reduce freshwater consumption in completions operations. However, minimizing environmental impact and a CE in O&G must be considered in conjunction with practices and technologies that increase, or at least maintain, O&G production. Declining resource production can not only constrain human flourishing (Epstein, 2014), but it also creates a

"resource drag" on developing economies as mineral extraction steadies and declines (Davis, 2010).

Reusing PW in treatments can alter operations and affect treatment implementation. However, recycling and reusing PW for operations can be economic for operators with development plans as few as 5 wells (Aro & Fowler, 2021). Figure 6.1 shows a treatment implemented using FW and figure 6.2 shows a treatment implemented using only PW. We see the differences in STP trends with the treatment using only PW having a near constant trend around 9000 psi.

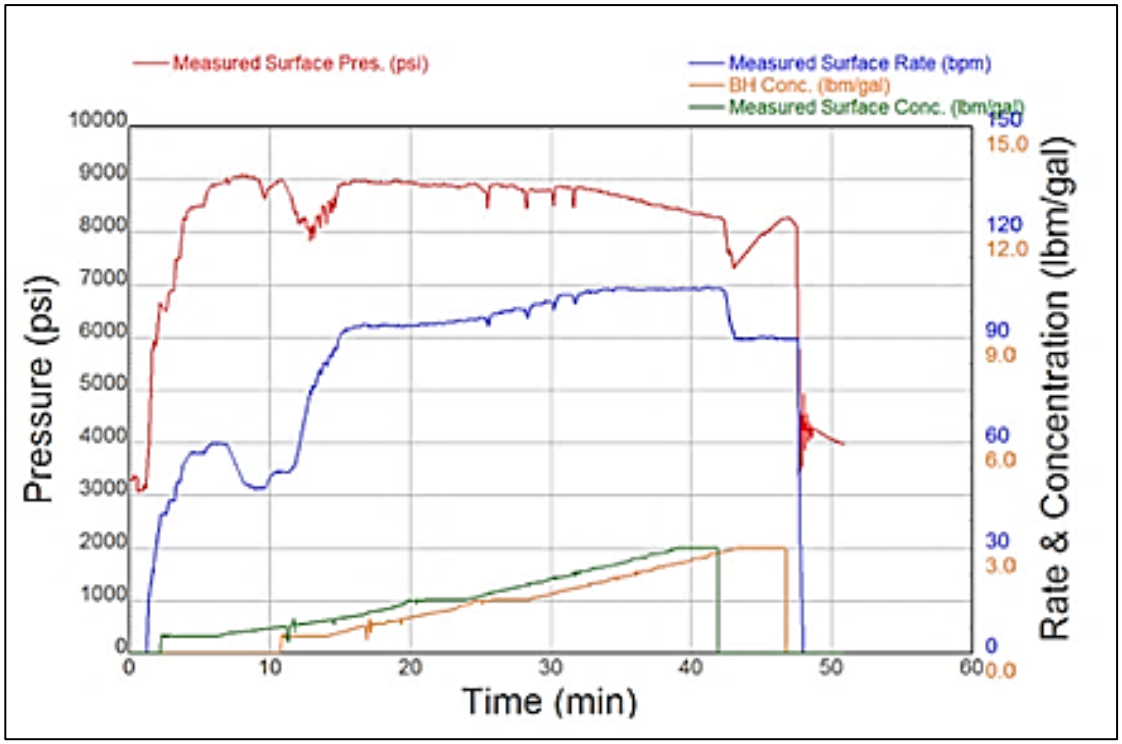

Figure 6. 1 - Treatment plot using only FW as base fluid and increasing rate to counteract STP decline

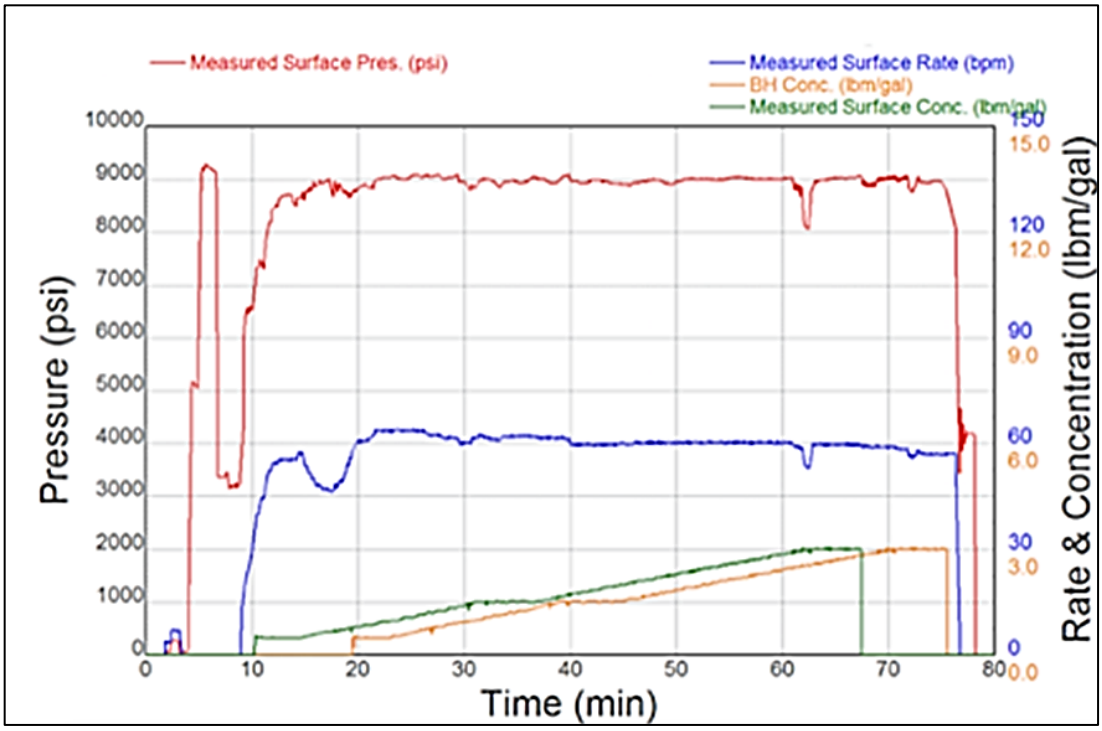

Figure 6. 2 - Treatment plot for using only PW as base fluid

This trend may be due to the increased perforation friction resulting from the increase in base fluid density of the treatment. This relationship is shown in equation 21 (Gustavo et al., 2016):

$$P_{pf} = \frac{0.2369Q^2\rho}{N_p^2 D_p^4 C_d^2}. \quad (21)$$

Where,

Q = Flow rate, volume/time, (bbl/min)

ρ = Fluid density, mass/volume (lb/gal)

N_p = Number of perforations, count

D_p = Diameter of perforations, L, (inches)

C_d = Coefficient of discharge, unitless

We see that fluid density will increase perforation friction and therefore increase the treating pressure. This may have implications when implementing CE into O&G operations and designs. Although there is plenty to critique about implementing circular economies, the results in table 6.3 show that, theoretically, efficient implementation of CE may be possible for the problem of PW reuse and recycling. More analysis of the effects of reusing PW for operations can be found in Appendix A.

6.1.3 Investigation of Trade-offs with Costs and Emissions

A quick comparison of the stages mentioned from above can be done by assuming a fuel efficiency factor of 16 hp-h/gal. The PW stage total volume pumped was 3838 bbl and the FW stage was 3865 bbl so the stage size is roughly equivalent. The stages are also close together in the lateral. Table 6.3 shows the treatment data needed to calculate HHP-H and the estimated diesel used to complete the stage.

- Treatment data and calculated parameters for fuel efficiency for PW and FW stages

Treatment	Average Rate (bpm)	Average Pressure (psi)	Pump Time (hours)	HHP	HHP-H	Diesel Used (gal)
PW only	58.4	8845	1.15	12660.5	14559.6	910
FW only	86.5	8718	0.783	18483	14478.4	905

The results from table 6.3 show that, theoretically, the fuel costs for both jobs is nearly identical with one treatment pumping at lower rates, higher pressures for a longer time (treatment using PW) and one treatment being pumped at a higher rate in less time. All else being equal, the treatments use the same amount of fuel and the resulting emissions are almost identical. However, there are implications regarding contracts that service

companies should be aware of and should build insurances for themselves. If an operator were to implement a program utilizing fresh water, requiring lower rates at higher pressures, there may be additional costs that fall on the service companies with pump maintenance, parts, and equipment.

After establishing that the fuel costs and HHP requirements are roughly equivalent in theory, engineers can now think at the margin in regards to design. In table 6.3, the fuel consumption is equivalent because the HHP-H between the two stages is roughly equivalent. However, by increasing the perforation standoff by 2 feet for each stage, we can expect reductions in average STP by roughly 469 psi (this is simply the marginal estimate of 234.5 in table 6.1 multiplied by two). So in theory, increasing the perforation standoff by two feet for both stages yields the expected results in table 6.4, ceteris paribus.

Table 6. 3 – Hypothetical fuel savings with reductions in STP using data from table 6.3

Treatment	Average Rate (bpm)	Average Pressure (psi)	Pump Time (hours)	HHP	HHP-H	Diesel Used (gal)
PW only	58.4	8376	1.15	11989	13788	862
FW only	86.5	8249	0.783	17489	13694	856

The reduction in average STP yields savings of 48 - 49 gallons of diesel per stage. Over a 40 stage well, this yields savings of 1920 – 1960 gallons of diesel. The cost savings realized will vary with the price of diesel. However, there are additional benefits from decreases in emissions. The Environmental Protection Agency (EPA) uses a conversion factor of 10.180x10^-3 metric tons CO_2/gallon of diesel (Environmental Protection Agency, 2023). So, this yields savings of 19.5 - 20 metric tons of CO_2 per well.

6.1.4 Contracts

In O&G, primary contract negotiation centers around cost and well service contracts will often contain parts for services and equipment (Thomas, 2013). Each party that controls their own personnel and equipment are best equipped with the knowledge necessary to minimize the risks associated with operations (Thomas, 2013). So, it may be beneficial to service company providers to incorporate STP considerations into their contracts, especially if design changes are implemented mid-treatment.

6.1.5 Simultaneous Fracturing Operations

The models constructed in this study may also have benefits in combination with simultaneous HF treatments. Over 200 wells have been completed using simultaneous HF treatments in the Bakken and Permian (Russel, Stark, Owens, Navaiz, & Lockman, 2021). Simultaneous treatment allows for separate treatment designs to be pumped on different wells concurrently (Russel, Stark, Owens, Navaiz, & Lockman, 2021). Having the ability to predict how STP of re-fracturing treatments will react to altering variables will help engineers better understand the wellbore integrity and surface equipment limitations and allow for any necessary changes to treatment design. This may become especially important in the case of simultaneously fracturing one re-fracture treatment and one treatment on a newly drilled and completed well.

6.1.6 Using Interaction Effects to Estimate Overall Effects

The results from the feature engineered model can provide estimates of overall effects between the interaction between perforation standoff and non-linear temporal

average STP. By taking the partial derivative of feature engineered model with respect to perforation standoff, we can estimate the overall effect of perforation standoff. This is shown in equation 22.

$$\frac{\partial avg\ stp}{\partial perf\ standoff} = -234.463 + 2.78 * Prev\ Avg\ STP^{0.5} \quad (22)$$

This relationship may be utilized in the field in the following manner. After pumping a stage, the engineer or operators on location should have post treatment data relatively quickly. After calculating the average STP for the stage, equation 22 can estimate the effects of adding one foot to perforation standoff before setting the plug for the next stage, since the calculated average STP is the previous stage average STP for the plug that is about to be set. So, for a stage with average STP of 7000 psi, the effect of adding one foot to perforation standoff would be $-234.463 + 2.78(7000)^{0.5} = -1.87$ psi. Additionally, for 7500 psi, the effect of adding one foot to perforation standoff would be $-234.463 + 2.78(7500)^{0.5} = 6.3$ psi

So, given the data set, for re-fractured wells in this area, there exists a point somewhere between 7000-7500 psi where we'd need to add 2 feet to perforation standoff to avoid temporal stress shadow effects. This analysis may be performed at any point during the job and recorded. The perforation standoff may be changed as necessary to mitigate any unnecessary temporal stress shadow effects that may affect treatment implementation.

6.2 Conclusions

- **In an attempt to answer research question 1, a dynamic multivariate regression model was constructed.** This research showed it is possible to create unique and area specific models to accurately predict average STP. Although this study focuses specifically on re-fractured wells, the same approach may apply to new drills provided there is data from offset wells. Initial models showed that stress shadow effects from the previous stage on a subsequent stage have a statistically significant effect on the average STP. According to the initial dynamic model developed in this study, a 1 psi increase in the average STP from the previous zone will create a 0.713 psi increase in average STP in the subsequent zone. This estimate seemed high but was roughly equivalent to estimates from literature (Roussel, 2017). However, this model appeared to suffer from endogeneity since estimates of perforation standoff and acid volume did not match what physics based models and domain knowledge propose. It was suspected that within well-fixed effects may be responsible for the suspected endogeneity.
- **Building off the pooled models and attempting to answer research question 2, unobservable wellbore and geologic properties that do not vary from stage to stage were accounted for using fixed effects (FE) multivariate regression models.** These models also allow for estimation of boundary effects from stress shadows from one stage to the next. These estimates may differ by formation due to inherent geologic, reservoir, and wellbore differences. However, having an estimate on the temporal dependence of pressure between stages may be an important design parameter for treatments, perforation design, and field

implementation. After accounting for FE, previous stage average STP, perforation standoff, and acid volume pumped were the only statistically significant predictors of average STP for a re-fracturing treatment. Based on our models, a marginal increase in average STP in one stage will yield a 0.537 psi increase in the subsequent stage. This is significantly smaller than our previous estimate. However, as with the pooled models, marginal effects from perforation standoff and acid volume were counterintuitive, indicating persisting endogeneity.

- In the initial and FE models, the amount of acid pumped had a positive and statistically significant effect on average STP. From field experience, it is often seen anecdotally that acid provides no pressure relief. However, we would not expect this effect to be positive and statistically significant. The positive effect may be due to interior perforations being opened, decreasing individual fracture width and also increasing perforation friction pressure as well as friction pressure along the fracture face. **So, in an attempt to create more accurate models, the inferential models constructed in this study were augmented with more complex feature engineered models that would also increase predictive power, answering research question 3.** The model utilizing feature engineering increases the predictive power substantially. The average % error was reduced from 5.8% in our original study to 2.1% per stage with the current predictive model. **Therefore, this confirms the hypothesis that marginal estimates can be obtained and accurate predictive models can be constructed to predict average STP for cost minimization purposes.** Prediction results are shown in table 7.2. Predictions were all converted to positive values and then averaged over the course of the well.

- After accounting for interaction effects between perforation standoff and non-linear temporal pressure dependencies and stage depth, the predictive model yielded inferential results that match physical reality. These marginal estimates for completion parameters were the following: The presence of a 3.5" liner has significant effects on average STP and increases expected average STP by roughly 500 psi for re-fractured wells in this area of the Williston basin; increasing perforation standoff by one lateral foot decreases expected average STP by roughly 230 psi for the subsequent stage; increasing average rate by one bpm increases expected average STP by roughly 25 psi for this area of the Williston basin; acid did not have any statistically significant effect on average STP. **Ultimately, the feature engineered model seemed to perform better at prediction and producing marginal estimates of the effects of completion parameters that more accurately represent the underlying physics. So, while although fixed models can be constructed, the feature engineered model produced better results than the fixed effects models. The feature engineered model also shows that after accounting for other treatment parameters and the interaction between non-linear stress shadow effects and perforation standoff, the stress shadow effects are not statistically significant while perforation standoff and the interaction is.** This result is unique in that the feature engineered model designed for prediction actually provided better inferences as well. This is usually not the case (Kuhn & Johnson, 2020), but is probably a product of the complexity of O&G reservoirs and operations.

- The approach in this study provides a framework that is reproduceable and allows for the estimation of marginal effects of completion parameters on treatment implementation. These are crucial in identifying high margin areas for cost minimization.

Appendix A: Using Machine Learning to Characterize Produced Water Reuse on Average STP during Hydraulic Fracturing

Recycling, reuse, and circular economics are gaining momentum in the O&G (O&G) industry. Although limited, circular economics does provide an answer to increasing pressure to reduce the environmental footprint from operations by attempting to internalize environmental externalities. One area of increasing interest is the reuse of produced water (PW) for completions operations. Not only will reusing produced water reduce the amount that must be disposed, but reusing PW may also reduce trucking costs and emissions and replace the use of freshwater (FW) in hydraulic fracturing. After extensive field trials, an operator in North Dakota has successfully incorporated large amounts of PW into their treatments at various stages in the treatment. Along with reducing FW consumption and reduced disposal, reuse of PW has also shown to act as a limited entry technique which may help perforation break down, equally distribute placement of fractures, and, therefore, increase production. This study constructs two multivariate regression models using field treatment data to characterize the extent to which PW affects the STP of treatments. The first is a difference-in-difference (diff-diff) model, usually used in policy analysis, using data from a hybrid stage that utilizes both produced and FW in the treatment. The second is a standard multivariate regression model using a dummy variable for PW. The results from both models indicate that the

use of PW has a positive and statistically significant effect on STP. The difference-in-difference model indicates that there are positive effects on STP beyond well and time fixed effects. Evidence from previous studies suggest this may be due to suspended solids, total dissolved solids, and residual oil inherent in produced water that affects injection pressure. This suggests that reusing produced water may have other implications for reservoir modeling as the practice becomes more common.

The idea of circular economies (CE) has been gaining momentum in academic literature. Although there are many definitions, the most frequently used definition is a combination of reduce, reuse, and recycle activities **(Kirchherr, Reike, & Hekkert, 2017)**. The idea has also touched the O&G industry as more rigorous practices in resource management have been adopted to minimize environmental impacts from O&G production **(Castilla, Zeuss, & Schmidt, 2021)**. One major concern in resource management is the reuse and recycling of PW. PW has generally been disposed of using disposal wells that are drilled into deep basement formations **(Tomomewo, et al., 2021)**. As environmental concerns grow with (HF) about fresh water (FW) consumption, there is now a desire to effectively reuse or recycle PW by incorporating it into operations. If PW can be effectively recycled, it may provide an opportunity to reduce environmental problems with disposal as well as reduce freshwater consumption in completions operations. However, minimizing environmental impact and a circular economy in O&G must be considered in conjunction with practices and technologies that increase and maintain production. Declining resource production can not only constrain human flourishing **(Epstein, 2014)**, but it also creates a "resource drag" on developing economies as mineral extraction steadies and declines **(Davis, 2010)**. This goal is

especially difficult as production from unconventional reservoirs and the HF treatments implemented in them are highly variable. Proppant distribution is generally distributed non-uniformly across perforation clusters and there exists a preference for heel and/or toe perforations as well as a low-side wellbore preference for fluid and proppant **(Roberts, Whittaker, & Paxson, 2020)**. These types of patterns make uniform treatment, and therefore optimal stimulation and production, problematic.

One technique to increase stimulation effectiveness, and thus increase well productivity, is limited entry (LE). Simply put, LE is the development of perforation friction through a reduction in the number of perforations **(Cramer, 1987)**. The resulting backpressure in the wellbore promotes treatment into multiple perforations of differing stress states (Cramer, 1987). Equation A1 shows the most common equation used to estimate perforation friction **(Gustavo, et al., 2016)**:

$$P_{pf} = \frac{0.2369Q^2\rho}{N_p^2 D_p^4 C_d^2}. \quad \text{(A1)}$$

Where,

Q = Flow rate, volume/time, (bbl/min)

ρ = Fluid density, mass/volume (lb/gal)

N_p = Number of perforations, count

D_p = Diameter of perforations, L, (inches)

C_d = Coefficient of discharge, unitless

Data suggest that STP (STP) should be maintained at the highest level possible within surface equipment constraints (Somanchi, Brewer, & Reynolds, 2017). This is especially important in pad to ensure breakdown and fracture extension and make certain that most perforations breakdown and begin taking fluid (Somanchi, Brewer, & Reynolds, 2017).

In practice and in literature, the previous focus has been primarily on limiting the number of perforations, limiting entry hole diameter, or increasing rate **(Gustavo, et al., 2016)** to increase perforation friction and thus, increasing STP above initiation pressure for each perforation **(Langorne & Rasmussen, 1962)**.

Increasing rate and/or reducing the number of perforations or size of perforations are all critical design parameters (Gustavo, et al., 2016). Rearranging equation A1 can also yield solutions of different variables. Cramer (1981) rearranges equation A1 to solve for Q, D_p, N_p, and C_d to provide solutions for these variables and how they will alter treatment. However, this process is dynamic, and these variables will change throughout stimulation. For instance, as perforations erode, D_p and C_d will increase which will decrease P_{pf} (Gustavo, et al. 2016). This may be compensated by an increase in Q which is meant to counteract these effects and maintain backpressure in the wellbore (Gustavo et al. 2016, Cramer 1987, Somanchi, Brewer, & Reynolds, 2017). It's also important to note that rate increases are ineffective during treatment if clusters have already been lost **(Somanchi, Brewer, & Reynolds, 2017)**.

Although it may seem straightforward, implementing effective LE in the field is difficult. Exceedingly high STP allows little room for error operationally. Pump rate must be constantly monitored to avoid screenouts or overpressuring. Having limited entry points into formation can also create problems when cleaning up the wellbore after a screenout. Another common problem is that LE treatment distribution tends to be very uneven during the slurry phase of treatment (Somanchi et al., 2017). Given the dynamic nature of HF treatments, it would be difficult to isolate any values, as there are numerous

variables changing at any time, to estimate individual effects of altering parameters and their effectiveness.

Looking at equation A1, it is apparent that increasing density will also increase $\boldsymbol{P_{pf}}$, although it has traditionally been considered fixed (excluding foamed fluids) (Cramer, 1987). This study was undertaken after observing the noticeable differences in STP between HF treatments that incorporate PW as a base fluid. Qualitative investigation of two HF stages shows noticeable differences in the trends of STP and, specifically, STP was maintained at the maximum level without increasing rate while using PW as a base fluid. The treatment plot for stages 26 and 29, respectively, for a single well are shown in figure A.1 and figure A.2. Figure A1 shows a treatment using only FW as a base fluid for the treatment and figure A.3 shows the corresponding chemical plot. It's important to note that PW was used in flush. Figure A.2 shows a treatment plot using only PW as the base fluid while figure A.4 is the corresponding chemical plot. The proximity of the stages within the well will help to alleviate some uncertainty about any heterogeneities that may exist in formation or along the wellbore. Therefore, it is assumed that these stages are equivalent in these regards. Although specific chemical names could not be included in the chemical plots, the blue line represents FW High Viscosity Friction Reducer (HVFR) concentration, the green line represents surfactant concentration, and the orange line represents PW compatible HVFR concentration. All chemicals are measured in concentration ratios of gal/1000gal.

Although the trends observed in figure A.1 and figure A.2 are straightforward, the evidence is anecdotal and needs to be accompanied by a model. With so many moving parts, it can be hard to dissect the treatments and draw any useful conclusions about the

effectiveness of altering any design parameters. Ramirez and Iriarte (2021) noted this problem using high frequency time series treatment data when comparing a neural network and domain knowledge-based algorithm to identify the breakdown pressure of formation during a HF treatment.

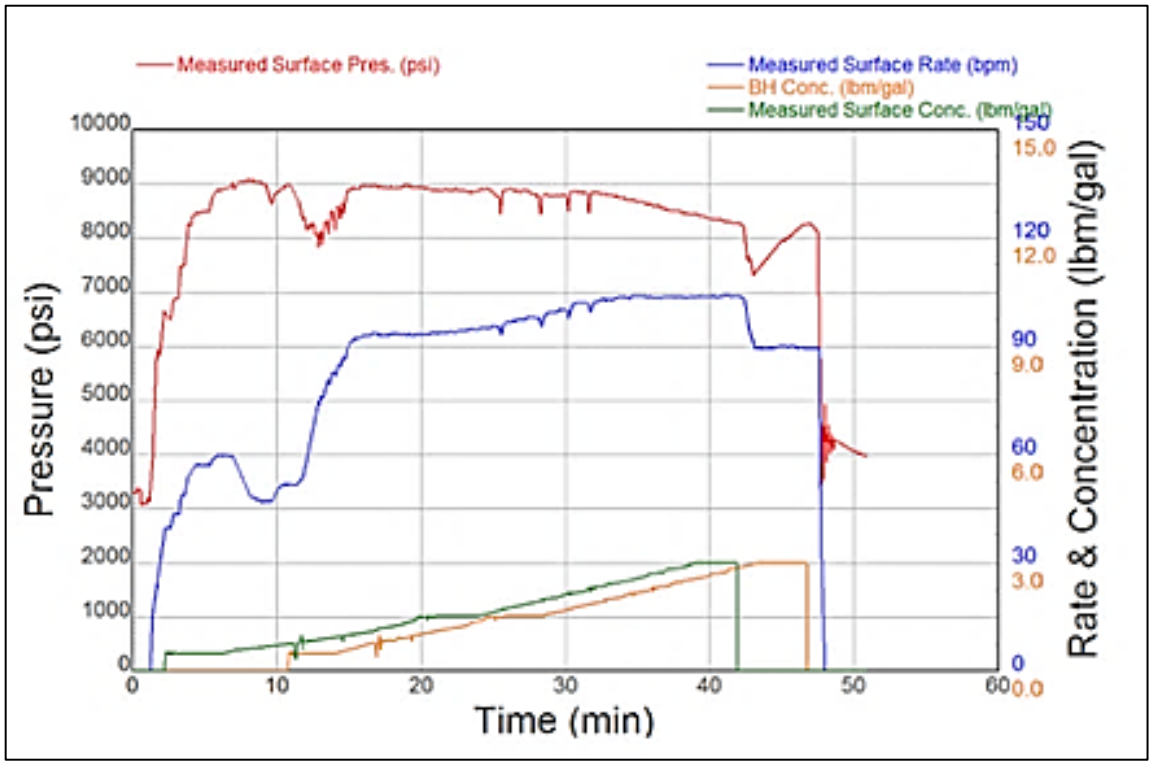

Figure A. 1 - Stage 29 treatment plot using only FW as base fluid and increasing rate to counteract STP decline

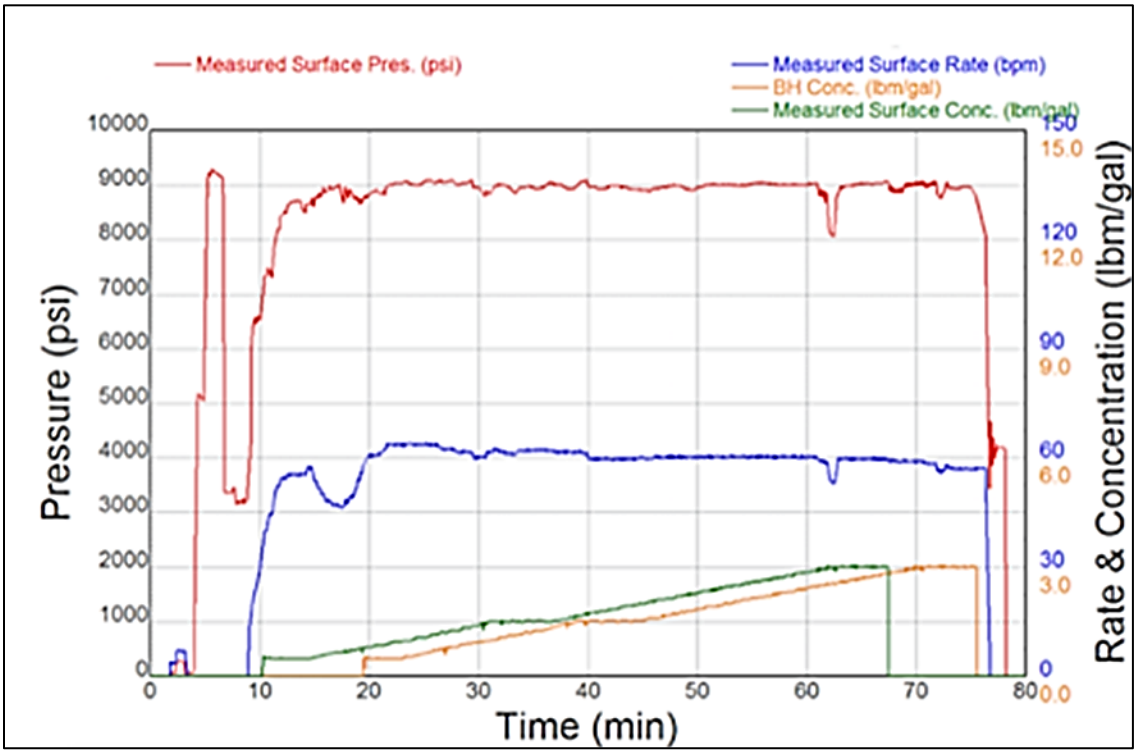

Figure A. 2 - Stage 26 treatment plot for using only PW as base fluids

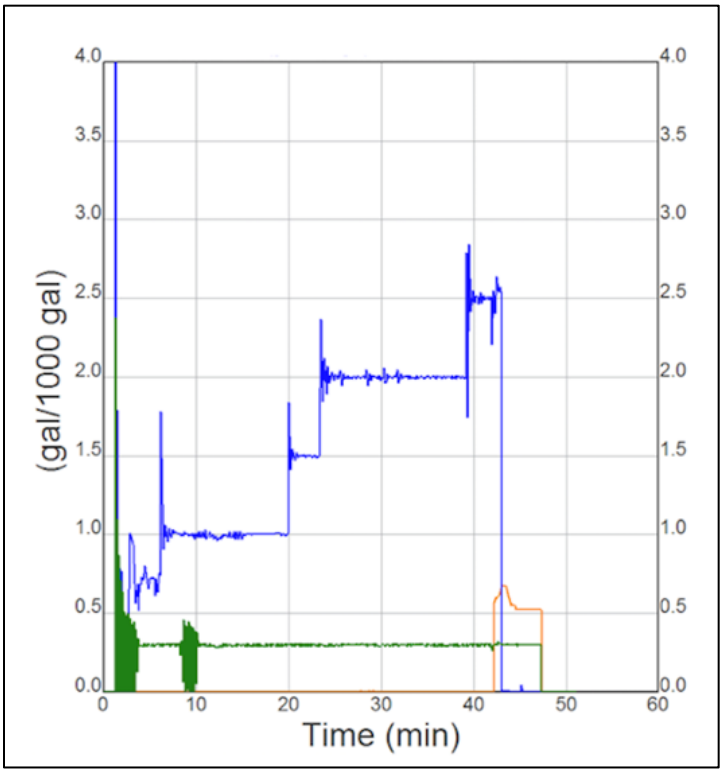

Figure A. 3 - Stage 29 chemical plot using only freshwater FR fluid system

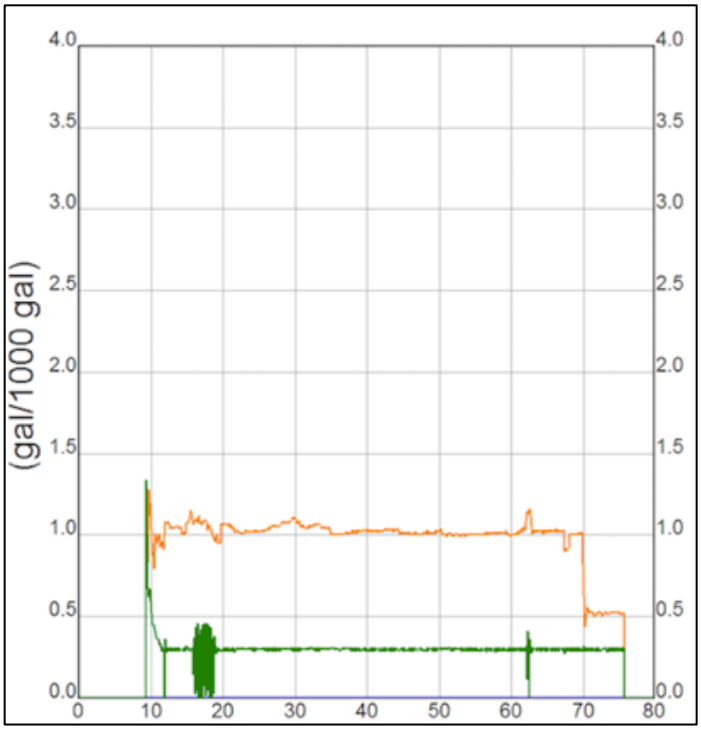

Figure A. 4 - Stage 26 chemical plot using PW compatible FR fluid systems

Estimating the effect of PW on STP is the primary motivation for the models constructed in this study. To reduce costs and lessen the environmental footprint from HF operations, the operator has also inadvertently added another way in which LE can be achieved in the field by altering the density parameter in the perforation friction equation. This will be important to incorporate into treatment and surface equipment designs as PW becomes more common in HF treatments. There have been other studies using multivariate regression to characterize HF treatments. Kroschel, Rabiei, & Rasouli (2022a) used multivariate regression models to investigate the marginal effects of design parameters on average STP.

The time series data for the study was collected from the treatment for stages 25 and 29 on a single well. The construction of a difference-in-difference (diff-diff) model for HF treatments is not quite as straightforward as in social sciences, mainly because of

the control group. In social sciences, the treated and control groups are viewed side by side over time. However, since two treatments cannot be viewed side by side in the same well over the same time period, a control stage was selected that was close to the treatment stage. As stated earlier, due to the proximity of the stages in the wellbore, it is assumed that reservoir and geomechanical properties will be the same in both treatments. This assumption allows for the use of stage 29 as a control when constructing the diff-diff model. The longitudinal data used for the multivariate model was from stage 25 in the well. The treatment data was analyzed raw only excluding data for pressure tests.

The swap from PW to FW was then marked in the control stage at the same point in the time series data. Again, these stages were not viewed side by side over the same period, but it is assumed that the stages would not have behaved differently had they been pumped at the same time.

Because of a unique tank and manifold design, the operator was able to switch between PW and FW "on the fly", meaning the water source can be switched without shutting down the frac pumps. This allows for the utilization of as much PW as possible even if there is insufficient volume to pump an entire stage. This unique design also allows for the utilization of longitudinal data where we can observe the effects of switching water sources on treatment of an individual stage over time. The treatment plot for this design is shown in figure A.5 with the corresponding chemical plot shown in figure A.6. When comparing with figures A.1 and A.3, we see the same trend of a relatively constant STP when pumping PW and a general decline when using FW. The chemicals and corresponding colors shown in figure A.6 are the same as figures A.3 and

A.4. Figure A.6 shows that PW was pumped for the first half of the stage and FW was pumped for the second half of the stage.

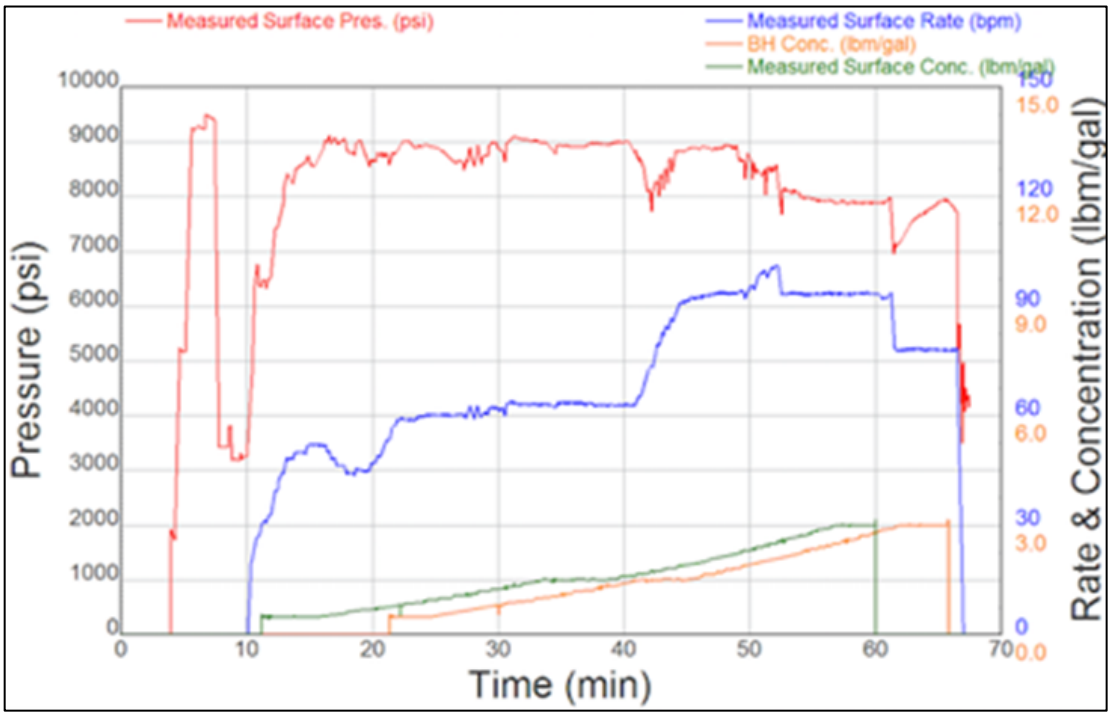

Figure A. 5 - Stage 25 treatment plot highlighting the effect of PW on STP

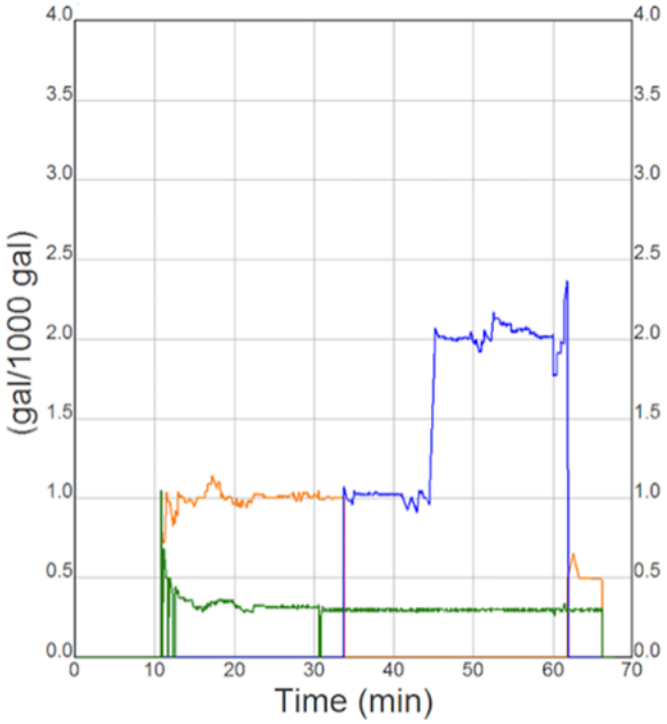

Figure A. 6 - Stage 25 chemical plot highlighting the switch between fluid systems during stage

This study will compare two econometric models that attempt to characterize the difference in STP using PW as a base fluid for HF treatments instead of FW. First, a diff-diff model will compare the differences between stages 25 and 29 comparing the treatment stage pumped with PW and FW (stage 25) and a control stage pumped solely with FW (stage 29). Diff-diff models are generally used in social sciences to attempt to identify the effects from policy actions using a group that received a particular treatment and comparing those results to a control group. Next, a general multivariate regression model was created using the longitudinal data from stage 25 that contains both PW and FW using a dummy variable for PW use. Due to the stochasticity and complexity of unconventional reservoirs and HF treatments, simple, qualitative models should be

applied instead of more detailed models **(Cleary, et al., 1993)** because more complex models are not necessarily likely to be more accurate, but they are more likely to contain mistakes due to assumptions **(Bailey, 2017)**. Much like social systems, these flexible statistical models may better capture the complex relationships in HF than more detailed, complex models. To the author's knowledge, there is no study attempting to use diff-diff models to characterize and investigate the effects of altering parameters mid-stage on STP.

As stated earlier, diff-diff models are generally used in social sciences to evaluate the effect of enacted policy. However, these models can be used anytime that there is a specific and identifiable point when a treatment is enacted on a treatment group. Then, these results are compared against a reasonable control group. After controlling for other factors, it is assumed that the differences in the differences of the means of each group are due to the treatment **(Bailey, 2017)**. Eq. A2 shows the general form for a diff-diff model and is the base for the model constructed in this study **(Bailey, 2017)**:

$$Y_{it} = \beta_0 + \beta_1 Treated_i + \beta_2 After_t + \beta_3 (Treated_i * After_t) + \epsilon_{it}. \quad \text{(A2)}$$

Where $Treated_i$ equals 1 for a treated zone (stage 25) and 0 for a control zone (stage 29). $After_t$ equals 1 for all observations after swapping base fluids for both the treated and control zones, and $Treated_i * After_t$ represents the interaction of treatment and control zones before and after the fluid swap during the treated zone **(Bailey, 2017)**. There exists an average STP that the control stage will experience which is captured by β_0. The average STP of the treated well is then captured by $\beta_0 + \beta_1 Treated_i$. Notice

this can be more, less, or equal to the average STP of the control well depending on if the β_1 coefficient is positive, negative, or zero respectively. It is important to note the difference between this experiment and the way in which a diff-diff model is set up in a policy analysis. In a policy analysis, the effects between the treatment and the control group are measured by taking a difference of the differences in the means after a treatment is prescribed, over the same time period. However, for this experiment, the effects that we are trying to estimate take place before the "treatment" takes effect. In other words, we did not observe the control and treated stages, use PW on the treatment stage and then observe the effects afterwards. The process was in reverse. So, since the "treatment" or PW will be used in the beginning of the stage, the switch to FW will act as the treatment.

The $After_t$ term attempts to account for time effects that would exist in the absence of treatment. This equals 1 for all time periods for both the control and treatment wells after the treatment takes effect (i.e. after the switch to PW in the treated stage). The key coefficient is β_3 which will equal 1 in the treatment well after treatment takes affect and 0 for the control well after the treatment takes effect (since the control well is denoted with a 0 via the $Treated_i$ variable). This coefficient tells us the additional effects due to treatment beyond the preexisting conditions for the treated and control stages **(Bailey, 2017)**.

The model also allows for the control of other variables. The models in this study will also control for pump rate and sand concentration which will also influence STP from friction, hydrostatic pressure, and perforation erosion.

Next, we will construct a more general multivariate regression model simply by using a binary variable to indicate when PW is being pumped. In the data set, 1 will indicate if PW is being used and 0 will indicate FW and the results will therefore be relative to FW. The general form of the equation is shown in Eq. A3 **(Bailey, 2017)**:

$$Y_i = \beta_0 + \beta_1 X_1 + \cdots + \beta_n X_n + \varepsilon. \quad \text{(A3)}$$

Table A.1 shows the regression results for the diff-diff model and table A.2 shows the results from the longitudinal data set. Both models estimate statistically significant effects on STP from using PW as base fluid.

Table A. 1 - Regression results from the diff-diff model.

	Dependent variable:
	STP (psi)
Treatment	1,087.703***
	(26.599)
After	-925.263***
	(24.507)
Proppant Concentration (ppg)	-44.451***
	(14.870)
Total Pump Rate (bpm)	38.847***
	(0.516)
Treatment*After Interaction	264.257***
	(37.501)
Constant	5,406.875***

	(38.695)
Observations	5,002
R^2	0.729
Adjusted R^2	0.729
Residual Std. Error	591.927 (df = 4996)
F Statistic	2,686.602*** (df = 5; 4996)
Note:	*p**p***p<0.01

Table A. 2 - Regression results from the longitudinal model.

	Dependent variable:
	STP (psi)
Proppant Concentration (ppg)	103.956***
	(19.793)
Total Pump Rate (bpm)	28.395***
	(1.209)
Produced Water as Base Fluid (binary)	1,074.255***
	(41.186)
Constant	5,829.634***
	(86.773)
Observations	3,439
R^2	0.230
Adjusted R^2	0.229
Residual Std. Error	775.409 (df = 3435)
F Statistic	342.165*** (df = 3; 3435)
Note:	*p**p***p<0.01

When comparing the results for the two models, it appears the diff-diff model (table A.1) not only explains more variance in the data with an R^2 of 0.729 (compared to 0.23 in the multivariate model), but the coefficients and their respective signs seem appropriate given what is known about wellbore dynamics. Overall, the diff-diff model seems to work better in characterizing the effects of PW on HF stimulation.

Referring to table A.1, the diff-diff model estimates 5406.875 psi as the expected average STP for the control stage. The average expected STP for the treatment stage is 5406.875 + 1087.703 = 6494.578 psi. This is the result we would expect as we observed qualitatively that the STP was higher in the treated stage. The β_3 coefficient of 264.257 estimates the effects above and beyond those not fixed within the stages over time.

The coefficient estimates for $After_t$ is -925.263. This captures the change in the dependent variable, STP, in the treatment and control stage after switching to FW. This makes sense intuitively and appears to align with the qualitative investigation of the treatment plots for the treated and control stages in figures 1 and 5, respectively. However, when comparing to the treatment plots, the negative effect on STP in the control stage may be due to things like perforation erosion and may be due to a difference in density in the treatment stage.

For proppant concentration, we see that the coefficient estimate is -44.451 indicating that a one ppg increase in sand concentration will yield an expected 44.451 psi decrease in STP. This makes sense intuitively as we would expect higher sand concentrations to increase hydrostatic pressure, thus decreasing STP, and increased sand concentration would increase perforation erosion thus decreasing STP.

The coefficient estimate for total pump rate is 38.847 indicating an expected 38.847 psi increase in STP with a one bpm increase in total pump rate. This is expected due to friction.

Overall, these results align with the qualitative investigation and knowledge of wellbore dynamics. The β_3 coefficient estimate of 264.257 estimates the effects above and beyond those not fixed within the stages over time. These effects are after switching to FW in the treatment and control stages. This indicates a positive and statistically significant effect on STP for the treatment stage after switching to FW in the middle of an HF treatment. This suggests there may be other processes (chemical or mechanical) that are taking place in the reservoir due to PW re-entering formation. This result is backed by some evidence in literature from disposal well injection of PW and will be discussed in the next section.

Although the general multivariate regression model results provide an estimate of the expected increase in STP due to using PW, the coefficient interpretation for proppant concentration does not make sense. Although we would expect the coefficient for proppant concentration to be negative (i.e. STP should decrease with an increase in proppant concentration), the coefficient of 28.395 for pump rate is reasonable and we would expect an increase in pump rate to increase STP, all other things begin equal. The PW binary coefficient tells us that if we use PW as a base fluid (PW Binary = 1), we expect to see a 1074.255 psi increase in STP. This estimate seems to be in line with what was observed qualitatively comparing figure A.1 and figure A.2. Although it's possible that increases in proppant concentration may in fact increase STP, the model may also be

endogenous. Endogeneity leads to the deduction of causal relationships that may be spurious and can be defined broadly as correlation between independent variables and the error term **(Bailey, 2017)**.

Maybe the most important note is that with some forethought and prior planning, diff-diff models may provide a tool to help answer tough questions that the O&G industry faces. Since each well is a unique experiment, it is difficult, if not impossible, to replicate the experiments and draw conclusions based on any replicable results. However, if reasonable controls are in place prior to operations starting, it may be possible to design stages, wells, and operations so that they are able to be reproduced as best as possible. As a hypothetical example, two wells on the same pad that are drilled in the same formation may be HF in parallel with each other where one well is used as a control for experimental stages on another. Another example may be if there are multiple pads with multiple wells in an area and two wells are drilled in the same formation, they may be HF in parallel with each other to provide a treatment well and a control well with sufficient distance between the two to alleviate concerns about stress shadows effects on treatment. This type of study may also be very useful if something such as fiber optics are available. With the industry focus shifting towards cost reduction **(Barree, 2020)**, it will become increasingly important to plan these experiments in conjunction with operations if we are to make practical and meaningful engineering advances.

Water costs in the Bakken are upwards of USD 5 per bbl of freshwater with disposal costs around USD 9 (Ba Geri et al. 2019). These high prices reflect the growing concern about FW scarcity as well as opportunity costs. Considering the economic cost (including the opportunity cost), it is most likely more efficient to utilize as much PW as possible in

lieu of FW. The PW could also be used to support coal mining and irrigation with improved quality **(Tomomewo, et al., 2021)** and may provide another source of reuse. More efficient reuse of PW may reduce overall transportation costs and emissions associated with PW disposal thus further reducing the environmental impact from operations. Dow et al. (2022) identified the opportunity to reduce emissions from rig activities by increasing operational efficiencies. This will become increasingly important as hydrocarbons will rely heavily on O&G to improve the quality of life for the world's expanding population in the future (Dow et al, 2022).

There may be additional cost considerations with PW reuse. Higher STP tends to limit pump rate and therefore increase stage pump time. This may be a significant source of increase in treatment cost that could offset gains made from reusing PW.

Concerns have been raised about the depletion of surface and groundwater for HF operations **(Hausman & Kellogg, 2015)**. This water is not being used in other industries within the state of North Dakota, particularly agriculture, creating an opportunity cost, or the cost of forgoing use in other sectors. Because the FW is normally procured from multiple sources, there is also a cost imposed on those by simply being downstream from a private entity who may sell the FW on their land to an O&G operator. The state of North Dakota recognizes the following hierarchy for the distribution of FW as a resource **(Kurz, et al., 2016)**:

1. Domestic use
2. Municipal use
3. Livestock use
4. Irrigation use

5. Industrial use
6. Fish, Wildlife, and other recreation use

This hierarchy represents a list of opportunity costs for FW use. It's apparent the state of North Dakota values FW use in sectors more than Industrial use, so there may be significant opportunity costs associated with use of FW in HF. This is further evidence that reducing FW consumption for HF, while maintaining production, provides further economic benefit.

Fluid treatment of PW is out of the scope of this study as data was not collected for PW quality and composition. More research would need to be performed to investigate whether different qualities of PW fluid may change the additional effects estimated by the diff-diff model. This may be of interest as it is reasonable to assume that different formations with varying geological properties may react differently to treatments with PW as a base fluid. Previous research suggests that treating pressure and water quality are linked to each other **(Ochi, Dexheimer, & Corpel, 2013)**. No matter the degree of treatment prior to re-injection, PW will always contain residual oil that will affect reservoir interface plugging and could create relative permeability issues (Ochi, Dexheimer, & Corpel 2013, Rossini et al. 2020). Rossini et al. (2020) identify total suspended content as the third most important factor affecting bottom hole fracturing pressure in their PW re-injection models **(Rossini, et al., 2020)**. Therefore, the additional effects estimated by the models may be plugging due to residual solids in the PW. A need to augment these models with production and fiber optic data follows logically from the previous work and the models constructed in this study. If production increases were

realized, further research would need to be performed to investigate whether the increases were due to LE or the use of PW.

- PW can be effectively reused in HF treatments. With unique manifold, tank, and valve designs, it's possible to incorporate PW into treatments at any point so PW volumes on surface will not be a binding constraint on operations. Unique approaches and effectively incorporating PW into HD treatments will become essential to reduce the environmental impact from operations. These designs are discussed in Appendix B;
- Using the tank and manifold layouts in Appendix B, it is also possible to incorporate PW into wireline operations and realize further substitution of FW for PW. The manifold and tank layouts are discussed in Appendix B. Appendix C proposes a simple mass balance algorithm used in the field for spotting acid during WL runs that decreases the amount of fluid needed. This produces further cost savings.
- It is apparent from the perforation friction equation that increasing base fluid density increases perforation friction and can counteract the effects of declining STP throughout the stage due to things like increased perforation erosion. This assertion is backed by qualitative investigation. This assertion is also backed by the models constructed in this study. Therefore, increasing rate or limiting the number of perforation or restricting perforation diameter are not the only ways to counteract declining STP during treatment;

- The diff-diff model constructed in this study estimates additional effects, other than stage and time fixed effects, on STP from treatments using PW. These additional effects may be caused by an interaction of PW with formation. PW injection models in literature suggest that suspended solids in PW may also have an effect on bottom hole fracturing pressure. These may be the additional effects captured by the diff-diff model constructed in this study. Further analysis should be done on water treatment and the effects of PW quality in HF fluids has when interacting with the reservoir;
- There may be substantial gains from reusing PW in terms of a circular economy by reducing disposal costs. Using PW for HF substitutes FW and provides an opportunity for this resource to be used in other sectors of the economy, thus not incurring the opportunity cost of using FW for HF treatments;

Appendix B: Produced Water Tank and Manifold Layout for Disposal

Recycling and reuse of produced water from oil and gas (O&G) operations is of interest to the state of North Dakota, onshore USA, and companies alike. By reusing produced water, companies can decrease their disposal, trucking, and freshwater consumption costs. North Dakota has an interest in recycling and reusing produced water (PW) because of the impact freshwater (FW) consumption can have on local economies and environmental spillovers of produced water disposal, both of which may dampen public interest in O&G extraction.

There is a conflation in the O&G industry with the use of PW to augment FW use in support of operations and the total replacement of FW with produced water in operations. There also seems to be an implicit assumption that to effectively incorporate PW into operations, a salt tolerant fluid system or friction reducer is necessary and sufficient. However, because stages like a wellbore flush in treatments and wireline pump down operations do not require a fluid system to maintain its integrity for proppant carrying capacity, these stages do not require such a fluid system. It is therefore possible to alter tank arrangements, manifold layouts, and valve placement to use PW in these instances. These changes also allow the flexibility to use FW in instances of produced water shortages due to the irregularity of delivery to location. Furthermore, in colder temperatures, these layouts have the added benefit of allowing continual circulation of FW to avoid freezing.

The designs in this study have already been implemented by certain operators in North Dakota that the author has consulted for. While the author realizes that the tank

arrangements in this study will not alter fluid compatibility and do not allow complete substitution of FW for PW, they do allow for the possibility to substitute FW for PW where fluid compatibility is not an issue. These designs also allow operators to use FW when necessary to avoid shutting down operations, which can be costly. Shutting down operations can also be dangerous in winter. Thus, there are parts of a frac design and completions operations where manifold and tank arrangements are sufficient to substitute FW for PW while also providing the added flexibility of switching between fresh and produced water on-the-fly.

The practice of disposing of PW is of importance to the operators that stand to accrue cost savings as well as the state of North Dakota. Drilling, completions, and production operations all produce large amounts of wastewater that is primarily disposed of through saltwater injection wells, mostly into the Dakota formation (Energy and Environmental Resource Center, 2016). The Energy and Environmental Research Center (EERC) estimates that saltwater disposal volumes have increased by 341% from 2008 to 2014 and may increase another 328% between 2014 and 2035 (EERC, 2021).

The opportunity cost of water use is increasingly important in parts of the world where water is scarce (Al-Muntasheri, 2014). Some previous studies have already been undertaken investigating the use of flowback water in frac operations. The produced water from the Bakken is highly saline with a TDS range of 170,000-350,000 ppm (Tomomewo, et al., 2020). Samples from the Bakken formation have also shown a pH and specific gravity of 4.56 and 1.19, respectively (Griffin, Poppel, Siegel, & Weijers,

2014). There have also been field trials with salt-tolerant friction reducers (FRs) with some success (Wilson, 2017).

It's also important to note that along with technologies such as salt-tolerant FRs, effective programs must also be implemented to allow drilling and completions operations to operate efficiently (Paktinat, O'neill, & Tulissi, 2011). This may include things like "stimulation design, execution, and effectiveness" (Griffen et al. 2014). When considering "execution" and "effectiveness" in a stimulation design, this not only includes the technical aspects like fluid and frac design but should also include the ability to execute operations effectively as well. For instance, the randomness of PW being recovered can lead to supply issues that can be costly to operations. This may slow down or even temporarily shut down operations. This can be costly and dangerous during winter operations in places like the Williston basin. Therefore, it is also necessary to have the ability to use FW as well as PW during stimulation treatments. The layouts in this study make switching between PW and FW possible and will not interrupt stimulation operations and should allow almost all pump down operations to be performed with PW. While these advancements don't completely eliminate the use of fresh water, significant decreases in consumption can be made.

This study aims to provide a water transfer surface design, tank arrangements, and valve placement that may be utilized on (frac) locations that will incorporate PW into frac designs and wireline pump down operations so it may be reused and recycled on site. It's also important to note that the term "produced water" includes water from flowback operations as well as any water cut from production.

This study provides three layouts that can be used on frac locations that are capable of disposing of PW during completions operations. These designs provide a starting point and can be altered when necessary. Instances of changes may include things like space limitation and equipment arrangements. Layout 1 provides a design for incorporating PW into wireline pump down operations. Layout 2 can be used during frac operations. Finally, layout 3 will combine the two that can be used in temperatures above freezing or as pad size and layout allows. For each layout, it is recommended to place the tanks at the edge of location and utilize a drive-over for the suction or discharge hoses. The drive-over simply protects the hoses while allowing traffic to be able to flow through location. There are numerous ways these can be built and the operator should use different materials as they see fit. Typically, these have been constructed with rig mats and steel plates, although some service companies offer pre-built drive-overs.

One of the main problems with incorporating PW into frac operations has been the resources necessary to store enough fluid for twenty-four hour operations (Boschee, 2012). However, this need not be the case. Problems may arise if an operation relied solely on PW to pump entire stages. The randomness of flowback rates, water cuts, frac operations, truck travel, and weather conditions can create bottlenecks and shortages which would require operations to start and stop unnecessarily due to insufficient produced water volumes to pump an entire frac stage. A salt tolerant fluid system is not sufficient to continue operations in these instances. However, if an operator were able to switch between fresh and produced water effectively, safely, and quickly, much of the costs associated with these bottlenecks would be diminished or eliminated. It's also important to note that with efficiency gains and an increased number of stages being

pumped and perforated per twenty four hour period, it may not be necessary for an operator to pump an entire frac stage with produced water to effectively reuse it. If the rate of reuse on location through wireline plug pump downs, wellbore flushes, brine caps on the wellhead, etc. exceeds the rate from flowback and water cuts, there is no need for a salt-tolerant fluid system to effectively reuse produced water.

Figure B.1 shows the tank, valve, and manifold layout to be utilized for wireline pump down operations. The tanks that are generally used are 400 bbl upright tanks with 4 in. hose connections and hoses. Connections to the service company pump down equipment will vary and the operator should consult with the service company as to the best setup so achieve sufficient rate for pump downs. It should be noted that it is best practice to offload PW through a getty box and into the tanks through a line into the top of the tanks. A getty box is an enclosed apparatus that contains a connecting point for trucks to offload. This allows the connection to be broken without spilling any residual fluid in the line. It's also best practice to leave the tanks open to each other to equalize at the manifold or other equalizing hoses between the tanks. By doing this, the tanks are more easily equalized, thus minimizing any mistakes caused by human error. If it is necessary to switch to FW, it simply requires isolating the produced water tanks at the manifold and opening the two valves on the manifold to fresh water. This should be performed prior to starting the pump down depending on PW levels.

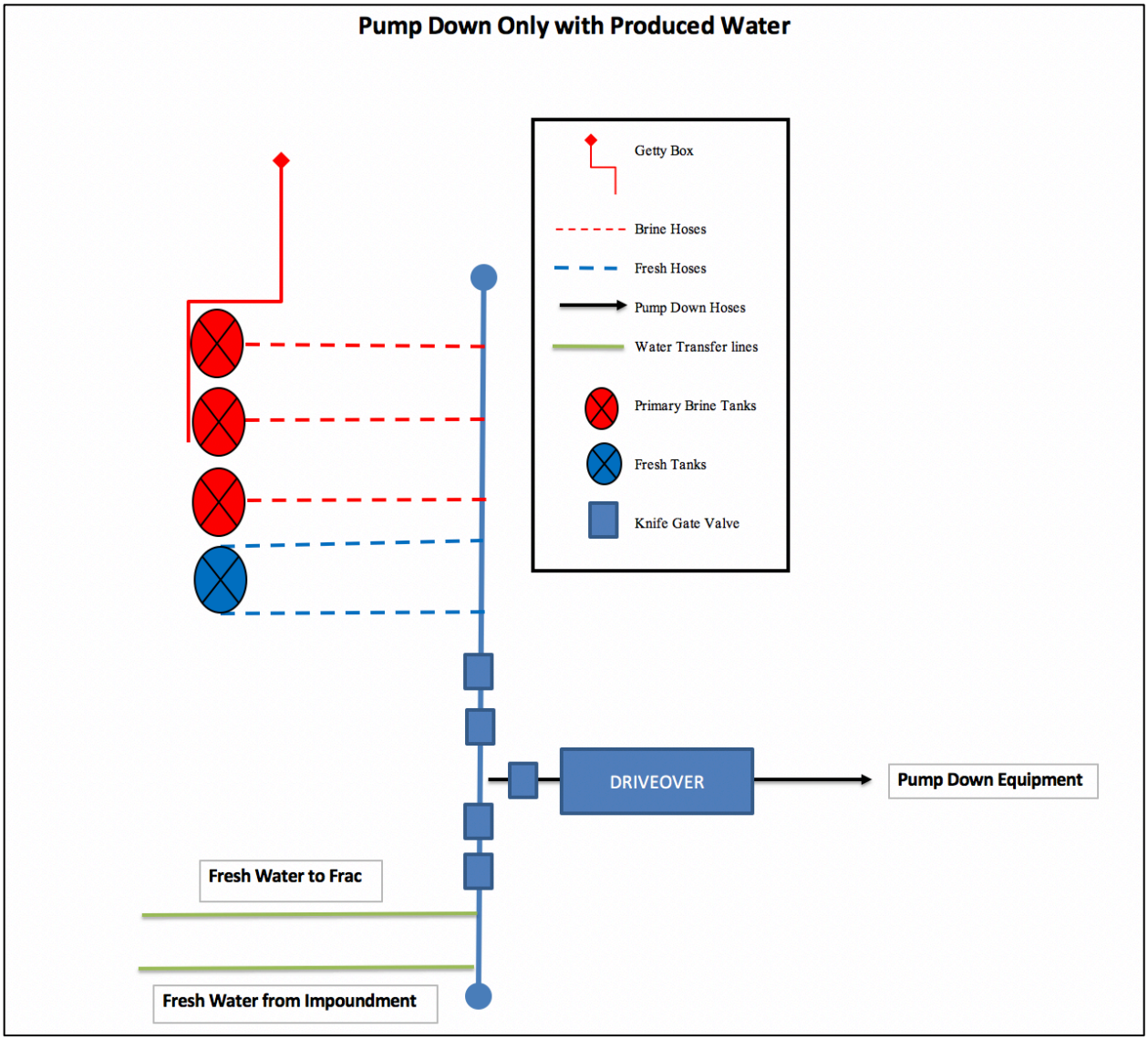

Figure B. 1 - Layout 1 Depicting pump down manifold design with ability to switch between fresh and produced water.

This layout also has the benefit of being able to continuously circulate the FW through transfer lines from the fresh water impoundment.

Heavier brine is also needed on locations during winter operations to prevent wellheads and equipment from freezing. It is recommended to use the tank furthest from the drive-over for this purpose. This enables the entire manifold to be filled with heavy brine and reduces the probability of having a dead spot of FW that can create ice plugs.

Figure B.2 shows the layout that enables operators to substitute FW for PW in flush, pad volumes, low sand concentration stages, or any other place the operator may want to cut out FW from the treatment design. Usually, a drive over is not necessary with this design because the blender will be spotted at the end of the manifolds with the rest of the frac equipment along-side of it. If a drive-over is necessary in some situation, it may be used in the same way as in figure B.1. There are some important features to note that will help operators and service companies during operations:

- FW transfer lines should be rigged in such a way that they are able to circulate through the entire FW manifold to avoid any ice plugs. The size of the suction hoses from water transfer to the frac equipment will depend on the connections the frac equipment has. Operators should consult with the service company to ensure sufficient rate depending on treatment design.
- PW tanks may be equalized to each other or through the manifold. Care should be taken when offloading produced water because the rate of offload may be greater than the rate at which the tanks equalize. This will prevent any tanks from overflowing.
- The brine manifold should be tied into the FW running to the blender at a point furthest from the blender so it may be brined up when needed.

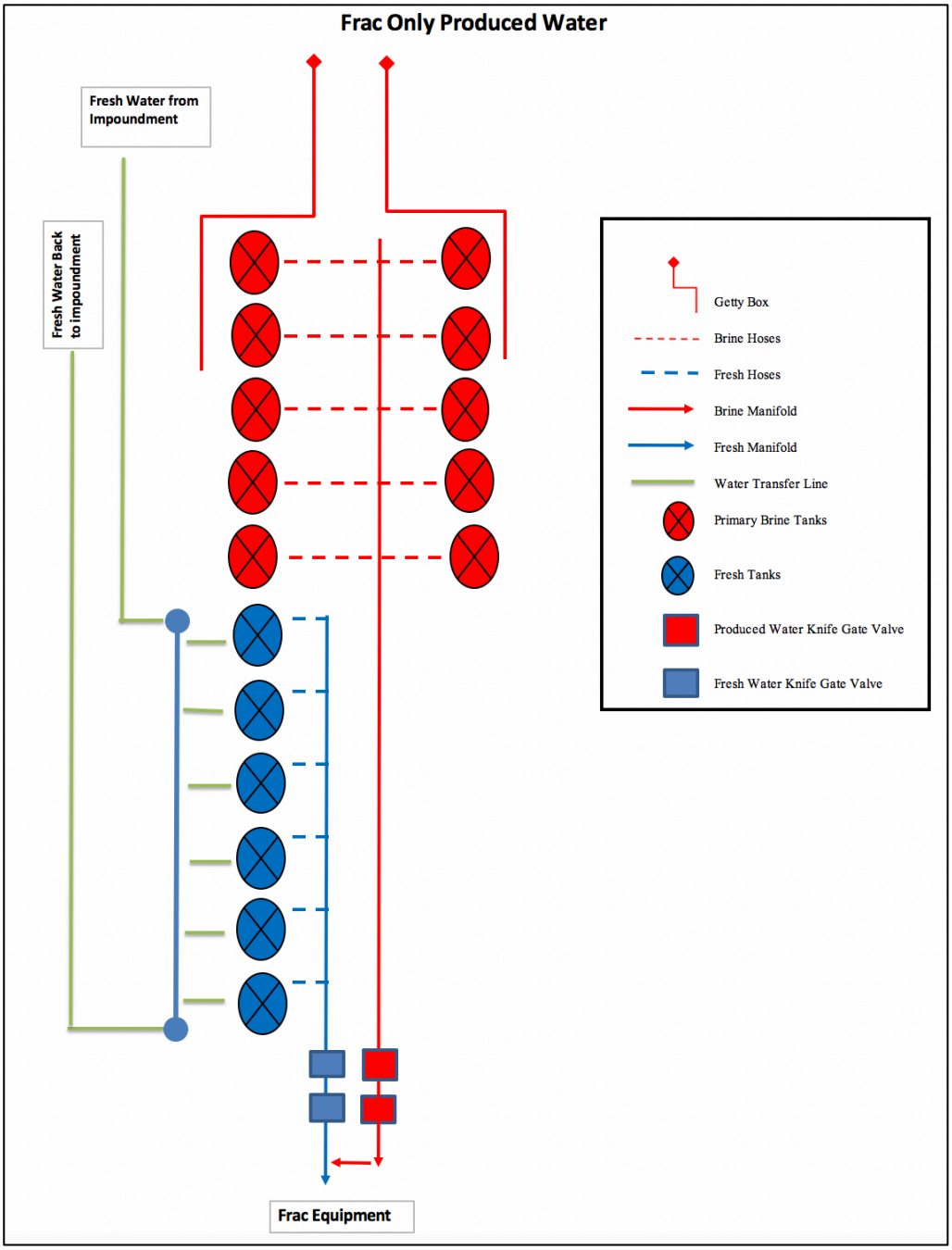

Figure B. 2 - Layout 2 Depicting Manifold and Valve Design for Pumping PW with Stimulation Equipment.

Using the layout in figure B.2 provides the service company and operator a simple way to swap between FW and PW as the design or volumes allow. Since hauling PW to

location from production facilities or other operations is irregular, this design allows for flexibility and does not require one or the other to pump a frac stage. The two gate valves on the FW manifold and PW manifold allow the service company to easily swap between the two as needed. If the switch needs to be made on-the-fly, open one set of valves before closing the other to ensure there is no loss of suction at the blender. For example, if a stage is being pumped with FW and the service company wishes to switch to PW for flush, open the produced water knife gate valves and then close the fresh water valves to maintain prime.

Figure B.3 shows a combined layout for both frac and wireline pump down that will work in warmer weather but will most likely cause too many problems in winter. The pump down common manifold is tied into the frac PW and FW manifolds. This design does provide the benefit of having consolidated operations. This layout will require increased monitoring of the PW tanks as they are common for both frac and wireline and the operations may interfere with one another.

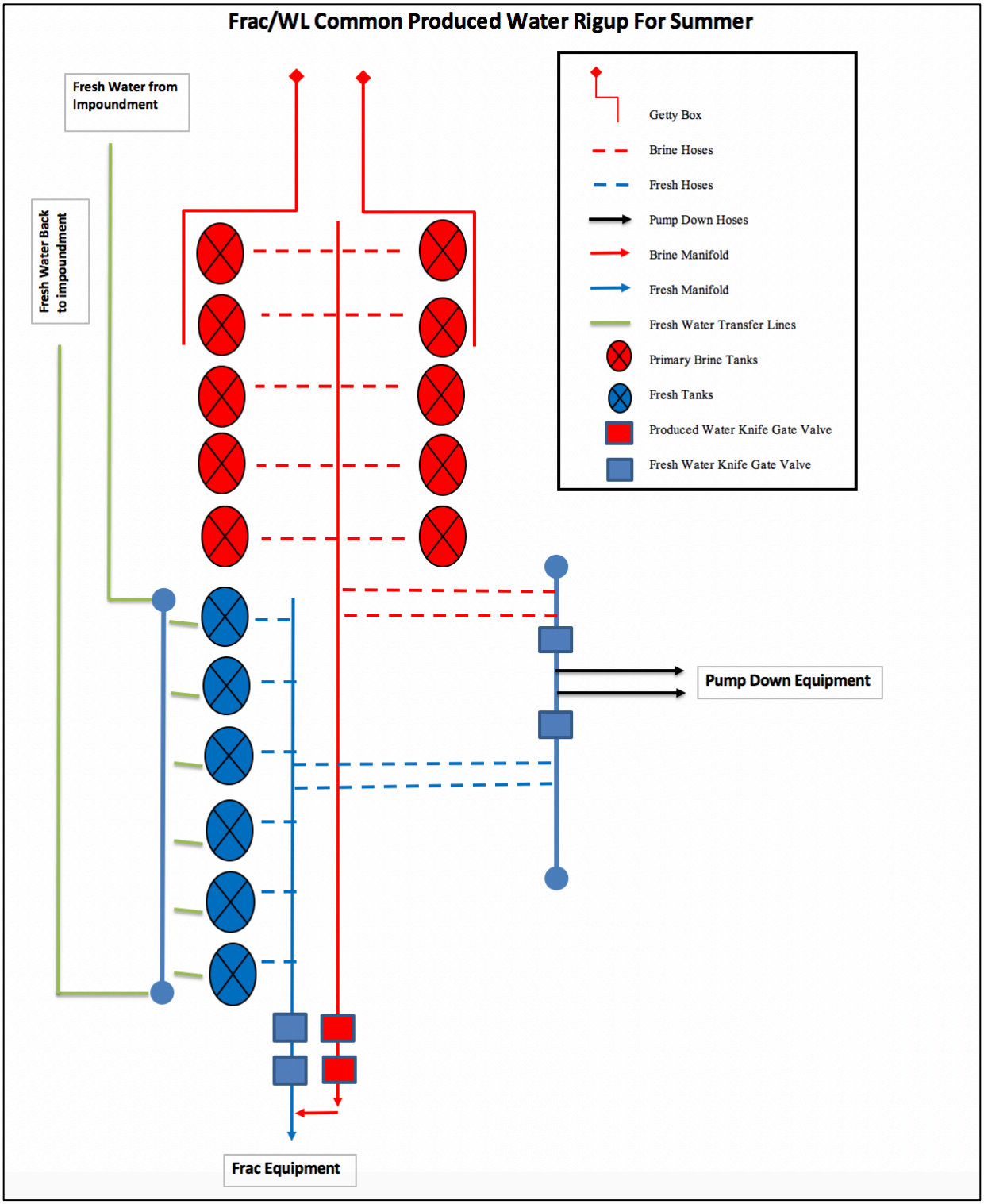

Figure B. 3 - Layout 3 Depicting Manifold and Valve Design for Pumping Produced Water with Stimulation Equipment.

The above three designs allow operators to efficiently dispose of produced water without having to sacrifice downtime in operations due to insufficient produced water volumes. However, this process does require the continuous monitoring of tank volumes and freshwater lines. If the process is executed effectively, operators may see decreases in well costs by reducing disposal fees and providing sustainable solutions to freshwater consumption and truck traffic (Barnes, et al., 2015).

First, substituting fresh water for produced water in frac designs and pump down operations can reduce freshwater costs. Consider a wellbore design for a well drilled in the middle Bakken with 7 in. 32 # casing string from 0 ft. at the tubing head to 10,500 ft. with liner top at 10,500 ft. and a 4 ½ in. 13.5# cemented liner to a TD of 20,000 ft. MD. This wellbore will have a vertical capacity of 0.0360 bbl/ft and a liner capacity of 0.0149 bbl/ft. Also assume each frac stage has an interval length of 220 ft. So, the flush volume will decrease by the volume equivalent of total length of wellbore from plug to plug, or 220 ft. *(0.0149 bbl/ft) = 3.3 bbls. Using this logic, the wellbore volumes for each type of wellbore is summarized in table B.1. A second wellbore design will also be considered which only differs in that there is 4 ½ tie back ran from liner top to surface, thus decreasing the vertical volume. The flush volumes for each design are summarized in table B.1. Note, these volumes do not include the volume of surface lines.

Table B. 1 - Example volumes and cost savings for 7 in. and 4.5 in. wellbore designs.

Stage Number	TD (feet)	Lateral Length (feet)	7 in. Flush Volume (bbl)	4 ½ in. Flush Volume (bbl)	Cost saving per flush for 7 in. (USD)	Cost saving per flush for 4 ½ in. (USD)
1	20,000	9,500	519.6	298.0	649.4	372.5
2	19,780	9,280	516.3	294.7	645.3	368.4
3	19,560	9,060	513.0	291.4	641.2	364.3
4	19,340	8,840	509.7	288.2	637.1	360.2
5	19,120	8,620	506.4	284.9	633.0	356.1
6	18,900	8,400	503.2	281.6	629.0	352.0
7	18,680	8,180	499.9	278.3	624.9	347.9
8	18,460	7,960	496.6	275.1	620.8	343.8
9	18,240	7,740	493.3	271.8	616.7	339.7
10	18,020	7,520	490.0	268.5	612.6	335.6
11	17,800	7,300	486.8	265.2	608.5	331.5
12	17,580	7,080	483.5	261.9	604.4	327.4
13	17,360	6,860	480.2	258.7	600.3	323.3
14	17,140	6,640	476.9	255.4	596.2	319.2
15	16,920	6,420	473.7	252.1	592.1	315.1
16	16,700	6,200	470.4	248.8	588.0	311.0
17	16,480	5,980	467.1	245.6	583.9	306.9
18	16,260	5,760	463.8	242.3	579.8	302.8
19	16,040	5,540	460.5	239.0	575.7	298.7
20	15,820	5,320	457.3	235.7	571.6	294.6
21	15,600	5,100	454.0	232.4	567.5	290.6
22	15,380	4,880	450.7	229.2	563.4	286.5
23	15,160	4,660	447.4	225.9	559.3	282.4
24	14,940	4,440	444.2	222.6	555.2	278.3
25	14,720	4,220	440.9	219.3	551.1	274.2
26	14,500	4,000	437.6	216.1	547.0	270.1

27	14,280	3,780	434.3	212.8	542.9	266.0
28	14,060	3,560	431.0	209.5	538.8	261.9
29	13,840	3,340	427.8	206.2	534.7	257.8
30	13,620	3,120	424.5	202.9	530.6	253.7
31	13,400	2,900	421.2	199.7	526.5	249.6
32	13,180	2,680	417.9	196.4	522.4	245.5
33	12,960	2,460	414.7	193.1	518.3	241.4
34	12,740	2,240	411.4	189.8	514.2	237.3
35	12,520	2,020	408.1	186.5	510.1	233.2
36	12,300	1,800	404.8	183.3	506.0	229.1
37	12,080	1,580	401.5	180.0	501.9	225.0
38	11,860	1,360	398.3	176.7	497.8	220.9
39	11,640	1,140	395.0	173.4	493.7	216.8
40	11,420	920	391.7	170.2	489.6	212.7
Totals			**18225.2**	**9363.2**	**22781.5**	**11704.0**

We can see from table B.1 that the cost savings accrued (assuming 1.25 USD purchase price per barrel of FW which is within the range outlined by Boschee (2012)) by substituting FW for PW can yield significant cost savings per well.

The totals provided at the bottom of table B.1 show that even substituting fresh water for PW solely for flush would yield a savings of over 18,000 bbls of FW for one 7 in. wellbore design. These savings can scale rapidly if multiple operators just use PW for flush. The savings are further amplified if most pump down operations also utilize produced water; thus, for one well on one pad, there is potential to save over 18,000 bbls of PW which may be redistributed to other sectors of the economy.

Estimating the amount of PW that may be reused during wireline plug pump down is harder to estimate. The amount of fluid required to pump a bottom hole assembly (BHA) down to plug setting depth is highly dependent on the pump-down efficiency (Walton, Nichols, & Fripp, 2019). Factors like plug design, friction in the wellbore, weight of BHA, residual proppant in wellbore, line tension, and bypass rates all affect the pump down efficiency (Walton et al. 2019). From the field experience of the author, it may take 400-450 bbl to pump a 3.68 in. OD plug to a depth of about 20,000 ft from liner top, which is usually around 10,500 ft. So, if PW is used for these plug pump downs, the operator should see even more savings in FW costs.

There are some tradeoffs associated with this practice that may be worth mentioning. First, an operator runs the risk of pumping PW that may be contaminated with organic content as it is transferred from facilities and trucks on site. So, operators should treat this water the same as any other FW source.

Next, the water that is brought from formation to surface is also very hot, usually in excess of 150°F. Now while this provides additional benefits to operations in the winter, it is a safety hazard to consider. Proper precautions should be taken when handling any of the hoses or working around the tanks.

There are also considerations with water quality. These designs would be hard to implement if there were any H2S present in either flowback or PW. Again, special safety considerations would have to be taken into account if this were the case and operators should implement them accordingly.

Although not a problem with PW from production, solids could potentially pose a problem when reusing flowback water directly (Boschee, 2014). This is usually mitigated during flowback through filtration at the separating tank to avoid any solids disposal costs. If an operator wishes to reuse fluid from flowback operations after a screenout, there are numerous service companies that offer on-site filtration systems that are relatively inexpensive compared to trucking and disposal fees. Tank arrangement, manifold design, and valve placement can provide the augmentation of fresh water and produced water for stimulation operations.

Fluid compatibility is not an issue during wellbore flush, wireline pump down, and any stage that doesn't require sand carrying capacity. Therefore, fluid compatibility is not an issue during these stages and operations. Relying solely on PW and a salt-tolerant fluid system may cause shut downs in operations due to the irregularity of flowback, frac, water cuts, and weather. Therefore, a back-up plan is needed for these instances. The layouts in this study provide a back-up plan. Avoiding operational shut downs can prevent a positive feedback loop for compounding problems during winter operations. Operators may save thousands of barrels of FW, even by simply substituting FW for PW using wellbore flushes and wireline pump down operations.

Nomenclature

TD = Total Depth, feet

MD = Measured Depth, feet

OD = Outer Diameter, inches

= linear pound per foot of casing/tubing, lbm/ft

Appendix C: Simple Mass Balance Algorithm for Spotting Acid During a Wireline Run

The oil and gas (O&G) industry has recently seen an increase in the development and deployment of machine learning and data mining techniques to become more efficient. Although these techniques can streamline operations and reduce human error through automation and prediction, they cannot yet develop wisdom and must, therefore, be augmented. This novel study presents such wisdom through an algorithm developed in the field for spotting acid during wireline operations. Although this is just one of many algorithms used in field operations, many of them are not documented as they lack data or were not developed through a series of complex equations and are thus, cast aside. However, as the O&G industry and field operations begin to automate, these practices will become vital resources that can help machine learning models built by data scientists gain wisdom. The algorithm developed in this study presents a process to place acid in the wellbore during wireline operations. Combining wireline and acid spotting operations can save pump time, make operations more efficient, and thus decrease costs. Combining these operations was also made possible by the development of a hydrochloric replacement acid that is non-corrosive on wireline equipment and necessary for the operation. Although the process was developed on two horizontal wellbore designs in the Williston basin, the process is robust and will work on any wellbore design. Additional operational considerations regarding the wellbore design are also presented.

This domain knowledge or wisdom contained in human capital rather than data sets has been steadily decreasing in the O&G industry. Between January 2015 and November

2021, the number of employees in the O&G extraction sector decreased from around 200,000 to about 140,000 (U.S. Bureau of Labor Statistics, 2021). This is a major loss of human capital and anecdotal and domain knowledge. Therefore, it must be captured and managed to successfully develop practical results from ML and automation. Symptoms of knowledge mismanagement include good ideas and best practices that are not effectively dispersed leading to repetition of past events and operations (Van Der Spek 2017). These symptoms can increase overall costs and may be quantified as the "cost of ignorance" (Van Der Spek 2017).

The methodology for spotting acid during wireline operations set forth in this paper was developed though field operations on two different types of wellbore construction: the first is a 7" intermediate string with a 4.5" cemented liner and the second was a 4.5" tieback with a cemented 4.5" liner. The calculations to place acid are the same in both instances, and should be the same for any wellbore configuration, but the operational issues will be different.

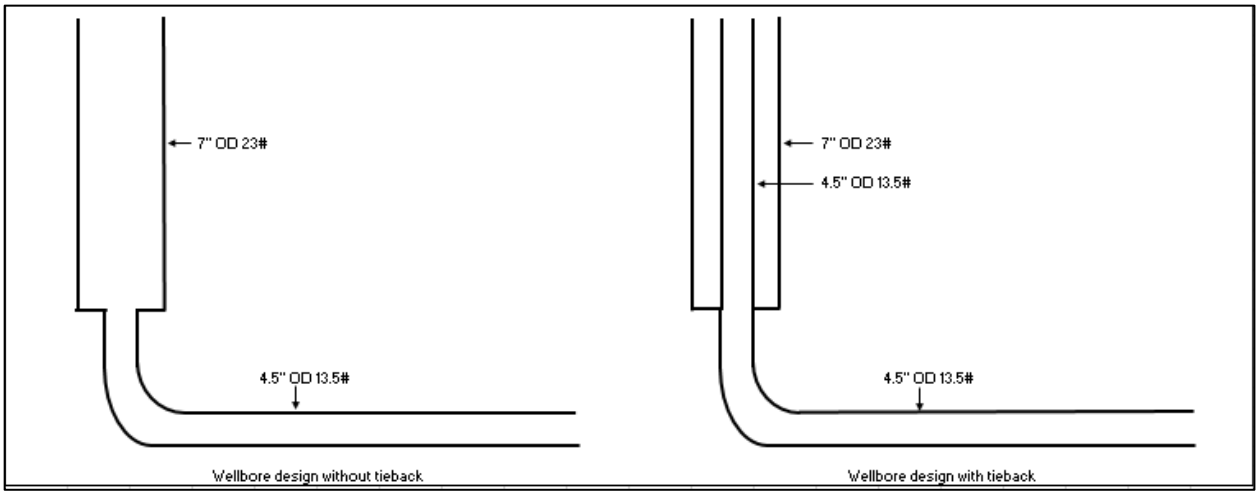

Figure C. 1 - Schematic of wellbore designs when developing acid spotting algorithm

Figure C.1 shows the wellbore designs present when developing the algorithm in this study, although this method should work on any wellbore design.

The method developed in this study was developed out of the need for: 1) decreased freshwater (FW) consumption and 2) decrease pump time on the frac spread to decrease stage costs. Traditionally, the wireline tool string is pumped down hole to the desired depth using separate pumps from the frac spread. After the isolation plug is set, wireline pulls out of the hole and shoots the perforations while moving or "on the fly". After the well is handed over to the frac, the frac spread spearheads acid (traditionally hydrochloric acid) and displaces it to clean up the perforations, clean up skin caused by cement, or aid in breaking down formation.

The compatibility of newly developed hydrochloric acid replacement (HCR) that is non-corrosive to wireline equipment allows them to interact in the wellbore at the same time (Yocham, Allison, & Schwartz, 2021). This is the critical reason why the operations had to previously be kept separate. Compatibility allows the acid and tool string to run in the hole at the same time and on the wireline side of operations. Being able to spot acid during wireline operations allows for the acid to sit a small volume away from perforations once the frac spread begins to pump. This essentially eliminates one wellbore volume of fluid and thus decreases stage pump time and freshwater consumption.

Yocham et al. (2021) conducted a case study on the deployment of acid during wireline operations. They noted significant efficiency gains in decreased time to reach designed pump rate and a decrease in the total stage pump time (Yocham et al. 2021),

thus decreasing costs. This is possible because most pump down charges are included in packages and thus there is no cost associated with this pump time. Although Yocham et al. (2021) established the efficiency gains by spearheading acid during wireline operations, there is no formal method of performing the operation in the field.

This study proposes a methodology to effectively place or spot acid during wireline operations at any desired point in the wellbore. This is formalized in equation C.1:

$$-BF_{bbl} + P_{PD} = F - SF \tag{C.1}$$

Where,

BF_{bbl} = Number of barrels to pump before/after starting pump down (bbl)
P_{PD} = Previous zone pump down volume (bbl)
F = Wellbore flush volume to plug depth (bbl)
SF = Desired safety factor (volume before plug depth) (bbl)

More importantly,

$$-BF_{bbl} = F - SF - P_{PD} \tag{C.2}$$

Equation C.2 calculates how much fluid volume is to be pumped before/after pumping acid. The sign convention is such that a (-) value for BF_{bbl} indicates fluid volume (including acid) before starting pump down and (+) indicates fluid volume (excluding acid) after starting pump down. This process is necessary because the fluid required to

pump a tool string downhole depends on the drag and is not necessarily equal to the wellbore volume at the desired depth while pumping acid in the wellbore depends on casing capacities. This procedure requires performing at least one pump down prior to implementing this method. This will establish approximately how much fluid it will take to pump before the tool string reaches the desired depth. This process assumes the difference in volume between the previous and current pump down is negligible. This also assumes surface line volume is negligible.

For example, if the previous pump down volume was 400 bbl, the flush volume to plug setting depth is 300 bbl, and a 50 bbl safety factor is desired, then equation C.2 becomes $-BF_{bbl} = 300 - 50 - 400 = -150$ bbl, or $BF_{bbl} = 150$ bbl. This means that 150 bbl of fluid needs to be displaced during the pump down before pumping acid. So, if acid is clearing the well head at 150 bbl on the flowmeter after starting to pump down the wireline tool string, we can expect to pump approximately 250 bbl ($P_{PD} - BF_{bbl}$) before shutting down. This should put the acid 50 bbl short of the flush volume to the plug setting depth (300 bbl). Once operations are lined out, pump down volumes tend to decrease by a few bbl per stage.

Conversely, if the previous pump down volume was 300 bbl and the flush volume to the plug were 400 bbl, with the same 50 bbl safety factor, equation C.2 yields $-BF_{bbl} = 400 - 50 - 300 = 50$ bbl, or $BF_{bbl} = -50$ bbl. This would mean that you will need to pump 50 bbl of fluid (including acid) before starting to pump down the tool string. So procedurally, pumping 50 bbl (including acid) then a 300 bbl pump down will leave the acid 50 bbl short of the volume to the plug (400 bbl).

Example 2 also represents the operation after encountering an inflection point in the well as plug setting depths progress towards the heel of the well where $P_{PD} < F$. For example, if $P_{PD} = 50$ bbl, F = 200 bbl, and SF = 50 bbl, then equation C.2 becomes

$$-BF_{bbl} = 200 \text{ bbl} - 50 \text{ bbl} - 50 \text{ bbl}$$

Or

$$B_{BF} = -100 \text{ bbls}$$

This indicates that 100 bbl needs to be displaced before pumping the tool string downhole. This makes logical sense: If it will only take 50 bbl to pump the tool string to the desired depth, but the volume that the acid has to travel is 200 bbl, then it will need to be further down the wellbore before pumping the tool string to the desired depth.

The above process also requires using a ball in place method, where the ball for the frac plug is secured onto the seat of the plug before being installed on the setting tool. The best practice is to set the plug, start the rate assist using a pump, and wait for pressure to start to increase to a point below the pop-off pressure. The pressure should start to fall off after successful firing of guns and the rate assist can be resumed.

As the perforation depths move towards the heel of the well, there will come an inflection point where $P_{PD} < F$. This is the case shown in *Example 2* (although this may be the case for other operations as well). For example, if $P_{PD} = 50$ bbl and F = 200 bbl, then equation (2) becomes

$-BF_{bbl} = 200\ bbl - 50\ bbl - 50\ bbl$

Or

$B_{BF} = -\ 100\ bbls$

This indicates that 100 bbl needs to be displaced behind the acid before pumping the tool string downhole. This makes logical sense: If it will only take 50 bbl to pump the tool string to the desired depth, but the volume that the acid has to travel is 200 bbl (since it's a fluid and subject to the wellbore volume), then it will need to be further down the wellbore before pumping the tool string to the desired depth.

It's worth noting here that this operation should be performed at the discretion of the operator depending on wellbore construction and the assumed level of risk. The first option is to pump the acid before wireline runs in the hole. The second option is to pump the acid while wireline is in the vertical section of the wellbore. Although this may be marginally quicker, there is an increased level of risk if the area between the ID of the casing and the OD of the tool string is sufficiently small. Problems may arise if the pumps are shut down too quickly creating large water hammers and surges.

Combining wireline and acid spotting operations can yield significant savings in pump time by allowing the desired pump rate to be achieved more quickly and thus decreases overall stage pump time (Yocham et al. 2021). Although previous studies have investigated the cost savings, there is no established algorithm to properly execute pumping acid during a wireline run.

This study provides a methodology for placing HCR acid during wireline pump down operations to any desired depth in the wellbore. This will help formalize the process and disseminate the necessary knowledge to decrease the cost of ignorance (Van der Spek, 2017).

The methodology developed in this study will work for any wellbore design, thus providing a standard procedure for spearheading acid in the wellbore during wireline operations.

Appendix D: Supplementary Figures

Figure D. 1 - Static water tanks holding fluid to be used for stimulation. Photograph was taken in Wyoming.

Figure D. 2 - Monoline with low and high pressure circuits. Photograph taken in Wyoming.

Figure D. 3 - Frac pumps hooking up to missile, converting low pressure fluid to STP. Photograph was taken in North Dakota.

Figure D. 4 - Fluid end of a positive displacement frac pump. Photograph was taken in Wyoming.

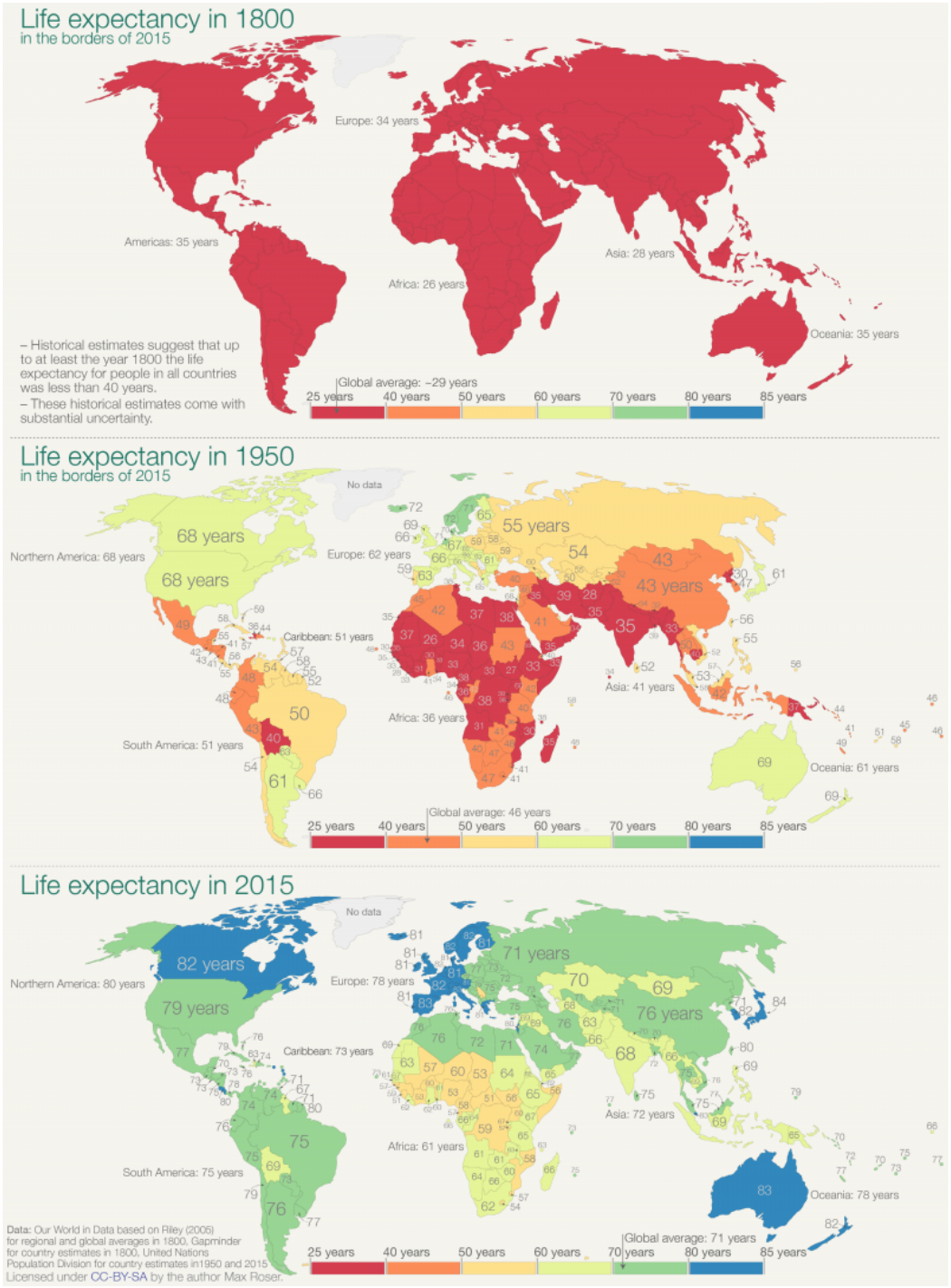

Figure D. 5 - Graphic of changes in global life expectancy since 1800 (Roser et al., 2013)

References

Çambel, A. (1993). *Applied Chaos Theory: A Paradigm for Complexity.* San Diego, CA: Academic Press.

Abobaker, E., Elsanoose, A., Khan, F., Rahman, M., Aborig, A., & Noah, K. (2021). A New Evaluation of Skin Factor in Inclined Wells with Anisotropic Premability. *Energies, 14, 5585*, https://doi.org/10.3390/en14175585.

Al-Muntasheri, G. (2014). A Critical Review of Hydraulic Fracturing Fluids Over the Last Decade. *Paper presented at the SPE WEstern North American and Rocky Mountain Joint Meeting, Denver, Colorado, USA, 17-18 April. SPE-169552-MS.*, https://doi.org.10.2118/169552-MS.

Aloini, D., Dulmin, R., Mininno, V., Stefanini, A., & Zerbine, P. (2020). Driving the Transition to a Circular Economic Model: A Systematic Review on Drivers and Critical Success Factors in Circular Economy. *Sustainability, 12(24)*.

Aro, D., & Fowler, S. (2021). Turning Produced Water into an Asset: A Delaware Basin Case History. *Paper presented at the SPE Hydraulic Fracturing Technology Conference and Exhibition, Virtual, 4-6 May*.

Bailey, M. (2017). *Real Econometrics: The Right Tools to Answer Important Questions.* New York: Oxford University.

Bangert, P. (2021). The Necissity for Collaboration Between Data Scientists and Domain Experts. *Paper presented at SPE Symposium: Artificial Intelligence - Towards a Resilient and Efficient Energy Inudstry, Virtual, 18-19 October*, https://doi.org/10.2118/208634-MS.

Barati, R., & Alhubail, M. (2021). *Unconventional Hydrocarbon Resources: Techniques for Reservoir Engineering Analysis.* Hoboken, New Jersey: Aermican Geophysical Union and John Wiley & Sons, Inc.

Barba, R., Allison, J., & Villarreal, M. (2022). *A Comparison o LAtest Generation Frac New Well and Refrac Results with Evidence of Refrac Reorientation.* Paper presented at the SPE/AAPG/SEG Unconventional Resources Technology Conference, Houston, Texas, USA, June.

Bardi, U., & Lavacchi, A. (2009). A Simple Interpretation of Hubbert's Model of Resource Exploitation. *Energies, 2*, 646-661. doi: 10.3390/en20300646.

Barnes, C., Marshall, R., Mason, J., Skodack, D., DeFosse, G., Smith, D., . . . Cecchini, M. (2015). The New Reality of Hydraulic Fracturing: Treating Produced Water is Cheaper than Using Fresh. *Paper presented at th eSPe Annual Technical*

Conference and Exhibition, Houston, Texas, USA, 28-30 September. SPE-174956-MS., https://doi.org/10.2118/174956-MS.

Barree, R. (2020). Potential Issues with Limited Entry Horizontal Wells. *Paper presented at the Unconventional Resource Technology Conference Austin, Texas, USA, 20-22 July*, https://doi.org/10.15530/urtec-2020-1024.

Baumol, W. (2002). *The Free Market Innovation Machine: Analyzing the Growth Miracle of Capitalism.* Princeton, New Jersey: Princeton University Press.

Behrmann, A., & Nolte, K. (1998). Perforating Requirements for Fracture Stimulations. *Presented at the SPe International Symposium on Formation Damage Control, LAfayette, LA, USA, 18-19 February*, SPE-59480-PA. https://doi.org/10.2118/59480-PA.

Bhat, B., Lui, S., Lin, Y., Sentmanat, M., Kwon, , J., & Akbulut, M. (2021). Supramolecular Dynamic Binary Complexes with pH and Salt-Responsive Properties for use in Unconventional Resrvoirs. *PLoS ONE 16(12)*, https://doi.org/10.1371/journal.pone.0260786.

Borenstein, S., & Kellogg, R. (2021). Challenges of a Clean Energy Transition and Implications for Energy Infrastructure Policy. *Rebuilding the Post-Pandemic Economy, ed. Melissa S. Kearney and Amy Ganz (Washington D.C.:Aspen Institute Press)*, 234-271.

Boschee, P. (2012). Handling Produced Water from Hydraulic Fracturing. *Oil & Gas Fac 1 (01): 22-26. SPE-0212-0022-OGF.* , https://doi.org/10.2118/0212-0022-OGF.

Boschee, P. (2014). Produced and Flowback Water Recycling and Reuse Economics, Limitations, and Technology. *Oil & Gas Fac 3 (01), 16-21*, SPE-0214-0016-OGF.

Brady, K., Atwell, B., Wulterich, K., Breyer, J., Braim, M., & Pack, C. (2022). *Completions Optimization and Fracture Evaluation of Infill Wells: A Southern Delaware Basin Case Study.* Paper presented at the Unconventional Resource Technology Conference, Houston, Texas, USA, 20-22 June.

Castilla, A., Zeuss, M., & Schmidt, M. (2021). Crcular Economy in the Oil and Gas Exploration and Production; Resource Recovery from Drill Cuttings and other Oily Wastes. *Paper presented at the Abu Dhabi International Petroleum Exhibition & Conference, Abu Dhabi, UAE, November 2021*.

Chen, B., Xiong, R., Li, H., Sun, Q., & Yang, J. (2019). Pathways for sustainable energy transition. *Journal of Cleaner Production*, volume 228, 1564-1571.

Cheung, C., He, G., & Pan, Y. (2020). Mitigating the Air Pollution Effect? The REmarkable Decline in the Pollution-Mortality Relationship in Hong Kong. *Journal of Environmental Economics and Management, 101, 102316*.

Cleary, M., Johnson, D., Kogsboil, H., Owens, K., Perry, K., de Pater, C., . . . Tambini, M. (1993). Field Implementation of Proppant Slugs To Avoid Premature Scrren-Out of Hydraulic Fractures With Adequate Proppant Concentration. *Paper Presented at the SPE Rocky Mountain Regional Low Permability Reservoirs Symposium, Denver, Colorado, U.S.A., 12-14 April*.

Cramer, D. (1987). The Application of Limited Entry Techniques in Massive Hydraulic Fracturing Treatments. *Paper presetned at SPE Prodution and Operations Symposium, Oklahoma City, Oklahoma, 8-10 March.* .

Cross, T., Niederhut, D., Cui, A., Sathaye, K., & Chaplin, J. (2021). Quantifying the Diminishing Impact of Completions Over Time Across the Bakken, Eagle Ford, and Wolfcamp Using Multi-Target Machine Learning Model and SHAP Values. *Paper prepared for presentation at the Unconventional Resources Technology Conference, Houston, Texas, USA, 26-28 July*.

Cross, T., Niederhut, D., Saythaye, K., Darnell, K., & Crifasi, K. (2020). Deriving Time-Dependent Scaling Factors for Completions Parameters in the Williston Basin using a Multi-Target Machine Learning Model and Shapley Values. *Paper presnted at the Unconventional Resources Technology Conference, Austin, Texas, USA, 20-22 July*.

Damjanac, B., Maxwell, S., Pirayehgar, A., & Torres, M. (2018). Numerical Study of Stress Shadowing Effect on Fracture initiation and Interaction Between Perforation Clusters. *Presented at the Unconventional Resrouces Technology Conference, Houston, Texas, USA, 23-25 July*, URTEC-2901800-MA. https://doi.org10.15530/URTEC-2018-2901800.

Darabi, H., Zhai, X., Kianinejad, A., Ma, A., Castineira, D., & Toronyi, R. (2020). Augmented AI Framework for Well Performance Prediction and Opportunity Identification in Unconventional Reservoirs. *Paper presented at the International Petroleum Technology Conference, Dhahran, Sadi Arabia, 13-15 January*.

Davis, G. (2010). The Resource Drag. *International Association for Energy Economics. Working Paper.*

Dendup, N. (2022). Returns to grid electricity on firewood and kersoene: Mechanism. *Jouranal of Environmental Economics and Management, 111*.

Dhanya, P., & Geethalakshmi, V. (2023). Reviewing the Status of Droughts, Early Warning Systems and Climate Services in South India: Experiences Learned. *Climate, 11(3), 60*.

Egenhoff, S. (2017). The Lost Devonian Sequence - Sequence Stratigraphy of the Middle Bakken Member, and the Importance of Clastic Dykes in the Lower Bakken Member Shale, North Dakota, USA. *Marine and Petroleum Geology, 81* , 278-293.

Egenhoff, S., & Fishman, N. (2013). Traces In The Dark - Sedimentary Processes and Facies Gradients In The Upper Shale Member Of The Upper Devonian-Lower Mississippian Bakken Formation, Williston Basin, North Dakota, U.S.A. *Journal of Sedimentary Research*, v. 83, 803-824. DOI: 10.2110/jsr.2013.60.

EIA. (2022). *Bakken Region Drilling Productivity Report.* Energy Information Administration.

Energy and Environmental Resource Center. (2016). *A Review of Bakken Water Management Practices and Potential Outlook. Final Report.* Grand Forks, North Dakota: University of North Dakota.

Energy Information Administration. (n.d.). Retrieved from https://www.eia.gov/finance/data.php

Energy Information Administration. (2022, December 14). Retrieved from U.S. Energy-Related Carbon Dioxide Emissions, 2021: https://www.eia.gov/environment/emissions/carbon/

Energy Information Administration. (2023a, 4 6). *Energy and the Environment Explained*. Retrieved from https://www.eia.gov/energyexplained/energy-and-the-environment/where-greenhouse-gases-come-from.php

Energy Information Administration. (2023b, 4 3). *U.S. Energy Facts*. Retrieved from https://www.eia.gov/energyexplained/us-energy-facts/imports-and-exports.php,

Energy Information Administration. (2023c, 4 3). *Petroleum and Other Liquids*. Retrieved from https://www.eia.gov/dnav/pet/hist/LeafHandler.ashx?n=PET&s=MCRFPUS2&f=A

Energy Information Administration. (2023d, 5 3). *Short Term Energy Outlook*. Retrieved from https://www.eia.gov/outlooks/steo/

Environmental Protection Agency. (2023, 03 03). *Greenhouse Gases Equivalencies Calculator - Calculations and References*. Retrieved from United States Environmental Protection Agency Website: https://www.epa.gov/energy/greenhouse-gases-equivalencies-calculator-calculations-and-references

Epstein, A. (2014). *The Moral Case for Fossil Fuels.* New York: Penguin Group.

Fishman, N., Egenhoff, S., Boehlke, A., & Lowers, H. (2015). Petrology and diagenetic history of the upper shale ember of the Late-Devonian-Early Mississippian Bakken Formation, Williston Basin, North Dakota. *The Geological Society of America, Special Paper 515*.

Fry, J., & Paterniti, M. (2014). Production Comparison of Hydraulic Fracturing Fluids in the Bakken and Three Forks Formations of North Dakota. *Paper presented at the SPE Western North American and Rocky Mountain Joint Regional Meeting, Denver, Colorado, USA, 16-19 April*.

Gaswirth, S., Marra, K., Cook, T., Charpentier, R., Gautier, D., Higley, D., & Whidden, K. (2013). Assesment of Undiscovered OIl Re-sources in the Bakken and Three Forks Formations, Williston Basin Province, Montana, North Dakota, and South Dakota. *USA.Geol. Surv. Fact. Sheet 2013, 3013.* .

Gerhard, L., & Anderson, S. (1988). Geology of the Williston Basin (United States portion). *Sloss, L.L., ed, Sedimentary over - North American Craton, U.S., Boulder Colorado, Geological Society of America, The Geology of North America, v. D-2*.

Gorucu, S., Shivastava, V., & Nghiem, L. (2021). Numerical Simulation of Proppant Transport in Hydraulically Fractured Reservoirs. *Presented at the SPE Reservoir Simulation Conference, On-Demand, 26 October*, SPE-203927-MS. https: //doi.org/10.2118/203927-MS. .

Griffin, L., Poppel, B., Siegel, J., & Weijers, L. (2014). Technical Implementation and Benefits of Use of Produced Water in Slickwater and Hybrid Treatments in the Bakken Central Basin. *Paper presented at the SPE Western North American and Rocky Mountain Joint Meeting, Denver Colorado, USA, 17-18 April*, SPE-169497-MS. https://doi.org/10.2118/169497-MS.

Grisso, R., Perumpral, J., Vaughan, D., Roberson, G., & Pitman, R. (2010). *Predicting Tractor Diesel Fuel Consumption.* Virginia Polytechnic Institute and State University, Communications and Marketing, College of Agricultural and Life Sciences. Virginia Cooperative Extension.

Gustavo, A., Ugueto, C., Huckabee, , P., Molenaar, , M., Wyker, , B., & Somanchi, , K. (2016). Perforation Cluster Efficiency of Cemented Plug and Perf Limited Entry Completions; Inisghts from Fiber Optic Diagnostics. *Paper presented at the SPE Hydraulic Fracturing Technology Conference, The Woodlands, Texas, 9-11 February.*

Hadjdu, T., & Hajdu, G. (2022). Climate change and the mortality of the unborn. *Journal of Environmental Economics and Management, 118*, https://doi.org/10.1016/j.jeem.2022.102771.

Harrell, J. F. (2015). *Regression Modeling Strategies: With Application to Linear Models, Logistic and Ordinal Regression, and Survival Analysis, Second Addition.* New York: Springer Series in Statistics.

Hausman, C., & Kellogg, R. (2015). Welfare and Distributional Implications of Shale Gas. *Brrokings Papers on Economic Activity*.

Haustveit, K., Elliot, B., & Roberts, J. (2022). Empirical Meets Analytical-Novel Case Study Quantifies Fracture Stress Shadowing and Net Presure Using Bottom Hole Pressure and Optical Fiber. *Paper presented at the SPE Hydraulic Fracturing Technology Conference and Exhibition, The Woodlands, Texas, USA, 1-3 February*.

Havens, J. (2012). *Mechanical Properties of the Bakken Formation.* Golden, Colorado: Master Thesis.

Havens, J., & Batzle, M. (2011). Minimum Horizontal Stress in the Bakken Formation. *Paper presented at the 45th US Rock Mechanics/Geomechanics Symposium, San Francisco, California, U.S., 26-29 June.*

He, J., Pei, Y., & Sepehrnoori, K. (2021). Impact of Stochastic Hydraulic Fracture Properties on Stress Reorientation in Shale Gas Reservoirs. *Paper presented at the 55th US Rock Mechanics/Geomechanics Symposium, Housten Texas, USA 20-23 June*, ATMA-2021-1340.

Heck, T., LeFever, R., Fischer, D., & LeFever, J. (2002). *Overview of the Petroleum Geology of the North Dakota Williston Basin.* Bismarck, North Dakota: North Dakota Geologic Survey.

Hlavac, M. (2022. https://CRAN.R-project.org/packages=stargazer). Stargazer: Well-Formatted Regression and Summary Statistics Tables. Socail Policy Institute, Bratislava, Slovakia.

Huang, J., & Datta-Gupta, A. (2017). Optimization of Hydraulic Fracture Development and Well Performance Using Limited Entry Perforations. *Paper presented at the SPE Oklahoma City Oil and Gas Symposium, Oklahoma City, Oklahoma, USA, 27-31 March*.

Hubbert, M. (1962). Energy Resrouces. A Report to the Committee on Natural Resources;. *National Academy of Scinces, National Research Council: Washington DC, USA*, 54.

Huchton, J., Mallory, C., Calvin, J., Alberts, M., Rogers, S., Romines, G., . . . Boyer, J. (2020). Enhancement of Production and Economics Through Design Optimization in the STACK. *Paper presented at the SPE Hydraulic Fracturing Conference and Exhibition, The Woodlands, Texas, 4-6 February*.

Hume, D. (1985). *A treatise of human nature (E.C. Mossner, Ed.).* Penguin Classics.

Huntington-Klein, N. (2022). *The Effect: An Introduction to Research Design and Causality.* Boca Raton: CRC Press.

Huntington-Klein, N. (2022). *The Effect: An Introduction to Research Design and Causality.* Boca Raton, FL: CRC Press.

Josifovic, A., Roberts, J., Corney, J., & Davies, B. (2016). Reducing the environmental immpact of hydraulic fracturing through design optimisation of positive displacement pumps. *Energy, 115*, 1216-1233.

Kirchherr, J., Reike, D., & Hekkert, M. (2017). Conceptualizing the circular economy: An analysis of 114 definitions. *Resources, Conservation & Recycling.* , http://dx.doi.org/10.1016/j.resources.2017.09.005.

Kroschel, J., Rabiei, M., & Rasouli, V. (2022a). Modeling Temporal Dependence of Average Surface Treating Pressure in the Willison Basin Using Dynamic Multivariate Regression. *Energies, 15, 2271*, https://doi.org/10.3390/en15062271.

Kroschel, J., Rabiei, M., & Rasouli, V. (2022b). Accounting for Fixed Effects in Re-fracturing Using Dynamic Multivariate Regression. *Energies, 15, 5451*, https://doi.org/10.3390/en15155451.

Kuhn, M., & Johnson, K. (2020). *Feature Engineering and Selection: A Practical Approach for Predictive Models.* Boca Raton: CRC Press.

Kurz, B., Stepan, D., Glazewski, K., Stevens, B., Doll, T., Kovacevich, J., & Wocken, C. (2016). *A Review of Bakken Water Management Practices and Potential Outlook.* Grand Forks, North Dakota: Energy and Environmental Research Center.

Langorne, K., & Rasmussen, J. (1962). Better Completion by Controlled Fracture Placement Limited-Entry TEchnique. *American Petroleum Institute*.

Li, C., Han, J., LaFollette, R., & Kotov, S. (2016). Lessons LEarned From Refractured Wells: Using Data to Develop an Engineered Approach to Rejuvenation. *Paper presented at the SPE Hydraulic Fracturing Technology Conference, The Woodlands, Texas, USA, 9-11 February*.

Ling, K., Zeng, Z., He, J., Peng, P., Zhou, X., Liu, H., . . . Gosnold, W. (2013). *Geomechanical Study of Bakken Formation for Improved Recovery.* Grand Forks, North Dakota. DOE Award No. DE-FC26-08NT0005643: U.S. Department of Energy - National Energy Technology Laboratory.

List, J. A. (2022, June 9). "Coffee with John List on the Voltage Effect". (N. N. Economics, Interviewer)

Loewenstein, G., & Thaler, R. (1989). Anomalies: Intertemporal Choice. *Jouranl of Economic Perspecives. Volume 3, Number 4*, 181-193.

Luo, Y., Guo, J., Zeng, F., Lu, C., Wang, R., Guan, B., . . . Shan, X. (2022). Optimization Method of Hydraulic Fracturing Design for Horizontal Shale Gas Well Based on Artificial Neural Network and Genetic Algorithm. *Paper presented at the 56th U.S. Rock Mechanics/Geomechanics Symposium, Santa Fe, New Mexico, USA, June*, https://doi.org/10.56952/ARMA-2022-0364.

Maldonado, R., & Aoun, A. (2019). A Data-Driven Evaluation of Hydraulic Fracturing Performance in the Hassi Messaoud Field, Algeria. *Paper presented at the SPE Western Regional Meeting, San Jose, CA, USA 23-26 April; SPW-195294-MS*, https://doi.org/10.2118/195294-MS.

Manchanda, R., & Sharma, M. (2012). Impact of Completion Design on Fracture Complexity in Horizontal Wells. *Paper presented at the SPE Annual Technical Conference and Exhibition, San Antonio, Texas, USA, 8-10 October*.

Meissner, F. (1978). Petroleum Geology of the Bakken Formation, Williston Basin, North Dakota and Montana. *Montana Geol. Symp.*, 220-227.

Mohaghegh, S. (2016). Fact-Based Re-Frac Candidate Selection and Design - A Case Study in Application of Data Analytics. *Presented at the Unconventional Resource Technology Conference, San Antonio, TX, USA, 1-3 August*, URTeC - 2433427-MS. https://10.15530/URTEC-2016-2433427.

Mohaghegh, S. (2019). Shale Descriptive Analytics: Which Parameters are Controlling Production in Shale. *Paper presented at the SPE Annual Technical Conference and Exhibition, Calgar, Alberta, Canada, 20 September - 2 October*, SPE-196226-MS. https://doi.org/10.2119/196226-MS.

Mohaghegh, S., Gaskari, R., & Maysami, M. (2017). Shale Analytics: Making Production and Operational Decisions Based on Facts: A Case Study in Marcellus Shale. *Paper presented at the SPE Hydraulic Fracturing Technology Conference and Exhibition, The Woodlands, TX, USA, 24-26 January*, URTeC-2433427-MS. https://doi.org/10.15530/URTEC-2016-2433427.

NASA. (2022, October 25). *Methane 'Super Emitters' Mapped by NASA's New Earth Space Mission*. Retrieved from https://www.nasa.gov/feature/jpl/methane-super-emitters-mapped-by-nasa-s-new-earth-space-mission

NASA. (2023, February 14). *What is the greenhouse effect?* Retrieved from https://climate.nasa.gov/faq/19/what-is-the-greenhouse-effect/.

Newell, R., Prest, B., & Sexton, S. (2018). The GDP-Temperature Relationship: Implications for Climate Change Damages: Working Paper 18-17 REV. *Resource for the Future*.

Niu, W., Lu, J., & Sun, Y. (2021). A Production Prediction Method for Shale Gas Wells Based on Multiple Regression. *Enegies, 14, 1461*, https://doi.org/10.3390/en14051461.

Nordhaus, W. D. (2017). Revisting the social cost of carbon. *Proceedings of the National Academy of Sciences*, 114(7) 1518-1523.

Ochi, J., Dexheimer, D., & Corpel, V. (2013). Produced Water Re-Injection Design and Uncertainties Assesment. *Paper presented at the SPE European Formation Damage Conference and Exhibition, Noordwijk, Netherlands, 5-7 June*.

Ostadhassan, M. (2013). *Geomechanics and Elastic Anisotropy of the Bakken Formation, Williston Basin.* Grand, Forks, North Dakota: University of North Dakota, PhD dissertation.

Ostadhassan, M., Zeng, Z., & Zamiran, S. (2012). *Geomechanical Modeling of an Anisotropic Formation - Bakken Case Study.* Chicago, Illinois: 46th US Rock Mechanics/Geomechanics Symposium.

Oyekale, A., & Molelekoa, T. (2023). Multidimensional Indicator of Energy Poverty in South Africa Using the Fuzzy Set Approach. *Energies 16(5)*, https://doi.org/10.3390/en16052089.

Paktinat, J., O'neill, B., & Tulissi, M. (2011). Case Studies: Impact of High Salt Tolerant Friction REducers on Fresh Water Conservation in Canadian Shale Fracturing Treatments. *Paper presented at the Canadian Unconventional Resources Conference, Calgary, Alberta, Canada, 15 November*, SPE-149272-MS. https://doi.org/10.2118/149272-MS/.

Peirce, A., & Bunger, A. (2015). Interference Fracturing: Nonuniform Distributions of Perforation Clusters That Promote Simultaneous Growth of Multiple Hydraulic Fractures. *SPE J. 20, 384-395. https://doi.org/10.2118/172500-PA*.

Peterson, J. (2017). The Psychological Significance of the Biblical Stories. *Lecture 7: Walking with God - Noah and the Ark*.

Ponomareva, I., Galkin, V., & Martyushev, D. (2021). Operational method for determining bottom hole pressure in mechanized oil producing wells, based on the application of multivariate regression analysis. *Petroleum Research, volume 6, issue 4*, 351-360. https://doi.org/10.1016/j.ptlrs.2021.05.010.

R Core Team. (2022. https://www.R-project.org/). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria.

Rassenfoss, S. (2021). A Robot Takes Over the Drilling Floor. *J Pet Technology (Published online 1 December 2021)*.

Rignol, V., & Bui, B. (2020). The Key Factors for Re-fracturing Success: A Simulation Study. *Presented at the 54th US Rock Mechanics/Geomechanics Symposium, Golden, CO, 28 June-1 July*.

Ritchie, H., Rosado, P., & Roser, M. (2022). *Natural Disasters*. Retrieved from Published online at OurWorldInData.org. Retrieved from: 'https://ourworldindata.org/natural-disasters' [Online Resource] .

Ritchie, H., Roser, M., & Rosado, P. (2020). *CO2 and Greenhouse Gas Emissions*. Retrieved from Published online at OurWorldInData.org. Retrieved from: 'https://ourworldindata.org/co2-and-greenhouse-gas-emissions' [Online Resource].

Roberts, G., Whittaker, J., & Paxson, T. (2020). *Proppant Distribution Observations from 20,000+ Perforation Erosion Measurements.* Paper presented at the SPE Hydraulic Fracturing Technology Conference and Exhibition, The Woodlands, Texas, USA, 4-6 February.

Roser, M., Ortiz-Ospina, E., & Ritchie, H. (2013). *Life Expectancy*. Retrieved from Published online at OurWorldInData.org. Retrieved from: 'https://ourworldindata.org/life-expectancy' [Online Resource].

Rossini, S., Roppoli, G., Mariotti, P., Simona, R., Manotti, M., Viareggio, A., & Biassoni, L. (2020). Produced Water Quality Impact on Injection Performance: Predicting Injectivity Decline for Waterflood Design. *Paper presented at the International Petroleum Technology Conference, Dhahran, Saudi Arabia, 13-15 January*.

Roussel, N. (2017). Analyzing ISIP Stage-by-Stage Excalation to Determine Fracture Height and Horizontal-Stress Anisotropy. *Paper presented at the SPE Hydraulic Fracturing Technology Conference and Exhibition, The Woodlands, Texas USA 24-26 January*, SPE-184865-MS.

Roussel, N., & Sharma, M. (2011). Optimizing Fracture Spacing and Squencing in Horizontal-Well Fracturing. . *Presented at the SPE International Symposium and Exhibition on Formation Damage and Control, Lafayette, LA, USA 10-12 February*, SPE-127986-PA. https://doi.org/10.2118/127986-PA.

Russel, D., Stark, P., Owens, S., Navaiz, A., & Lockman, R. (2021). Simultaneous Hydraulic Fracturing Improves Completion Efficiency and Lowers Costs Per Foot. . *Presented at the SPE Hydraulic Fracturing Technology Conference and Exhibition, Virtually, 4–6 May*.

Sakhardande, R., & Devegowda, D. (2021). Data-Driven Causal Analyses of Parent-Child Well Interactions for Well Spacing Decisions. *Paper presented at the SPE Hydraulic Fracturing Technology Conference and Exhibition, Virtual, 4-6 May*.

Schlumberger. (2023, 3 31). *Schlumberger Energy Glossary*. Retrieved from https://glossary.slb.com/en/terms/s/skin

Seibert, M., & Rees, W. (2021). Through the Eye of a Needle: An Eco-Heterodox Perspective on the Renewable Energy Transition. *Energies 2021, 14, 4508, https://doi.org/10.3390/en14154508*.

Shammam, F., Alkinani, A., Al-Hameddi, A. D.-N., & Al-alwani, M. (2021). Assesment of Production Gain From Refracturing Wells in the Major Shale Plays in the United States. *Presented at 55th Annual U.S. Rock Mechan-ics/Geomechanics Symposium, Virtual*.

Shukla, P. ,. (2022). *Summary for Policymakers. In: Climate Change 2022: Mitigation of Climate Change.* Intergovernmental Panel on Climate Change.

Smith, M., & Bustin, R. (1995). Sedimentology of the Late Devonian and Early Mississippian Bakken Formation, Williston BAsin. *Seventh International Basin Symposium*.

Somanchi, K., Brewer, J., & Reynolds, A. (2017). *Extreme Limited-Entry Design Improves Distribution Efficiency in Plug-and-Perforate Completions: Insights from Fiber-Optic Diagnostics.* Paper presented at the SPE Hydraulic Fracturing Technology Conference and Exhibition, The Woodlands, Texas, 24-26 January. .

Song, M., Zhang, J., Liu, X., Zhang, L., Hao, X., & Li, M. (2023). Developments and Trends in Energy Poverty Research - Literature Visualization Analysis Based on CiteSpace. *Sustainability, 15(3), 2576*, https:doi.org/10.3390/su15032576.

Stotz, L. (1938). *History of the Gas industry.* New York: Stettiner Bros.

Supran, G., Rahmstorf, S., & Oreskes, N. (2023). Assessing ExxonMobil's globals warming projections. *Science, Vol 379, Issue 6628*, DOI: 10.1126/science.abk0063.

Thomas, A. (2013, December). Service Contracts In the Oil and Gas Industry. Cleveland, Ohio: Cleveland State University.

Timmins, C., & Vissing, A. (2022). Environmental justice and Coasian bargaining: The role of race, ethnicity, and income in lease negotations for shale gas. *Journal of ENvironmental Economics and Management, 114*.

Tomomewo, O., Drystad-Cincotta, N., Mann, D., Ellafi, A., Alamooti, M., Srinivasachar, S., & Nelson, T. (2020). Proposed Potential Mitigation of Wastewater Disposal Through Treated Produced Water in Bakken Formation. *Paper presented at the 54th U.S. Rock Mechanics/Geomechanics Symposium, Physical Even Cancelled, 28 June - 1 July*, ARMA-2020-1254. ISBN 978-0-9794975-5-1.

Tomomewo, O., Mann, M., Ellafi, A., Jabbari, H., Tang, C., Ba Geri, M., . . . Iroko, A. (2021). Creating Value for the High-Saline Bakken Produced Water by Optimizing its Viscoelastic Properties and Proppant Carrying Tendency with High Viscosity Friction Reducers. *Paper presented at the SPW Western Regional Meeting, Virtual, April*.

U.S. Bureau of Labor Statistics. (2021, December 14). *Oil and Gas Extraction Sector Employment Statistics.* Retrieved from https://data.bls.gov/timeseries/CES1021100001?amp%253bdata_tool=XGtable&output_view=data&include_graphs=true.

U.S. Department of Energy. (2020). *U.S. Oil and Natural Gas: Providing Energy Security and Supporting Out Quality of Life.* U.S. Department of Energy.

U.S. Energy Information Administration . (February 2011). *Performance Profiles of Major Energy Producers 2009.* Washington D.C.

Van der Spek, R. (2017). Reduce the Costs of Ignorance Through Knowledge Management. *Paper presented at the Abu Dhabi International Petroleum Exhibition & Conference, Abu Dhabi, UAE, November*, https://doi.org/10.2118/188602-MS. .

Walton, Z., Nichols, M., & Fripp, M. (2019). Frac Plug Pump Down Efficiencies and Techniques. *Paper presented at the SPE Annual Technical Conference and Exhibition, Calgary, Alberta, Canada, 30 September - 2 October*, SPE-196210-MS. https://doi.org/10.2118/196210-MS.

Wan, X., Ge, J., Li, W., & Pu, H. (2018). Potential Thermoelastic and Poroelastic Stresses Effects during the Fracture Propagation of Hydraulic Fracturing Treatments in Horizontal Bakken Wells. *Paper presented at the 52nd US Rock Mechanics/Geomechanics Symposium, Seattle, Washington, USA, 17-20 June*.

Wan, X., Rasouli, R., Damjanac, B., Torres, M., & Qiu, D. (2019). Numerical Simulation of Integrated Hydraulic Fracturing, Production and Refracturing Treatments in the Bakke Formation. *Paper presented at the 53rd US Rock Mechanics/Geomechanics Symposium, Seattle, Washington, USA, 17-20 June*.

Wang, S., & Chen, S. (2018). Insights to fracture stimulation design in unconventional reservoirs based on machine learning modeling. *Journal of Petroleum Science and Engineering, 174 (2019)*, 682-695. https://doi.org/10.1016/j.petrol.2018.11.076.

Wang, Z., Hou, J., Hao, H., Wang, C., & Wang, L. (2022). Using the Multiple Linear Regression Method for CO2 Flooding Evaluation in the Daqing Oilfield. *Frontiers Energy Research, 30 June.*, https://doi.org/10.3389/fen-rg.2022.929606.

Weller, Z., Hamburg, S., & von Fischer, J. (2020). A National Estimate of Methane Leakage from Pipeline Mains in Natural Gas Local Distribution Systems. *Environ. Sci.Technol. 54, 14, 8958-8967*, https://doi.org/10.1021/acs.est.0c00437.

White House. (2021, 12 8). *Fact Sheet: The Biden-Harris Administration is taking Action to Restore and Strengthen American Democracy*. Retrieved from https://www.whitehouse.gov/briefing-room/statements-releases/2021/12/08/fact-sheet-the-biden-harris-administration-is-taking-action-to-restore-and-strengthen-american-democracy/

Wikipedia. (2021, January 16). *Spurious Relationship*. Retrieved from https://en.wikipedia.org/wiki/Spurious_relationship

Wikipedia. (2023, 5 3). *Hirchens's razor*. Retrieved from https://en.wikipedia.org/wiki/Hitchens%27s_razor

Wilson, A. (2017). New Salt-Tolerant Friction-Reducer System Enables 100% Reuse of Produced Water. *J Pet Technol 69 (12): 76-80/SPE-1217-0076-JPT*, https://doi.org/10.2118?1217-0076-JPT.

Woodford, S. (2021, 2 3). *Rationality Rules*.

World Bank. (2023, 3 6). Retrieved from https://data.worldbank.org/indicator/NY.GDP.MKTP.CD?name_desc=false

World Health Organization. (2022, 11 28). *Household Air Pollution*. Retrieved from who.int: https://www.who.int/news-room/fact-sheets/detail/household-air-pollution-and-health

Wu, K., & Olson, J. (2015). Simultaneous Multifracture Treatments: Fully Coupled Fluid Flow and Fracture Mechanics for Horizontal Wells. *SPE J. 20, 337-346*, doi: https://doi.org/10.2118/167626-PA.

Yocham, K., Allison, D., & Schwartz, M. (2021). Case Study of a Wireline Deployable Spearhead Acid in the Denver-Julesburg Basin. *Paper presented at the Unconventional Resources Technology Conference, Houston, TexS, 26-28 July*, URTeC 5046.

Zoback, M. (2010). *Reservoir Geomechanics.* New York, New York: Cambridge University Press.