

EXPERIMENTAL AND COMPUTATIONAL ANALYSIS OF OIL FIELD OPERATIONS

by

Kamga L. Ngameni

Copyright by Kanga L. Ngameni 2021

All Rights Reserved

A thesis submitted to the Faculty and the Board of Trustees of the Colorado School of Mines in partial fulfillment of the requirements for the degree of Doctor of Philosophy (Operations Research with Engineering).

Golden, Colorado

Date _____

Signed: _____

Kamga L. Ngameni

Signed: _____

Dr. Alexandra Newman
Thesis Advisor

Signed: _____

Dr. Jennifer L. Miskimins
Thesis Advisor

Golden, Colorado

Date _____

Signed: _____

Dr. John Berger
Professor and Head
Department of Mechanical Engineering

ABSTRACT

This dissertation presents results of an experimental study that was done using a novel apparatus from the Halliburton Advanced Perforation Laboratory and a statistical analysis derived from those experimental studies. We examine both the subsurface component of unconventional reservoirs to improve fluid flow, and also the surface component of oilfield operations. Regarding the latter, we propose to optimize operations by reducing non-productive time and minimizing infrastructure and operational costs associated with drilling and fracturing.

Historically, Darcy's equation had been used to predict hydrocarbon production (Darcy 1856); however, it is inaccurate in unconventional reservoirs. This study was conducted in order to validate the Barree - Conway model (2004), which yields accurate flow predictions for all flow regimes in porous media, and examines output data such as pressures, temperature, and mass flow rate given a proppant pack design and specific fluid flow rate. Data were collected from various experimental studies conducted at high-flow rates through various proppant packs and using different experimental procedures. This study identifies the most important parameters associated with non-Darcy flow behavior. Statistical analysis of non-Darcy flow reveals that the most important predictors of mass flow rate are the particular material used, fluid velocity, fluid viscosity, and apparent permeability. We compare the results of the statistical analysis to previous work done in unconventional reservoir studies. These results show that statistical analysis can recommend which parameters to prioritize during the hydraulic fracturing design process. Although this work was conducted under experimentally controlled conditions, it is hypothesized that the results will hold for unconventional reservoirs. Following this experimental and predictive work, we introduce an optimization, or a prescriptive, model that improves efficiency and reduces costs for surface operations. The model considers drilling, fracturing, and production of 71 pads and a total of 729 wells, and identifies a schedule of operations to minimize idle time while adhering to precedence and resource

constraints. Using a specialized algorithm, we solve a large-scale model to inform not only a schedule that minimizes makespan but also one that can identify bottleneck resources.

CONTENTS

Abstract	iv
List of Figures	vii
List of Tables	ix
Chapter 1 INTRODUCTION	3
1.1 Research Contribution.....	5
1.2 Dissertation Layout.....	5
Chapter 2 EXPERIMENTAL STUDY ON NON-DARCY FLOW IN A PROPPANT PACK USING NITROGEN GAS UNDER HIGH PRESSURE.....	6
2.1 Introduction.....	7
2.2 Minimum Permeability Plateau.....	9
2.3 Experimental Design of the Laboratory Apparatus	10
2.4 System Design	10
2.5 Methodology	13
2.6 Experimental Results.....	15
2.7 Discussion	16
2.8 Conclusions and Future Work	19
2.9 Nomenclature	21
Chapter 3 STATISTICAL ANALYSIS OF EXPERIMENTAL STUDIES OF NON- DARCY FLOW IN PROPPANT PACKS	22
3.1 Introduction.....	23
3.2 Darcy vs. Non-Darcy flow.....	23
3.3 Past Experimental Studies	25
3.4 Colorado School of Mines Experimental Apparatus.....	26
3.5 Halliburton Apparatus	28
3.6 Collection of Statistical Data	30
3.7 Review of Statistical Analyses in Unconventional Reservoirs.....	30
3.8 Assumptions and Definitions.....	31

3.9 Results and Discussion	33
3.10 LASSO Analysis:	33
3.11 Conclusions.....	39
3.12 Nomenclature	41
Chapter 4 REDUCING NON-PRODUCTIVE TIME IN OIL FIELDS	42
4.1 Introduction.....	42
4.2 Literature Review.....	45
4.3 Activities in Oilfield Operations	46
4.4 Current Scheduling and Proposed Optimization Model.....	49
4.5 Results.....	53
4.6 Conclusions.....	57
Chapter 5 CONCLUSION	58
Bibliography	60
Appendix A SUPPLEMENTAL DATA FOR CHAPTER 2	64
Appendix B R CODE.....	67
Appendix C SUPPLEMENTAL FORMULATION FOR CHAPTER 4.....	78
Appendix D COPYRIGHT IMAGES.....	80

LIST OF FIGURES

Figure 2.1	Validation of non-Darcy flow equations for Ottawa 20/40 sand under 3000 psi confining stress. From Lopez-Hernandez (2007).....	10
Figure 2.2	Dimensionless plot with a wide range of Reynolds numbers with different proppant tested using the B & C model. From Lopez-Hernandez (2007) ...	11
Figure 2.3	Proppant setup	13
Figure 2.4	Actual setup of the proppant with 4.4 in. length (12 cm) with 3/2-in. inside diameter.....	14
Figure 2.5	New fixture for the proppant pack: proppant is inside a 3/8-in. (0.9525 cm) inside diameter	15
Figure 2.6	(a) Apparent permeability and mass flow rate as a function of the differential pressure and (b) The mass flow rate and axial pressure drop as a function of time. Experimental runs generate 3,480 psi differential pressure.....	17
Figure 2.7	(a) Forchheimer plot showing the value of the pseudo-Reynolds number, and (b) Mass flow rate and apparent permeability as a function of the differential pressure. The red circles represent a stagnation of the data during experiments.	18
Figure 2.8	Vena Contracta behavior as a function of the Venturi effect	19
Figure 3.1	Validation of non-Darcy flow equations for Ottawa 20/40 sand under 3000 psi confining stress. From Lopez-Hernandez (2007).....	26
Figure 3.2	CSM experimental set up	27
Figure 3.3	Proppant pack showing different pressure measurement points used at CSM	28
Figure 3.4	Final design of the proppant pack showing the initial rubber sleeves and the inseted material used in order to reduce the size of the proppant pack.	29
Figure 4.1	An oil and gas sedimentary basin	43
Figure 4.2	Unconsolidated, or hydrocarbon, formation which contains rock grains, and drilling casing pipe that enters the formation in an horizontal manner, (Chen et al., 2019).....	44

Figure 4.3	Oil and gas separation to the midstream (Lea Jr and Rowlan, 2019)	44
Figure 4.4	Typical onshore drilling equipment layout (Courtesy of International Snubbing Services)	47
Figure 4.5	Example of a well being drilled horizontally (Chen et al., 2019)	48
Figure 4.6	Example of a fracturing crew at an oil field completion site(Courtesy of Liberty Oilfield Services).....	49
Figure 4.7	Example of flowback pipe	49
Figure 4.8	Geographic locations of the pads and wells	51
Figure 4.9	Utilization for each of the three resource types over the time horizon employing two rigs.....	54
Figure 4.10	Utilization for each of the three resource types over the time horizon employing three rigs	55
Figure 4.11	Cumulative pad production over the time horizon using two rigs.....	55
Figure 4.12	Cumulative pad production over the time horizon using three rigs	56
Figure A.1	Figure showing the highest differential pressure as a function of the mass flow rate	64
Figure A.2	Figure showing the highest differential pressure as a function of the mass flow rate	64
Figure A.3	Figure showing the Pseudo-Reynolds number as a function of the inverse permeability.....	65
Figure D.1	Approval email from Henry Lopez for Figures 2.1, 2.2, 3.1, 3.2, and 3.3 ...	81
Figure D.2	Approval email from Sally Benson for Figure 4.1	81
Figure D.3	Approval email for Figure 4.2.....	82
Figure D.4	Approval email from Rightlink for Figure 4.3	82
Figure D.5	Approval email from rightlink for Figure 4.5	83
Figure D.6	Approval email from Liberty Oilfield for Figure 4.4, 4.6, and 4.7	84

LIST OF TABLES

Table 3.1	LASSO regression values	35
Table 3.2	LASSO regression values showing in bold the most significant variables.	35
Table 3.3	This table shows the most significant variables from the random forest methodology; these match those determined by LASSO.....	37
Table 3.4	This table shows the characteristics of all data points available from the experimental studies in R.	37
Table 3.5	This table reflects the characteristics of all numerical values, as well as variables used in the analysis.	37
Table 3.6	This table shows the initial run using the linear regression model with some NA's associated with variables that need to be removed	38
Table 3.7	Table showing the significant code from the variables of the experimental study	38
Table 3.8	Table showing the most significant variables in LASSO	39
Table 4.1	The Production schedule contains various activities	51
Table 4.2	Resource constraints in our production scheduling model	52
Table 4.3	Input file types and their descriptions	53

ACKNOWLEDGEMENT

I want to extend my sincerest gratitude to my advisors: Pr. Alexandra Newman and Pr. Jennifer Miskimins, for the guidance and support throughout my time at the Colorado School of Mines. Pr. Newman, I am so grateful for you, thank you for giving me this chance and for never giving up on me; it has been the most difficult but also the most rewarding process. Pr. Miskimins - I know it has been challenging and I would like to thank you for following me all the way to the end. We started in the Master program together and now, here we are! I have no words to thank you for trusting me from that first hydraulic fracturing class. You both have been everything a student could wish for, you never hesitated to provide the resources I needed in order to complete my research work and write my dissertation.

I also want to thank my committee members, Pr. Soutir Bandyopadhyay, Pr. Tulay Flamand, and Pr. Derrick Hudson for your valuable feedback and for taking the time to correct my papers and dissertation. Thank you for stepping in when needed. I appreciate you all!

I am very grateful for the FAST (Fracturing Acidizing, and Stimulation Technologies) consortium at the Colorado School of Mines. Pr. Miskimins and the member companies have helped me since the start of my Master degree.

I am grateful for the ORwE (Operations Research with Engineering) graduate program at the Colorado School of Mines, directed by Pr. Newman. I am thankful to be part of this group and for the family or friends and colleague I found in that group. Thank you for your help and encouragements even in times where I despaired the most.

I want to extend many thanks to Amanda Rebol, Dr. Akshay, and Pr. Marcos for accepting to step in when needed. You guys made this possible.

Finally, I want to thank you maman, papa, Dorine, Hermine and Julien Kamgang, and all my relatives and friends who have helped me accomplish this and many other goals in my life.

DEDICATION

To God: for the gift of life, for giving me the opportunity to cross path with so wonderful people in and out of the Colorado School of Mines, and for giving me the strength to keep moving forward no matter how hard it got.

To my mother, Magne Bernadette: for devoting her life to my education, for her daily rosary prayers, and for never giving up on me no matter how difficult I might have been.

A ma mere, Magne Bernadette: Merci pour avoir consacree ta vie a mon education, merci pour tes prieres, et surtout merci pour avoir cru en moi meme quand je ne le faisais pas.

Alla mia amata Dorine, grazie di essere sempre stata al mio fianco, con forza, dignità ed onore. Di aver sempre creduto in me, e cosa più importante grazie per la luce e la motivazione data alle nostre vite: Zahra

To the family Kamgang: Thank you for the support, the continuous encouragement, for believing that there is light at the end of the tunnel.

To Binam of Colorado and K3T: Thank you as you have set the bar high for all. This helped me challenged myself while learning from all of you. You are both amazing and always seeking excellence in your members.

To my Dad, maman Christine, my brothers, and sisters for your encouragement along this challenging road.

CHAPTER 1

INTRODUCTION

Hydraulic fracturing is the most commonly used technique applied to unconventional reservoirs for oil recovery. As of 2016, nearly 70% of wells completed in the United States were hydraulically fractured (EIA, 2018). Hydraulic fracturing consists of injecting fluid at high pressure into the reservoir and keeping the rock open by inserting proppant. However, historically industry has failed to consistently implement procedures based on sound engineering knowledge and analysis. Specifically, it is critical to consider the following:

1. Hydraulic fracturing operations must use adequate proppant.
2. Fluid flow in the corresponding reservoirs does not subscribe to a Darcy regime because of the Reynolds number.
3. Because hydraulic fracturing is expensive, both the crews and the equipment need to be utilized efficiently in order to reduce non-productive time and cost.

Hydraulic fracturing, despite its challenges, has been very successful in the United States, providing the country with energy independence (EIA, 2018) to the point of surpassing Saudi Arabia in global oil production in 2018. Despite past recessions, the US has always been able to find a way to produce oil from unconventional reservoirs cheaply and maintain pace with demand. The 2020 COVID pandemic has reduced global demand for oil to the point of negative oil prices in May 2020. Specifically, Western Texas Intermediate contract prices traded at -37 USD for May delivery (Bloomberg Finance LP, April 2020). Overcoming turbulent economic times will require extensive engineering analysis in unconventional reservoirs, although the US is expected to resume its pre-COVID activities, and energy consumption is expected to return to pre-COVID levels. It is crucial to be more proactive in unconventional field development and operation.

Unconventional reservoirs are known to have a higher flow rate after fracturing occurs, relative to the flow in conventional reservoirs. In this case, the Darcy equation is no longer valid (Lopez-Hernandez, 2007; Lai, 2010). Rather, the Barree - Conway model (2004)

provides an improvement. However, the minimum permeability, as described in the Barree - Conway model, has not been entirely validated. Part of this dissertation consists of validating the minimum permeability plateau; this work modifies the apparatus in the Halliburton Advanced Perforation Flow Laboratory (APFL) in Alvarado, TX. The corresponding experiments produced output that a statistical analysis uses to determine the most significant variables for predicting fluid flow. These findings help prioritize the importance and associated costs of hydraulic fracturing operations. Therefore, the second paper focuses on a statistical analysis of the production mass flow rate of nitrogen gas in a proppant pack using various proppant types while changing pressures and temperatures. By using existing data, in combination with novel experimental results, the analysis builds on extant studies of significant variables for the production of hydrocarbon reservoirs. Specifically, we determine that the material used, the apparent and pack permeability, and the fluid viscosity are the most significant variables when predicting mass flow rate. While subsurface applications inform better engineering analysis, it is important to examine surface operations as well, as they are costly owing to the labor associated with drilling and fracturing. An optimization model reduces non-productive time and schedules activities in a cost-effective manner.

The objectives of this work are to present the most recent experimental studies validating the Barree-Conway (2004) model, to perform statistical analysis related to fluid flow, and to optimize surface operations. The ultimate goal is to present a unifying, cost-saving approach to unconventional oil and gas operations.

The following steps accomplish our objectives:

1. Review the relevant literature pertaining to the Barree-Conway (2004) model and adapt the experimental apparatus of the Halliburton Advanced Perforation Laboratory to the needs of our study.
2. Acquire experimental data to justify the minimum permeability plateau as described by the Barree - Conway model.

3. Perform a statistical analysis on the obtained data in order to find the most significant variables to predict the mass flow rate and its effect in an unconventional reservoir.
4. Optimize surface activities in order to reduce non-productive time and costs during oilfield operations.

1.1 Research Contribution

This research started with the understanding of fluid flow in unconventional reservoirs. The Barree-Conway (2004) model has been proven to be more accurate than Darcy's (1856) for unconventional reservoirs in which fluid flow rates are higher due to the Reynolds number. Second, statistical analysis provides a better understanding of significant variables related to mass flow in unconventional reservoirs. Third, we reduce non-productive time in surface operations by using an optimization model to design and schedule production.

1.2 Dissertation Layout

This dissertation is organized into the following chapters. Chapter 1 introduces the research objectives and contribution. Chapter 2 consists of a paper that will be submitted to the *Journal of Petroleum Science and Engineering*, which discusses experimental work on the Barree-Conway model and the results obtained from those experiments. Chapter 3 consists of statistical analysis that determines significant variables affecting the mass flow rate under experimental conditions. That paper is submitted to the *Journal of Petroleum Science and Engineering*. Chapter 4 includes the optimization model as it pertains to surface operations. This paper will be submitted to the *INFORMS Journal of Applied Analytics*. Chapter 5 concludes and presents future research.

CHAPTER 2

EXPERIMENTAL STUDY ON NON-DARCY FLOW IN A PROPPANT PACK USING NITROGEN GAS UNDER HIGH PRESSURE

This paper will be submitted for publication in *Journal of Natural Gas and Engineering*

Kamga L Ngameni¹, Jennifer Miskimins², Jacob McGregor³

Abstract

This paper presents the results of the latest experimental studies conducted at high flow rates through a proppant pack in order to validate the minimum permeability plateau as described by the Barree and Conway (2004) model. The experimental studies were conducted at Halliburton's Advanced Perforating Flow Laboratory at the Jet Research Center located in Alvarado, TX, in a proppant pack under high pressure with a confining stress to generate a differential pressure large enough for validation purposes. The results of these experiments, although they do not entirely verify the hypothesis of the minimum permeability plateau, show a pseudo-Reynolds number improvement from

$$1,440 \frac{g}{cm^2 - s - cp} \text{ to } 4,030 \frac{g}{cm^2 - s - cp}.$$

¹Operations Research with Engineering Graduate Program, Colorado School of Mines, Golden, CO 80401

²Petroleum Engineering Department, Colorado School of Mines, Golden, CO 80401

³Halliburton Jet Research Center, Alvarado, TX 76009

2.1 Introduction

Historically, Darcy's law has been used to describe fluid flow in porous media (Darcy, 1856b). Darcy's law shows a linear relationship between Darcy velocity (Muskat, 1946), and fluid potential, as shown in Equation 2.1:

$$-\frac{\partial P}{\partial L} = \frac{\mu v}{k_d} \quad (2.1)$$

where

$$\frac{\partial P}{\partial L} = \text{potential flow gradient, } \frac{atm}{cm}$$

$$\mu = \text{viscosity, cp, } \frac{g}{100cm - sec} \left[\frac{Mass}{Length - Time} \right]$$

$$v = \text{superficial velocity, } \frac{cm}{sec}, \left[\frac{Length}{Time} \right]$$

$$k_d = \text{permeability, Darcies, } [Length^2]$$

Deviation from this linear relationship is known as non-Darcy flow (Forchheimer, 1901). During the last century, there has been strong evidence that non-Darcy flow occurs in many subsurface systems, including oil or gas wells, and water or gas injection wells (Al-Otaibi et al., 2010). Non-Darcy flow significantly affects the productivity of gas wells and high-rate oil wells (Miskimins et al., 2005). In studies of non-Darcy flow, Forchheimer (1901) first observed that, at high flow velocities, the relationship between pressure gradient and fluid velocity is no longer linear. In an attempt to describe this nonlinear relationship, Forchheimer added a quadratic term (Forchheimer, 1901), as seen in Equation 2.2, and a cubic term (Equation 2.3) to Darcy's linear form. Many authors have observed the limitations of the Forchheimer equations, as they do not describe the entire range of potential fluid velocities in producing-well situations. In 2004, Barree and Conway (B&C) used apparent permeability in their fluid flow equation, which takes into account a

minimum permeability (Equation 2.4) and is able to cover all ranges of Reynolds numbers, defined as the ratio of inertial forces to viscous forces. Specifically, this equation is not only valid for low flow rates, but also for the highest flow rates displayed under field conditions. Therefore, in this study, the non-Darcy flow equation (Equation 2.5), as described in the B&C model, is used.

$$-\frac{\partial P}{\partial L} = \frac{\mu v}{k_d} + \beta \rho v^2 \quad (2.2)$$

$$-\frac{\partial P}{\partial L} = \frac{\mu v}{k_d} + \beta \rho v^2 + \gamma \rho v^3 \quad (2.3)$$

where,

$$\beta = \text{inertial flow parameter, } \frac{atm - 100sec^2}{g}$$

$$\rho = \text{fluid density, } \frac{g}{cm^3}, \frac{Mass}{Length^3}$$

$$\gamma = \text{factor of the Forchheimer cubic equation}$$

Considering the wide range of Reynolds numbers encountered in fluid flow through porous media, the B&C model (Equation 2.4) suggests that the concept of an apparent permeability (k_{app}) to describe Darcy (linear) and non-Darcy (nonlinear) flow in porous media can be employed in a single equation (Barree and Conway, 2004). Of all the equations tested, the B & C model is the only one able to satisfy all ranges of the Reynolds numbers achieved in laboratory studies (Lopez-Hernandez, 2007; Lai et al., 2010). Figure 2.1 contrasts the experimental behavior with the theoretical behavior of Forchheimer quadratic, Forchheimer cubic and B & C equations.

$$k_{app} = k_{min} + \frac{(k_d - k_{min})}{(1 - R_e^F)^E} \quad (2.4)$$

$$-\frac{\partial P}{\partial L} = \frac{\mu v}{k_{app}} \quad (2.5)$$

$$R_e = \frac{\partial v}{\mu T} \quad (2.6)$$

where,

k_{app} = apparent-rate dependent permeability, Darcies

k_d = constant Darcy permeability, Darcies, $\frac{cm - g}{100sec^2 - atm}$

k_{min} = minimum permeability at high rate, Darcies

R_e = Reynolds number, dimensionless

T = characteristic length, 100/cm, $\left[\frac{1}{L}\right]$

F, E = exponents, dimensionless

Despite the fact that the B&C model closely matches the experimental data for all ranges of tested Reynolds numbers, the plateau behavior at high Reynolds numbers is a hypothesis that has been only partly validated by Lopez-Hernandez (2007), Lai et al. (2010), and Aljalahmah (2014).

Figure 2.1 shows that the plot of the Forchheimer quadratic equation deviates upward from the empirical data, while the Forchheimer cubic equation exhibits concave down behavior relative to the empirical data. The only plot that closely matches all empirical data is that given by the B & C model (shown in red), which is why it is tested in the current experimental study.

2.2 Minimum Permeability Plateau

The minimum permeability plateau can be defined as the minimum value the apparent permeability reaches when flowing through small-pore space at extremely high velocities, or Reynolds numbers. Several experimental studies have attempted to reach the minimum permeability plateau (Aljalahmah, 2014). Unfortunately, past experiments (Lopez-Hernandez, 2007; Lai et al., 2010; Aljalahmah, 2014) were not conducted at Reynolds numbers high enough to justify the use of the minimum permeability plateau in

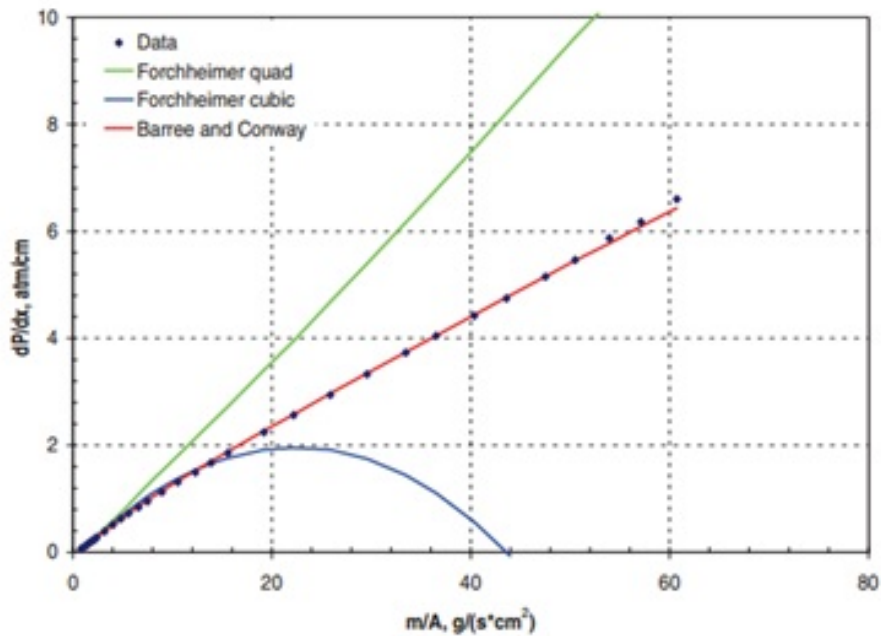


Figure 2.1: Validation of non-Darcy flow equations for Ottawa 20/40 sand under 3000 psi confining stress. From Lopez-Hernandez (2007)

the B & C equation, which is a function of the apparent permeability, where the apparent permeability is a function of the minimum permeability. Figure 2.2 shows a series of experiments that were conducted at the Colorado School of Mines, in which the minimum permeability was not fully reached due to limitations of the experimental setup. However, based on the recommendations made by the latest study, a modified commercial system, located at Halliburton’s Advanced Perforating Flow Laboratory (APFL), allowed for an increase in differential pressure.

2.3 Experimental Design of the Laboratory Apparatus

The experimental apparatus at the Advanced Perforation Flow Laboratory was used in an attempt to reach the hypothesized permeability plateau.

2.4 System Design

To maximize the effectiveness of the experimental setup, it is important to ensure that the parameters are adaptable to the setup, allowing the permeability to reach its lowest point.

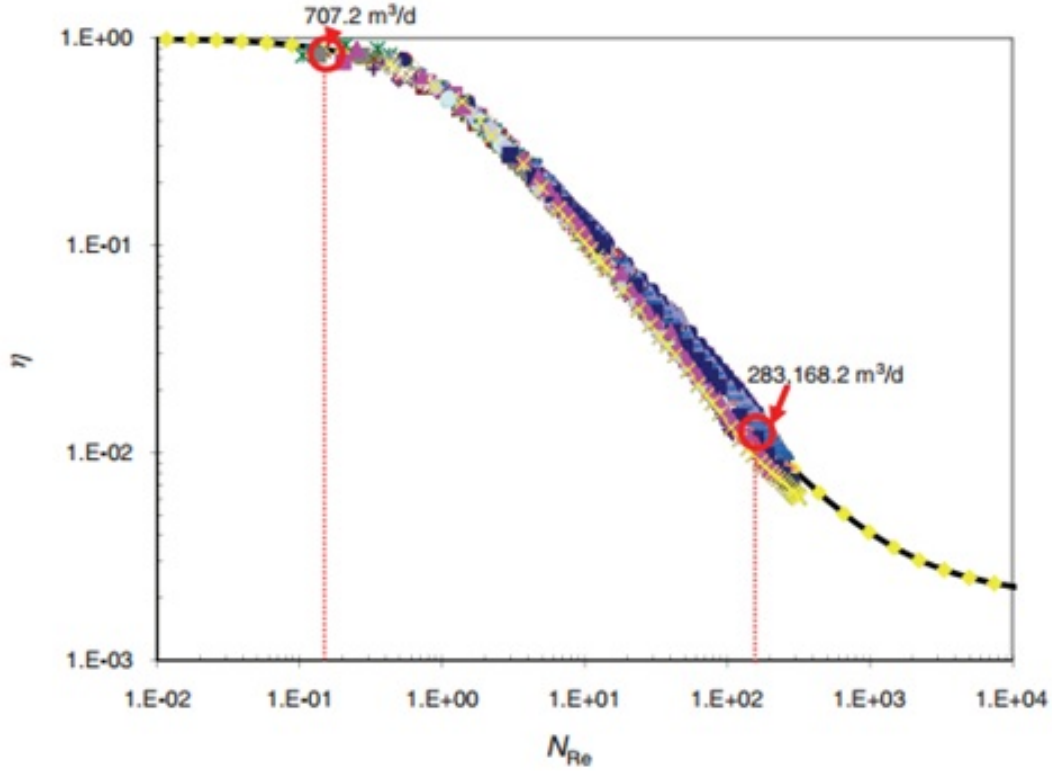


Figure 2.2: Dimensionless plot with a wide range of Reynolds numbers with different proppant tested using the B & C model. From Lopez-Hernandez (2007)

a. Proppant Selection

Aljalahmah (2014) used resin-coated sand, which serves two primary functions:

1. It improves the effective strength of the proppant; and,
2. It prevents the generation and migration of fines into the wellbore.

Based on the actual market, the demands for fracturing sand, resin-coated sand, and ceramic proppants are not identical (Beckwith et al., 2011). Basic sand is in much higher demand and is used more frequently than resin-coated or ceramic sand. Thus, the choice of proppant in this experimental setup does not depend on market demand or the economics.

The grains of these ceramic proppants are artificially manufactured to be rounder (more spherical) and to have a higher resistance to crushing forces than

natural sands tend to display. During an experimental run, Lopez-Hernandez (2007) was able to test various sets of proppants. Therefore, this study does not repeat such experiments.

In this work, a ceramic proppant is used, as that is the strongest proppant the laboratory presently possesses. The ultimate goal of this experiment is to reach the minimum permeability plateau without crushing the proppant.

b. Fluid selection

The choice of fluid is primarily based on the fluid's ability to reach a high Reynolds number. This simple definition explains why a low-viscosity fluid is preferred in order to obtain higher Reynolds numbers. Typically, gases have been used in flow-through porous media because a high Reynolds number can be reached much more quickly than when a liquid is used. Other factors that influence the choice of fluid are:

- High volumes of gas consumption;
- Availability of the fluid in the Jet Research Center;
- Safety regulations; and,
- Cost of modification of the unit.

The laboratory where the experiments are conducted has been using nitrogen gas which was the choice for these experiments as well.

c. Proppant Pack Design

The proppant is confined in a proppant pack with dimensions 15 inches in length by one inch in diameter. In order to avoid end effects, which are due to the pressure in relation to the length of the apparatus, Lopez-Hernandez (2007) used a proppant pack with a length greater than seven inches. A confining pressure up to 10,000 psi is then applied. Thereafter, the flowing differential pressure along the proppant pack can be measured.

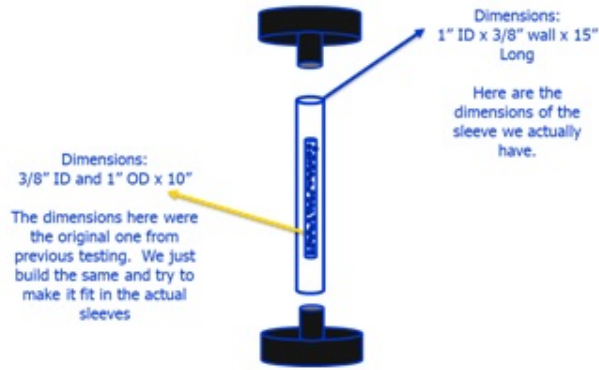


Figure 2.3: Proppant setup

The type of proppant pack selected for this experimental study was based on the design of the apparatus, and on calculations made by Lopez-Hernandez (2007), done in conjunction with flow-rate calculations and the performance of the laboratory setup already in place. The proppant pack design is dependent on the magnitude of the Reynolds numbers that must be reached during the experiments.

The first experiments demonstrate the need for a decrease in diameter and length of the proppant pack, because the pressure drop during those experiments was not able to capture a differential pressure between the inlet and outlet valve of the experimental setup.

The inner diameter of the bladder was constrained to be one inch. Therefore, in order to attain a 3/8-inch inside diameter and a 10 inches in length, the setup shown in Figure 2.3 was used.

2.5 Methodology

This section discusses the methodology used to adapt the experimental setup at the Jet Research Center and the selection of the parameters to ensure that they are compatible with the setup and the physics.

The approach used in the experimental runs was based on previous work and on trial-and-error of the new experimental apparatus. The methodology includes the following:

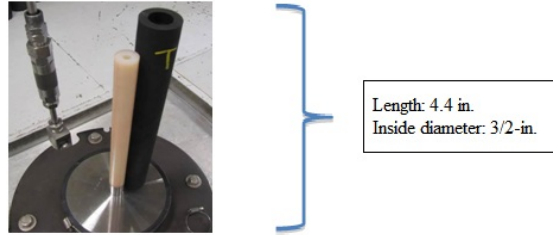


Figure 2.4: Actual setup of the proppant with 4.4 in. length (12 cm) with 3/2-in. inside diameter

The components used for laboratory testing are as follows:

- Prepared proppant pack;
- Inclusion of the proppant in the pack;
- Seal of the proppant pack under specific conditions;
- Flow of the nitrogen through the proppant pack while increasing the stress and differential pressure; and,
- Recording of the pressure before and after runs.

The proppant is tested at a mass flow rate of 4.8g/s and this allowed for the generation of a differential pressure of 3,250 psi as shown in Figure 6a.

- Confection of the Proppant Pack

The initial proppant pack, whose outer layer was made from a Tygon™ material with a diameter of 3/2-inch and a length of 4.4 inches (Figure 2.4), was not able to generate a high-pressure drop. The initial results indicated the need to reduce the diameter and the length of the proppant pack to the following dimensions: 3/8-inches for the inside diameter and 4.4 inches length (Figure 2.5). Plastic material was used to reduce the inside diameter to 3/8-inch; this secondary material was inserted into the proppant pack, as seen in Figure 2.6, and generated a differential pressure of 3,780 psi.

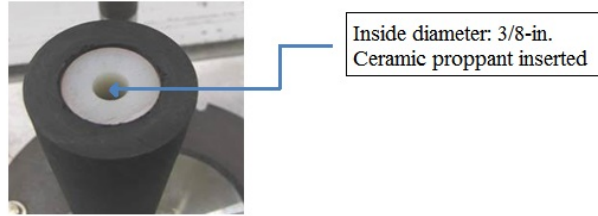


Figure 2.5: New fixture for the proppant pack: proppant is inside a 3/8-in. (0.9525 cm) inside diameter

- Inclusion of proppant and sealing the proppant pack under specific conditions

The ceramic proppant, which is strong enough to resist high pressure, was inserted into the new fixture of the proppant pack, as shown in Figure 2.5. To prevent the proppant from flowing out of the pack and into the downstream flow loop piping, the pack was sealed with 100-mesh paper. Flowing the nitrogen through the proppant pack, while increasing the stress and differential pressure (to a maximum of 10,000 psi) ensures robustness of the experimental results. Data regarding pressure, mass flow rate, and permeability were collected every 5 seconds over a time horizon of 55 minutes.

- Limitation of the apparatus

The apparatus cannot exceed 10,000 psi differential pressure and is unable to accommodate pressure transducers along its length. Experimental runs were conducted with pressure transducers at the top and bottom of the apparatus. This is a limitation because, unlike in previous experimental runs, we had pressure transducers across the proppant pack.

2.6 Experimental Results

The results of the experiment (seen below) showed that we reached a differential pressure of 3,480 psi, which is considerably higher than the 2,680 psi achieved in previous experimental runs, yielding results that are closer to, although still not at, the minimum permeability plateau. All figures derive from different experimental runs. Figure 6a plots the mass flow rate on a log-log scale and Figure 6b axial pressure drop. The red plot of the

apparent permeability on the secondary axis shows improvement from previous work. The pseudo-Reynolds number used to determine the how fast the gas nitrogen flows in the proppant pack was increased from

$1,440 \frac{g}{cm^2 - s - cp}$ to $4,030 \frac{g}{cm^2 - s - cp}$. This is an improvement from previous results. However, obtaining a higher differential pressure might allow confirmation and validation of the experiments. Figure 6b shows on the x-axis the elapsed time of the experimental runs, the primary y-axis provides the mass flow rate, and the secondary y-axis depicts pressure drop. Although there is an increase in the differential pressure that is matched by an increase in mass flow rate, the differential pressure increases to 3,450 psi.

Figure 2.6a shows the Forchheimer plot and a maximum pseudo-Reynolds number of $4,030 \frac{g}{cm^2 - s - cp}$. Figure 2.6b shows an apparent permeability as a function of the differential pressure. It also seems that the experimental setup precludes obtaining data above a certain range. The results in Figure 2.6a suggest that the fluid flow is being choked near a pseudo-Reynolds number of $4030 \frac{g}{cm^2 - s - cp}$

2.7 Discussion

This section discusses the proposed reason precluding the experimental setup from reaching higher differential pressures. It is believed that “choke flow” affects the potential to reach high pressures.

Choke flow is associated with the venturi effect and is described as the flow slowing when gas or fluid is passing through a constriction Gerhart et al. (2020). It is often encountered when there is a change in the flow path cross-sectional area. As a compressible fluid passes through a restriction, changes occur in both velocity and pressures. The fluid might start upstream, at higher pressure, which decreases as the fluid gains velocity through the restriction. After passing the *vena contracta* (Figure 2.8), the fluid begins to fill the cross-sectional area of the pipe and, as it does, it slows down and regains pressure as the Mach number reaches a value of 1. The pressure change can no longer communicate and plateaus at a given value.

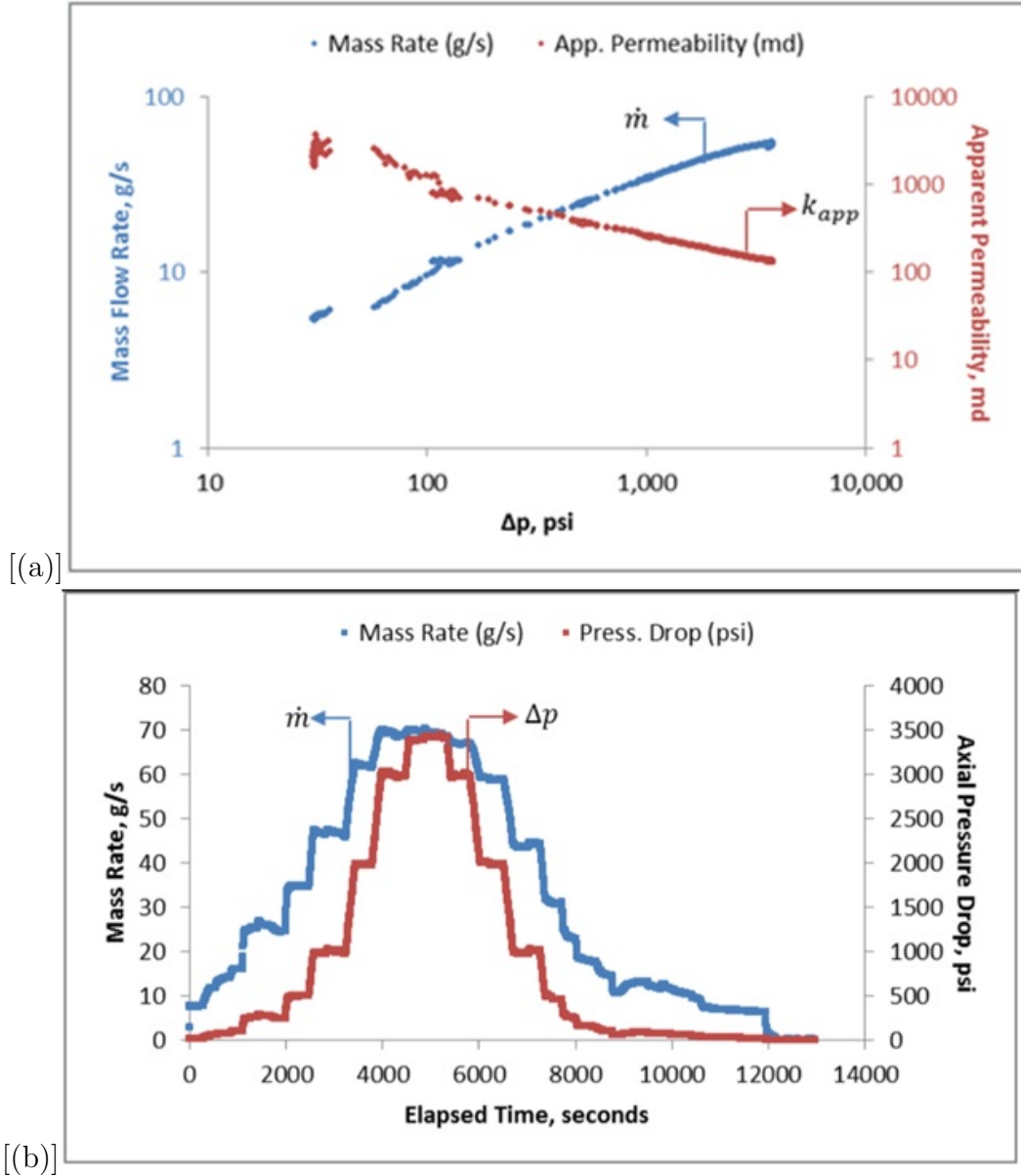


Figure 2.6: (a) Apparent permeability and mass flow rate as a function of the differential pressure and (b) The mass flow rate and axial pressure drop as a function of time. Experimental runs generate 3,480 psi differential pressure.

The described phenomenon is similar to the minimum permeability plateau described by the B&C model. It states, “in normal porous media flow, the fluid must accelerate through the relatively narrow pore throats and can decelerate as it enters the larger pore bodies. The repeated changes in velocity give rise to head losses associated with inertial effects. The inertial losses are responsible for the decrease in apparent permeability

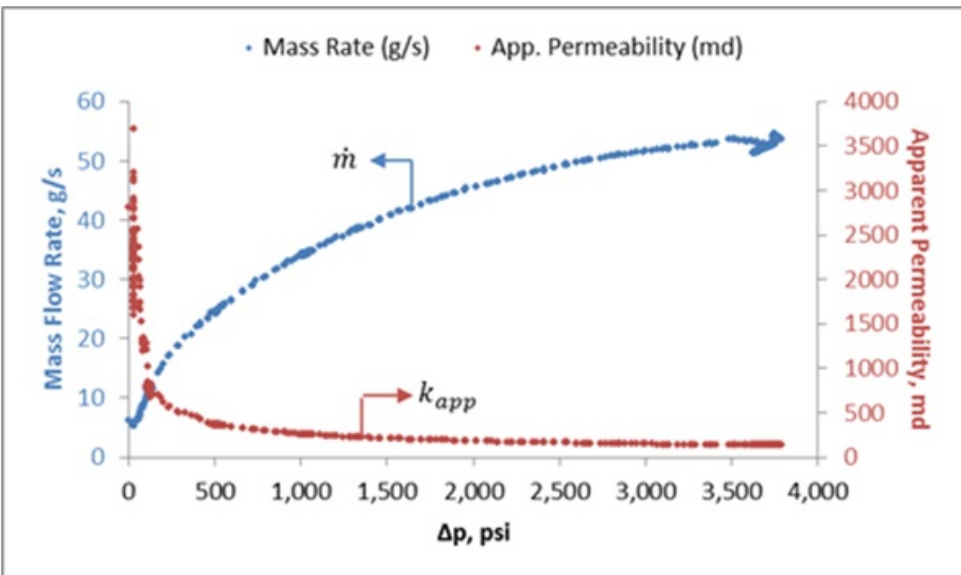
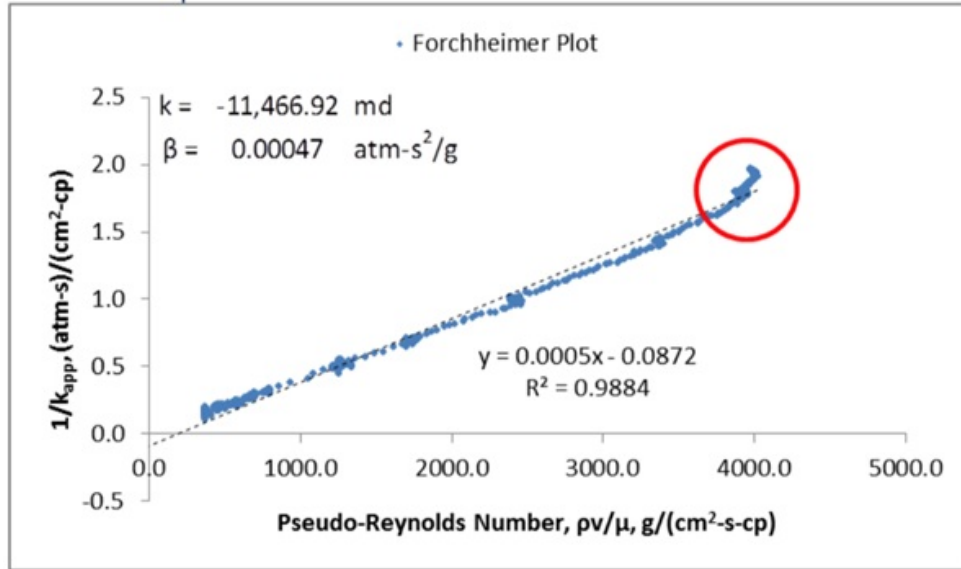


Figure 2.7: (a) Forchheimer plot showing the value of the pseudo-Reynolds number, and (b) Mass flow rate and apparent permeability as a function of the differential pressure. The red circles represent a stagnation of the data during experiments.

associated with non-Darcy flow” (Lopez 2004).

When the fluid flows through a restricted area, it gets choked. In our setup, the pores between the proppant in the pack represent that restricted area, yielding a strong possibility of the nitrogen gas being choked in the pack under high pressure. After running the experiments for several minutes, as the mass flow rate of nitrogen gas increased, the

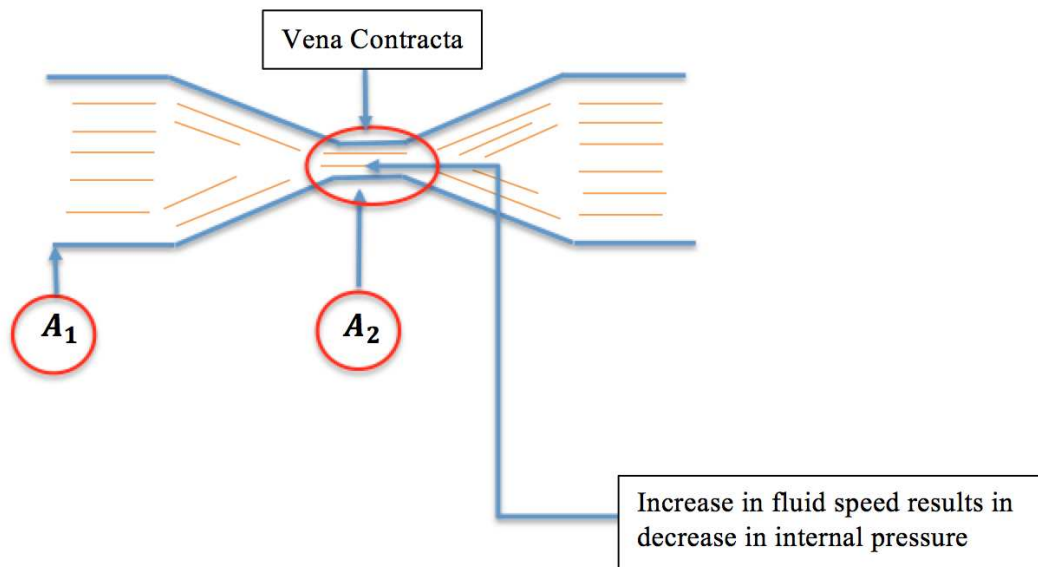


Figure 2.8: Vena Contracta behavior as a function of the Venturi effect

observed stagnant data might represent the choke flow. The *vena contracta* shows how the flow might be contracted in certain areas and expanded in others.

2.8 Conclusions and Future Work

Non-Darcy flow is an important physical concept in reservoir types such as:

- Gas production wells for which local velocities can be high;
- High-rate oil wells; and,
- Hydraulically fractured wells.

A deviation with respect to the Darcy and Forchheimer equations has been observed in the fluid flow at high Reynolds numbers. The deviation appears to be based on the existence of a minimum permeability plateau. The mechanism responsible for the minimum permeability plateau has been studied at the Colorado School of Mines for the last 10 years and this research confirms the direction taken from previously cited experiments.

As of the time of this writing, the Barree and Conway model is the only one that has been proven to be valid for a wide range of Reynolds numbers (Aljalalmah, 2014) and therefore most applicable to our industry. However, it has not been fully implemented in

most simulation software in the oil and gas industry, stemming from the uncertainty of being able to reach the minimum permeability plateau, as described in the Barree and Conway model.

The experiments performed in the Halliburton laboratory have achieved the highest possible differential pressure of 3,400 psi and the lowest minimum permeability to date. Physics and the experimental setup preclude the setup from reaching higher differential pressures. Finding a method to bypass the choke flow will improve the experimental results.

2.9 Nomenclature

Symbol	Description
A	: area open to flow, cm^2
CSA	: cross-sectional area, cm^2
d	: diameter, cm
$\frac{\partial P}{\partial L}$: pressure gradient, $\frac{atm}{cm}$
ECS	: effective confining stress, psi
E	: coefficient in the BC non-Darcy flow model
F	: coefficient in the BC non-Darcy flow model
k_{app}	: apparent permeability, <i>Darcy</i>
k_d	: Darcy permeability, <i>Darcy</i>
k_{min}	: minimum permeability, <i>Darcy</i>
L	: length, <i>inches</i> or cm
Re	: Reynolds number
μ	: viscosity, <i>cp</i>
β	: inertial flow parameter, $\frac{atm - s^2}{gr}$
ρ	: fluid density $\frac{gr}{cm^3}$
v	: superficial velocity, $\frac{cm}{sec}$
T	: characteristic length, $\frac{100}{cm}$

CHAPTER 3
STATISTICAL ANALYSIS OF EXPERIMENTAL STUDIES OF NON-DARCY FLOW
IN PROPPANT PACKS

This paper has been submitted for publication to the *Journal of Petroleum Science and Engineering*

Kamga L Ngameni¹, Jennifer L Miskimins², Soutir Bandyopadhyay³

Abstract

The objective of this study is to determine the most important variables associated with non-Darcy flow behavior. Statistical analysis of non-Darcy flow reveals that the most important predictors of mass flow rate are the particular material used, fluid velocity, fluid viscosity, and apparent permeability. Data were collected from various experimental studies conducted at high-flow rates through different proppant packs and using different experimental procedures. The study was conducted in order to validate the Barree and Conway (2004), which yields accurate flow predictions across all flow regimes in porous media. This study examines output data such as pressures, temperature, and mass flow rate given a proppant pack design and specific flowing fluid. Historically, Darcy's equation (Darcy, 1856a) has been used to predict hydrocarbon production; however, it is inaccurate in most reservoirs. Using results based on non-Darcy flow experimentation provides a different perspective to the hydraulic fracturing process.

The results of the statistical analysis were compared to previous work done in unconventional reservoir studies. These results show that statistical analysis can recommend the variables to prioritize during the hydraulic fracturing design process. Although this work was conducted under experimentally controlled conditions, it is hypothesized that the results will hold for field conditions.

¹Operations Research with Engineering Graduate Program, Colorado School of Mines, 1500 Illinois Street, Golden, CO 80401

²Petroleum Engineering Department, Colorado School of Mines, 1500 Illinois Street, Golden, CO 80401

³Applied Mathematics and Statistics Department, Colorado School of Mines, Golden, CO 80401

3.1 Introduction

This paper focuses on a statistical analysis of the production mass flow rate of nitrogen gas in a proppant pack using various proppant types while changing pressures and temperatures. By using existing data, in combination with novel experimental results, the analysis builds on extant studies of significant variables for the production of hydrocarbon in reservoirs.

Well operators seek to increase the mass flow rate of hydrocarbon production. This study consists of statistical analysis of experimental work that was done in the Halliburton Advanced Perforation Flow Laboratory (APFL) and at the Colorado School of Mines (CSM), which experimented with variables such as pressure, temperature, and mass flow rate using different types of proppants. The use of propped hydraulic fracturing is a common method to enhance oil production flow rates (Mohammed et al., 2017). Thus, the importance of variables associated with the proppant type to the hydraulic fracturing process should be investigated.

In practice, for processes involving hydraulic fractures, (Barree and Conway, 2004) (hereafter, referred to as the B&C model) show that the velocity at which the flow propagates is non-Darcy. Therefore, this study uses data from non-Darcy flow and specifically data pertaining to the Barree and Conway (2004) model. Although the existence of non-Darcy flow has been confirmed and experimental work on the Barree and Conway model has been done, very little statistical work examines the nature of the Barree and Conway model and associated experimental results.

3.2 Darcy vs. Non-Darcy flow

To Darcy's linear equation (Equation 3.1), Forchheimer added a quadratic term (Forchheimer, 1901), and a cubic term (Equation 3.2) when he observed the limitation of Darcy's equation within a certain Reynolds number range (Forchheimer, 1901).

$$-\frac{\partial P}{\partial L} = \frac{\mu v}{k_d} \tag{3.1}$$

where

$$\begin{aligned} \frac{\partial P}{\partial L} &= \text{potential flow gradient, } \frac{atm}{cm} \\ \mu &= \text{viscosity, cp, } \frac{g}{100 \text{ cm-sec}} \left[\frac{\text{Mass}}{\text{Length-Time}} \right] \\ v &= \text{superficial velocity, } \frac{cm}{sec}, \left[\frac{\text{Length}}{\text{Time}} \right] \\ k_d &= \text{permeability, Darcies, } [Length^2] \end{aligned}$$

$$-\frac{\partial P}{\partial L} = \frac{\mu v}{k_d} + \beta \rho v^2 + \gamma \rho v^3 \quad (3.2)$$

where,

$$\begin{aligned} \beta &= \text{inertial flow parameter, } \frac{atm-100 \text{ sec}^2}{g} \\ \rho &= \text{fluid density, } \frac{g}{cm^3}, \left[\frac{\text{Mass}}{\text{Length}^3} \right] \\ \gamma &= \text{factor of the Forchheimer cubic equation} \end{aligned}$$

The limitations of both the Darcy and the Forchheimer equations have been observed at higher velocities where the relationship between fluid velocity and pressure gradient deviate from linearity and do not describe the entire range of potential fluid velocities in downhole situations.

In 2004, Barree and Conway defined apparent permeability in their fluid flow equation, which takes into account a minimum permeability (Equation 3.3) and is able to cover all ranges of Reynolds numbers, defined as the ratio of inertial forces to viscous forces (Barree and Conway, 2004). Specifically, this equation is not only valid for low flow rates but also for the highest flow rates displayed under field conditions. Therefore, in this study, the non-Darcy flow equation, as described by the B&C model (Equations (3.3) – (3.5)), is used.

$$k_{app} = k_{\min} + \frac{(k_d - k_{\min})}{(1 - R_e^F)^E} \quad (3.3)$$

$$-\frac{\partial P}{\partial L} = \frac{\mu v}{k_{app}} \quad (3.4)$$

$$R_e = \frac{\rho v}{\mu T} \quad (3.5)$$

where,

k_{app} = apparent-rate dependent permeability, Darcies, $\frac{cm-g}{100 \text{ sec}^2 - atm}$

k_d = constant Darcy permeability, Darcies, $\frac{cm-g}{100 \text{ sec}^2 - atm}$

k_{\min} = minimum permeability at high rate, Darcies, $\frac{cm-g}{100 \text{ sec}^2 - atm}$

R_e = Reynolds number, dimensionless

T = characteristic length, 100/cm, $[\frac{1}{L}]$

F, E = exponents, dimensionless

Because of the wide range of Reynolds numbers encountered in fluid flow through porous media, the B&C model (Equation 3.3) suggests an apparent permeability (k_{app}) to describe Darcy (linear) and non-Darcy (nonlinear) flow in porous media with a single equation (Barree and Conway, 2004). Considering all previous experimental work, the B&C model is the only one able to satisfy all tested ranges of Reynolds numbers (Lopez-Hernandez 2007; Lai 2010; Aljalalmah 2014). Figure 3.1 contrasts the experimental behavior for one type of proppant with the theoretical behavior of Forchheimer quadratic, Forchheimer cubic, and B&C equations. All three of the approaches account for data points associated with low flow rate, but only the B&C model accounts for the data across the entire measured range.

3.3 Past Experimental Studies

The data set used for this study is derived from past experimental studies. The data set exists from the work done by Lopez-Hernandez (2007) who started by building a custom apparatus to acquire data for a variety of proppant types and sizes across a wide range of Reynolds numbers. During his research, Lopez-Hernandez (2007) was not only able to build the first apparatus to run flow tests involving a wide range of Reynolds numbers at CSM, but also able to validate the

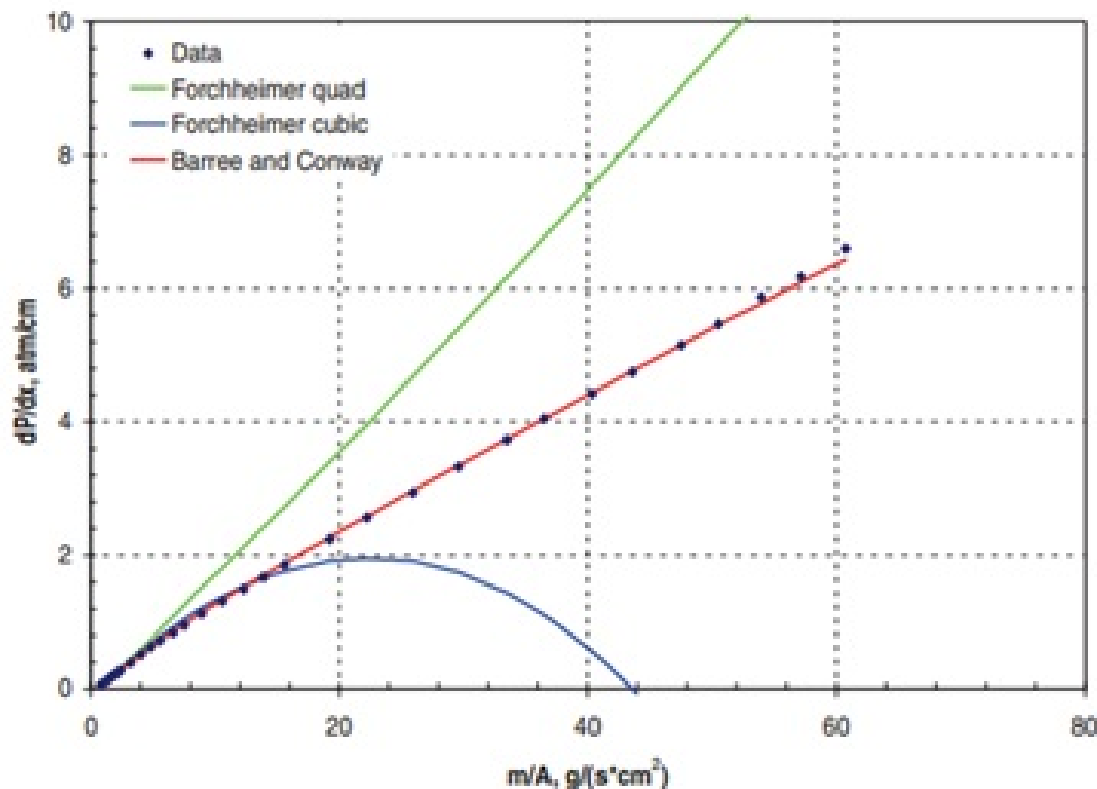


Figure 3.1: Validation of non-Darcy flow equations for Ottawa 20/40 sand under 3000 psi confining stress. From Lopez-Hernandez (2007).

Barree and Conway (2004). Aljalahmah (2014) conducted further experiments to validate the minimum permeability. Chapter 2 in this dissertation represents the highest differential pressures attained up to the time of this writing using Halliburton’s APFL. During testing, trends from previous work are observable and generate a larger data set that can be used for statistical analysis.

3.4 Colorado School of Mines Experimental Apparatus

The type of cell used in the Lopez-Hernandez (2007) and Aljalahmah (2014) experiments was similar. Aljalahmah (2014) used the apparatus that was originally built by Lopez-Hernandez (2007) and employed a maximum confining pressure of 6,000 psi.

Lopez-Hernandez (2007), Aljalahmah (2014), and Chapter 2 used liquid nitrogen in order to run their experiments. Upon release from the tank, the nitrogen generated high volumes of low-viscosity gas required to conduct the experiments by allowing for high differential pressure



Figure 3.2: CSM experimental set up

generation. The nitrogen at the Colorado School of Mines was stored in six cylinders with a capacity of 40,000 cc each and a maximum differential pressure of 6,000 psi. The CSM apparatus was able to run at 6,000 psi and at a flow rate of 2,000 cc/s. Lopez-Hernandez (2007), Aljalalmah (2014) and Chapter 2 of this dissertation all used gases in the flow tests through porous media as the objective is to obtain high Reynolds numbers. The viscosity of most gases is significantly lower than the viscosity of liquids. Nitrogen gas was used in all the experimental studies.

This flow system consisted of a pressure regulator panel, proppant pack cell, confining stress setup and mass flow rate measurements (Figure 3.2). The maximum pressure generated by the confining stress pump is 10,000 psi, and the proppant pack cell assembled by Lopez-Hernandez (2007) was 26 cm long with different locations where the pressure assemblies were planted and it had an inside diameter (I.D of 3/8-in) (Figure 3.3). It was made of a Hassler-sleeve Tygon material, which allows for the measurement of the pressure at different points of the proppant pack; the stress can be uniformly applied around the cell. In this apparatus, Lopez-Hernandez was able to test eight different types of proppant during his studies.

All variables measured and recorded during the tests were raw data; therefore, it was necessary to do some variable manipulation such as calculation of mass flow rates. The overall



Figure 3.3: Proppant pack showing different pressure measurement points used at CSM

purpose of the experimental study of Lopez-Hernandez was to validate the existing flow equation for a wide range of Reynolds numbers. This novel experimental apparatus designed at the CSM has been able to achieve Reynolds numbers ten times higher than those found in previous work (Lopez-Hernandez, 2007), and he was therefore able to identify new patterns beyond the one of Forcheimer that were not observed in previous studies. The results show that the B&C model are valid for the ranges and variables tested (Lopez-Hernandez, 2007). However, the study suggests that regarding the validation of the plateau behavior at high Reynolds numbers, it might be important to reach Reynolds numbers of 25,000 or more, which was not feasible with the current setup in the CSM lab.

Subsequently, Aljalalmah (2014) extended the experiments to increase the differential pressure and validate the existence of the minimum permeability. He was able to gather more data points but was still not able to reach the minimum permeability plateau. He concluded that his research confirmed the correlation of Lopez-Hernandez (2007) and B&C (2004). Chapter 2 of this dissertation experiments using a different apparatus but similar testing methodology also confirmed the correlations of past work. His experiment improves on the minimum permeability, as there is a possibility of differential pressure increase. This shows that the data points follow the same trend and are reliable for statistical analysis.

3.5 Halliburton Apparatus

Unlike the CSM apparatus, the Halliburton setup was used because it had the potential to reach a differential pressure of 10,000 psi, allowing the permeability to reach a lower point, ultimately the minimum permeability, as described in the B&C model. Based on the amount of pressure that could be attained, an initial step was to ensure that the proppant would not crush. Therefore, the choice of proppant was mainly based on its ability to resist high-pressures. The

grains of ceramic proppant have a higher resistance to crushing forces than natural sands and thus were used in these experimental runs. Another important factor in these experiments is the choice of fluid like Lopez-Hernandez (2007) and Aljalahmah (2014), this experiment used nitrogen gas, as it is the best choice for the ability to reach high Reynolds number as it is a low-viscosity fluid. The proppant pack design itself consisted of a Tygon™ material with a length of 4.4-in and a diameter of 0.375-in (Fig. 3.4). Unfortunately, the APFL does not allow the use of an outside diameter less than 4.5-in. Therefore, an insert to create the 0.375-in ID was used.

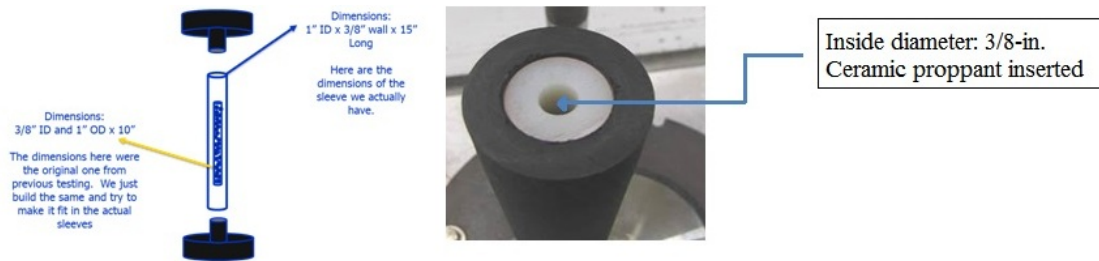


Figure 3.4: Final design of the proppant pack showing the initial rubber sleeves and the inseted material used in order to reduce the size of the proppant pack.

The approach used in the APFL experimental runs was derived from Lopez-Hernandez (2007) and modified for the APFL experimental apparatus. The methodology includes first a preparation of the proppant pack. The aim is to make sure the proppant pack is the right size for experimental slots. Then, a good seal of the proppant pack helps contain the pressure and create large differential pressure when flowing the nitrogen gas through the pack. The 20/40 ceramic proppant, which is strong enough to resist high pressure, was inserted into the new 0.375-in fixture of the proppant pack. Flowing the nitrogen through the proppant pack while increasing the stress and differential pressure allowed for the collection of data regarding pressure, mass flowrate, and permeability every five seconds over a time period of 55 seconds.

Despite numerous attempts, the Halliburton apparatus was not able to exceed a differential pressure of 4,000 psi despite the fact that the system could theoretically be run at a differential pressure up to 10,000 psi. That limitation only permitted a maximum pseudo-Reynolds number of $4,030 \frac{g}{cm^2-s-cp}$.

3.6 Collection of Statistical Data

The data set used for this analysis came from different sources. Lopez-Hernandez (2007) built a custom experimental set up to acquire data for a variety of proppant types and sizes across a wide range of Reynolds numbers. Aljalahmah (2014) used Lopez-Hernandez's experimental setup on a greater variety of proppant types used a setup adapted from both Aljalahmah (2014) and Lopez-Hernandez at Halliburton's Lab under an increased differential pressure. Despite the different locations of those experimental setups, all researchers had similar objectives, input parameters, and independent variables. The goal in all cases was to validate minimum permeability (Barree and Conway, 2004) using a proppant pack with nitrogen gas as the fluid flow. The experiments that are the subject of this paper additionally allow for a statistical analysis of mass flow rate using the Barree and Conway model as a basis for the analysis.

3.7 Review of Statistical Analyses in Unconventional Reservoirs

Over the last decade, several studies involving statistical analysis were performed with an objective to have a better understanding of independent variables with the greatest effect on oil well completion, as well as on reservoir simulation, both of which influence production of hydrocarbons. Modeland et al. (2011) performed statistical analysis of the effect of completion methodology on production of the Haynesville shale. Their study considered the effects of pressure, temperature, and lithology in order to analyze the production rates in the formation. They also found that production was impacted by proppant placement strategies that affect fracture conductivity. They concluded that in order to improve production, the Haynesville completions must focus on proppant placement that provides sustained conductivity. It is expected that factors such as proppant amount, mesh size, proppant types, and proppant concentration impact the conductivity.

Saldungaray et al. (2013) performed a statistical analysis that showed that the most important variables for optimizing completion of an oil well were lateral length, fracture stage spacing, stimulation fluid and proppant volume. Erturk et al. (2015) used a commercial simulator to investigate the effect of the completion of a reservoir on production performance of shale and other unconventional reservoirs. Their study shows the impact of the number of stages on production, noting that more stages might not always yield higher production.

No studies in the literature show statistical work done on the experimentation of the non-Darcy flow equations used in a reservoir; rather, they included relationships that assume Darcy flow, which do not hold for the case of many reservoirs including unconventional reservoirs.

The literature shows that the largest amount of work was performed at specific sites involving specific reservoirs; very few studies are derived from experimental work. However, most of the past studies have in common the analysis of proppant in unconventional reservoirs. Generally, during the unconventional reservoir studies, wells are grouped in several clusters (Sinha et al., 2016) such as:

- Geological setting (the presence of fault and petrophysical property gradation)
- Reservoir parameters (reservoir fluid types and reservoir pressure)
- Geomechanical parameters (local stress regimes, compaction, and zipper fracture effects)
- Completion design (proppant volumes, perforated length, and surface pumping rates)
- Well spacing and refracturing.

As with the above clusters, this experimental study focuses on clusters of the reservoir characteristics of pressure, temperature, permeability, viscosity, and proppant type, as they are the most observable in experimental work.

3.8 Assumptions and Definitions

The statistical analysis performed in this work makes three main assumptions: (i) despite the data stemming from different sources, the data are accurate within some random noise; (ii) the associated calculations were performed properly; and, (iii) the experimental conditions were consistent.

The following terms used in the statistical analysis are defined as follow:

Data Compilation: The compilation process consists of combining data from different experimental studies that were done in the past.

Variable Selection: Variables are the items or properties that are measured, controlled, or manipulated in the research. Variables are usually classified into two major types - independent variables and dependent variables. The definition of an independent variable or a predictor

variable is a variable that is being manipulated in an experiment in order to observe the effect on a dependent variable or an outcome variable.

Initial Screening: There are multiple issues to consider when performing the initial screening of the data. These include accuracy of the data, missing data, the impact of outliers, and more complicated issues such as the merging of data from different sources or data collected with different experimental devices. The first step in screening the data is to conduct a simple descriptive analysis of all the variables. To address the issue of missing data, multiple options are available, the first of which is simply to remove the observation from the sample. The second is to use an estimation procedure, and the third is to use a more complex method such as one employing a missing data correlation matrix (Tague et al., 2000). If the researcher chooses to eliminate or estimate the value, the impact must be considered on the eventual analysis.

Correlation and Regression: Once screened for unacceptable values, correlation analysis is conducted and, based on the result; several strong correlations can quickly be observed (Tague et al., 2000).

Regression Analysis: This is a predictive modeling technique, which investigates the relationship between a dependent (target) and independent variables (predictor); this technique is used for forecasting, time series modelling, and finding the causal effect relationship between variables.

Linear Regression Equation: This equation establishes a relationship between a dependent variable (Y) and one or more independent variables (X) using a best-fit straight line (also known as regression line), represented by Equation (3.6).

$$Y = b_0 + b_1X_1 + b_2X_2 + \dots + b_pX_p + \varepsilon \quad (3.6)$$

where,

Y = predicted or expected value

$X_1 \dots X_p$ = independent or predictors

b_0 = Value of Y (intercept)

$b_1 \dots b_p$ = estimated regression coefficient

ε = error terms

LASSO (Least Absolute Shrinkage and Selection Operator) Regression: this is a type of linear

regression that uses a shrinkage system where all the data are shrunk towards a central point. This is best used in a simple model with variables likely to create overfitting. This helps improve the prediction accuracy and the model fit.

Random Forest: In this case, a decision tree is used to plot the significance of a variable. Although this is not a classification problem, Random Forest will help confirm other analysis. Random Forest can be used to rank the importance of variables in a regression or a classification problem.

3.9 Results and Discussion

Lopez-Hernandez (2007), Aljalalmah (2014), and Chapter 2 of this dissertation provided data for the statistical analysis, yielding a total of 3,501 data points and 49 variables, appropriately formatted for use in R. After cleaning the data, linear regression analysis quantified the relationship between a single dependent variable and one or more explanatory variables. However, before the use of the linear regression analysis, it is important to be able to isolate the most important variables of the experimental study. This was initially done using a LASSO regression, and subsequently using Random Forest and Linear Regression. The primary objective of this statistical study is to use LASSO regression to determine important variables then derive an equation.

3.10 LASSO Analysis:

The LASSO regression analysis was done from the experimental data points collected. LASSO is a form of regularized regression, which avoids overfitting by adding information, and thereby shrinking the variables values in the model. With more than ten variables, overfitting can be prevalent.

The LASSO regression uses the following equation:

$$\sum_{i=1}^n \left(Y_i - \sum_{j=1}^p X_{ij} \beta_j \right)^2 + \lambda \sum_{j=1}^p |\beta_j|$$

While the normal regression uses the following equation, the difference between the LASSO and the normal regression is in the approximation of the λ value because if it is equal “zero,” then it reverts to the standard regression equation and the least squares approach.

$$\sum_{i=1}^n \left(Y_i - \sum_{j=1}^p X_{ij} \beta_j \right)^2$$

with:

λ = Strength of the penalty

β = Regression model factor

The steps taken in order to develop and run the LASSO analysis are the following:

1. Creation of a model matrix using dummy variables that is used to approximate lambda – response variable. This trains the model because it helps to identify the value of lambda. As the value of lambda increases, bias increases while when lambda decreases, variance increases. The estimate becomes more stable. If lambda goes to infinity, the slope terms will be equal to zero while the intercept will not.
2. Approximation of the lambda value by setting the coefficient to 1 and by using a built-in command. The approximation of the lambda value in the LASSO regression is an important factor. The analysis gave a value of 0.001.
3. Generation of training and test data via cross-validation, using the lambda value.

From the steps and techniques described above, the most important independent variables as the LASSO regression indicated were:

- Material used
- Apparent permeability
- Permeability
- Fluid viscosity

Table 3.1 shows the R-squared values of the training and test data during the LASSO regression analysis.

Table 3.1: LASSO regression values

TRAINING DATA	
R-Square	RMSE
0.99	0.278
TEST DATA	
R-Square	RMSE
0.99	0.304

After having identified the most important variables, the random forest techniques, although not a classification analysis, can help confirm the initial analysis with LASSO. LASSO identifies the most important variables by shrinking the least important to zero. Table 3.2 shows in bold the variables that are significant.

Table 3.2: LASSO regression values showing in bold the most significant variables.

LASSO REGRESSION SIGNIFICANT VARIABLES			
	Term Definition	Unit	
(Intercept)			18.5
Test.No.	Number of tests run	dimensionless	0.0
Initial pressure	Pressure starting at proppant pack	psi	0.0
P3...Pack.flow.Press	Pressure in the proppant pack	psi	0.0
Final pressure	Pressure exiting the pack	psi	0.0
Confining pressure	Pressure confining the pack	psi	0.0
Pack Temperature	Temperature of proppant pack	(°F)	0.1
Pack Temperature	Temperature of proppant pack	(°F)	0.1
Initial.Tank.Press.	Inlet tank pressure	psig	0.0
Final.Tank.Press.	Final tank pressure	psig	0.0
Initial.Tank.Temp.	initial tank temperature	(°F)	0.0
Final.Tank.Temp.	Final tank temperature	(°F)	0.0
Time.Elaspsed..DT..sec.	Time elapse between runs	sec	0.0
Z-factor Initial	Initial Compress ability factor	dimensionless	65.5
Z-factor Final	Final Compress ability factor	dimensionless	-36.4
CSA..cm ² .	Cross sectional area of pack	(cm ²)	-10.2

Z1	Compress ability factor	dimensionless	-22.8
Fluid Density	Density of nitrogen fluid	(gm/cc)	18.3
Initial fluid viscosity	Viscosity of nitrogen fluid	cp	-135.9
CSA..cm ² ..1	Cross sectional area after runs	(cm ²)	8.2
Z5	Compressibility factor	dimensionless	-16.8
Final fluid density	Fluid density after runs	(gm/cc)	5.6
Final fluid viscosity	Fluid viscosity after runs	cp	0.2
SIMPSON.Rules...	Calculated Simpson rules	dimensionless	0.0
Kapp	Avg Apparent permeability	milliDarcy	0.0
Viscosity (CP)	Average Viscosity	cp	-266.9
Mass.CSA	Calculated mass as function of CSA	(gm/cc)	0.5
Vavg	Average velocity	(cm/sec)	0.0
Pseudo.Reynolds	Pseudo reynolds number	dimensionless	0.0
Kapp.1	Calculated apparent permeability	milliDarcy	0.2
X		dimensionless	0.0
dP.dX..atm.cm.	Pressure over the length		-3.9
V..cm.sec.	Velocity	(cm/sec)	0.0
X1.perm.app		milliDarcy	0.0
Name of the tester	Person who ran the test		-1.8
Material used	Proppant type		0.0

Random Forest: The underlying principle of random forest analysis is to build multiple decision trees that merge together to produce an accurate prediction. Random forest is a classification technique. However, despite the fact that the problem is not technically under classification, in this case, the data are assumed uncorrelated and this process separates the data into different classes.

The initial step is to select the number of trees for the analysis. The software R recommends an initial analysis with 500 trees, to be increased if needed. The mass flow rate as a function of all other experimental variables was used and it was observed that with no more than 400 trees,

the error was reduced to zero. This showed that no more than 400 trees were necessary to produce no error. The most significant independent variables are shown in Table 3.3.

Table 3.3: This table shows the most significant variables from the random forest methodology; these match those determined by LASSO.

Mean Decrease Gini	
Material Used	124,313.4
Apparent Permeability	415,946.2
Rock Permeability	482,782.9
Viscosity	333,560.7

Linear regression: Linear regression analysis was also performed using R and the dataset with the characteristics given in Table 3.4. Those data still undergo a variable selection in an attempt to match the other two methods.

Table 3.4: This table shows the characteristics of all data points available from the experimental studies in R.

Data Available	
Number of data points	Number of variables
3,541	49

The analysis initially considered the entire dataset. During the manipulation of the data, the software could not process anything other than numerical values. Therefore, the names of the tester as well as the names of the material used during testing were assigned numerical values. The above changes reduced the total data set as seen in Table 3.5. This did not alter the testing capabilities but just combined some repetitive information.

Table 3.5: This table reflects the characteristics of all numerical values, as well as variables used in the analysis.

Data Available	
Number of data points	Number of variables
3,507	49

The initial cleaning of the data shows that there are still some missing values, as well as outliers (Table 3.6). Therefore, the cleaning process needed to be done in an iterative manner so as to systematically remove all the missing values and the outliers. The software R has an application that transforms raw data into consistent data and improves the content of statistical

statements based on the data, as well as their reliability. Data cleaning may profoundly influence the statistical statements based on the data. It is then important to make sure this process is done properly so as not to influence the variables in a way that might alter the results.

Table 3.6: This table shows the initial run using the linear regression model with some NA's associated with variables that need to be removed

FIRST RUN OF THE MODEL THAT STILL SHOWS SOME NA VARIABLES TO BE ELIMINATED						
		Unit	Estimate	Std. Error	T-Value	Pr(> t)
(Intercept)			-56.97	8.642	-6.593	
Test.No.	Number of test ran	Dimensionless	0.00957	0.000927	10.329	<
InitialPressure	Pressure starting at proppant pack	psi	0.004527	0.000506	8.95	<
P3...Pack.flow.Press	Pressure in the proppant pack	psi	-0.001065	0.000191	-5.57	0
Final Pressure	Pressure exiting the pack	psi	-0.002272	0.000363	-6.261	0
Confined Pressure	Pressure confining the pack	psi	0.000008	0.000008	0.975	0.329404
Pack Temperature	Temperature of proppant pack	(oF)	0.1086	0.01324	8.205	0
Initial.Tank.Pressure	Inlet tank pressure	psig	0.00068	0.000116	5.849	0
Final.Tank.Pressure	Final tank pressure	psig	-0.002806	0.000157	-17.899	<
Initial.Tank.Temperature	initial tank temperature	(oF)	-0.002006	0.000604	-3.319	0.000913
Final.Tank.Temperature	Final tank temperature	(oF)	0.00249	0.000387	6.439	0
Time.Elapsed	Time elapse between runs	sec	-0.001904	0.000185	-10.274	<
Initial Z-factor	Initial Compressibility factor	Dimensionless	-5.92	3.307	-1.79	0.073538
Final Z- factor	Final Compressibility factor	Dimensionless	-4.095	4.205	-0.974	0.330173
Cross sectional area	Cross sectional area of pack	(cm2)	-8.253	0.4437	-18.602	<
Z1	Compressibility factor	Dimensionless	18.43	9.482	1.944	0.052026
r1..gm.cc.	Density of nitrogen fluid	(gm/cc)	17.71	6.108	2.899	0.003763
m1..cp.	Viscosity of nitrogen fluid	cp	-2946	452.4	-6.511	0
CSA..cm2..1	Cross sectional area after runs	(cm2)	4.327	0.5234	8.267	<
Z5	Compressibility factor	Dimensionless	105.1	21.89	4.8	0.000002
r5..gm.cc.	Fluid density after runs	(gm/cc)	54.99	10.84	5.072	0
m5..cp.	Fluid viscosity after runs	cp	315.6	878.9	0.359	0.719543
SIMPSON.Rules...	Calculated simpson rules	Dimensionless	0.0163	0.00033	49.339	<
Kapp	Avg Apparent permeability	milliDarcy	0.000007	0.000072	0.102	0.918577
Average Viscosity	Average (Avg) Viscosity	cp	-832.8	876.3	-0.95	0.341975
Mass.CSA	Calculated mass as function of CSA	(gm/cc)	0.6294	0.00344	182.985	<
Vavg	Average velocity	(cm/sec)	-0.02904	0.000639	-45.47	<
Pseudo.Reynolds	Pseudo reynolds number	Dimensionless	-0.003895	0.000108	-36.006	<
Kapp.1	Calculated apparent permeability	milliDarcy	0.7349	0.01524	48.217	<
X		Dimensionless	0.00457	0.000126	36.208	<
dP.dX..atm	Pressure over the length		-2.007	0.06508	-30.838	<
V..cm.sec.	Velocity	(cm/sec)	-0.001708	0.00016	-10.686	<
X1.perm.app	Permeability	milliDarcy	0.00116	0.000623	1.862	0.062657
Name of the tester	Person who ran the test	unitless	-0.08874	0.01582	-5.608	0
Material used in testing	Proppant type	unitless	0.001435	0.000277	5.184	0

Once the data is cleaned, the subsequent step is to generate a predictive model, which yields the mass flow rate as a function of the pressure, temperature, viscosity, material used, and permeability. Table 3.7 provides the indicators associated with the important variables.

Table 3.7: Table showing the significant code from the variables of the experimental study

Significance of Code	
Highly significant	‘***’
Not significant	‘*’

The variables with the fewest “stars” might not have an impact on the mass flow rate. This analysis determines the variables with the most impact on the mass flow rate, shown in Table 3.8. The significant variables match those from the LASSO and the random forest analysis.

Table 3.8: Table showing the most significant variables in LASSO

Significant Variables in the Analysis					
	Estimate	Std. Error	T Value	Pr(> t)	Significance
(Intercept)	-64.5	2.311	-27.912	>2-E16	***
Material used	-0.2182	0.0394	-5.538	>2E-16	***
Apparent permeability	-0.2928	0.009	-32.533	>2E-16	***
Rock permeability	3.453	0.0369	93.478	>2E-16	***
Viscosity	2663	78.55	33.906	>2E-16	***

Residual standard error	10.53
Multiple R-squared	0.7959
Adjusted R-squared	0.7955

Viscosity and permeability have traditionally been important factors in predicting the mass flow rate in hydraulically fractured reservoirs. The software R allows for the identification of the most significant variables and their corresponding coefficients via any of the three methods. The mass flow rate linear regression equation is:

$$Y = -64.5 - 0.21 * X_1 - 0.293 * X_2 + 3.45 * X_3 + 2663 * X_4 \quad (3.7)$$

where

Y = Mass flow rate

X_1 = Material used

X_2 = Apparent permeability

X_3 = Rock permeability

X_4 = Viscosity

3.11 Conclusions

A statistical methodology, independent of the academic discipline, has been developed to determine important factors in non-Darcy flow. This methodology is based on laboratory testing. These statistical techniques can be applied to various hydraulically fractured jobs in any field or

area given certain characteristics and the availability of relevant data. Completion costs can be reduced and trends can be determined using the methodology described in this paper.

The goal of this study was to quantify the most important variables as given by the laboratory experiments conducted in an effort to validate the Barree-Conway model. Linear regression was found to be the most suitable approach, confirming past statistical work that suggests certain pressures are required to achieve a given mass flow rate. The most important variables are:

- Material used
- Apparent permeability
- Pack permeability
- Viscosity

The study also produced an equation that can be used in order to identify the mass flow rate based on specific reservoir characteristics.

Regardless of whether or not the flow is Darcy or non-Darcy, similar conclusions have been reached about the type of independent variables associated with the mass flow rate. The magnitudes of the coefficient are higher in some instances and smaller in others depending on the field and laboratory testing from previous work that has been done in similar settings. These variables are significant not only for strengthening the Barree-Conway model but also in the critical recovery of hydrocarbon during the hydraulic fracturing process. As such, the study contributes to the ongoing improvement of knowledge regarding unconventional extraction processes.

3.12 Nomenclature

Symbol Description

A : area open to flow, cm^2

CSA : cross-sectional area, cm^2

d : diameter, cm

$\frac{\partial P}{\partial L}$: pressure gradient, $\frac{\text{atm}}{\text{cm}}$

ECS : effective confining stress, psi

E : coefficient in the BC non-Darcy flow model

F : coefficient in the BC non-Darcy flow model

k_{app} : apparent permeability, Darcy

k_d : Darcy permeability, Darcy

k_{min} : minimum permeability, Darcy

L : length, inches or cm

Re : Reynolds number

μ : viscosity, cp

β : inertial flow parameter, $\frac{\text{atm} - \text{s}^2}{\text{gr}}$

ρ : fluid density $\frac{\text{gr}}{\text{cm}^3}$

v : superficial velocity, $\frac{\text{cm}}{\text{sec}}$

T : characteristic length, $\frac{100}{\text{cm}}$

CHAPTER 4

REDUCING NON-PRODUCTIVE TIME IN OIL FIELDS

This paper is planned for submission to *INFORMS Journal on Applied Analytics*

Kamga L Ngameni¹, Alexandra Newman², Jennifer Miskimins³

Abstract

Oilfield operational planning determines a long-term production schedule, often to minimize makespan. For a time horizon of between months and years, optimization models seek a sequence of activities with daily or weekly fidelity to satisfy precedence constraints as well as resource constraints on the amount of material required versus its availability to execute the activities. With algorithmic advances, as well as those in oilfield planning software such as the one used by Enersight and Schlumberger, we are able to solve instances with a decade-long horizon at daily fidelity. The resulting objective, repeatable, and defensible schedules inform production companies regarding supervisory decisions. We implement our solutions on an existing site.

4.1 Introduction

Energy is a fundamental resource of the industrialized world. At the time of this writing, the forms of energy commonly used in North America are oil, natural gas, and electricity. However, because the sources of oil and natural gas are commonly located at a distance from their use, it is important to plan production operations in advance (Raleigh et al., 1965). The cost of oil and gas production has been a significant factor affecting the competitiveness of the industry. In particular, North America has one of the highest costs of hydraulic fracturing production internationally. It is therefore important to identify potential cost-cutting measures in the areas of exploration, drilling and production, manufacturing, transportation, and sales. The production process, which is commonly executed in phases, starts with exploration for hydrocarbon-bearing rock formations, and the review of geological maps in order to identify major sedimentary basins (Benson, 2005).

Once a promising geological structure has been found, to confirm the presence of

¹Operations Research with Engineering Graduate Program, Colorado School of Mines, Golden, CO 80401

²Operations Research with Engineering Graduate Program, Colorado School of Mines, Golden, CO 80401

³Petroleum Engineering Department, Colorado School of Mines, Golden, CO 80401

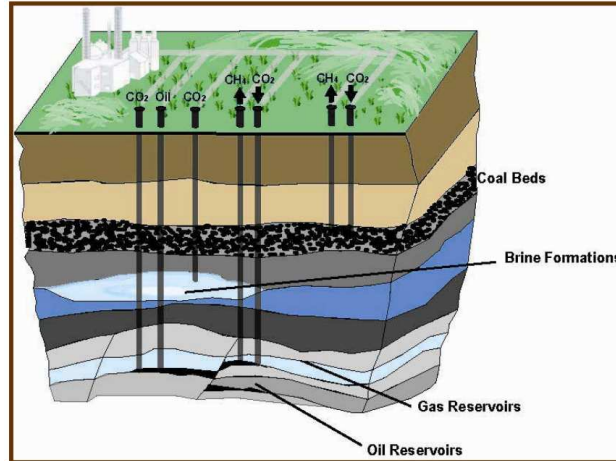


Figure 4.1: An oil and gas sedimentary basin

hydrocarbons and the thickness and internal pressure of a reservoir, engineers drill exploratory boreholes, commonly known as “wildcats.” The goal is to drill the fewest possible wildcats in a given field to gain the most information possible as it is costly to drill them; Monte-Carlo simulation methods based on predictions of the geology help effect this. Following exploration drilling, engineers gain further information regarding the size and the extent of the field by drilling either “outstep” or “appraisal” wells. This drilling establishes the number of wells required, the size of the oilfield, and whether more seismic data is needed to determine a more complete characterization. Eventually, the development, or production, wells are established.

An engineering team develops a small reservoir using one or more of the appraisal wells; a larger reservoir requires the drilling of multiple additional production wells. The goal is to minimize land requirement and infrastructure cost associated with a *pad*, or location in which several wells can be drilled. Most new commercial oil and gas wells are initially *free flowing*; that is, the underground pressures drive the liquid and gas up the wellbore to the surface. The rate of flow depends on factors such as the properties of the reservoir rock, the underground pressures, the viscosity of the oil, and the ratio of oil to gas present in the subsurface reservoir. Once the oil cannot reach the surface unaided, some form of additional lift is required. At the time of this writing, it is common to inject gas or water in order to maintain reservoir pressure and improve production rates and ultimate recovery. The most commonly used method in North America for stimulating production is the *hydraulic fracturing technique*, which requires significant machinery

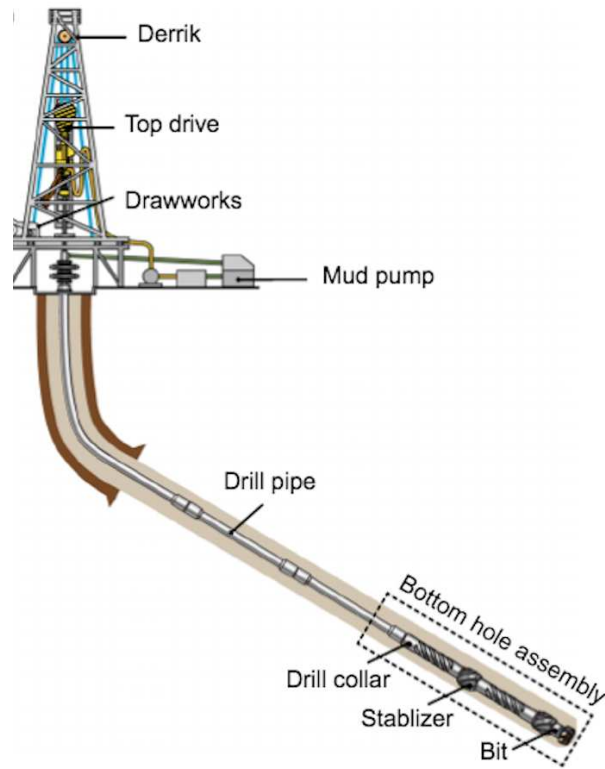


Figure 4.2: Unconsolidated, or hydrocarbon, formation which contains rock grains, and drilling casing pipe that enters the formation in an horizontal manner, (Chen et al., 2019)

and manpower. Once the hydrocarbon reaches the surface, it is routed to a production facility depending on fluid type; for example, heavy oil has a higher viscosity than lighter oil. Here, oil, gas, and water are separated.

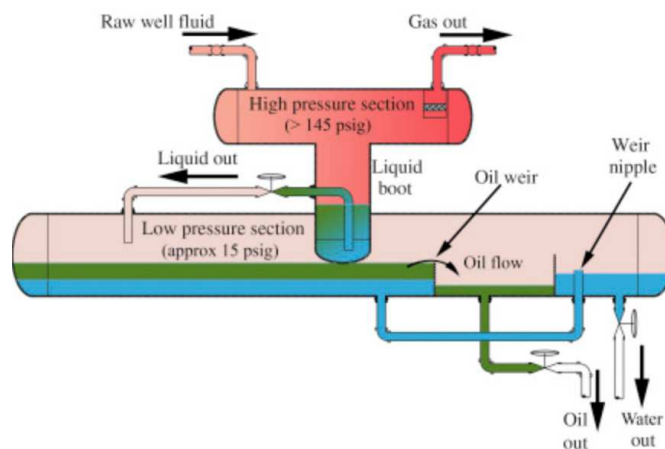


Figure 4.3: Oil and gas separation to the midstream (Lea Jr and Rowlan, 2019)

After two to four decades, when all oil and gas contained in the fields have been transferred to production facilities and the commercial life of the wells has been exhausted, the final step is decommissioning, during which the site is restored to environmentally sound conditions, and measures are introduced to encourage site re-vegetation.

Therefore, decisions in operations scheduling consist of determining when various activities should occur over the course of a horizon of one-to-many years. These activities prepare and, subsequently, retire an area from which oil and gas is produced, and consist of: (i) drilling the area and inserting explosives; (ii) blasting the rock (or fracturing it via mechanical means) to permit oil and gas to be extracted, i.e., hydraulically fractured; (iii) completing activities, i.e., protecting the borehole; and, (iv) transporting oil and gas to the surface for further processing and eventual sale. An exploration company generally possesses an objective of maximizing net present value and reducing non-productive time. Constraints consist of rules governing the order in which activities can occur, and the number of resources over a time span that can be used to conduct these activities. This paper contributes to the area of operations research in petroleum engineering by minimizing makespan associated with the logistics involved in bringing a well to production.

4.2 Literature Review

Applications of operations research in the oil and gas industry are numerous. Companies such as Exxon Mobil, Total, Shell, and Chevron have employed in-house operations research analysts since the 1970s (DiCarlo et al., 2019) to employ techniques such as linear programming for refinery planning, or optimizing the conversion of crude oil into gasoline, jet fuel, diesel fuels, and lubricants. Dynamic programming (Rardin, Ch 9) has been used for resource allocation and, more specifically, has been applied to human resource management in the oil and gas industry (Ghaeli, 2019). It has also been used to determine how to allocate resources among a set of unconventional oil and gas projects to maximize total expected profits, considering their feasibility, geological reserves, and estimates, and also to determine an appropriate well setting to maximize net present value while adhering to constraints such as maximum watercut and maximum liquid production rates (Wen et al., 2011). In unconventional reservoirs, the configuration of a multi-well pad has been considered to be most efficient because it reduces the

environmental impact while maximizing oil production (Wang and Chen, 2017); however, it can lead to significant non-productive time. It remains challenging to simultaneously determine hydraulic fracture operational variables such as the proppant which is used to keep the fractures open, the type and quantity of liquid injected in order to create adequate fractures of each well, and well placement (Wang and Chen, 2017). Other optimization questions entail determining the best way to produce the field and maximize profit by selecting water versus gas to maintain high field pressure or performing more hydraulic fracturing in order to create greater permeability associated with a commensurate initial production rate.

Other relevant techniques include queuing theory, decision analysis, and simulation. Queuing theory analyzes the delay associated with waiting in line by examining behavior during peak periods and the service process (Alhanai et al., 2002) and has been used in the oil and gas industry in order to calculate the return on investment of a coastal refinery to minimize the cost of keeping tankers waiting. It has been used to analyze vessel traffic flow based on the number and speed of vessels delivering petroleum products internationally (Roy et al., 2016), and employed to assess oil price volatility (Hannan and Bridwell, 2012). Decision analysis can integrate petroleum engineering with geology, process design, risk analysis, and finance (Walls et al., 1995). Operations analysts make the decision with the help of managers in the specific area or location of the field (Keefer et al., 2000). Other than in the exploration stage, decision analysis is used in: (i) bidding and pricing, (ii) environmental risk, (iii) product and project selection, (iv) technology choice, and (v) strategy. Simulation determines how to develop an oilfield using reservoir properties, and how to estimate the performance of oil and gas reservoirs (Mydland et al., 2020). For instance, in a tight unconventional reservoir, it is used to capture the fracture complexity in the reservoir caused by the intensive hydraulic-fracturing process to find ways to produce the most oil at the lowest costs (Settari et al., 1980).

4.3 Activities in Oilfield Operations

This background is based on a typical project within the United States, and includes information related to top-hole and lateral drilling, hydraulic fracturing, completions, flowback, and production. A drilling rig and a crew must first complete top-hole activities by maintaining vertical section of the wells via a costly and inefficient process to remove overburden pressure. For

areas in which multiple vertical wells exist in close proximity, the potential for collision with another well is greatly increased when the vertical section, or top hole, cannot be drilled straight. If the top hole is deep enough, it might require placement and cementing of the casing. Extreme tortuosity of the top hole results in high cementing cost. However, the aim is to reduce cost and reduce non-productive time (Gaines et al., 2013).

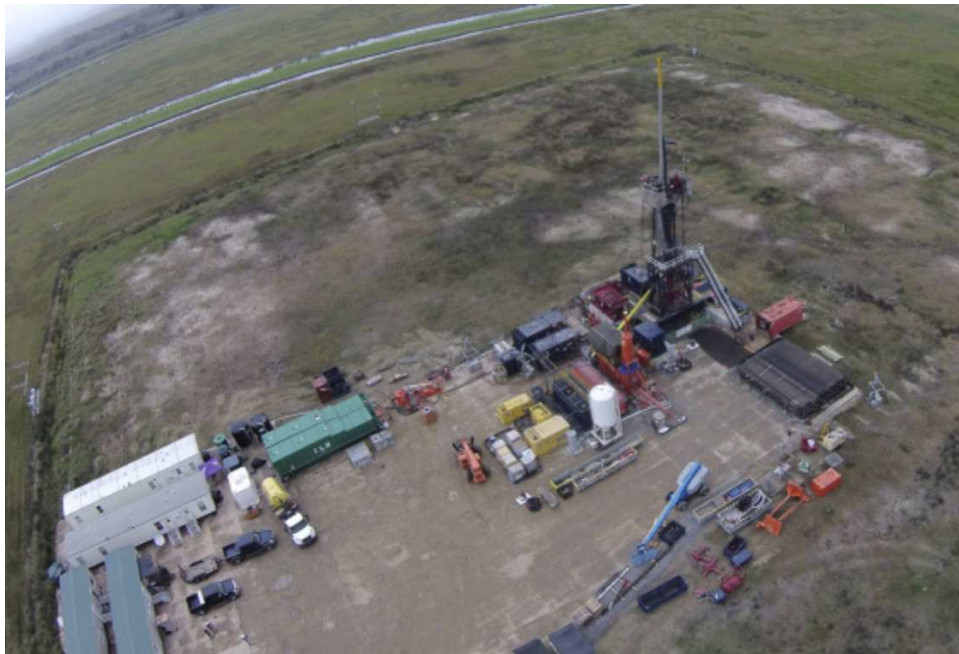


Figure 4.4: Typical onshore drilling equipment layout (Courtesy of International Snubbing Services)

The next step of the operation uses drilling crews to build the lateral part of the hole, which is often the most time-consuming part and requires the most accuracy. Depending on the wells, it might take an average of two weeks to build laterally, while the top hole might only take two days. A *drilling rig* resource is used to perform this activity. The number of drilling rigs varies depending on the size of the field. The typical well drilled can have vertical depth up to 10,000 feet and lateral length up to 4,000 feet (Johnson Jr et al., 1999). These two parts must be drilled consecutively and continuously. For our case study, we assume that drilling costs are sunk; however, these costs, which we associate with strategic decisions outside the scope of our work, can vary greatly depending on the challenges associated with the drilling.

Fracking preparation activities must be performed after drilling and do not require any

associated resources, but cannot commence until 30 days after the drilling activities are finished. If there are several wells on the same pad, all fracking preparation activities must happen simultaneously beginning 30 days after the drilling activities. The fracking preparation consists of three key components:

1. Cellar preparation: the wellhead is fitted for fracturing operation
2. Toe preparation: the well itself is prepared for the actual operations
3. Water preparation: delivery and associated storage is constituted

Following fracking preparation, the crew starts the completion activities that include moving the well into production, and incorporates the steps necessary to transform a drilled well into a producing one. These steps include: casing, cementing, perforating, gravel packing and installing a production tree. Specifically, casing (consisting of steel pipe that is joined together to make a continuous hollow tube) prevents the well from closing in on itself after the drilling fluids have been removed, and protects the wellstream from outside elements, such as water and sand (Figure 4). Cementing the well includes pumping cement slurry into it to displace existing drilling fluids and to fill the space between the casing and sides.

Flowback does not require any resources and takes five days to complete. All flow back must be performed at the same time on all the pads. This process allows fluids to flow from the well

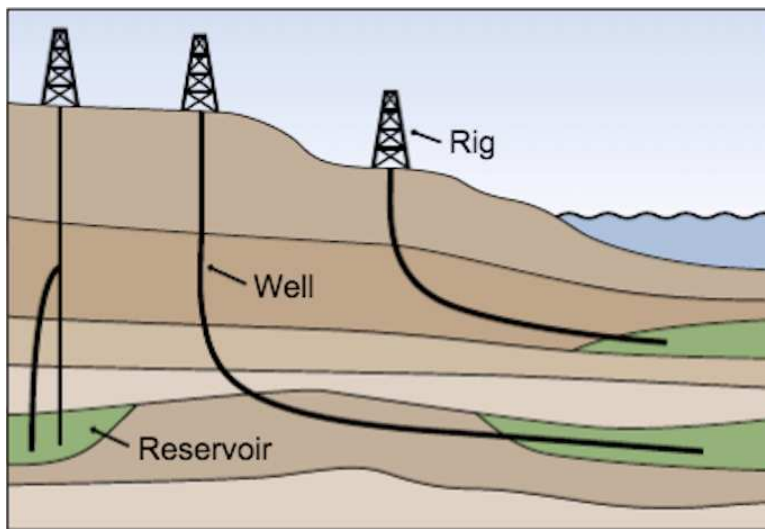


Figure 4.5: Example of a well being drilled horizontally (Chen et al., 2019)



Figure 4.6: Example of a fracturing crew at an oil field completion site(Courtesy of Liberty Oilfield Services)



Figure 4.7: Example of flowback pipe

following the hydraulic fracture treatment, either in preparation for a subsequent phase of treatment or in preparation for cleanup and returning the well to production. Production follows flowback; once the well has been put in production, the sequence of activities for that well is complete. Crews dedicate their time to routine maintenance and check-up, which is not a limiting factor in resource consumption. If unforeseen events cause a decline in oil well production, engineers can reassess the operational plan.

4.4 Current Scheduling and Proposed Optimization Model

At the time of this writing, the oil and gas industry does not possess a consistent set of detailed optimization techniques to schedule oilfield operations. Day-to-day execution of activities

in the unconventional oil and gas industry have traditionally been unpredictable, leading to inefficiencies such as the over-employment of (sometimes inexperienced) crews in the field. Depending on the size of the projects, some extractive resource companies have hired professionals in order to help with activity scheduling (Kuchta et al., 2004; O’Sullivan and Newman, 2014; Brickey et al., 2021). However, smaller oil and gas companies involved in hydraulic fracturing tend to forgo the employment of formal optimization techniques owing to up-front costs and lack of awareness of the field. Instead, a greedy approach is often followed that focuses on executing activities to obtain a high production rate at the onset of operations; this leads to fracturing more wells than is optimal. Our goal is to formulate a formal optimization model and to use customized software to solve it in order to improve the scheduling of operations and reduce non-productive time over the lifespan of the operation; the model minimizes makespan subject to precedence requirements and resource constraints.

In order to create a representative optimization model, we delineate the number of pads, and the number wells on each pad, as well as their geographic location. Pads, whose sizes vary, are located in different areas of the oilfield. Some pads may contain 20 wells while others may only contain 12. We consider a field possessing 71 pads containing a total of 729 wells, and a time horizon of 10 years with daily fidelity. Figure 4.8 shows where pads are geographically positioned in the field related to our case study.

Activities consist of construction of the pad, drilling of the well, fracturing preparation, well completion, flowback, and production in the field, and corresponding duration of each. Each activity is associated with a duration and a resource. The duration of each activity must be respected, though there can be idle time between the completion of one activity and the commencement of its successor. On the other hand, sometimes a lag is required; for example, it requires two days to move a drilling resource from a well to another, and one week to move the same resource from one pad to another. Table 4.1 provides detail of the individual activities and their durations:

The following resources exist: (i) a construction crew, which is required for the construction of the pad, (ii) two drilling rigs, which are required to drill the top hole and the lateral section of the well, and (iii) one completion crew, which is used for the completion of activities on the pad.

Table 4.1: The Production schedule contains various activities

Activity type	Description	Duration [days]
Pad construction	Building the pad in order to drill the wells on it	15
Well drilling	Drilling the top, followed by the lateral, hole	2, 14
Fracturing preparation	Readying equipment for hydraulic fracturing and completion	5
Well completion	Casing, cementing, perforating and installing a production tree	5
Flowback	Release of fluid used in fracturing	5
Production	Generation of oil or gas	0

However, not all activities require resources. Specifically, fracturing preparation, flowback, and production reflect a “piggybacking” on the completion of a prior activity. For example, once a well is ready after completion, the same crew might finish the flowback work (as part of the completion work), then turn on the pump and depart the site, which requires negligible work for the flowback activity on the part of a crew. Table 4.2 provides detail.

Given these sets and parameters, we construct an optimization model by formulating the following components:

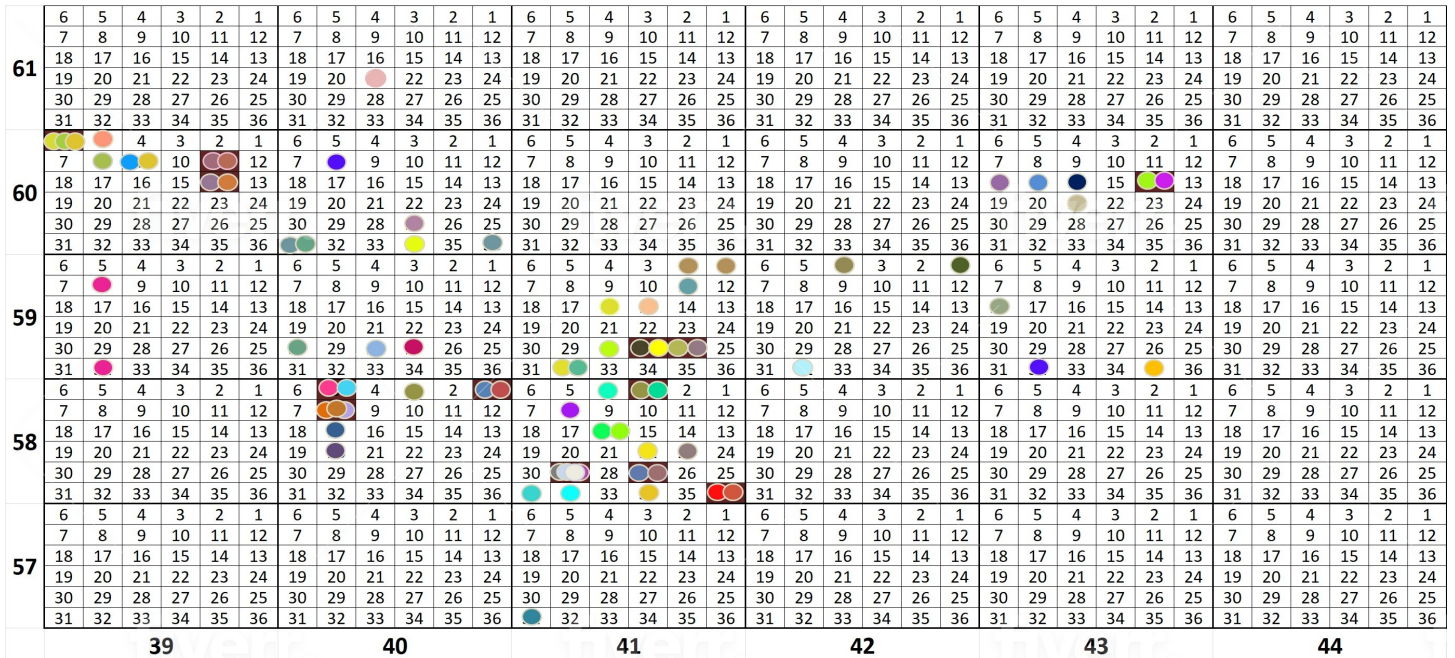


Figure 4.8: Geographic locations of the pads and wells

Table 4.2: Resource constraints in our production scheduling model

Resource	Purpose	Relevant Activities
Drilling rigs	Drill vertical and horizontal well sections	Well drilling
Construction crew	Required for constructing pads	Pad construction
Completion crew	Employed after fracturing the well to obtain oil	Completion

1. Variables: The date (in our case, day) on which each activity should start
2. Objective Function: The makespan, or time required to complete all activities (as given by their durations and any necessary lags), which we minimize
3. Constraints: Activity sequence based on predecessor activities that must be completed before a given activity can begin; and, the operational limitations on the amount of labor and infrastructure that can be employed during each time period as a function of the amount used based on the activities executed in each time period.

As presented here, this model possesses the structure of a RESOURCE-CONSTRAINED PROJECT SCHEDULING PROBLEM (Johnson, 1967; Talbot, 1982; Patterson, 1984; Kolisch et al., 1995) in that it schedules activities according to precedence and limitations on shared resources between activities. In practice, instances tend to be very large, possessing thousands of activities or more, and at least as many time periods; our instance is no exception. In order to expedite solutions, Bienstock and Zuckerberg (2010) develops an algorithm to solve the linear programming relaxation of a specialized version of the problem. Muñoz et al. (2018) present an adaptation of the algorithm in the afore-mentioned reference that incorporates activity durations, auxiliary variables and algorithmic speed-ups. While some authors tailor this general approach to incorporate additional considerations in underground mine production scheduling, e.g., the transition from open pit to underground operations (King et al., 2017) or the inclusion of ventilation and refrigeration when working underground (Ogunmodede et al., 2021), we use a straight-forward implementation on a standard model (similar to Brickey et al. (2021)), but alter the objective to minimize makespan, rather than to maximize net present value; we provide the mathematical formulation in the Appendix. The algorithm itself, coded as OMP (Rivera et al., 2015), requires the five input files given in Table 4.3 with their associated descriptions:

Table 4.3: Input file types and their descriptions

File type	Description
Blocks (*.blocks)	Lists all activities and their associated daily resource consumption
Problem (*.prob)	Controls how the overall model is constructed, and contains the time horizon over which to solve the model, the problem size, and parameter values
Precedence (*.prec)	Lists the predecessors for each activity
Delay (*.delay)	Provides the lag corresponding to each precedence between activities
Mapp (*.mapp)	Contains the mapping from an identifier given in the Excel database to the corresponding identifier used by the mathematical optimization software

4.5 Results

We run the problem instance over a time horizon of 16 years to minimize the makespan associated with completing all activities in the oil field subject to precedence and resource constraints. The baseline solution is associated with an operational plan that requires 16 years to execute. We measure the utilization as the quotient of the number of resource-days over which a resource is used and the total number of resource-days a resource is available. (So, for example, with two drill rigs, there are 365×2 drill rig-days available over the course of a year assuming that operations can run continuously throughout the year.) With an average utilization of 23%, completion crews are not the binding resource. The construction crew is highly utilized, in general, but only for the first three years during which the associated activities are continuously executed; after this, the construction crew is not needed, indicating that, when taken over the course of the time horizon, this is also not a binding resource. However, the drilling resources are binding because the utilization is close to 100% throughout the horizon. Drilling activities start upon construction of the pads and are executed, on average, across the entire time horizon, 352 days out of each year. Figure 4.9 details resource usage for the three different types throughout the time horizon of our study.

There are three consecutive years (i.e., 6, 7, and 8) in which there are no completions (and, correspondingly, no production) because the associated drilling activities require a total of sixteen days, each, which is far greater than the amount of time required to execute a completion activity

(i.e., five days each). Figure 4.11 depicts cumulative production levels allowed based on resource use. Noting that drilling is the binding resource, we test a scenario in which we increase the number of available drilling rigs to three, holding all other inputs constant. The modified instance generates a schedule that is able to finish the project in 11 years instead of the original 16 years, shortening the time required to complete all activities by five years.

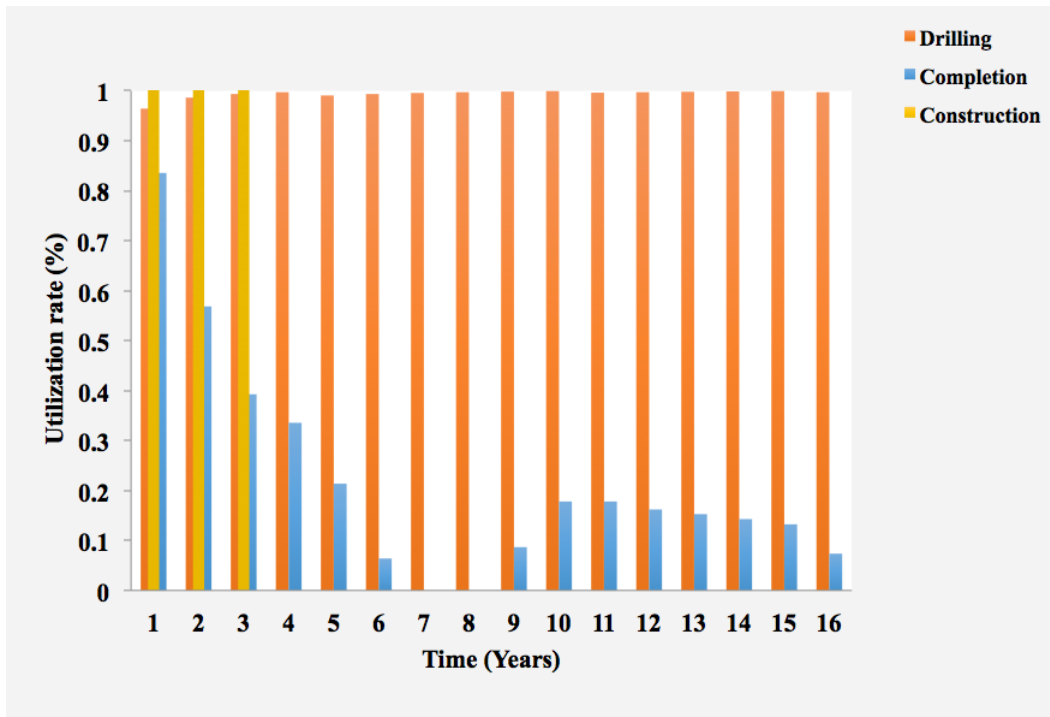


Figure 4.9: Utilization for each of the three resource types over the time horizon employing two rigs

Adding more rigs, we realized that the completions crew is a lot more utilized than when working with only two rigs. At this time, we see that the extra rig allows us to complete the drilling faster while still utilizing all drilling resources adequately, as seen in Figures 4.10 and 4.12. Therefore, we recommend that the company to work with three rigs as the cost of one more rig might be much lower than the cost of working with two rigs for 16 years and having an idle completion crew, instead of only 11 years.

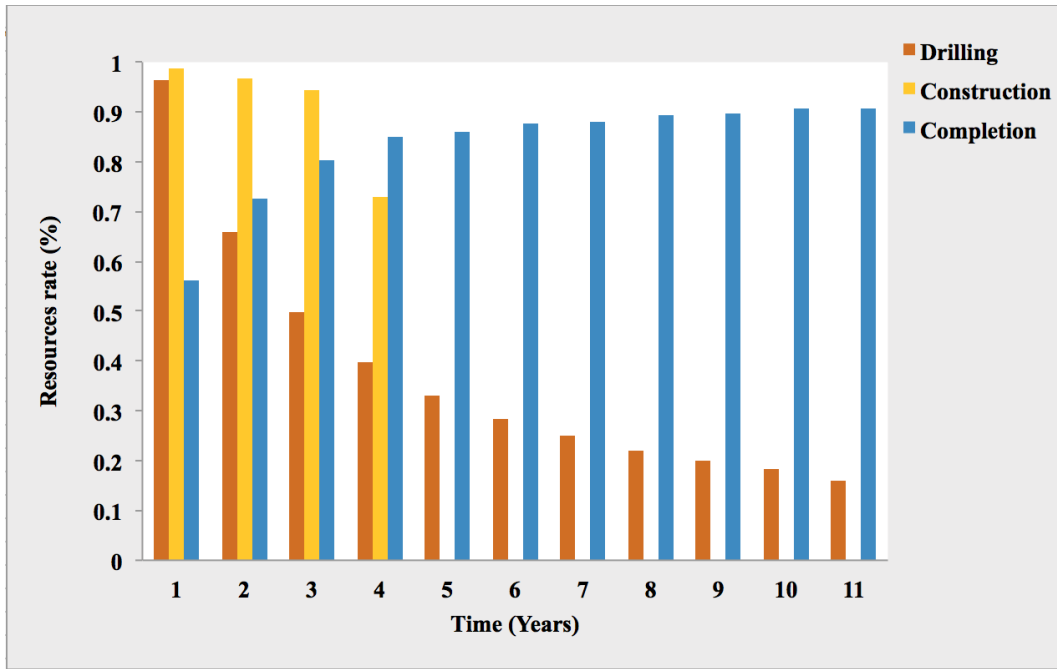


Figure 4.10: Utilization for each of the three resource types over the time horizon employing three rigs

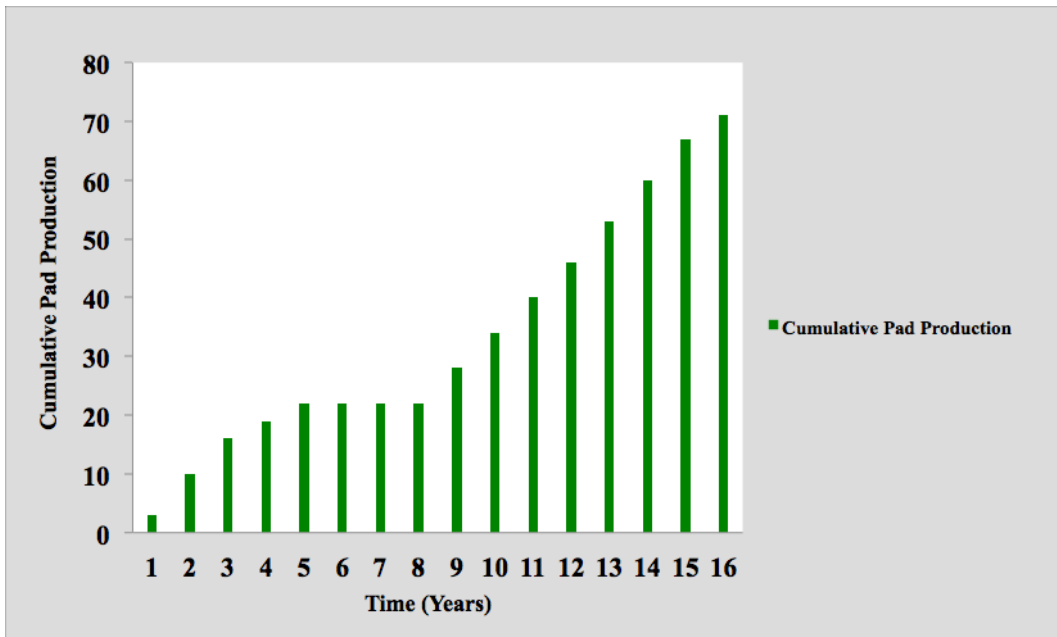


Figure 4.11: Cumulative pad production over the time horizon using two rigs

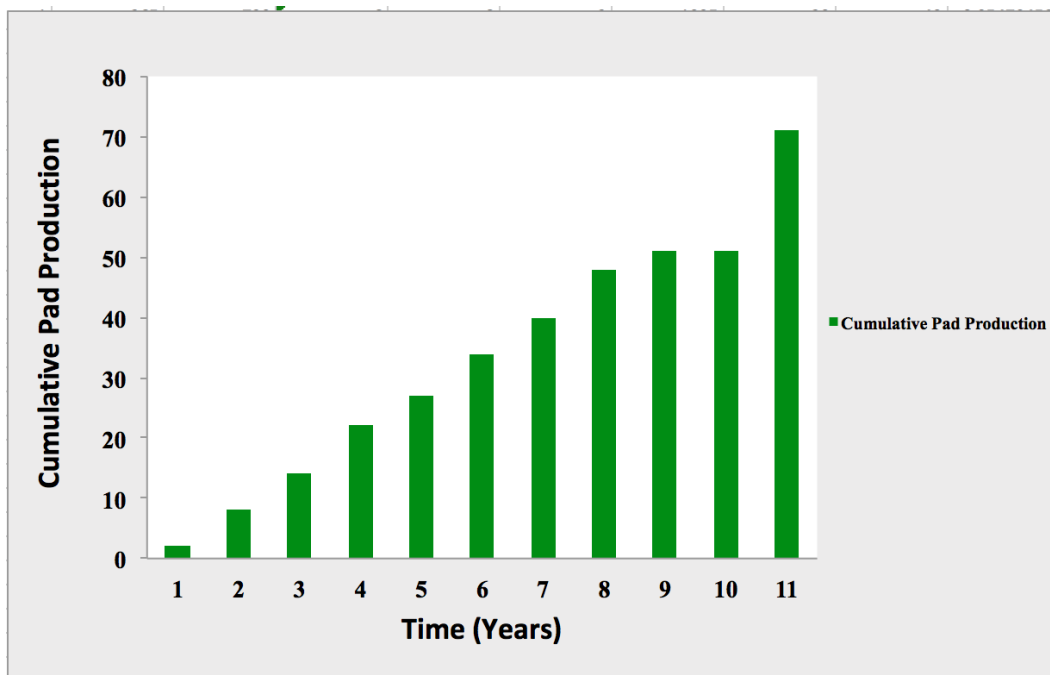


Figure 4.12: Cumulative pad production over the time horizon using three rigs

4.6 Conclusions

We report on the creation of an optimal operating plan in an unconventional oilfield. We use a procedure that has gained some traction in the extractive resource industry, but, more specifically, in mining operations. By drawing the appropriate analogies between that industry and oil and gas, we are able to: (i) determine that a formal optimization technique helps reduce non-productive time and thus save money, (ii) create a production schedule that determines the binding resources, and (iii) generate a blueprint regarding how to optimize the time and resources needed in this type of project. We are able to show that with the current optimization model, we can schedule activities on 71 pads having a total of 729 wells over 16 years. Additionally, extra resources, i.e., drilling rigs, might reduce operational time by five years to 11 years.

Our modeling paradigm of the resource-constrained production schedule problem, and its associated solution procedure, can be applied in a variety of other settings such as supplier selection, servicing companies, plant engineering and construction, and information systems.

Acknowledgements

We would like to thank the following people for their support of this project: Amanda Rebol of Kimmeridge; Oluwaseun Ogunmodede and Jennifer DiCarlo of Colorado School of Mines; Akshay Chowdu of South Dakota School of Mines; and, Marcos Goycoolea from the Universidad Adolfo Ibáñez for their insights and technical help.

CHAPTER 5

CONCLUSION

Quantitative modeling in the oil and gas industry is becoming increasingly common. This dissertation seeks a better understanding through empirical studies of the associated operations, and then employs several quantitative modeling techniques. Specifically, Chapter 2 provides an empirical study of the fluid flow in unconventional reservoirs, examining the Barree-Conway model. As of the time of this writing, this model is the only one that has been proven to be valid for a wide range of Reynolds numbers (Aljalalmah, 2014) and is therefore most applicable to the industry. However, it has not been fully implemented in most simulation software. We observe a deviation with respect to the Darcy and Forchheimer equations at high Reynolds numbers, which appears to be based on the existence of a minimum permeability plateau. The mechanism responsible for the minimum permeability plateau has been studied at the Colorado School of Mines for the last 10 years and this research confirms a trend observed in previously cited experiments.

Using data from these empirical experiments, we then introduce a statistical study to determine important factors in non-Darcy flow. The corresponding statistical techniques can be applied to various hydraulically fractured projects given availability of relevant data. This study quantifies the significant variables to achieve a given mass flow rate; these are:

- Material used
- Apparent permeability
- Pack permeability
- Viscosity

The study thereby produces a regression equation that can be used in order to identify the mass flow rate based on specific reservoir characteristics. Regardless of whether or not the flow is Darcy or non-Darcy, other studies produce similar conclusions regarding the independent variables associated with the mass flow rate; the analysis in any case helps to characterize the critical recovery of hydrocarbons during the hydraulic fracturing process.

Finally, we use an existing, but specialized, algorithm applied in other extractive resource settings to so-called RESOURCE-CONSTRAINED PROJECT SCHEDULING PROBLEMS. By casting the determination of an optimal operating plan in an unconventional oilfield as such, we can minimize makespan while adhering to precedence and resource constraints. We demonstrate for our case study that makespan covers a 16-year horizon, but when more resources are available, the makespan can be reduced significantly. In this way, we not only determine an optimal operating strategy for a given set of conditions, but we can also suggest improvements to resource allocation that result in significant reduction in time to completion.

Future research could expand the applicability of statistical and optimization modeling techniques in the oil and gas industry to include other experimental setups and more detailed models. For example, regression models for independent variables other than mass flow rate could be developed, and scheduling in oilfields to maximize net present value, consider different types of resource constraints, or different operational strategies (e.g., “must-start-by-a-certain-time” activities) could be formulated and solved.

BIBLIOGRAPHY

- Al-Otaibi, A. M., Y.-S. Wu, et al. (2010). Transient Behavior and Analysis of non-Darcy Flow in Porous and Fractured Reservoirs According to the Barree and Conway Model. In *SPE Western Regional Meeting*. Society of Petroleum Engineers.
- Alhanai, W. T. et al. (2002). Management and Control of the Inventory of Problematic Wells; A Stochastic Process Approach. In *Abu Dhabi International Petroleum Exhibition and Conference*. Society of Petroleum Engineers.
- Aljalalahmah, F. A. (2014). *Experimental Characterization of the Barree and Conway (2004) Single-phase non-Darcy Flow Model in Various Hydraulic Fracturing Sands*. Ph. D. thesis, Colorado School of Mines. Arthur Lakes Library.
- Barree, R. and Conway (2004). Beyond Beta Factors: A Complete Model for Darcy, Forchheimer, and Trans-Forchheimer Flow in Porous Media. In *SPE Annual Technical Conference and Exhibition*. Society of Petroleum Engineers.
- Beckwith, R. et al. (2011). Proppants: Where in the world. *Journal of Petroleum Technology* 63(04), 36–41.
- Benson, S. M. (2005). Carbon Dioxide Capture and Storage in Underground Geologic Formations. In *The Pew Center on Global Climate Change and the National Commission on Energy Policy: “The 10-50 Solution: Technologies and Policies for a Low-Carbon Future”*.
- Bienstock, D. and M. Zuckerberg (2010). Solving LP Relaxations of Large-Scale Precedence Constrained Problems. In *International Conference on Integer Programming and Combinatorial Optimization*, pp. 1–14. Springer.
- Brickey, A., A. Chowdu, A. Newman, M. Goycoolea, and R. Godard (2021). Barrick’s Turquoise Ridge Gold Mine Optimizes Underground Production Scheduling Operations. *INFORMS Journal on Applied Analytics* 51(2), 106–118.
- Chen, K.-D., J.-Q. Chen, D.-F. Hong, X.-Y. Zhong, Z.-B. Cheng, Q.-H. Lu, J.-P. Liu, Z.-H. Zhao, and G.-X. Ren (2019). Efficient and high-fidelity steering ability prediction of a slender drilling assembly. *Acta Mechanica* 230(11), 3963–3988.
- Darcy, H. (1856a). The Public Fountains of the City of Dijon. *Victor Dalmont, Paris, France*.
- Darcy, H. P. G. (1856b). *Les Fontaines publiques de la ville de Dijon. Exposition et application des principes à suivre et des formules à employer dans les questions de distribution d’eau, etc.* V. Dalamont.
- DiCarlo, J., A. Eustes, G. Steeger, et al. (2019). A History of Operations Research Optimization in the Petroleum Industry. In *SPE Annual Technical Conference and Exhibition*. Society of Petroleum Engineers.
- Erturk, M. C., C. Sinayuc, et al. (2015). Production Performance Analysis of Unconventional Gas Reservoirs with Different well Trajectories and Completion Techniques. In *SPE Middle East Unconventional Resources Conference and Exhibition*. Society of Petroleum Engineers.
- Forchheimer, P. (1901). Wasserbewegung durch Boden. *Verein deutscher Ingenieure*, 45–50.
- Gaines, M. J., P. Neil, and D. Herrington (2013). Novel Drilling System Offers Rotary Steerable

- System Performance with Real Time Survey in Top Hole Sections While Reducing Overall Costs. In *SPE Annual Technical Conference and Exhibition*. Society of Petroleum Engineers.
- Gerhart, P. M., A. L. Gerhart, and J. I. Hochstein (2020). *Fundamentals of Fluid Mechanics*. Wiley.
- Ghaeli, M. (2019). A Dynamic Programming Approach for Resource Allocation in Oil and Gas Industry. *Journal of Project Management* 4(3), 213–216.
- Hannan, B. and L. Bridwell (2012). Oil Price Volatility: How Queuing Theory can Explain and Warn of Impending Problems in our Energy System. *Pace University* 10(2).
- Johnson, T. J. R. (1967). *An Algorithm for the Resource Constrained Project Scheduling Problem*. Ph. D. thesis, Massachusetts Institute of Technology.
- Johnson Jr, J., G. W. Schein, T. Moser, S. Hayes, H. D. Liles, S. W. Webb, et al. (1999). High Efficiency Drilling-A Novel Approach for Improved Horizontal and Multi-Lateral Drilling. In *SPE Mid-Continent Operations Symposium*. Society of Petroleum Engineers.
- Keefer, D. L., J. L. Corner, and C. W. Kirkwood (2000). *Decision Analysis Applications in the Operations Research Literature, 1990-1999*. Department of Management Systems, University of Waikato.
- King, B., M. Goycoolea, and A. Newman (2017). Optimizing the Open Pit-to-Underground Mining Transition. *European Journal of Operational Research* 257(1), 297–309.
- Kolisch, R., A. Sprecher, and A. Drexel (1995). Characterization and Generation of a General Class of Resource-Constrained Project Scheduling Problems. *Management Science* 41(10), 1693–1703.
- Kuchta, M., A. Newman, and E. Topal (2004). Implementing a Production Schedule at LKAB's Kiruna Mine. *Interfaces* 34(2), 124–134.
- Lai, B. (2010). *Experimental Measurements and Numerical Modeling of High Velocity Multiphase non-Darcy Flow Effects in Porous Media*. Ph. D. thesis, Colorado School of Mines.
- Lai, B., J. L. Miskimins, et al. (2010). A New Technique for Accurately Measuring Two-Phase Relative Permeability Under non-Darcy Flow Conditions. In *SPE Annual Technical Conference and Exhibition*. Society of Petroleum Engineers.
- Lea Jr, J. F. and L. Rowlan (2019). *Gas well deliquification*. Gulf Professional Publishing.
- Lopez-Hernandez, H. D. (2007). *Experimental Analysis and Macroscopic and Pore-Level Flow Simulations to Compare non-Darcy Flow Models in Porous Media*. Ph. D. thesis, Colorado School of Mines. Arthur Lakes Library.
- Miskimins, J. L., H. D. J. Lopez, R. D. Barree, et al. (2005). Non-Darcy Flow in Hydraulic Fractures: Does it Really Matter? In *SPE Annual Technical Conference and Exhibition*. Society of Petroleum Engineers.
- Modeland, N., D. Buller, and K. K. Chong (2011). Statistical Analysis of the Effect of Completion Methodology on Production in the Haynesville Shale. In *North American Unconventional Gas Conference and Exhibition*.
- Mohammed, O. Q., R. Kassim, L. K. Britt, and S. Dunn-Norman (2017). Combining Statistical

- Analysis with Simulation to Optimize Unconventional Completions-Upper and Lower Montney Formations, Canada. In *Unconventional Resources Technology Conference, Austin, Texas, 24-26 July 2017*, pp. 705–719. Society of Exploration Geophysicists.
- Muñoz, G., D. Espinoza, M. Goycoolea, E. Moreno, M. Queyranne, and O. R. Letelier (2018). A Study of the Bienstock–Zuckerberg algorithm: Applications in Mining and Resource Constrained Project Scheduling. *Computational Optimization and Applications* 69(2), 501–534.
- Muskat, M. (1946). The Flow of Homogeneous Fluids Through Porous Media. Technical report.
- Mydland, S., C. H. Whitson, M. L. Carlsen, M. M. Dahouk, and I. Yusra (2020). Black-Oil and Compositional Reservoir Simulation of Gas-Based EOR in Tight Unconventionals. In *Unconventional Resources Technology Conference, 20–22 July 2020*, pp. 2745–2774. Unconventional Resources Technology Conference (URTeC).
- Ogunmodede, O., P. Lamas, A. Brickey, G. Bogin, and A. Newman (2021). Underground Production Scheduling with Ventilation and Refrigeration Considerations. Submitted.
- O’Sullivan, D. and A. Newman (2014). Extraction and Backfill Scheduling in a Complex Underground Mine. *Interfaces* 44(2), 204–221.
- Patterson, J. H. (1984). A Comparison of Exact Approaches for Solving the Multiple Constrained Resource, Project Scheduling Problem. *Management Science* 30(7), 854–867.
- Raleigh, J., D. Flock, et al. (1965). Operation Engineering is not Newton’s Research in Petroleum Production. *Journal of Canadian Petroleum Technology* 4(01), 1–4.
- Rardin, R. L. (2017, Ch 9). *Optimization in Operations Research*, Volume 166. Prentice Hall Upper Saddle River, NJ.
- Rivera, O., M. Goycoolea, E. Moreno, and D. Espinoza (2015). The OMP guide.
- Roy, D., A. Gupta, and R. B. De Koster (2016). A non-Linear Traffic Flow-Based Queuing Model to Estimate Container Terminal Throughput with AGVs. *International Journal of Production Research* 54(2), 472–493.
- Saldungaray, P. M., T. Palisch, R. Shelley, et al. (2013). Hydraulic Fracturing Critical Design Parameters in Unconventional Reservoirs. In *SPE Unconventional Gas Conference and Exhibition*. Society of Petroleum Engineers.
- Settari, A. et al. (1980). Simulation of Hydraulic Fracturing Processes. *Society of Petroleum Engineers Journal* 20(06), 487–500.
- Sinha, S., D. Devegowda, B. Deka, et al. (2016). Multivariate Statistical Analysis for Resource Estimation in Unconventional Plays Application to Eagle Ford Shales. In *SPE Eastern Regional Meeting*. Society of Petroleum Engineers.
- Tague, J. et al. (2000). Multivariate Statistical Analysis Improves Formation Damage Remediation. In *SPE Annual Technical Conference and Exhibition*. Society of Petroleum Engineers.
- Talbot, F. B. (1982). Resource-Constrained Project Scheduling with Time-Resource Tradeoffs: The Nonpreemptive Case. *Management Science* 28(10), 1197–1210.

- Walls, M. R., G. T. Morahan, and J. S. Dyer (1995). Decision Analysis of Exploration Opportunities in the Onshore US at Phillips Petroleum Company. *Interfaces* 25(6), 39–56.
- Wang, S. and S. Chen (2017). A Novel Bayesian Optimization Framework for Computationally Expensive Optimization Problem in Tight Oil Reservoirs. In *SPE Annual Technical Conference and Exhibition*.
- Wen, Z., L. J. Durlofsky, B. Van Roy, K. Aziz, et al. (2011). Use of Approximate Dynamic Programming for Production Optimization. In *SPE Reservoir Simulation Symposium*. Society of Petroleum Engineers.

APPENDIX A

SUPPLEMENTAL DATA FOR CHAPTER 2

Calculations Made for the Graphs

In order to make the graphs used in this chapter, we were able to use the direct output of the Halliburton AFPL and in some instances, we needed to make calculations for certain graphs.

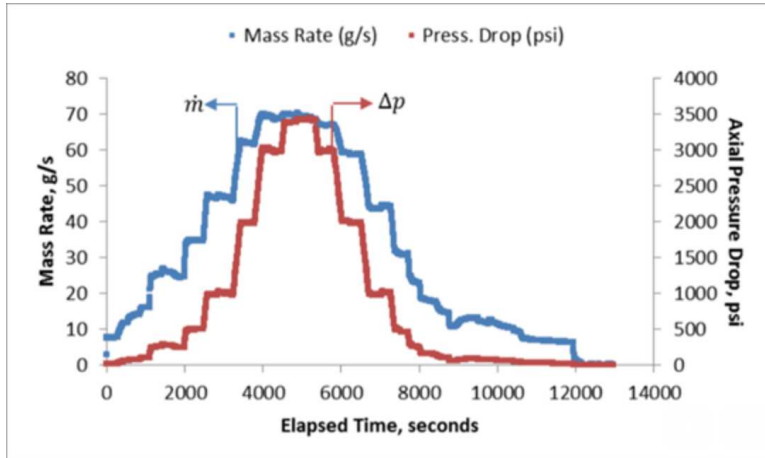


Figure A.1: Figure showing the highest differential pressure as a function of the mass flow rate

The Figure 1 was made using the \dot{m} from the output system at the APFL.

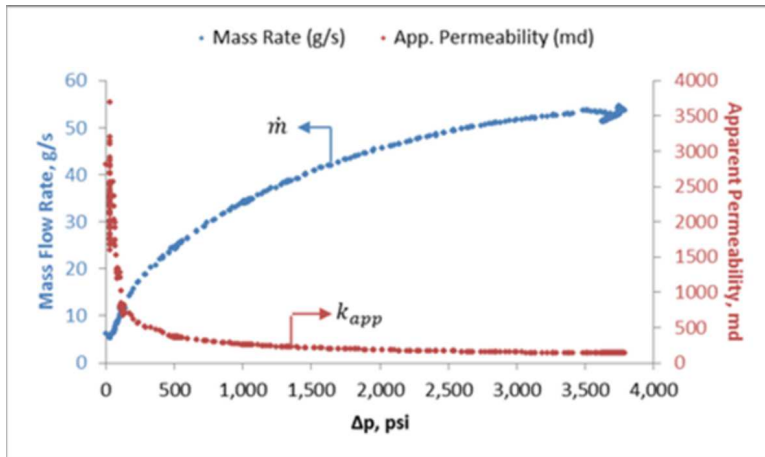


Figure A.2: Figure showing the highest differential pressure as a function of the mass flow rate

The Figure 2 shows the apparent permeability as calculated using the formula as seen below:

$$k_{app} = q_{sc}(\mu B)_e L / (A \cdot \Delta m(p))$$

where,

$$\Delta m(p) = m(p_i) - m(p_o)$$

$$m(p)|_{T=\bar{T}} = (\overline{\mu B})_e \int_{p_h}^p \frac{1}{\mu B} dp + p_b$$

Figure 2 is graphed as a function of the axial pressure drop and the axial pressure.

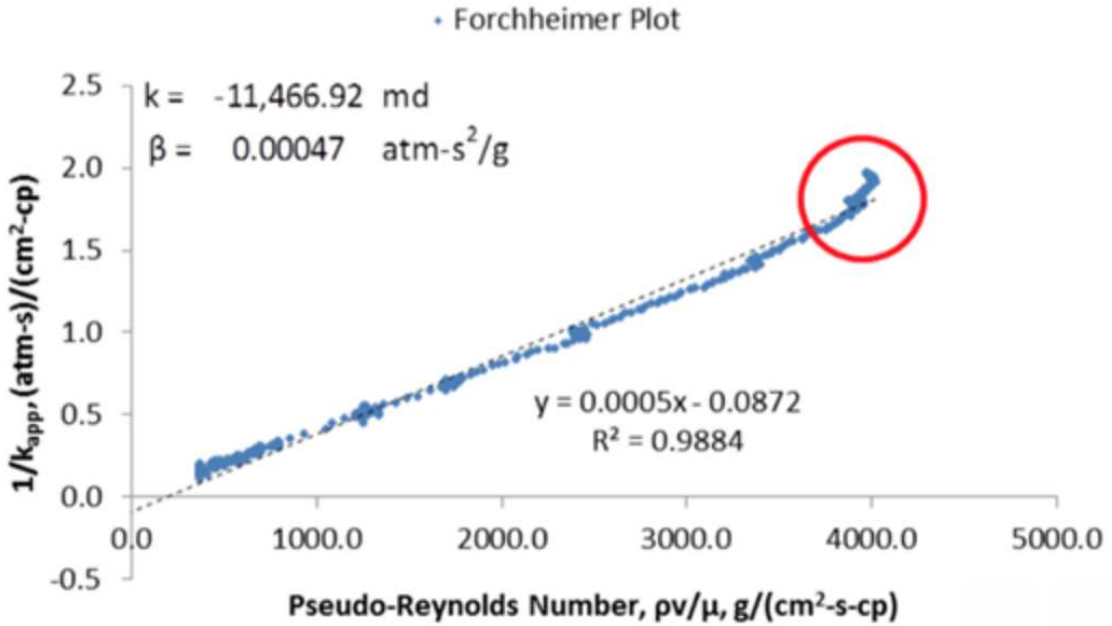


Figure A.3: Figure showing the Pseudo-Reynolds number as a function of the inverse permeability

Figure 3 shows the classic Forchheimer plot where we have the reciprocal of the apparent permeability $1/k_{app}$ as a function of the core-averaged pseudo-Reynolds number $\bar{\rho}v/\bar{\mu}$

To compute the core-average pseudo-Reynolds number, we used the following equality:

$$\bar{\rho}v/\bar{\mu} = \dot{m}/(\bar{\mu}A)$$

where,

$$\begin{aligned}\bar{T} &= (T_i - T_o) / 2 \\ \bar{p} &= \frac{2}{3} \left(p_i + p_o - \frac{p_i p_o}{p_i + p_o} \right) \\ \bar{\rho}(\bar{p}, \bar{T}) &= \frac{\bar{p} M_w}{z(\bar{p}, \bar{T}) R \bar{T}} \\ \bar{\mu}(\bar{T}, \bar{\rho}) &= \mu(\bar{T}, \bar{\rho}(\bar{p}, \bar{T})) \\ \frac{\Delta m(p)}{L} &= \frac{\mu}{k_d} v + \beta \rho v^2 + \gamma \rho^2 v^3\end{aligned}$$

The area was calculated using the following formula: $A = \pi D^2 / 4$

$$\bar{T} = \frac{1}{n} \sum_{j=1}^n (T_i, T_o)_j = \text{average of all inlet and outlet temperatures}$$

$$B = \rho_{sc} / \rho$$

$$\rho(p, \bar{T}) = \frac{p M_w}{z(p, \bar{T}) R \bar{T}}$$

$$q_{sc} = \dot{m} / \rho_{sc}$$

APPENDIX B

R CODE

```
1  library(randomForest)
2  require(caTools)
3  library(dplyr)
4  library(caret)
5  library(data.table)
6
7
8  setwd("/Users/kamgalois/Desktop")
9  data = read.csv("dataFlow.csv", sep = ";")
10
11 str(data)
12
13 data <- na.omit(data)
14
15 nameTest <- data$Name.Tester
16 material <- data$Material.Used
17
18 data <- data %>%
19   select(-c("Name.Tester", "Material.Used"))
20 data <- data.frame(apply(data, 2, function(x) as.numeric(sub(",",".",x))))
21
22 data$Name.Tester <- nameTest
23 data$Material.Used <- material
24
25 data$Name.Tester <- as.numeric(as.factor(data$Name.Tester))
26 data$Material.Used <- as.numeric(as.factor(data$Material.Used))
27
28 str(data)
29
30 data$0..Test.No. <- NULL
```

```

31 data$Mass.Flowrate.m..gm.sec. <- NULL
32
33 #-----#
34 # Significane of all parameters in general
35 #-----#
36
37 linear_model_all = lm(Mass.Flowrate.m..gm.sec..1 ~ ., data=data)
38 s_all <- summary(linear_model_all)
39 s_all
40
41 # We eliminate variables with NAs:
42
43 naTo_remove <- c("Initial.Tank.Temp..TTi..oR.", "Final.Tank.Temp..TTf..oR
44 .", "P5..psia."
45 , "P5..atm.", "ECS..psig..1", "P5..psia.", "P5..atm.", "ECS..psig..1", "P.Psia
46 .", "ECP.psia.",
47 "Axial.Pressure", "ravg..gm.cc...Cubic.Average", "CSAavg..cm2.")
48
49 data <- data %>%
50 select(-naTo_remove)
51
52 str(data)
53
54
55 linear_model_nNA1 = lm(Mass.Flowrate.m..gm.sec..1 ~ ., data=data)
56 s_nNA_1 <- summary(linear_model_nNA1)
57 s_nNA_1
58
59 # We eliminate variables with NAs:
60 #-----#
61 # This part is the same as the last one
62 # We eliminate variables that had NAs
63
64 naTo_remove_2 <- c("P1..psia.", "P1..atm.", "ECS..psig.")
65 data <- data %>%

```

```

62 select(-naTo_remove_2)
63 str(data)
64
65 linear_model_nNA2 = lm(Mass.Flowrate.m..gm.sec..1 ~ ., data=data)
66
67 s_nNA_2 <- summary(linear_model_nNA2)
68 s_nNA_2
69 #sink("significanceGeneral.txt")
70 #print(s_nNA_2)
71 #sink()
72 p_values_G <- as.data.frame(s_nNA_2$coefficients[,4])
73
74 setDT(p_values_G, keep.rownames = TRUE)[]
75
76 # Then we change the names of columns
77
78 colnames(p_values_G) <- c("parameter","p_value")
79
80 # Then we order the data Table of p_values from smallest to biggest
  value
81 p_values_G <- p_values_G[order(p_values_G$p_value, decreasing = FALSE),]
82
83
84 jpeg("graph.jpg", width = 10000, height = 5000)
85 ggplot(p_values_G,aes(x = parameter, y = log(1/p_value), color =
  parameter, fill=parameter))
86 + geom_bar(stat = "identity") +
87 ggtitle("Significane of the different parameters on predicting the flow"
  ) +
88 theme(legend.text = element_text(size = 100),legend.title = element_text
  (size =100),
89 axis.text=element_text(size=100), axis.text.x = element_text(angle = 90,
  hjust = 1))

```

```

90 dev.off()
91
92
93
94 # Now you go to the place where the dataset is, you will find the jpg
    file
95 # The p_value, the smallest, the more significant the variable the first
    is the most
96 # significant, then the second etc
97 # Now in the graph if you see in line 176, we computed log(1/p_value),
    which means
98 # smaller p_value gives higher log(1/p_value),
99 # I did this to get a graph, where the biggest bar in the barplot, means
    that
100 # the variable is more significant
101
102 # We will work with this model under the name: General Model
103
104 #-----#
105 # We take just the variables: - "Mass flowrate", "Material used", "Kapp"
    AND "mavg (cp)"
106 # We call this model ImportantModel
107 #-----#
108 # In this part, we are doing the exact same thing but just taking the
    four variables
109 varTotake <- c("Mass.Flowrate.m..gm.sec..1", "Material.Used", "Kapp", "
    Kapp.1", "mavg..cp.")
110 data_Important <- data %>%
111 select(varTotake)
112 str(data_Important)
113
114 linear_model_import = lm(Mass.Flowrate.m..gm.sec..1 ~ ., data = data_
    Important)

```

```

115
116 s_import <- summary(linear_model_import)
117 s_import
118
119 # Flow rate = -64.89 - 0.21*Material.Used - 0.294*Kapp + 3.44*Kapp.1 + 2
      687*navg..cp.
120 # That's the equation
121
122 # =====> the result is of course different, we didn't use as many
      predictors as before,
123 # and I don't know if you understand the math
124 # behind a linear model, but it tries to minimize the error function
      which is the sum
125 # of the
126 #  $(Y_i - (a_1*Variable_1 + a_2*Variable_2 + \dots + a_N*Variable_N))^2$ 
      ////////////////
127 # This minimization is done using the values of
128 # the variables we considered, and by trying the coefficients a_1...a_N
      that give us
129 # the minimum value of the error
130 # So if you consider different Variables (Variable_1...Variable_N) you
      will get
131 # different coefficients, hence a different result.
132
133 #sink("significanceGeneral.txt")
134 print(s_import)
135 #sink()
136
137 # Plot significance, we plot the log(1/p_value), that way we have a
      barplot with values
138 # from the biggest to
139 # Smallest, we chose this transformation because the smaller p_value the
      better,

```

```

140 # so a "more significant variable"
141 # is one with the biggest value on the barplot
142 p_values_import <- as.data.frame(s_import$coefficients[,4])
143 setDT(p_values_import, keep.rownames = TRUE)[]
144 colnames(p_values_import) <- c("parameter", "p_value")
145
146 jpeg("file.jpg")
147 ggplot(p_values_import, aes(x = parameter, y = log(1/p_value), color =
148   parameter,
149   fill=parameter)) +
150   geom_bar(stat = "identity") + ggtitle("Significane of the different
151     parameters
152     on predicting the flow")
153 dev.off()
154
155 # Testing linear prediction model
156 #-----#
157 # I just didn't remove this line, but to explain: this does not split
158 # the data,
159 # it just creates a vector assigning TRUE to
160 # 75% of it, and FALSE to the remaining 25%
161 # sample = sample.split(data$Mass.Flowrate.m..gm.sec..1, SplitRatio = .7
162 #   5)
163 # =====> The 75% is just a choice, I could've used 80%, or 70%.
164 # General Model
165 #-----#
166
167 # This line is set to have the same randomness
168 # When the data partition is done in R, there is a randomness factor, I

```

```

    mean normally
169 # if we run the splitting lines
170 # two consecutive time we would not get the same result, but when we set
    the number
171 # in set.seed() we can have the same results
172 # 123 is also just a choice, you can put every number you want
173 set.seed(123)
174 training.samples <- data$Mass.Flowrate.m..gm.sec..1 %>%
175 createDataPartition(p = 0.75, list = FALSE)
176
177 # We take the samples of training and assign them to train.data
178 train.data <- data[training.samples, ]
179
180 # We take the samples of testing and assign them to test.data
181 test.data <- data[-training.samples, ]
182
183 # Build the model
184 # These small point refers to all except the flowrate
185 linear_modelG <- lm(Mass.Flowrate.m..gm.sec..1 ~., data = train.data)
186
187 # Make predictions and compute the R2, RMSE and MAE
188 predictions <- linear_modelG %>% predict(test.data)
189
190
191 # true_Test is just a vector contains the actual values of the Flow in
    the data, in order
192 # to compare them with predictions after
193 # the [-which(is.na(predictions))] is used to remove NAs from this
    vector
194 true_Test <- test.data$Mass.Flowrate.m..gm.sec..1[-which(is.na(
    predictions))]
195 predictions <- predictions[-which(is.na(predictions))]
196

```

```

197 data.frame( R2 = R2(predictions, true_Test),
198 RMSE = RMSE(predictions, true_Test),
199 MAE = MAE(predictions, true_Test))
200
201 # RMSE = sqrt(sum(trueValue - predictedValue)), The smaller the better
202 # R2 is an assessment of the model, closer to 1 means better model
203
204 # As I said before, R2 after testing is actually excellent, the
    interpretation here
205 # is that the variables (predictors) used
206 # are perfect to predict the flow
207
208
209 # ImportantModel
210 #-----#
211
212 # We do the same thing as for the general model
213 train.dataI <- data_Important[training.samples, ]
214 test.dataI <- data_Important[-training.samples, ]
215 # Build the model
216 linear_modelI <- lm(Mass.Flowrate.m..gm.sec..1 ~., data = train.dataI)
217
218 # Make predictions and compute the R2, RMSE and MAE
219 predictionsI <- linear_modelI %>% predict(test.dataI)
220 true_TestI <- test.dataI$Mass.Flowrate.m..gm.sec..1[-which(is.na(
    predictionsI))]
221 predictionsI <- predictionsI[-which(is.na(predictionsI))]
222
223 data.frame( R2 = R2(predictionsI, true_TestI),
224 RMSE = RMSE(predictionsI, true_TestI),
225 MAE = MAE(predictionsI, true_TestI))
226
227

```

```

228 # Here as I explained before, we are doing the same thing but with just
      the variables you
229 # specified in particular.
230 # The result of course would be different, because we have just a subset
      of the variables.
231 # You could see it in the R2 result, which is different
232
233 #-----#
234 # Random Forest, (although it is not a classification problem)
235 #-----#
236
237 # General Model
238 #-----#
239
240 # In here we build the model
241 rf_ModelG <- randomForest(Mass.Flowrate.m..gm.sec..1 ~., data = train.
      data)
242
243 # Random forest does not give an equation, as I explained in the pdf
      file.
244 # We make the predictions
245 predictionsRF <- rf_ModelG %>% predict(test.data)
246 true_TestRF <- test.data$Mass.Flowrate.m..gm.sec..1[-which(is.na(
      predictionsRF))]
247 predictionsRF <- predictionsRF[-which(is.na(predictionsRF))]
248
249 # Then we see the accuracy
250 data.frame( R2 = R2(predictionsRF, true_TestRF),
251            RMSE = RMSE(predictionsRF, true_TestRF),
252            MAE = MAE(predictionsRF, true_TestRF))
253
254 # Random forest models are not used to see what variable influences the
      most, it is used

```

```
255 # only for prediction
256 # And for Classification, not this type of problems, but what it CAN do
    is that it can predict
257 # the output of the Flow
258 # given the entry variables. You give values of the predictors, and it
    returns a prediction
259 # of the value of the flow
260
261 # Random forest models are called black box models, they do not give an
    equation
262
263 # In this model, I used all of the variables, with just the training set
    to do the ML
264 # testing of how the model is could
265 # Basically, we do the model, we predict and compare with the actual
    values everytime
266 # The random forest technique is a quiete complex model, I will give you
    the basic ideas.
267 # First random forest is a model we get from a model called Decision
    trees.
268 # Decision trees gives a tree of judgement over the values of the data,
    for example
269 # It says if variable_1 > 50 and Variable_2 < 20 I will have class 1
270 # if variable_1 <= 50 and Variable_2 < 20 I will have class 2
271 # Etc ...
272 # Random forest is a technique which builds many decision trees, makes
    the prediction,
273 # and every tree votes for the class it got
274 # The result with the most votes is the one predicted.
275
276
277
278 # ImportantModel
```

```

279 #-----#
280
281 rf_ModelI <- randomForest(Mass.Flowrate.m..gm.sec..1 ~., data = train.
    dataI)
282
283 # We have data is your excel sheet, right?
284 # so let's take a sample
285
286
287 predictionsRF_I <- rf_ModelI %>% predict(test.dataI)
288 true_TestRF_I <- test.data$Mass.Flowrate.m..gm.sec..1[-which(is.na(
    predictionsRF_I)]]
289 predictionsRF_I <- predictionsRF_I[-which(is.na(predictionsRF_I)]]
290
291 data.frame( R2 = R2(predictionsRF_I, true_TestRF_I),
292   RMSE = RMSE(predictionsRF_I, true_TestRF_I),
293   MAE = MAE(predictionsRF_I, true_TestRF_I))
294
295
296
297 ###=====> The results are different in terms of model accuracy,
    you can see it
298 # from R2 and RMSE
299 # The result here in random forest is just the accuracy of the model,
    you cannot
300 # get any other information
301
302 write.csv(data, "puromycin_data.csv")
303

```

Listing B.1: R code

APPENDIX C

SUPPLEMENTAL FORMULATION FOR CHAPTER 4

Appendix - Formulation

The model formulation that OMP solves is as follows.

Indices and sets:

- $a \in \mathcal{A}$ an activity within the set of all activities
- $\bar{a} \in \bar{\mathcal{A}}_a$ an activity \bar{a} within the set of predecessor activities to activity a
- $r \in \mathcal{R}$ a resource within the set of resources, such as number of drill rigs and crew,
whose limits are enforced on a daily basis
- $t \in \mathcal{T}$ a day within the set of daily time periods

Parameters:

- q_{ra} daily consumption of resource r associated with completing activity a [people, number]
- \bar{r}_{rt} maximum amount of resource r available on day t [people, number]
- d_a duration of activity a [days]
- $d_{\bar{a}}$ duration (including mandatory delay, or lag) of activity a [days]

Decision variables:

- X_{at} 1 if activity a is completed by the end of time t , 0 otherwise

$$\text{(RCPSP)} \quad \min \left(\max_{a \in \mathcal{A}} \left\{ \sum_{t \in \mathcal{T}} t \cdot (X_{at} - X_{a,t-1}) \right\} \right) \quad (\text{C.1a})$$

$$\text{sst} \quad X_{a,t-1} \leq X_{at} \quad \forall a \in \mathcal{A}, t \in \mathcal{T} \quad (\text{C.1b})$$

$$X_{a|\mathcal{T}|} = 1 \quad \forall a \in \mathcal{A} \quad (\text{C.1c})$$

$$X_{at} \leq X_{\bar{a},t-d_{\bar{a}}} \quad \forall a \in \mathcal{A}, \bar{a} \in \bar{\mathcal{A}}_a, t \in \mathcal{T} \quad (\text{C.1d})$$

$$\sum_{a \in \mathcal{A}} \frac{q_{ra}}{d_a} (X_{at} - X_{a,t-d_a}) \leq \bar{r}_{rt} \quad \forall r \in \mathcal{R}, t \in \mathcal{T} \quad (\text{C.1e})$$

$$X_{at} \text{ binary} \quad \forall a \in \mathcal{A}, t \in \mathcal{T} \quad (\text{C.1f})$$

The objective (C.1a) minimizes makespan, which is expressed as the time period in which the last activity in the sequence is completed. Constraints (C.1b) ensure that once an activity is completed at time $t-1$, it remains completed for all future time periods $t, \dots, |\mathcal{T}|$, and constraints

(C.1c) require that all activities are completed by the end of the time horizon; this latter constraint is usually absent from models that consider the maximization of net present value in their objectives, but is required to preclude the trivial solution for those that minimize makespan. Constraints (C.1d) enforce precedence between an activity a and its predecessors \bar{a} , such that a cannot start unless \bar{a} starts sufficiently early that, when accounting for its duration, it is finished by the time a starts. Constraints (C.1e) constitute knapsacks and ensure that the amount of a resource of a particular type consumed by all activities on any given day cannot exceed the availability of said resource. All variables are required to be binary by constraints (C.1f). The instances were run on a MacBook Pro with a 2.6 GHz 6-core intel processor and 16 GB of RAM, and required just minutes of solution time.

APPENDIX D

COPYRIGHT IMAGES

Permissions for reuse of copyrighted materials (Figures D.1, D.2, D.3, D.4, D.5, D.6)

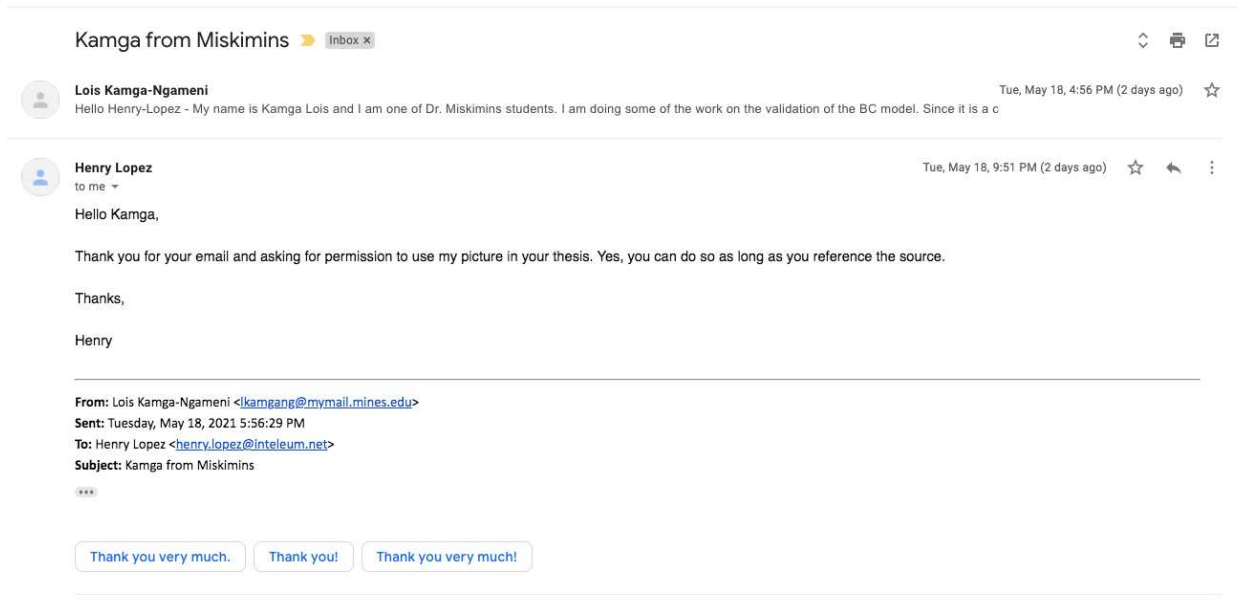


Figure D.1: Approval email from Henry Lopez for Figures 2.1, 2.2, 3.1, 3.2, and 3.3

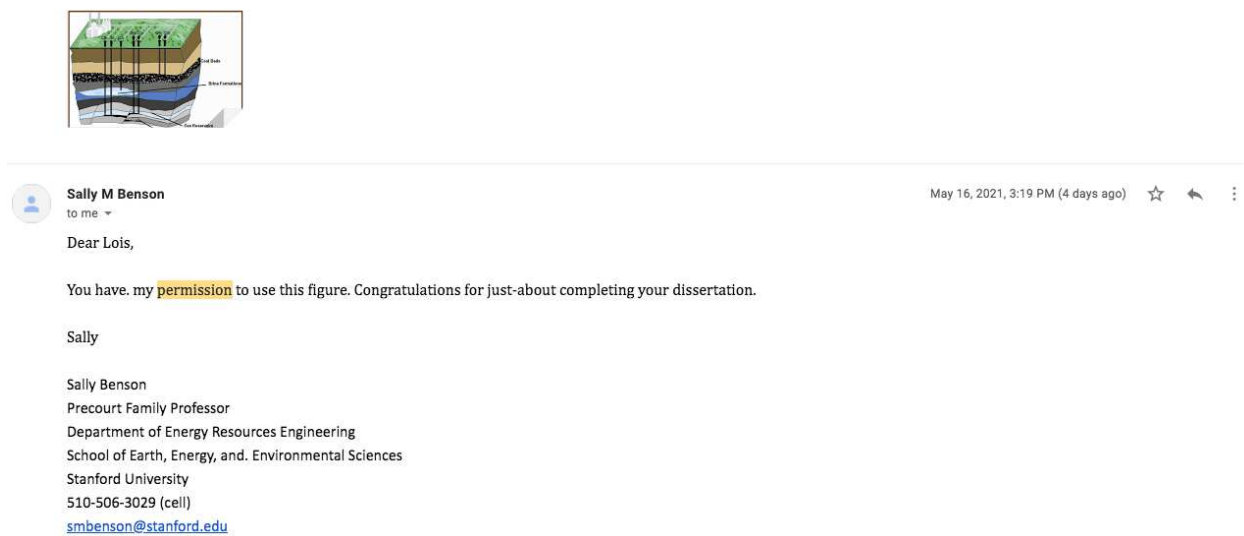


Figure D.2: Approval email from Sally Benson for Figure 4.1



刘佳鹏
to me, zhaozh, rengx ▾

Wed, May 19, 1:31 PM (5 days ago) ☆ ↶ ⋮

Hi Kamga,

Thank you very much for your email. You can re-use this figure with appropriate citation and acknowledgement by following the guidelines of your school.

Best
Jiapeng

Figure D.3: Approval email for Figure 4.2

Dear Mr. Kamga Ngameni,

Thank you for placing your order through Copyright Clearance Center's RightsLink[®] service.

Order Summary

Licensee:	Colorado School of mines
Order Date:	May 20, 2021
Order Number:	5073111140746
Publication:	Elsevier Books
Title:	Hydraulic Rig Technology and Operations
Type of Use:	reuse in a thesis/dissertation
Order Ref:	1
Order Total:	0.00 USD

View or print complete [details](#) of your order and the publisher's terms and conditions.

Sincerely,

Copyright Clearance Center

Figure D.4: Approval email from Rightlink for Figure 4.3

Dear Mr. Kamga Ngameni,

Thank you for placing your order through Copyright Clearance Center's RightsLink® service.

Order Summary

Licensee:	Colorado School of mines
Order Date:	May 17, 2021
Order Number:	5071560468493
Publication:	Elsevier Books
Title:	Gas Well Deliquification
Type of Use:	reuse in a thesis/dissertation
Order Total:	0.00 USD

View or print complete [details](#) of your order and the publisher's terms and conditions.

Sincerely,

Copyright Clearance Center

Figure D.5: Approval email from rightlink for Figure 4.5



Ely Lolon via libertyoilfieldservices.onmicrosoft.com
to me ▾

Sun, May 16, 3:18 PM (4 days ago) ☆ ↶ ⋮

Lois,

I've attached a picture of an LOS frac crew. Just acknowledge in your thesis that the picture is a courtesy of Liberty Oilfield Services if you use it.

We do not do flowback operations. You could contact one of the operators that do flowback.

Regards,

Ely



Figure D.6: Approval email from Liberty Oilfield for Figure 4.4, 4.6, and 4.7