Fall 2014

Data mining and statistical analysis of completions in the Canadian Montney formation

Mustafa Adil Al-Alwani

DATA MINING AND STATISTICAL ANALYSIS OF COMPLETIONS IN THE CANADIAN MONTNEY FORMATION

by

MUSTAFA ADIL AL-ALWANI

A THESIS

Presented to the Faculty of the Graduate School of the

MISSOURI UNIVERSITY OF SCIENCE AND TECHNOLOGY

In Partial Fulfillment of the Requirements for the Degree

MASTER OF SCIENCE IN PETROLEUM ENGINEERING
2014

Approved by

Shari Dunn-Norman, Advisor
Larry K. Britt
Runar Nygaard

ABSTRACT

This thesis documents a data-mining study and statistical analysis of well completion methods and their impact on production for more than 3300 horizontal wells in the Canadian Montney resource play.

The statistical software JMP is used to analyze well and production data for both horizontal Montney gas and oil wells, examining production trends with changes in completion parameters, such as the type of completion, fluid volume pumped, proppant load, number of fracture stages and completion costs. The analysis also provides a general understanding of average treatment characteristics, and how completions have changed with time for the Montney play.

Among the many results of this work, it is shown that there is a limit to adding stages to well completions in the Montney. While additional completed stages may increase cumulative recovery, the recovery per stage decreases after a point. This conclusion is consistent with recent findings (VISAGE and Jim Gouveia 2014). In addition, findings of the study clearly demonstrate that wells with the smallest frac fluid load recovery have the best cumulative recovery with time, and spending more for the completion translates into higher recovery.

This work is important as it is the first field-wide statistical review of wells completed in the Montney using large up to date dataset.

AKNOWLEDGMENTS

Foremost, I would like to express my sincere gratitude to the J. William Fulbright Foreign Scholarship Board and the U.S. Department of State for believing in me and granting me the honor of the Fulbright Scholarship. I am very proud to be a recipient of this life-changing award.

I would like to gratefully and sincerely thank Dr. Shari Dunn-Norman for approving me to join her research group and the trust she has honored me to work as a teaching assistant in her Artificial Lift class. Her guidance, understanding, patience and most importantly the selfless time and care she has offered made me feel like home and kept me going.

I would also like to thank my committee member Mr. Britt for teaching the Well Stimulation class during the Fall of 2013 and for all the help and data he provided throughout the research period. His knowledge was a tremendous help in the work.

Many thanks to Dr. Nygaard for agreeing to be a committee member in my research and his invaluable time and efforts in examining the thesis and all the constructive feedback.

A special thanks goes out to my research group members: Rashid Kassim, Chatetha Chumkratoke, Hussain Jawdat, Fen Yang, Abdulsalam Mohamed, Benjamin Pepple, Jinlin Zhang and Viriyah Thepporprapakorn for all the discussions and support.

Last but not least, a great thanks to my family. Words cannot express how grateful I am to my parents for all the support and encouragement. Without their prayers, I would not be here and sustain thus far.

TABLE OF CONTENTS

Page

ABSTRACT……………………………………………………………………………...iii

AKNOWLEDGMENTS iv

LIST OF ILLUSTRATIONS viii

LIST OF TABLES xiii

NOMENCLATURE xiv

SECTION

1. INTRODUCTION 1

1.1. MONTNEY PLAY 3

1.2. OBJECTIVES OF THE STUDY 6

2. LITERATURE REVIEW 7

3. METHODOLOGY 12

3.1. DATA MINING 12

3.2. BASIC STATISTICS 13

3.2.1. Mean. 14

3.2.2. Median. 14

3.2.3. Mode. 14

3.2.4. Standard Deviation. 14

3.2.5. Box Plot. 15

3.3. INTRODUCTION TO JMP 16

3.3.1. The JMP Data Table. 16

3.3.2. JMP Analyze Menu. 19

3.3.3. Data Visualization and Exploratory Data Analysis with JMP. 21

3.3.4. Data Selection and Management. 21

3.3.5. Linked Interactive Graphs and Analysis. 21

3.4. PHASES OF DATA MINING 25

3.4.1. Project Understanding Phase. 26

3.4.2. Dataset Understanding. 26

3.4.2.1 Fields in the Montney. 29
3.4.2.2 Major operators in the Montney. 30
3.4.2.3 Completion date. 32
3.4.2.4 Stimulation company. 33
3.4.2.5 Completion type. 35
3.4.2.6 Completed lateral length (m). 36
3.4.2.7 Number of stages. 36
3.4.2.8 Total fluids pumped (m^3). 38
3.4.2.9 Recovered load fluid (m^3). 39
3.4.2.10 Load fluid recovery percentage. 40
3.4.2.11 Normalized average fluid pumped per stage (m^3/stage). 41
3.4.2.12 Designed and pumped proppant (tonne). 42
3.4.2.13 Normalized average placed proppant per stage (tonne/stage). 43
3.4.2.14 Normalized proppant pumped per length (tonne/m). 44
3.4.2.15 Average fracturing spacing (m). 45
3.4.2.16 Fracture closure gradient (Kpa/m). 46
3.4.2.17 Avg. proppant concentration (lbs./gal) 47
3.4.2.18 Base fluid group. 48
3.4.2.19 Fracturing base fluid. 50
3.4.2.20 Fracturing fluid energizers. 50
3.4.3. Dataset Preparation and Validation Phase. 51
3.4.3.1 Validation of pumped fluid parameter. 51
3.4.3.2 Validation of completion cost data. 55
3.4.3.3 Validation of number of stages data. 58
3.4.3.4 Validation proppant data. 61
3.4.4. Final Parameters Selections and Normalizations. 65
4. ANALYSIS AND DISCUSSIONS 68
4.1. DEVELOPING GENERAL ANALYSIS TECHNIQUES 68
4.1.1. Statistical Techniques. 69
4.1.2. Simple Cross Plots and Heat Map Techniques. 84
4.2. ANALYSIS OF PARAMETERS OVER TIME 106

4.2.1. Completed Length (m). 106
4.2.2. Number of Stages. 106
4.2.3. Total Proppant Placed (tonne). 108
4.2.4. Total Pumped Fluid (M^3). 108
4.2.5. Type of Completion. 109
4.2.6. Proppant Concentration (lb/gal). 109
4.2.7. 18 Months Cumulative Gas Production (mmcf). 110
4.2.8. IP Gas (mcf/D). 110
4.2.9. Drilling and Completion AFE (K$). 111
4.2.10. Final Drilling and Completion Cost (K$). 111
4.2.11. Fracture Spacing (m). 112
4.3. COST ANALYSIS 112
5. CONCLUSIONS 115
5.1. PARAMETERS APPLICATIONS 115
5.2. FORMATION WIDE PERFORMANCE 116
6. FUTURE WORK 118
APPENDIX 119
BIBLIOGRAPHY 120
VITA 122

LIST OF ILLUSTRATIONS

Figure Page

1.1 Generalized Map Showing the Location of the Montney Formation in the Subsurface of Alberta and British Columbia 5

3.1 Box Plot Diagram (Kirkman,1996) 16

3.2 JMP Data Table 17

3.3 JMP Tables Menu 18

3.4 Histogram Distribution Example of Continuous Variable 20

3.5 Histogram Distribution Example of Nominal Variable 20

3.6 JMP Graph Builder Interface 23

3.7 Montney Data Distribution Histogram of the Top 20 Fields in Well Count 30

3.8 Data Distribution Histogram of Top 20 Operators in the Montney Dataset 32

3.9 Completion Date Distribution Histogram 33

3.10 Stimulation Companies in the Montney Data Distribution Histogram 34

3.11 Well Completion Type Data Distribution Histogram in the Montney Dataset 35

3.12 Completed Lateral Length Data Distribution in the Montney Dataset 36

3.13 Attempted and Actual Stages Number Data Distribution Histogram and Statistical Summary Table 37

3.14 Total Pumped Fluid Data Distribution Histogram and Statistical Summary 38

3.15 Load Fluid Data Distribution Histogram 39

3.16 Recovered Load Fluid (M3) Data Distribution Histogram 40

3.17 Load Fluid Recovery Percentage Data Distribution Histogram 41

3.18 Load Fluid Recovery Percentage Data Distribution Histogram (Shaded with Respect to Fluid Types) 41

3.19 Avg. Total Pumped Fluid per Stage (m^3/stage) Data Distribution Histogram 42

3.20 Total Designed and Placed Proppant Data Distribution Histogram 43

3.21 Normalized Avg. Placed Proppant per Stage Data Distribution Histogram 44

3.22 Normalized Total Proppant Pumped per Completed Lateral Length Data Distribution Histogram 45

3.23 Average Fracture Spacing Data Distribution Histogram 46

3.24 Avg. Fracture Closure Gradient Data Distribution Histogram 47

3.25 Average Proppant Concentration (lbs./gal) Data Distribution Histogram 48

3.26 Fracturing Fluid Base Groups that were Originally in the Dataset 49

3.27 Fracturing Fluid Base Groups after Modifying the Dataset 49

3.28 Hydraulic Fracturing Base Fluid Types Data Distribution Histogram 50

3.29 Fracturing Fluid Energizers Data Distribution Histogram 50

3.30 Using Box Plot Techniques to Identify Outliers in Total Fluid Pumped 52

3.31 Distribution Histogram with Box Plot of Total Pumped Fluids in the Montney Wells after Removing Outliers 52

3.32 Cross Plot of Total Fluid Pumped (M^3) Versus. Load Fluid (M^3) for All Wells in the Montney Grouped in Colors Based on Treatment Fluid Type 53

3.33 Cross Plot of Total Fluid Pumped (M^3) versus. Load Fluid (M^3) Grouped in Five Ranges of Completed Lateral Length (m) and Colored Basesd on Treatment Fluid Type 54

3.34 Cross Plot of Total Fluid Pumped (M3) versus Load Fluid (M3) for All Wells in the Montney Grouped in Five Ranges of Avg. Proppant Concentration (lbs./gal) ... 55

3.35 Cross Plot of AFE Completion Cost versus. Final Completion Cost 56

3.36 Cross Plot AFE completion Cost versus. Total End of Completion Cost Grouped based on Lateral length and Colored based on Operator Company 57

3.37 Cross Plot of AFE Completion Cost versus Final Completion Cost Grouped on Lateral length and Completion Type and Colored based on the Operating Company 58

3.38 Cross Plot Number of Attempted Stages versus Number of Actual Stages Grouped Based on Completion Type and Colored Based on Stimulation Company 59

3.39 Cross Plotting Number of Attempted Stages versus Number of Actual Stages 60

3.40 Cross Plotting Number of Attempted Stages versus Number of Actual Stages Grouped Based on Completion Type 61

3.41 Cross Plotting Total Designed Proppant versus Total Actual Pumped Proppant Grouped Based on Completion Type and Colored Based on Operating Company 62

3.42 Cross Plotting Total Designed Proppant versus Total Actual Pumped Proppant Grouped Based on Lateral Length and Completion Type and Colored Based on Operating Company 63

3.43 Cross Plotting Total Designed Proppant versus Total Actual Pumped Proppant Grouped Based on Treatment Fluid and Completion Type and Colored Based on Operating Company 64

3.44 Cross Plot Shows the Relation of Closure Gradient and Proppant Mass 64

4.1 IP Gas of Completion Types 73

4.2 IP Oil of Completion Types 73

4.3 IP Water of Completion Types 74

4.4 Completion Types 6 Months Gas Cumulative 74

4.5 Completion Types 12 Months Gas Cumulative 75

4.6 Completion Types 18 Months Gas Cumulative 75

4.7 Completion Types 6 Months Oil Cumulative 76

4.8 Completion Types 12 Months Oil Cumulative 76

4.9 Completion Types 18 Months Oil Cumulative 77

4.10 Completion Types 6 Months Water Cumulative 77

4.11 Completion Types 12 Months Water Cumulative 78

4.12 Completion Types 18 Months Water Cumulative 78

4.13 Completion Types 6 Months Gas Cumulative / Lateral Length 79

4.14 Completion Types 12 Months Gas Cumulative / Lateral Length 79

4.15 Completion Types 18 Months Gas Cumulative / Lateral Length 80

4.16 Completion Types 6 Months Gas Cumulative / Stage 80

4.17 Completion Types12 Months Gas Cumulative / Stage 81

4.18 Completion Type18 Months Gas Cumulative / Stage 81

4.19 Completions Type Cost 82

4.20 Completions Type Drilling Cost 82
4.21 Completion Types Proppant per Length 83
4.22 Completion Types Fluid per Length 83
4.23 Completion Cost versus Lateral Length 85
4.24 Recovered Load Fluid versus Pumped Load Fluid 85
4.25 Cumulative Gas Production versus Load Fluid Recovery 86
4.26 Example of Setting Color Scale for a Variable 88
4.27 Example of Setting the Color Scale in JMP 88
4.28 Cross Plot and Heat Map of 6 Months Cumulative Production and Load Fluid Recovery Percentage 89
4.29 Cross Plot and Heat Map of 12 Months Cumulative Production and Load Fluid Recovery Percentage 90
4.30 Cross Plot and Heat Map of 18 Months Cumulative Production and Load Fluid Recovery Percentage 90
4.31 Heat Map of 18 Months Cum. Gas Production versus. Gas IP 91
4.32 IP Gas versus. 6 Months Cum. Gas Production 92
4.33 IP Gas versus 12 Months Cum. Gas Production 93
4.34 IP Gas versus 18 Months Cum. Gas Production 94
4.35 IP Oil versus 6 Months Cum. Oil Production 94
4.36 IP Oil versus 12 Months Cum. Oil Production 95
4.37 IP Oil versus 18 Months Cum. Oil Production 95
4.38 IP Water versus. 6,12 &18 Months Cumulative Water Production 96
4.39 Normalized per stage 6 Months Cumulative Production versus Actual Stage Number 97
4.40 Normalized 12 Months Cumulative Production versus Actual Stage Number 98
4.41 Normalized 18 Months Cumulative Production versus Actual Stage Number 98
4.42 Actual Stage Number versus. 18 Months Cum Gas / Actual Stage Number 99
4.43 Heat Map of Normalized IP Gas/Stage versus. Actual Number of Stages 100
4.44 Heat Map of Normalized 6 Months Cumulative Gas/Stage versus Actual Number of Stages 100

4.45 Heat Map of Normalized 12 Months Cumulative Gas/Stage versus Actual Number of Stages 101

4.46 Heat Map of Normalized 18 Months Cumulative Gas/Stage versus Actual Number of Stages 101

4.47 Cross Plots 18 Months Cum. Gas /Stage versus Number of Stages (Cased, Open) Hole 102

4.48 The Effects of Total Proppant Placed on the Cross Plot of 18 Months Cum. Gas /Stage versus Number of Stages 103

4.49 The Effects of Load Fluid on the Cross Plot of 18 Months Cum. Gas /Stage versus Number of Stages 103

4.50 The Effects of Completed Lateral Length on the Cross Plot of 18 Months Cum. Gas /Stage versus Number of Stages 104

4.51 Cross Plot of 18 Months Cum. Gas /Stage versus Avg. Proppant / Stage 105

4.52 Cross Plot of 18 Months Cum. Gas /Stage versus Avg. Fluid / Stage 105

4.53 Completed Lateral Length Versus Time 106

4.54 Attempted Number of Stages Versus Time 107

4.55 Actual Number of Stages versus Time 107

4.56 Total Proppant Placed versus. Time 108

4.57 Total Pumped Fluid Versus Time (M^3) 108

4.58 Completion Type Versus Time 109

4.59 Proppant Concentration versus. Time 109

4.60 18 Months Cum. Gas versus. Time 110

4.61 IP Gas versus. Time 110

4.62 Drilling and Completion AFE versus. Time 111

4.63 Final Drilling and Completion Cost versus. Time 111

4.64 Fracture Spacing versus. Time 112

4.65 Heat Map Shows the Effects of Completion Cost on 18 Months Cum. Gas 113

4.66 Heat Map Shows the Effects of Drilling Cost on 18 Months Cum. Gas 114

4.67 Heat Map Shows the Combination of Final Drilling and Completion Cost 114

LIST OF TABLES

Table Page

3.1 JMP Graph Builder Main Drop Zones Descriptions........ 24
3.2 List of Original Well Parameters in the Dataset........ 27
3.3 List of Calculated Parameters in the Dataset........ 28
3.4 Montney Top 30 Fields in Well Count Reported in the Dataset........ 29
3.5 Montney Top 30 Major Operators in Well Count........ 31
3.6 List of Stimulation Companies in the Montney........ 34
3.7 Completion Type Well Percentage in the Montney........ 35
3.8 Closure Pressure Gradient Unit Conversion........ 47
3.9 List of Normalized Parameters Used in the Analysis........ 65
3.10 List of Design Parameters Used in the Analysis........ 66
3.11 List of Uncontrolled Parameters Used in the Analysis........ 66
4.1 Parameters Mean Values in Cased and Open Hole Wells........ 70
4.2 Parameters Median Values in Cased and Open Hole Wells........ 70

NOMENCLATURE

Symbol	Description
OHMS:	Openhole Multistage System
GIS:	Geographical Information System
CAPEX:	Capital Expediture
OPEX:	Operational Expenditure
s:	Sample Standard Deviation
$\sum$:	Summation
IQR:	Interquartile Range
EDA:	Exploratory Data Analysis
UWI:	Unique Well Identifier
AFE:	Authorized for Expenditure
N:	Number of Samples

1. INTRODUCTION

Unconventional reservoir systems (e.g., tight gas, shale gas, liquid-rich shales, and coalbed methane) have been defined as hydrocarbon accumulations which are difficult to be characterized and produced by conventional exploration and production technologies. (Mangha et al. 2012) Typically, these types of formations have low permeability. Shale plays are considered as ultra low permeability (.001 to .000001 mD).

Commercial production from low permeability shales is not feasible without hydraulic fracturing, as the natural rock has insufficient permeability for commercial flow. Hydraulic fracturing, which is the process of creating one or more cracks in the rock, greatly extends the drainage contact area with the reservoir, while providing highly conductive flowpaths to the wellore.

Advanced well completions, which combine multi-stage hydraulic fracturing treatments and horizontal well technology, are required to establish commercial rates from shale plays. Two types of completions systems, cased hole and open hole, are in prevalent use today. In the cased hole approach, the well is drilled and cased through the buildup section, then a liner is run and cemented in the lateral. Clusters of perforations are shot and the hydraulic fracturing treatment is pumped. This is referred to as a 'stage' of fracturing. Each stage is separated by a composite bridge plug, and all of these plugs are drilled out once all stages are fracture stimulated. In the openhole approach, the lateral is not cased and cemented. A liner equipped with openhole packers (mechanical or swellable) is run with ball activated sleeve systems. Once the packers are set, successively larger balls are dropped to shift the sleeves downhole, allowing each fracture stage to be pumped in an almost continuous operation. After all stages are stimulated, the balls flow back to surface with produced fluids, or may dissolve depending on the material used.

Advances in technology to produce and develop ultra-low permeability reservoirs such as shale gas reservoirs bring the difficulties and uncertainty associated with well performance characterization and analysis. The uncertainty is mainly due to the lack of complete understanding of the production mechanisms, factors controlling production rates, the physics of multistage completions and behavior of these reservoir systems there is also uncrtanty associated with establishing the long term production decline in these reservoirs.

The main issue facing most operators in the oil and gas industry is the capital and resources allocation (budget and people). The oil and gas industry's goal is to use the optimum completion practices to recover hydrocarbons from these resource plays. While the goal is the same, the approaches used among operators to achieve it can be very different. Some prefer a'trial and error' approach of drilling towards a solution. For example, an operator tested 800 wells in the Fayetteville Shale to understand well spacing. While this method can be fruitful, it may be very expensive. Others use empirical methods involving data mining of public and proprietary databases. Methods relying on analyzing well performance are preferred by operators who have access to high resolution rate and pressure information. Reservoir simulation of multi-fractured horizontal wells (MFHW), while extremely useful on a well-by-well basis, is still too complicated and time consuming.

Some of the specific difficulties in characterizing unconventional reservoirs (resource plays) include and are not limited to (Okouma Mangha et al. 2012):

- Inability to distinguish between hydraulic fractures and reservoir contributions from limited production/pressure history.
- Incomplete or limited knowledge about hydraulic fracture geometries in horizontal wellbores: bi-wing fractures, dentritic fractures and/or complex fracture geometry.
- Uncertainty of the stimulated-reservoir volume (SRV) contribution compared to the surrounding unstimulated reservoir volume.
- Lack of understanding of petrophysical/reservoir properties variations and their accuracy.
- Predominantly linear flow, as opposed to the conventional radial flow.
- Predominantly transient flow as opposed to the conventional boundary dominated flow.
- Pressure-dependent rock properties.
- Adsorption gas storage mechanics.

For these reasons, operators have been challenged to apply conventional analytical techniques in optimizing completions in shale plays. Statistical approaches have gained wide acceptance in trying to evaluate and understand the different fracture or completion applications and identify best practices. These approaches apply statistical analyses to large amounts of drilling, completion and production from MFHWs in a particular area, and require extensive data mining. Hence, data mining is one of the techniques that oil and gas industry is adapting to help in improving the quality of the wells productivity from the unconventional resources. As Paul Siegele the president of the Energy Technology Company at Chevron said “Information technology is enabling us to get more barrels of each asset.”

This thesis describes a project to apply data mining and statistical analysis to understand the well completion and stimulation effects on production performance, and provide a comparative means between different completion applications in the Canadian Montney shale formation. The study provides the first comprehensive, field-wide statistical review of the Montney shale, using publically available well data.

1.1. MONTNEY PLAY

The Montney formation resource play, which straddles the border between the Canadian provinces of British Colombia and Alberta, is considered by many to be the largest natural gas resource play in North America. (Wilson et al. 2011)

The Montney Formation’s marketable, unconventional petroleum potential was evaluated in a joint assessment by the National Energy Board, the British Columbia Oil and Gas Commission, the Alberta Energy Regulator, and the British Columbia Ministry of Natural Gas Development. The thick and geographically extensive siltstones of the Montney Formation are expected to contain 12,719 billion m^3 (449 Tcf) of marketable natural gas, 2,308 million m3 (14,521 million barrels) of marketable natural gas liquids (NGLs), and 179 million m3 (1,125 million barrels) of marketable oil (National Energy Board, British Columbia Oil & Gas Commission, Alberta Energy Regulator 2013).

The Montney Formation of Alberta and British Columbia has been the target of oil and gas exploration since the 1950s, with industry traditionally focusing on the

Montney's conventional sandstone and dolostone reservoirs. These conventional reservoirs are encased in siltstone, which represents a far greater volume of rock within the formation and also contains oil and gas. However, Montney siltstones remained undeveloped until 2005, when advances in horizontal drilling and multi-stage hydraulic fracturing made it possible to economically develop this extensive, unconventional siltstone resource. (National Energy Board, British Columbia Oil & Gas Commission, Alberta Energy Regulator 2013)

The Montney Shale is a hybrid of a shale reservoir and tight gas reservoir. The Montney Shale is rich in silt and sand, similar to tight gas, but the natural gas originates from the organic matter in the formation, making it a shale. The Montney is shallow and brittle, making hydraulic fracturing operations more successful than in some of the other Canadian shale basins. However, due to the presence of siltstone and sand throughout the formation, it has extremely low permeability and requires higher levels of fracture stimulation for successful extraction. (Petroleum Technology Alliance Canada & Science and Community Environmental Knowledge Fund 2012)

A generalized map showing the location of the Montney Formation in the subsurface of Alberta and British Columbia along with the major rock lithologies of the play is shown in Figure1.1

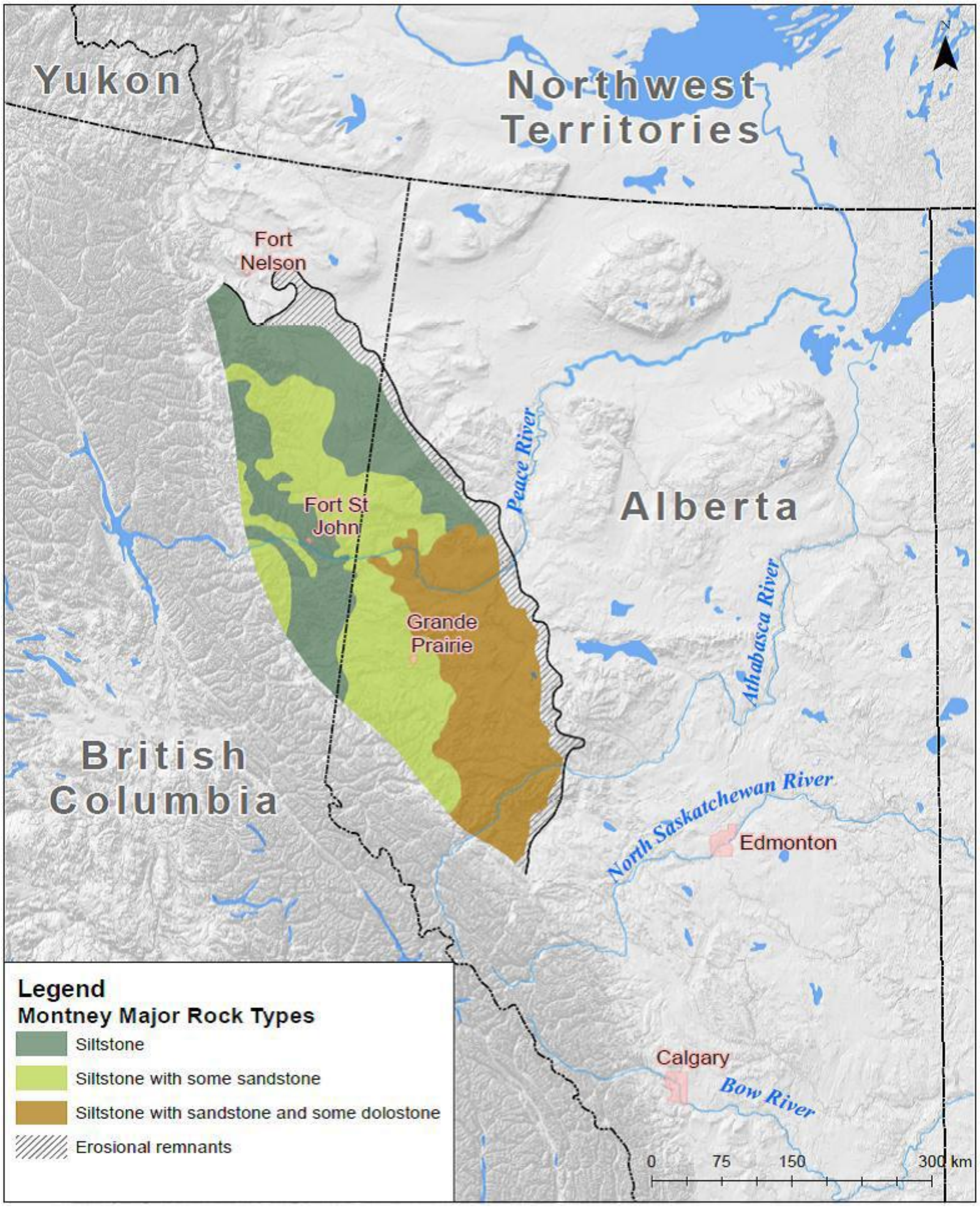

Figure 1.1 Generalized Map Showing the Location of the Montney Formation in the Subsurface of Alberta and British Columbia.Modified from the Geological Atlas of the Western Canada Sedimentary Basin (National Energy Board, British Columbia Oil & Gas Commission)

1.2. OBJECTIVES OF THE STUDY

This research has been conducted to study a large well dataset and develop a statistical review of the Montney unconventional resource play for better insights about the completion practices and trends that affect well performance. Fundamentally, the research will seek, either directly or indirectly, to answer questions such as:

- Is it better to complete Montney wells openhole or cased hole?
- What is the effect of increasing proppant and fluid volume on the wells
- production performance?
- How many stages should be used?
- What factors affect well performance?

Previous statistical studies in the unconventional resources reported that the impact of individual variables on the production outcome is often difficult to interpret with any degree of confidence when traditional linear regression methods are used because of the impacts of missing data, erroneous data, non-linear data and subtle interrelationships among variables (Lafollette et al. 2012b). Therefore, a secondary objective of this work is to provide quality data mining of the dataset, and a case study in the practical use of cross-plots to compare and distinguish the best practices in Montney resource play.

Previous work showed that the applications of practical data mining methods to a large shale dataset resulted in learning key lessons that were not apparent from small datasets (Lafollette et al. 2012b). Hence, it is expected that correlations and relationships identified in this research will lead to several useful conclusions, which may not have been readily discerned from more limited subsets of Montney wells.

2. LITERATURE REVIEW

Several studies have been found in the literature that address the effects of completion and stimulation methodologies on the well's production performance in the unconventional resources. Many of the authors used analytical and practical data mining approaches to evaluate well completion and fracture stimulation parameters to understand and gain insights about these complicated resources. This section reviews the historical work related to the practical applications of data mining and statistical analysis in North America’s unconventional resource plays.

In 2011, Wilson et al. conducted a comparative study to analyze two different multistage hydraulic fracturing technologies applied in the Lower Montney formation represented by cemented liner and openhole multistage system (OHMS) completions. The analysis, using simple averaging and plotting, was performed on field data from 15 wells that were divided into two separate geographical areas within the same field. The comparisons included production analysis, lateral lengths, number of stages, stage spacing, proppant volumes and pump rates. Additionally, operational time and cost comparisons were determined on a per-well and per-stage basis for both technologies.

Based on the analyzed field data, they concluded that the application of OHMS completion technology is best for the Lower Montney in the region of the play that was studied. The study also demonstrated that the application of this technology for the wells selected in the two study areas resulted in both greater initial production rates and overall cumulative production than cemented liner completed wells. Based on completion cost, they confirmed that both the average total cost of completion and the average cost per stage in conducting cemented liner jobs was higher than employing OHMS completions. Furthermore, less time was required to perform the fracture stimulation job when using OHMS technology as compared to cemented liners. (Wilson et al. 2011)

Another Montney play-wide performance analysis was carried out by Shell Canada Energy in 2012 to analyze the well performance histories of 74 producing multi-stage fractured horizontal wells using a common and consistent analytical framework. The study spanned five producing areas (A, B, C, D, and E), two different completion styles (50 versus. 100 m frac spacing), and three different initial production strategies (unrestricted through highly constrained, which might be related to the reservoir quality, i.e. lower rock quality will require higher drawdown). The main findings of this study were that wells completed at 50 m fracture spacing (using 30 tonnes of proppant per cluster) performed similarly to those with 100 m spacing (using 60 tonnes per cluster), and the 30-yr P50 predicted that the final recovery of the 50 m and 100 m spacing wells were very similar. The study also suggested which of the five producing zones had the highest well productivity and predicted recoveries. In addition, the study indicated wells that were produced without restrictions (high drawdown) showed the lowest productivity, highest completion resistance (skin) to flow, and lowest predicted final recoveries. (Okouma Mangha et al. 2012)

In 2012 Lafollette et al. performed a data mining of wells, hydraulic fracturing treatments and production parameters for horizontal wells in the north Texas Barnett Shale play for the wells completed between 2003 and 2009. The study used Geographical Information System (GIS) pattern recognition techniques in conjunction with more traditional statistical techniques to interpret trends in the dataset. They plotted the top 10% of the peak monthly production in the entire Barnett field, and based on that they identified a study area of interest with 2329 cased hole horizontal wells. In this work they realized that cross plot and regression analyses could be successfully applied to the analysis of production and well parameters if the wells are geographically grouped. They also concluded that wells with horizontal lengths of more than 3500-4500 feet are less efficient, which showed lower production per perforated foot than the shorter lengths. (Lafollette et al. 2011)

Follow up work was completed by the same authors (Lafollette et al) where they used merged reservoir quality proxies, well architecture, completion and stimulation data, that were listed along with the production data and placed in geographical perspective, for an improved understanding of hydraulic fracturing impacts. They modeled the well location and stimulation parameters to predict the maximum gas rate. They came up with six parameters based on the relative importance to the model and the most important variables were the true vertical depth (Mid-Perf TVDSS), y-direction path, total fracturing fluid volume used to treat the well, fracturing slurry average stage injection rate, the use of 20/40 mesh proppant and perforated lateral length. In this work, they concluded that when using the traditional linear regression methods, the impact of the individual variables on the production outcome is often difficult to interpret with any degree of confidence. Another conclusion was drawn from this study that the job volume and injection rate were an important predictor of the maximum gas rate in the Barnett, but a special caution must be applied to not extend the fracture and crack the Ellenberger water bearing zone which leads to a lower performing well than a smaller job that stays out of the water. (Lafollette et al. 2012b)

A study from the Bakken was conducted in 2012 by LaFollette et al. to analyze well and production data beginning with more than 400 wells in the greater Sanish-Parshall area. They used a combination dataset from the North Dakota Industrial Commission Oil and Gas Division, public data, and in-house proprietary data. The intention of the study was to show that the application of practical data-mining methods to an intermediate-size shale oil (light, tight oil) well data set could result in learning key lessons that may not be apparent when working with small datasets.

The authors of that study used Geographical Information System pattern recognition techniques, along with other data-mining techniques, to interpret trends in the data sets. The study was designed to search for relevant trends in the distribution of production results for wells completed with fracturing sleeves and packers, plugged and perforated, or complex completions to determine whether differences in productivity existed and needed to be factored into the completion recommendations. Trends examined in the project in addition to completion type, included treatment parameters such as fracturing fluid types and quantities, proppant types and quantities, number of completion stages and stage lengths, perforation cluster spacing and length, and calculated perforation friction drop. The most important conclusions that came out of this study were that the production efficiency decreases when the lateral length increases, and production per stage decreases when the stage counts in the lateral increase. The study also showed that decreasing the average proppant concentration appears to negatively affect productivity. (Lafollette et al. 2012a)

Griffin et al in 2013 also conducted a study on the Bakken in North Dakota to benchmark performance of completion and stimulation using a developed production and completion database of 1100 wells completed in the Central Basin from 28 operators.

The study used publicly available information from the North Dakota Industrial Commission's records augmented by additional completions information that was obtained directly from the operators. Completion performance benchmarking was performed using the developed database along with a Petra geological database, developed from all publicly available logs in the Central Basin. The authors applied multivariate analysis methods, which included geological input, to benchmark performance over most of the Williston Central Basin, and to compare the varied completion methods for the wells in the dataset.

The study outcomes provided insights about the studied area, a major conclusion was that: benchmarking resource play performance is complicated and production performance can vary dramatically over relatively small areas making simple evaluations difficult. The authors also concluded that reservoir quality is important and benchmarking the resource play without considering the reservoir properties can be very subjective and weighed heavily against preconceived ideas. The study showed that completions matter and that the higher cost of advanced completion techniques can be economically justifued when properly applied, because it appears that advanced completion designs create large reservoir contact areas (fractures and fracture networks) and effectively connect the contacted area back to the wellbore (conductivity). The economic evaluation of costs predicted that spending an additional $1-2 million per well for the advanced completions adds multiple millions of dollars revenue in the first year, and the additional costs are paid out in just a few months. The final conclusion derived from this study was that using the water cut as a primary indicator of reservoir quality helps in correlating wells to the geological resource model used for the Bakken in the Williston Central Basin. (Griffin et al. 2013)

Michael Roth and Roth in 2013 applied an analytical approach to optimize well spacing and completions in the Bakken/Three Forks plays. The objective of their study was to identify the optimum well spacing between the wells within the single and adjacent formation to eliminate the problem of frac communication between neighboring Three Forks and Middle Bakken as treatment fluid from the completed wells in Three Forks was being produced back by the adjacent Middle Bakken wells. An analytical technique was applied to the production and well parameters to combine the geological and engineering information in a multi-variate analysis and isolate the impact of individual parameters on the well performance. As an outcome from that study, insights were gained about the optimum well proximity for controlling well interactions and optimizing well recovery factors. (Roth & Roth 2013)

(Modeland et al. 2011) conducted a play-wide statistical analysis of the effects of completion methodology on production in the Haynesville shale. The study combined completion variables of 286 wells with public production data to construct a dataset used to develop cross plots between different completion strategies and well production performance. Trends of the cross plots showed that the Haynesville well production is heavily dependent on the geographic location and the total number of stages. Since the Haynesville shale is considered a softer rock than most of the North American shale plays, proppant concentration and placement strategy was shown to significantly impact the production and affect fracture conductivity. The main recommendations that came out of this study were to increase the number of effectively stimulated fractures along the lateral, and to design fracture jobs to improve the conductivity of the fractures in the adverse conditions of the Haynesville shale.

Unlike previous studies, this study analyzed more than 3300 horizontal wells in the Montney formation using JMP statistical software to compare the differences between the two completion types statistically and graphically, the applications over time and cost analysis was also included.

3. METHODOLOGY

The JMP Pro Statistical Discovery Software from SAS used in this study is introduced within this chapter to highlight the basic features and capabilities. This chapter will also introduce the concepts of data mining and lay out the procedures and steps that used to analyze the data. Finally, the first three data mining phases will be addressed in this chapter, along with the techniques followed in each phase to reach to the final clean dataset that used in the analysis.

3.1. DATA MINING

As early as 1984, John Naisbitt, a great American author and public speaker in the area of future studies wrote in his book Megatrends that “we are drowning in information but starved for knowledge”. This statement is especially true in unconventional completions, with the massive datasets and seemingly endless questions regarding which completion is best.

To understand the basic definition of Data Mining a few citations were selected to describe it based on many resources from the literature.

- “Data mining is the process of discovering meaningful new correlations, patterns and trends by sifting through large amounts of data stored in repositories, using pattern recognition technologies as well as statistical and mathematical techniques.” (The Gartner Group 2013).
- Another definition from MIT Press “Data mining is the analysis of (often large) observational datasets to find unsuspected relationships and to summarize the data in novel ways that are both understandable and useful to the data owner” (Hand et al. 2001).
- “The nontrivial process of identifying valid, novel, potentially useful, and ultimately understandable patterns in data.” (Fayyad et al.,1996).
- “Finding interesting structure (patterns, statistical models, relationships) in data bases”. (Fayyad, Chaduri and Bradley, 2003).

- "A knowledge discovery process of extracting previously unknown, actionable information from very large databases" (Zornes, 1996).
- "A process that uses a variety of data analysis tools to discover patterns and relationships in data that may be used to make valid predictions." (Edelstein,1999).

Many industries and especially the oil and gas industry have huge amount of datasets. In addition many organizations exist sololy to provide the service of collecting, organizing and analyzing the data. The tremendous growth in computing power and storage capacity has helped greatly in the ongoing remarkable growth in the field of data mining and helped recognize and understand hidden trends and correlation by studying huge datasets.

To make better decisions one needs to discover and understand the underlying patterns involved in the particular operation from the data. For example, it's not enough for a production engineer to know just the amount of oil and/or gas production from a field and the amount of catal expenditure (CAPEX) and operating expenditure (OPEX) for company in this highly competitive business environment. To increase recovery and achieve higher production the production engineer has to search for answers to the questions like: What would be the best stimulation design for a particular well? How to select the best candidate wells for stimulation? Which service company should be used more often for better results? How to balance the quality of an intervention job with the cost? And many other questions that can lead to higher production. ("Intelligent Solutions Inc." 2011)

3.2. BASIC STATISTICS

Statistics is a field of mathematics that pertains to data analysis. Statistical methods and equations can be applied to a dataset in order to analyze and interpret results, explain variations in the data, or predict future data.

The basic common statistics that used as part of this study will be introduced in the following sub-sections.(Andrew MacMillan et al. 2006)

3.2.1. Mean. The mean (also known as average), is obtained by dividing the sum of observed values by the number of observations, *n*. Although data points fall above, below, or on the mean, it can be considered a good estimate for predicting subsequent data points. The formula for the mean is given below as equation (1).

$$A = \frac{1}{n} * \sum_{i=1}^{n} x_i \tag{1}$$

3.2.2. Median. The median is the numerical value separating the higher half of a data sample from the lower half. The median of a finite list of numbers can be found by arranging all the observations from lowest value to highest value and picking the middle one. If there is an even number of observations, then there is no single middle value; the median is then usually defined to be the mean of the two middle values. The median can be used as a measure of location when a distribution is skewed, when end-values are not known, or when one requires reduced importance to be attached to outliers, e.g., because they may be measurement errors.

The median is useful if the data analyst is interested in the range of values that the system could be operating in. Half the values should be above and half the values should be below, so an idea about where the middle operating point can be figured out.

3.2.3. Mode. The mode is a statistical term that refers to the most frequently occurring number found in a set of numbers. The mode is found by collecting and organizing the data in order to count the frequency of each result. The result with the highest occurrences is the mode of the set. While the mean would incorporate the occasional outlying data.

3.2.4. Standard Deviation. The standard deviation gives an idea of how close the entire set of data is to the average value. Data sets with a small standard deviation have tightly grouped, precise data. Data sets with large standard deviations have data spread out over a wide range of values. The formula for standard deviation is given below as equation (2).

$$s = \sqrt{\frac{\sum(X - \bar{X})^2}{n - 1}} \tag{2}$$

where,

s = sample standard deviation

$\sum$= summation

$\bar{X}$= sample mean

n = number of scores in sample.

3.2.5. Box Plot. A box plot is one of the statistical techniques used in this research to help in refining the dataset and identifying outliers. This technique also referred as a box-and whisker diagram, it is a graph of dataset that consists of a line extending from the minimum value to the maximum value, and a box with lines drawn at the first quartile, Q_1; the median; and the third quartile, Q_3.

This simplest possible box plot displays the full range of variation from minimum to maximum, the likely range of variation which represented with the interquartile range (*IQR*), and a typical value (the median).

It is not uncommon that real datasets will display surprisingly high maximums or surprisingly low minimums called outliers. John Tukey has provided a precise definition for two types of outliers:

1. **Outliers:** are either 3×*IQR* or more above the third quartile or 3×*IQR* or more below the first quartile.
2. **Suspected outliers:** are slightly more central versions of outliers: either 1.5×*IQR* or more above the third quartile or 1.5×*IQR* or more below the first quartile.

If either type of outlier is present, the whisker on the appropriate side is taken to 1.5×*IQR* from the quartile (the "inner fence") rather than the max or min, and individual outlying data points are displayed as unfilled circles (for suspected outliers) or filled circles (for outliers). (The "outer fence" is 3×*IQR* from the quartile.) (Kirkman, 1996)

Figure 3.1 depicts the main parts of a Box Plot diagram. The diagram on the left represents a simple dataset with no outliers while the diagram on the right represents a more complicated dataset with existing outliers.

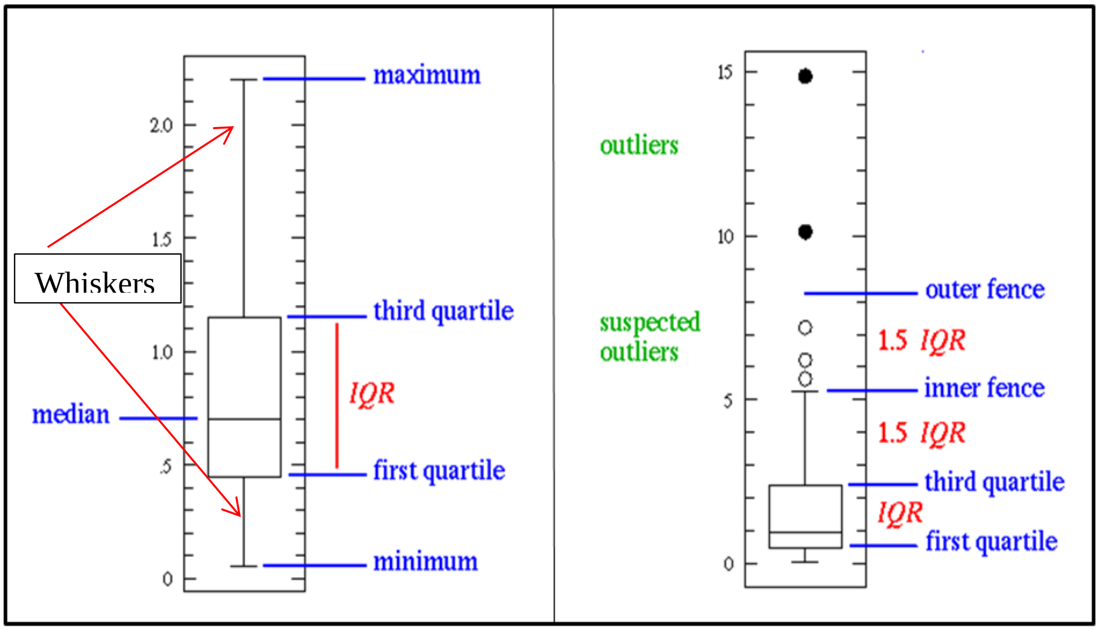

Figure 3.1 Box Plot Diagram (Kirkman,1996)

3.3. INTRODUCTION TO JMP

John Sall created *JMP* in 1989 as a tool for discovering information in data through visualization and graphics. *JMP* is designed to be a point-and-click, walk-up-and-use product that enables a user to discover more, interact more, and understand more. The correct graphs are integrated with the right analyses. Because *JMP* is task-oriented, not method-oriented, you do not need to be a professional statistician to use it. You only need to know what questions you wish to be answered.

The following sub-sections will introduce the JMP software main components and functions to explain the concepts behind the generated plots that were used in the dataset analysis along this study.

3.3.1. The JMP Data Table. Data to be processed in JMP must be in the form of a JMP data table. A data table is similar to a spreadsheet but the rows and columns have a special purpose. The data table looks like a spreadsheet with some enhancements.

Figure 3.2 is a snap shot from part of the JMP data table of the studied Montney dataset. The data table in this figure contains 60 columns and 3369 rows as indicated by the red arrows. Each column from the 60 columns represents a variable in the dataset, while each row represents an individual that is characterized by the parameters in each column. In this study each individual row represents a well.

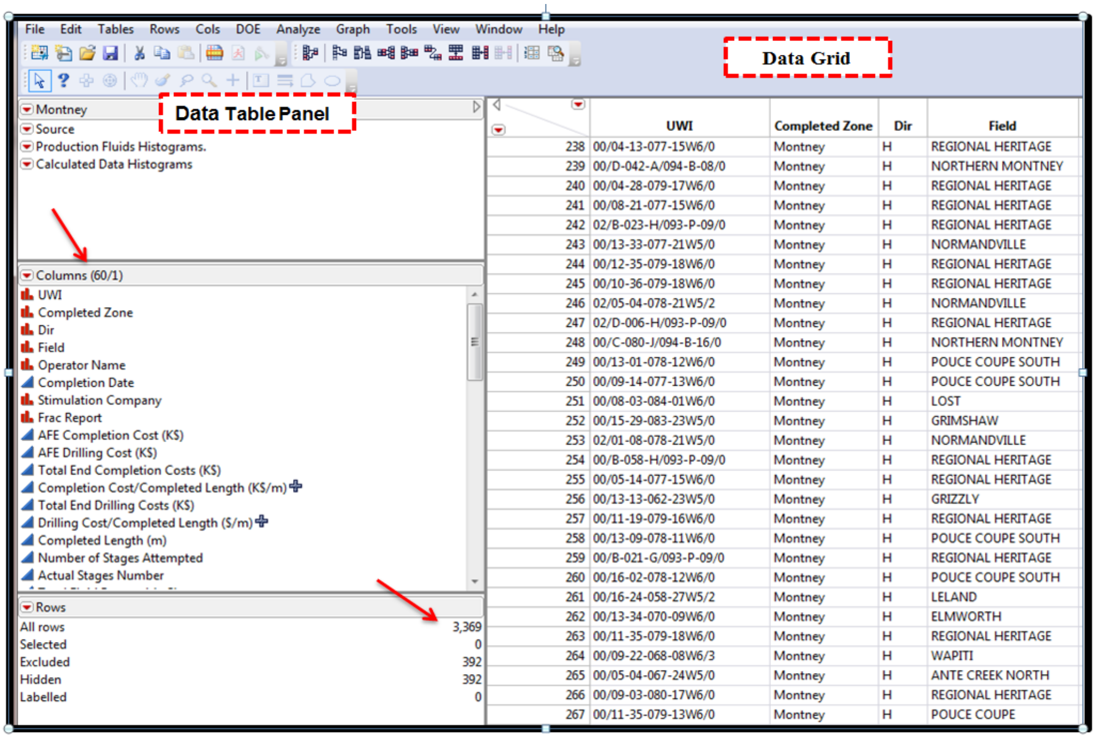

Figure 3.2 JMP Data Table

JMP Pro has an integrated data import wizard that can import the data from any saved format such as SAS, txt, csv, R… and xlsx. In this study the original data were saved in an excel spreadsheet with .xlsx format. In the import wizard, there are many flexible tools available to facilitate the data retrieval from the excel file in order to generate the new JMP data table. Once the JMP data table is opened, the data will be arranged in columns and rows and further editing may be applied to the data using the integrated tables tab, which is part of the main menu entry. Using the Tables Tab enables the users to apply several modifications and adjustments to the data columns. The users of JMP can request summary statistics by grouping columns, or subset the data by a specified column and sort the data in descending or ascending order. A stacking option is also available where data tables could be rearranged by stacking two or more columns into a single new column. New data tables can be created using Split Columns option by splitting one column into several columns. The transpose rows and columns option in the Tables Tab can create a new JMP table that is a transposed version of the active data table where the columns of active table are the rows of the new table, and its rows are the new table’s columns. Joining two or more JMP data tables can be accomplished using the Concatenate option in the Tables menu by combining rows from the two or more data tables into a new data table or rows could be appended to the first data table based on analyst preferences. Combining data tables by matching the values in one or more columns that exist in both data tables has been made easier by using the Join option. These were the main tools that used during the phase of exporting the data from the excel spreadsheet to the JMP data tables. Figure 3.3 depicts the Table menu program interface.

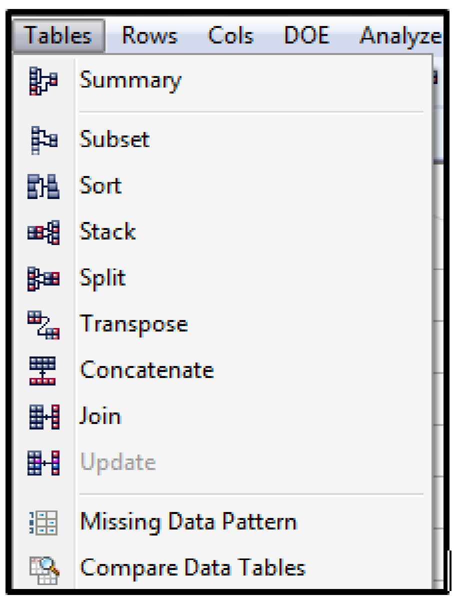

Figure 3.3 JMP Tables Menu

Once the data table is generated and the parameters arranged in the way that suits the objective study, new rows can be added to the data table and new columns can also be added to introduce a new parameter in the data table. These parameters could be in the form of mathematical equation to help in producing new variables based on the original dataset parameters. In this study many parameters were calculated in this way, by having the original production and some of the stimulation parameters normalized to different design and architecture parameters. It is worth mentioning that this is a brief introduction to JMP data handling, and many other options are available in the software that help in editing the format, organizing and handling all the data points in the data.

3.3.2. JMP Analyze Menu. Many types of analyses can be performed in JMP using the Analyze Menu. Several analysis platforms are available within this menu to help to understand and investigate the relationships between variables.

One of the most frequently used platforms for this study is The Distribution platform, which illustrates the distribution of a single variable using histograms. These variable distributions are examined within a generated report.

The report content for each variable changes depending on whether the variable is categorical (nominal or ordinal) or continuous. Continuous variable is any parameter that contains a numerical value, while nominal variable contains characters or names. The Distribution report window is interactive, clicking on a histogram bar highlight the corresponding data in any other histograms and in the data table.

Histograms visually display the data. For categorical variables, the histogram shows a bar for each level of the ordinal or nominal variable, while for continuous variables, the histogram shows a bar for grouped values of the continuous variable.

Figure 3.4 and Figure 3.5 shows and example of distribution histogram of continuous and nominal variables respectively. The value on top of each bar in these two figures represents the count number of individuals from the dataset that have the same value range in the continuous variable or the same category in the nominal and ordinal variables, the count number could be replaced by the percentage of individuals in the plotted parameter.

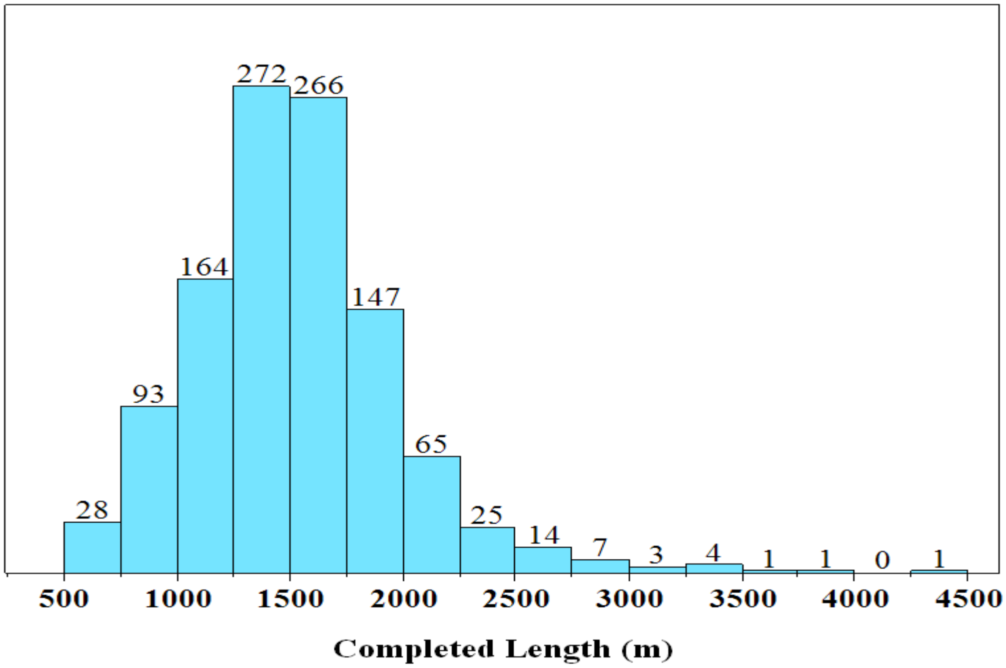

Figure 3.4 Histogram Distribution Example of Continuous Variable

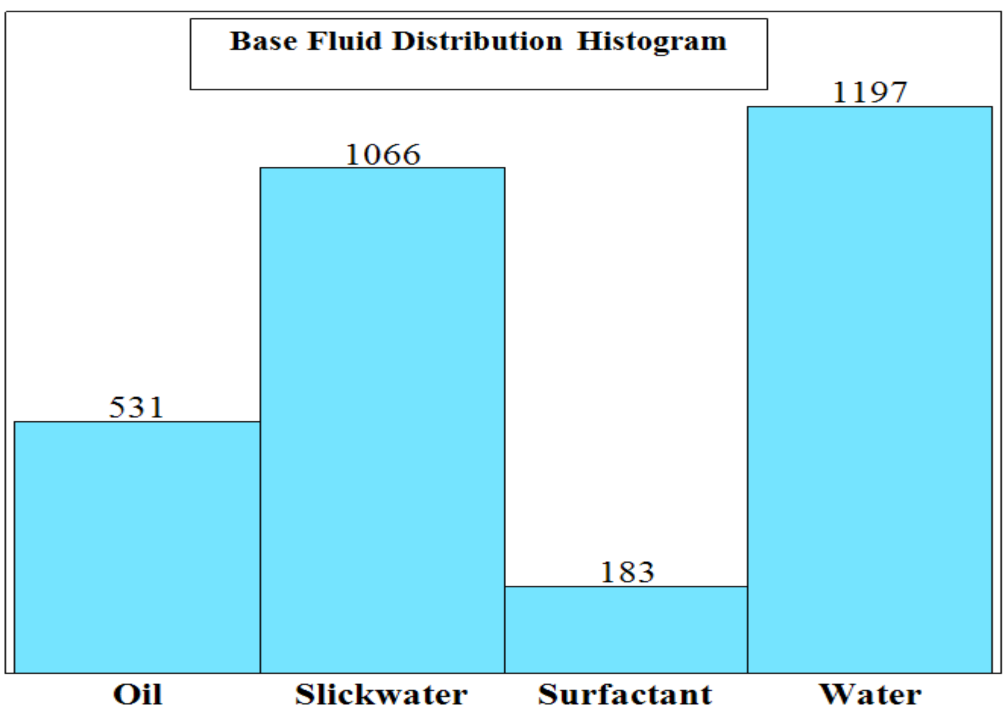

Figure 3.5 Histogram Distribution Example of Nominal Variable

3.3.3. Data Visualization and Exploratory Data Analysis with JMP. Exploratory Data Analysis (EDA) is a data analysis tool that can guide analysts in building useful models. In JMP, data visualization and EDA go hand in hand, giving the tools needed to make breakthrough discoveries and communicate results. Linking dynamic graphics with powerful statistics, JMP helps the analyst to construct a narrative and interactively share findings in ways the industry colleagues and decision makers can readily understand and act upon. (SAS Institute Inc. 2014)

3.3.4. Data Selection and Management. The collection and modification of data are the first and most important steps of the analytic journey. EDA helps find structure in data – whether in small samples or large volumes of data collected from many domains. JMP offers the tools needed to access, combine, filter and cleanse the data in preparation for data analysis. The interactive graphics and robust data analysis capabilities in JMP make it an ideal alternative to Excel for EDA and other types of statistical data analysis.

3.3.5. Linked Interactive Graphs and Analysis. The heart of JMP visuals are interactive graphs, supported by best analytics. Dynamic linking allows selections made on one graph or data table to be reflected in all graphs that are based on that table. The ability to view multiple graphs displaying the same selected data is one of the distinctive architectural underpinnings of JMP, which allows the analyst to explore the data and build on the analysis in multiple ways.

Perception is personal, and the open-ended nature of EDA means that analyst will develop his/her own style of analysis. JMP provides a wide repository of best-practice visualizations as part of the analysis output, so there are few limitations. Various tools allow to pan and probe these displays or zooming in for a closer look.

The innovative Graph Builder lets the data miner interactively build displays with multiple X and Y grouping variables, incorporating several types of graphs, including bar charts, histograms, line charts, heat maps and contour plots. Even with high-dimensional data, the data miner can find ways to see structure.

Added insight often comes from using multiple visualizations simultaneously, and dynamic linking and Data Filter capabilities in JMP make this approach especially useful.

Graph builder is the JMP tool used most frequently in this study during the phase of data validation and the analysis.

Figure 3.6 depicts the graph builder screen interface of JMP. As shown, there are several drop box regions for different variables. The analyst can interact with graph builder to create visualizations of the data by starting with drag and drop variables to place them where it is desired. Instant feedback encourages exploration and discovery of trends and behaviors based on how the variables are interacting between each other. The graph builder tool is flexible and inspires the data miner to change his/her mind and move variables to new positions to help in understanding the variables response to different case scenarios.

Graph builder is a powerful tool that helps in discovering and recognizing multi-dimensional relationships in the dataset with independent grouping variables for side-by-side or overlaid views.

Graph elements supported by graph builder include points, lines, bars, histograms, box plots heat maps, and contours. The underlying philosophy of Graph Builder is to visualize the dataset. To that end, the default visualization elements impose no assumptions, such as normality. Once the data are graphically represented, conclusions will be drawn directly, or decisions will be made where further analysis is needed to quantify relationships.

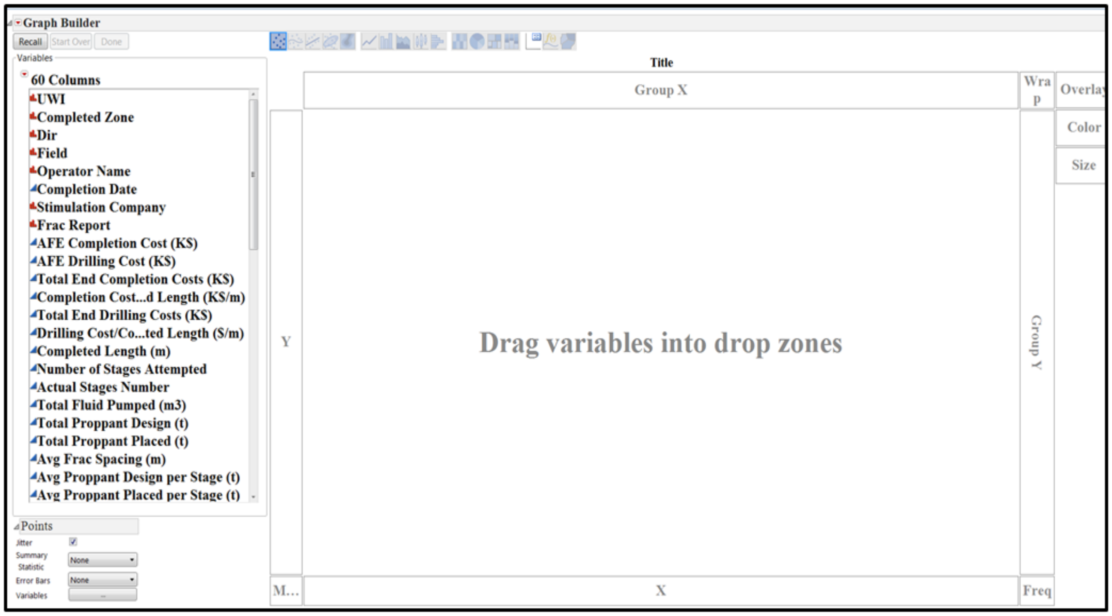

Figure 3.6 JMP Graph Builder Interface

The primary element in the Graph Builder window is the graph area which is the large open area shown in Figure 3.6. The graph area contains drop zones; variables from the Select Columns box on the left of Graph Builder window can be dragged and dropped into the preferred zones based on each zone function.

The following table describes the Graph Builder drop zones. The main drop zones function and descriptions within the graph builder are listed below in Table 3.1

Table 3.1 JMP Graph Builder Main Drop Zones Descriptions

JMP Graph Builder Drop Zones	
Drop Zone	**Description**
X, Y	Variables drop zone to assign the X or Y role.
Group X	Subsets or partitions the data based on the variable or variables that were selected. Displays the variable horizontally. Once a variable is placed there, no variable can be placed in Wrap.
Group Y	Subsets or partitions the data based on the variable or variables that were selected. Displays the variable vertically.
Map Shape	Drop variables there to create map shapes. If there is a variable in the Map Shape zone, the X and Y zones disappear.
Wrap	Subsets or partitions the data based on the variable or variables that were selected. Wraps the data horizontally and vertically. Once a variable is placed there, no variable can be placed in Group X.
Freq	Drop a variable there to use it as a frequency or weight for graph elements that use statistics, such as mean or counts.
Overlay	Groups the Y variables by the selected variable, overlays the responses, and marks the levels with different colors.
Color	The graph will be colored based on the drop variables. If a map or contour plot has been used, the map shapes or contours are colored. If the graph contains points, the points will be colored.
Size	Scales map shapes according to the size variable, minimizing distortion.
Legend	Shows descriptions of graph elements.

3.4. PHASES OF DATA MINING

Any given data mining project has a life cycle consisting of six phases listed as follows (Daniel T. Larose 2005):

1. Project Understanding Phase: This refers to the research understanding phase. At this stage, the research objectives and requirements should be enunciated clearly in terms of the research unit as a whole. Then the goals and restrictions should be translated into the preliminary strategy for achieving these objectives.
2. Data Understanding Phase: In this phase, the data is collected, and manipulated using exploratory data analysis in order to be familiarized with the data and discover initial insights, after that the data quality should be evaluated and if desired, the data may be divided into subsets contains actionable patterns.
3. Data Preparation Phase: In this phase, the final cleaned dataset is prepared to be used in the subsequent phases. The cases and the variables that are desired for the study should be identified and any data transformation and normalizations should be performed in order to have a clean raw dataset that is ready for analysis.
4. Modeling Phase: In this phase, the analysis is performed using the appropriate techniques and results are presented.
5. Evaluation Phase: The quality and the effectiveness of the analysis is evaluated in this phase before deploying applications for use in the field and a decision should be made regarding the proper use of the data mining results.
6. Deployment Phase: This is the final step of the data mining where a set of recommendations, or a report, will be generated to summarize the most significant outcomes of the research and address the limitations and the future improvements that are required for further rigorous model.

3.4.1. Project Understanding Phase. All the data mining stages were applied in this study starting with setting a well-defined objective. The objective is to use a pre-collected dataset that contains several parameters of more than 3300 horizontal wells in the Montney shale play to understand the effects of well design and completion strategies, coupled with the effects of hydraulic fracturing parameters, on well productivity performance. Setting this goal was the first step in the study and the dataset was prepared and organized to start the second phase of this research, which is represented by data understanding.

3.4.2. Dataset Understanding. Data understanding starts with an initial data collection and proceeds with activities to get familiar with the data in order to identify data quality problems, and to discover first insights into the data.

This phase of the study was accomplished by graphically representing the data. With the help of the JMP Pro Software Package, histograms were generated for each individual parameter. Based on the data distribution within the histograms an initial understanding of the data was gained which helped in the next phase of data preparation and cleaning.

A raw dataset was acquired from a commercial Canadian database. More than 3300 wells were in the dataset with different production, cost, completion and stimulation parameters.

The data initially acquired in the form of an excel spreadsheet (.xlsx) then it was exported and saved in the format of JMP (.jmp) in order to be able to use the JMP software for the data analysis and graphical representation.

Table 3.2 lists the original parameters included in the dataset. These parameters were classified into two groups: continuous and nominal parameters.

Initially, the dataset of wells completed in the Montney included 3369 wells. All the wells were horizontal with a completed lateral length ranging between 500 and 3000 meters. The reported completion strategy was either open hole or cased hole with no further information about thecompletion details whether the well completed barefoot, or with casing, liner or a pre-perforated or pre-slotted tubular. No details regarding the use of cement were available. The hudraulic fracturing delivery system (plug n perf or sliding sleeves) were also unknown. The wells were classified into either open or cased hole completions for the purpose of comparison and understanding the parameter distributions in the Montney dataset.

Table 3.2 shown below lists the parameters originally included in the dataset.

Table 3.2 List of Original Well Parameters in the Dataset

#	Parameter	#	Parameter
1	UWI (Unique Well Identifier)	20	Pumped Load Fluid (m3)
2	Field Name	21	Recovered Load Fluid (m3)
3	Operator Name	22	Completion Type (Open / Cased)
4	Completion Date	23	Base Fluid Group (Water / Oil)
5	Stimulation Company Name	24	Base Fluid (Slick Water, Surfactant, Water, Oil)
6	AFE Completion Cost (K$)	25	Energizers (CO2, N2, CO2/N2)
7	AFE Drilling Cost (K$)	26	IP Water (bwpd)
8	Total End of Completion Cost (K$)	27	IP Oil (bopd)
9	Total End of Drilling Cost (K$)	28	IP Gas (mcf/d)
10	Completed Lateral Length (m)	29	Water Production 6 Months Cum. (mbw)
11	Number of Stages Attempted	30	Water Production 12 Months Cum. (mbw)
12	Actual Number of Stages	31	Water Production 18 Months Cum. (mbw)
13	Total Proppant Designed (tonne)	32	Oil Production 6 Months Cum. (mbo)
14	Total Proppant Placed (tonne)	33	Oil Production 12 Months Cum. (mbo)
15	Total Fluid Pumped (m3)	34	Oil Production 18 Months Cum. (mbo)
16	Avg. Frac. Spacing (m)	35	Gas Production 6 Months Cum. (mmcf)
17	Avg. Proppant Placed per Stage (tonne/stage)	36	Gas Production 12 Months Cum. (mmcf)
18	Avg. Fluid Pumped per Stage (m3)	37	Gas Production 18 Months Cum. (mmcf)
19	Avg. Closure Gradient (Kpa/m)		

Table 3.3 represents a list of the parameters that were calculated and added to the dataset for each well. These parameters were calculated to provide normalization and aid in the comparison process.

Table 3.3 List of Calculated Parameters in the Dataset

#	**Parameter**
1	Completion Cost / Completed Lateral Length (K$/m)
2	Drilling Cost / Completed Lateral Length (K$/m)
3	Load Fluid Recovery Percentage
4	18 Months Cum. Gas Production / Completed Lateral Length (mmcf/m)
5	12 Months Cum. Gas Production / Completed Lateral Length (mmcf/m)
6	6 Months Cum. Gas Production / Completed Lateral Length (mmcf/m)
7	18 Months Cum. Gas Production / Actual Stages Number (mmcf/stage)
8	12 Months Cum. Gas Production / Actual Stages Number (mmcf/stage)
9	6 Months Cum. Gas Production / Actual Stages Number (mmcf/stage)
10	18 Months Cum. Gas Production / Avg. Frac. Spacing (mmcf/m)
11	12 Months Cum. Gas Production / Avg. Frac. Spacing (mmcf/m)
12	6 Months Cum. Gas Production / Avg. Frac. Spacing (mmcf/m)
13	Total Proppant Placed / Completed Lateral Length (tonne/m)
14	Total Fluid Pumped / Completed Lateral Length (m3/m)
15	Drilling AFE Cost - Drilling Final Cost ($)
16	Completion AFE Cost - Completion Final Cost ($)
17	Attempted Stages - Actual Stages (stage)
18	Total Proppant Designed - Total Proppant Placed (tonne)
19	Avg. Proppant Concentration (lbs./gal)
20	Avg. Total Fluid Pumped per Stage (m^3/stage)
21	Avg. Total Proppant Pumped per Stage (tonne/stage)

3.4.2.1 Fields in the Montney. There are 125 fields listed in the dataset but not all of these fields contain a large number of wells to be sufficient for statistical analysis on a field basis. The top 30 fields comrise 90% of the entire Montney wells. Table 3.4 lists the top 30 fields and their associated wellcount.

Table 3.4 Montney Top 30 Fields in Well Count Reported in the Dataset

#	Field Name	Number of Wells	Cumulative Well Percentage % in the Dataset
1	REGIONAL HERITAGE	1134	33.7
2	NORTHERN MONTNEY	374	44.8
3	KAYBOB SOUTH	185	50.3
4	POUCE COUPE SOUTH	182	55.7
5	KAYBOB	133	59.6
6	GIROUXVILLE EAST	116	63.0
7	GLACIER	101	66.0
8	NORMANDVILLE	87	68.6
9	ANTE CREEK NORTH	71	70.7
10	ANTE CREEK	61	72.5
11	ELMWORTH	51	74.1
12	STURGEON LAKE SOUTH	51	75.6
13	WASKAHIGAN	51	77.1
14	VALHALLA	47	78.5
15	FIR	42	79.7
16	KARR	40	80.9
17	SINCLAIR - ALTA	35	82.0
18	DIXONVILLE	33	82.9
19	SIMONETTE	33	83.9
20	RYCROFT	31	84.8
21	KAKWA	25	85.6
22	WORSLEY	21	86.2
23	POUCE COUPE	20	86.8
24	ALTARES	19	87.4
25	FOX CREEK	19	87.9
26	NIG CREEK	16	88.4
27	GRIMSHAW	15	88.8
28	RESTHAVEN	15	89.3
29	TANGENT	15	89.7
30	WAPITI	15	90.2
Total		3038	

Figure 3.7 is a histogram of the top 20 fields by well count, which that represents about 85% of the entire number of wells in the dataset. The dark shaded parts in the histogram pertain to the cased hole completed wells. It can be observed that cased hole dominate the wells completion type in Northern Montney field and more than 70% of the wells in Regional Heritage field.

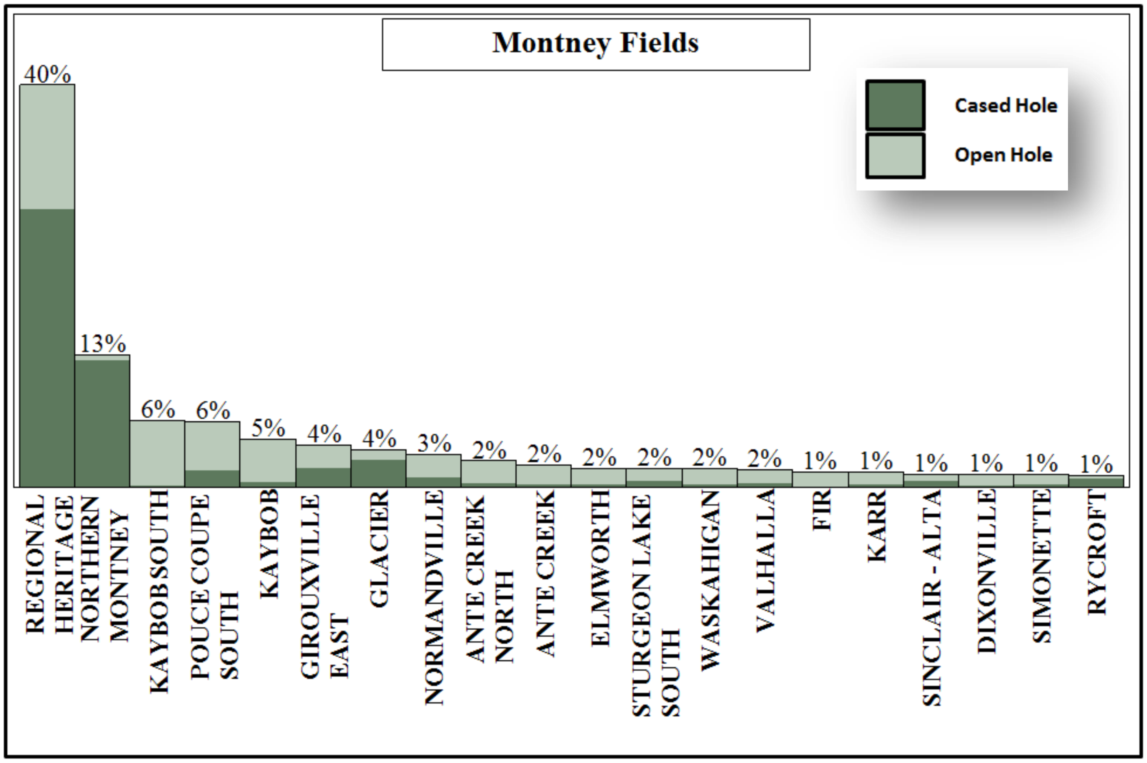

Figure 3.7 Montney Data Distribution Histogram of the Top 20 Fields in Well Count

3.4.2.2 Major operators in the Montney. There are 113 operators listed in the dataset but not all companies operates a large number of wells. The top 30 operators in the Montney by well counts, comprise 87% of the entire Montney wells in the dataset as illustrated in Table 3.5.

Figure 3.8 is a histogram of the top 20 operators by well counts, which represents about 79% of the entire number of wells in the dataset. Knowing the operators can help as a grouping factor during the analysis and comparisons between the parameters.

Table 3.5 Montney Top 30 Major Operators in Well Count

#	Operating Company	Number of wells	Cumulative Well Percentage % in the Dataset
1	EnCana Corporation	429	13
2	ARC Energy Trust	245	20
3	Shell Canada Limited	239	27
4	Murphy Canada Exploration Company	199	33
5	Progress Energy Resources Corp.	184	39
6	Talisman Energy	177	44
7	Trilogy Energy	155	48
8	Celtic Exploration Ltd.	146	53
9	Canadian Natural Resources	139	57
10	Galleon Energy Inc.	111	60
11	Birchcliff Energy Inc.	106	63
12	Advantage Oil & Gas Ltd.	94	66
13	Athabasca Oil Sands	74	68
14	Tourmaline Oil	72	70
15	Long Run Exploration	64	72
16	Paramount Resources Ltd.	54	74
17	Crew Energy Inc.	53	76
18	RMP Energy	43	77
19	Devon Canada Corporation	34	78
20	ConocoPhillips Canada Resource Corp.	33	79
21	NAL Oil & Gas Trust	32	80
22	Orleans Energy	29	81
23	Canbriam Energy Inc.	27	81
24	Cequence Energy	27	82
25	Guide Exploration	27	83
26	Storm Exploration	26	84
27	Huron Energy Corporation	25	85
28	Nuvista Energy	25	85
29	Daylight Energy Ltd.	23	86
30	Painted Pony Petroleum	23	87

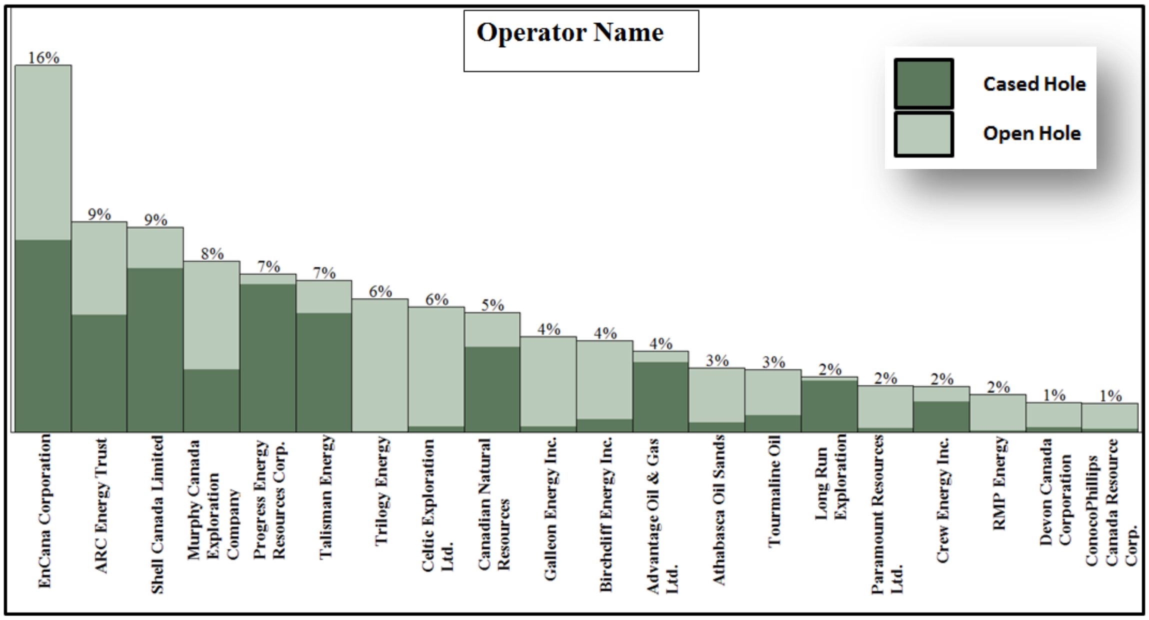

Figure 3.8 Data Distribution Histogram of Top 20 Operators in the Montney Dataset

Figure 3.8 shows that some of the operators mainly employ cased hole completion in their wells, for example Progress Energy Resources Corporation. Other operators such as Trilogy Energy tend to mainly complete their wells with open hole completions while the rest of the operators tried both completion technologies in their wells.

3.4.2.3 Completion date. Ninety three percent of the wells had a completion date in the data set and the majority of the wells were completed after 2008. There were 11wells reported to be completed between 1997 and 2004. These wells were removed from the dataset for the sake of consistency in technology, as these older completion might not reflect the same design and completion concepts as the newer wells.

To better understand trends in the types of completion methods employed over the time in Montney, a completion date histogram was prepared, as it is shown in Figure 3.9.

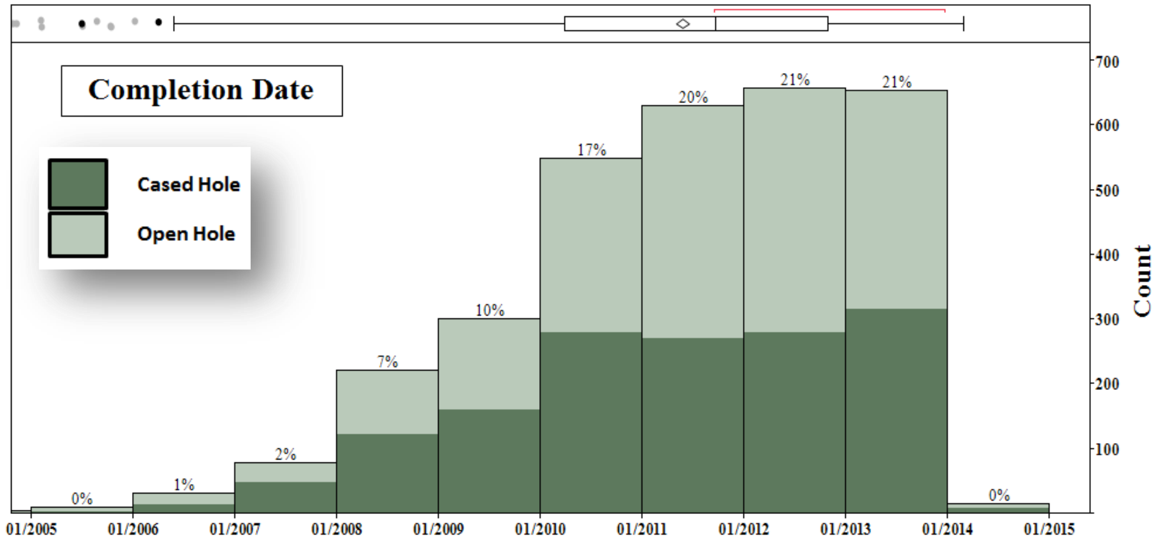

Figure 3.9 Completion Date Distribution Histogram

Figure 3.9 suggests that slightly more than half of the wells were completed with cased hole completion up to 2011. During 2011 and 2012 less than half of the wells were completed with cased hole completion. Then in 2013, the percentage of cased hole wells increased again to be just about 50% of the wells drilled during the past year.

3.4.2.4 Stimulation company. Fourteen different stimulation service companies were reported to be operating in the Montney, Calfrac, Trican, Halliburton, Canyon and Schulumberger are the most dominant stimulation companies, performing treatments for 74% of the wells in the data set.

Table 3.6 summarize the list of stimulation companies in the dataset along with the percentage of stimulated wells in the entire Montney dataset. The distribution histogram of the stimulation companies in the Montney data set is shown below in Figure 3.10. The dark shaded sections refer to the cased hole completion and it is shown that all the stimulation companies in the dataset were involved in stimulating both types of completions.

Table 3.6 List of Stimulation Companies in the Montney

Stimulation Company			
#	**Stimulation Company**	**Number of wells**	**Percentage (%)**
1	Calfrac	763	23.7
2	Trican	730	22.6
3	Halliburton	346	10.7
4	Canyon	326	10.1
5	Schlumberger	218	6.8
6	Sanjel	217	6.7
7	None	191	5.9
8	Baker Hughes	179	5.6
9	BJ Services	154	4.8
10	Unknown	69	2.1
11	Century	22	0.7
12	GasFrac	2	0.06
13	Other	2	0.06
14	Press Truck	2	0.06
15	Nabors	1	0.03
16	Nowsco	1	0.03
Total		**3223**	**100.0**

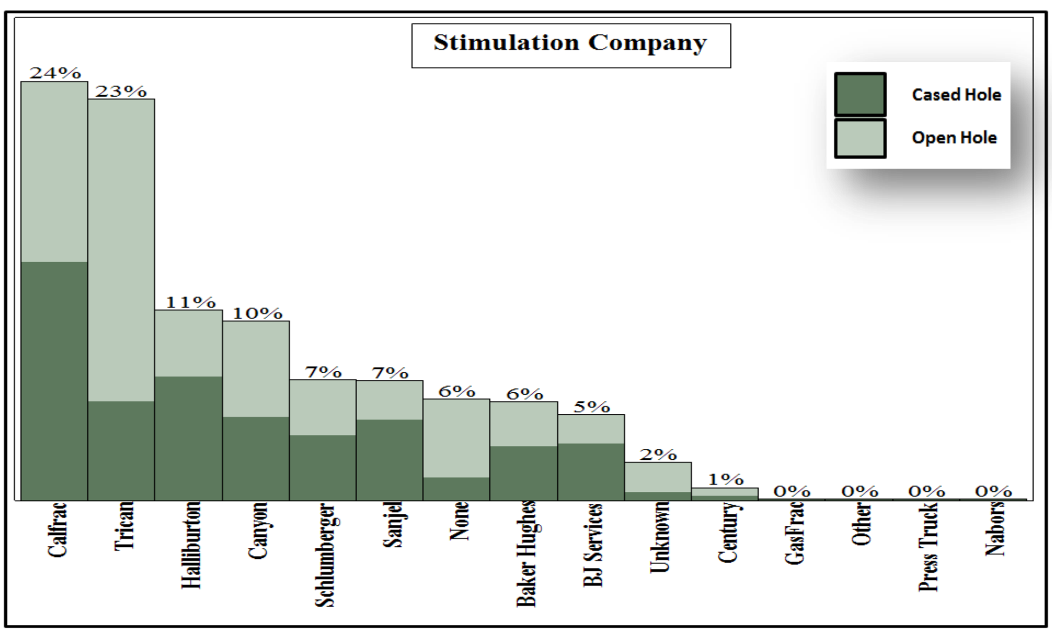

Figure 3.10 Stimulation Companies in the Montney Data Distribution Histogram

3.4.2.5 Completion type. A total of 47% of the reported wells in the dataset were completed with cased hole completion type and 53% with open hole. Table 3.7 lists the number of wells and the percentage for the two completion types. Figure 3.11 provides a graphical representation of cased hole and open hole completion type for the entire dataset. There is good sample density (number of wells) between the two completion types that helps in establishing reliable statistical analysis when comparing between the two completion methods.

Table 3.7 Completion Type Well Percentage in the Montney

Completion Type	**Number of Wells**	**Well Percentage (%)**
Cased	1521	47
Open	1692	53
Total	3213	100

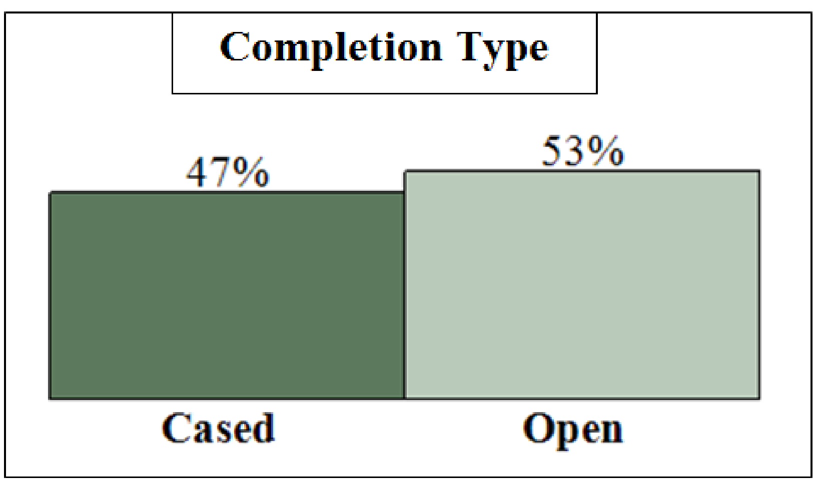

Figure 3.11 Well Completion Type Data Distribution Histogram in the Montney Dataset

3.4.2.6 Completed lateral length (m). Only 33% of the wells in the dataset used in this study have a reported completed lateral length. The lateral length parameter displayed a normal bell shape curve distribution with a mean value of about 1500 m, within a range between 500 and 3000 m. The same distribution applies to both types of completion, with a slight increase in cased hole completion wells as the lateral length increases (Figure 3.12.)

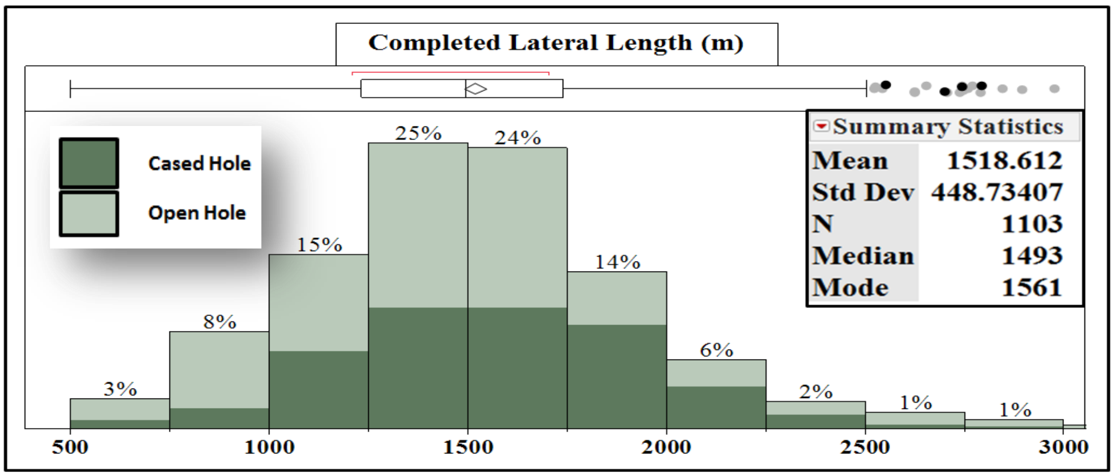

Figure 3.12 Completed Lateral Length Data Distribution in the Montney Dataset

3.4.2.7 Number of stages. The dataset used in the study differenciated between the "attempted" number of frac stages and the "actual" number of stages pumped, although only 37% of the wells had the attempted number of stages while 90% of the wells had the actual final number of stages. Knowing both the attempted and actual stage number helps in validating the data set reliability by cross plotting both parameters and identify the outliers as shown in the next phase of data mining (3.4.3.3). Having both parameters could also help in differentiating between various completion and design parameters to appreciate which method results in the least difference between attempted and actual parameter

Figure 3.13 shows the distribution histograms of the attempted and actual stage number for the wells in the Montney dataset. It can be observed that both measured values have the same distribution percentage, which reflects the good quality of the data set. The wells in the Montney were fracture stimulated with a wide range of stages, up to 32 stages. The most frequent number of stages (statistical mode) achieved in the Montney was 8 stages per well with 11% over the entire Montney play. The arithmetic mean for the actual number of stages across the dataset is 12 stages per well, but more than 50% of the wells were stimulated with a range of 7 to 14 stages. Case hole completion dominate the lower range of while the open hole completion starts to dominate at about 12 fracturing stages. The data set does not include the number of perforation clusters within each stage, and the analysis was performed on a per stage basis.

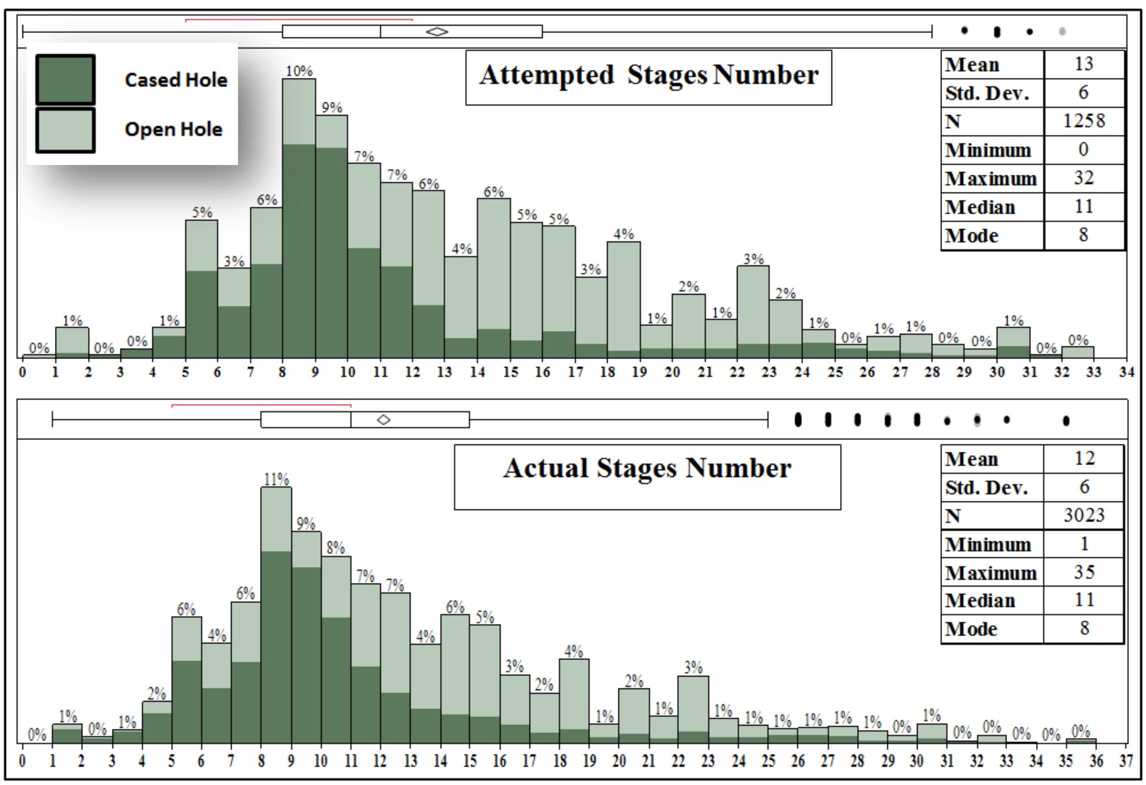

Figure 3.13 Attempted and Actual Stages Number Data Distribution Histogram and Statistical Summary Table

3.4.2.8 Total fluids pumped (m^3). Eighty percent of the wells have total pumped fluid volume reported in the dataset. More than 50% of the wells were treated with less than 2000 M^3 of fluids (520,000 Gal).

Figure 3.14 depicts the total pumped fluid distributions of 2698 wells in the Montney dataset. The data shows a wide spread in the range of actual treatment fluid volume up to 20,000 M^3(5,200,000Gal). A few scattered wells were reported to be treated with higher fluids volume up to a maximum of 51965 M^3. These high volume treatments might represent outliers in the data set, and further validation is needed to ensure that the analyzed parameters are representative and valid, as it discussed in section 3.4.3 of this chapter. The dark shaded parts represents the percentage of cased hole versus the open hole completions, higher fluid volume were pumped more frequently in the cased hole wells.

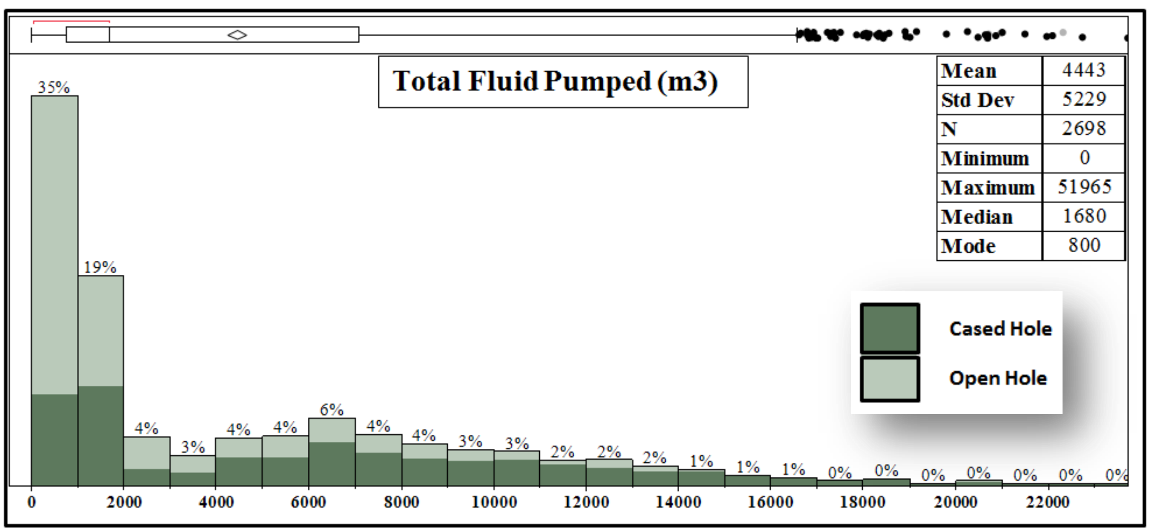

Figure 3.14 Total Pumped Fluid Data Distribution Histogram and Statistical Summary

Load fluid is another treatment fluid parameter reported in the dataset. Load fluid is a term used by the hydraulic fracturing industry to refer to the total amount of fluids that are pumped into the well. Having this parameter in the dataset helps in the data validation stage, as load fluid and total pumped fluids are almost the same parameter.

Thirty seven percent of the wells have the load fluid volume reported in the dataset. Figure 3.15 confirms that the load fluid and total pumped fluid are almost the same, by exhibiting the same distribution trend.

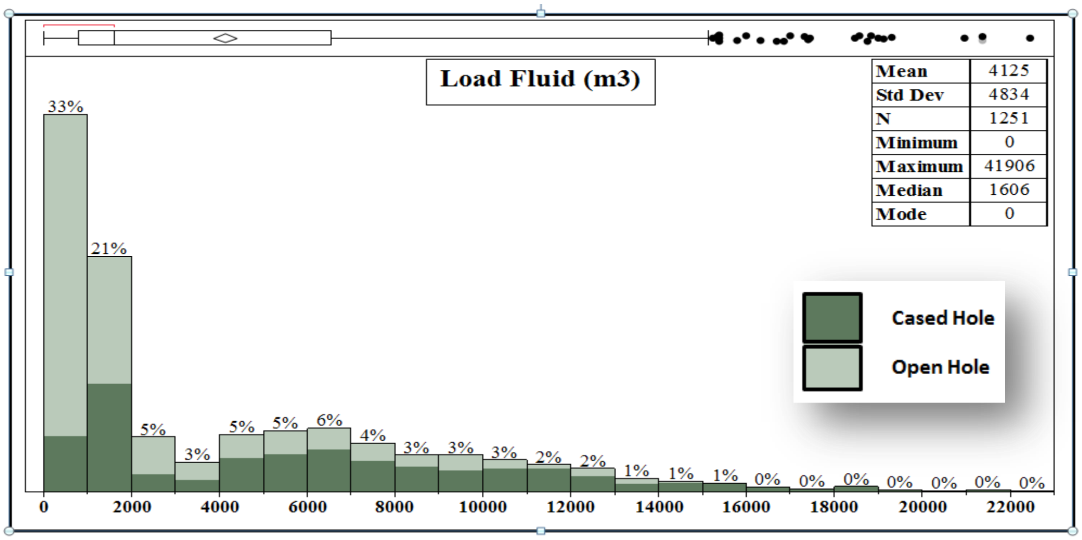

Figure 3.15 Load Fluid Data Distribution Histogram

3.4.2.9 Recovered load fluid (m^3). 37% of the wells had this parameter reported in the dataset. This parameter represents a measure of how much treatment fluid was recovered on the surface. This parameter was used to calculate the recovery fluid percentage, by dividing it by total actual pumped fluid. Figure 3.16 shows the distribution histogram of the recovered fluid volume. The distribution shows that 65% of the wells recovered less than 1000 m^3 (264,000 US Gal). It also shows that cased hole completion is more dominant in the higher recovery volume, but this can be correlated to the fact that the cased hole wells were usually treated with higher fluid volume. This is confirmed by analyzing the recover percentage parameters, as it is shown in the next sub-section.

It is important to note that the methods used to detect or classify the recovered load fluid are not explained in the dataset. Hence it is difficult to ensure the consistency of these measured data.

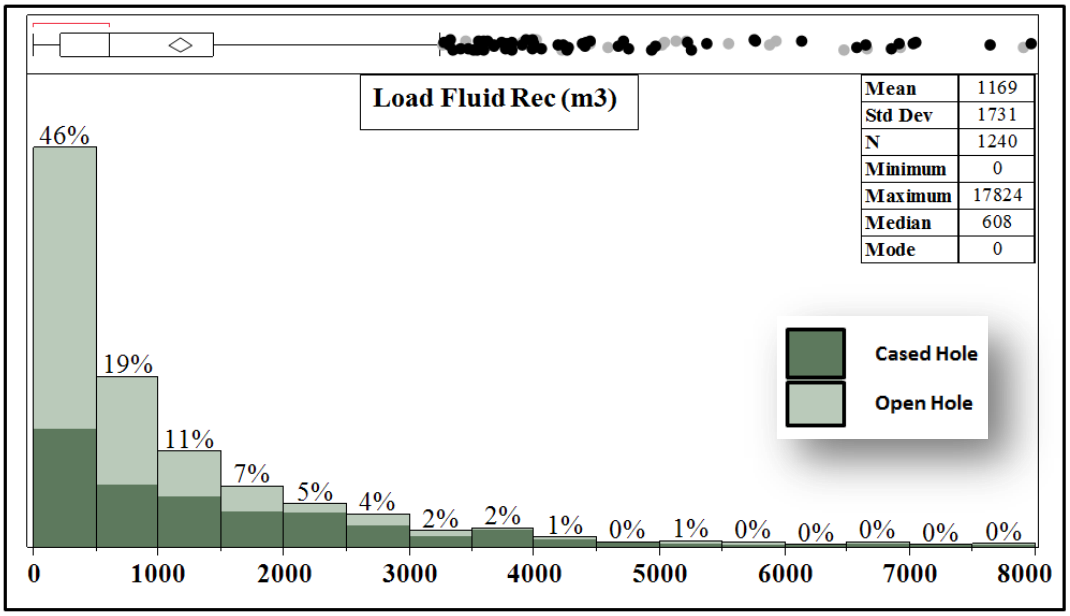

Figure 3.16 Recovered Load Fluid (M3) Data Distribution Histogram

3.4.2.10 Load fluid recovery percentage. This parameter was calculated for 37% of the wells in the original dataset by dividing the recovered load fluid volume by the total actual pumped fluid. Figure 3.17 shows the data distribution histogram of fracturing fluid recovery percent in the Montney. It is shown that the Montney typical recovery percentage lies between 0.1 to 0.35 (10 to 35 %) where half of the wells fall into this range of recovery. It is also shown that the cased hole wells were in a good conformance with this range. The data distribution also showed that 40% of the open hole wells have a higher value of recovered fluid percent. To understand the contributing parameters of higher recovery range, further analysis was performed by highlighting the histogram based on the type of fluid used in the treatment. Figure 3.18 shows the fluid recovery distribution histogram, but now with a highlighting that distinguishes the percentage of the wells that were treated with oil base fracturing fluids. It is clear that treatment with oil base fluid yields in higher recovery percentage and up to 100%.

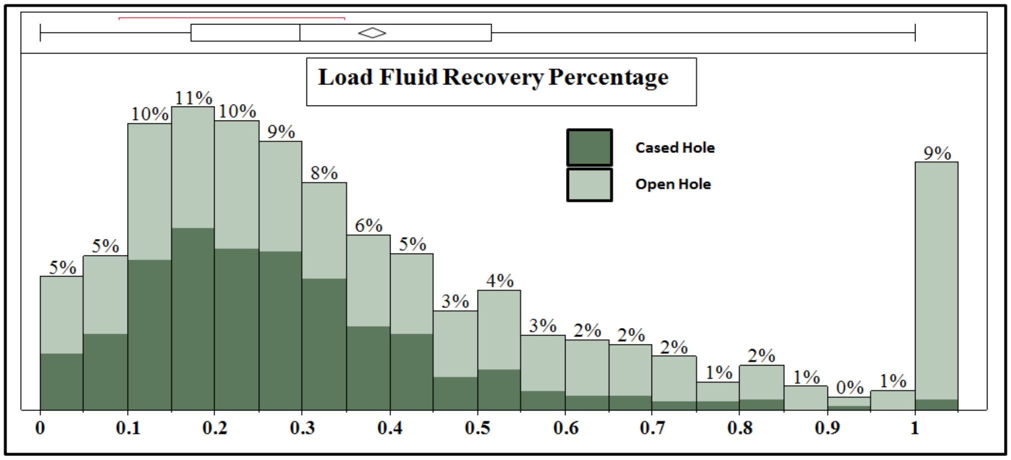

Figure 3.17 Load Fluid Recovery Percentage Data Distribution Histogram

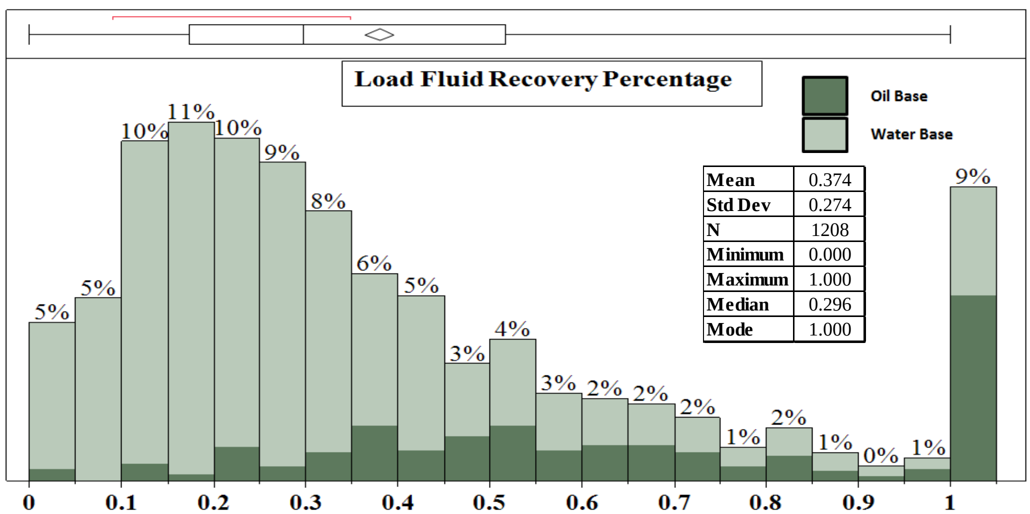

Figure 3.18 Load Fluid Recovery Percentage Data Distribution Histogram (Shaded with Respect to Fluid Types)

3.4.2.11 Normalized average fluid pumped per stage (m^3/stage). This parameter was calculated to obtain the normalized value of the total pumped fluid on a per stage analysis. The parameter was calculated for 80% of the wells based on the availability of total pumped fluids and actual number of stages in the dataset. The distribution histogram shown in Figure 3.19 indicates that most of the open hole wells were treated with less than 900 M^3 of fracturing fluid and 90% of these wells were actually treated with less than 500 M^3 (132,000 US Gal). Cased hole wells were treated with higher job volume per stage with more than 50% of the wells treated with over 500 M^3 (132,000 US Gal). This may be an indication that some operators are preferring to go with open hole completions to save in the upfront completion and stimulation costs.

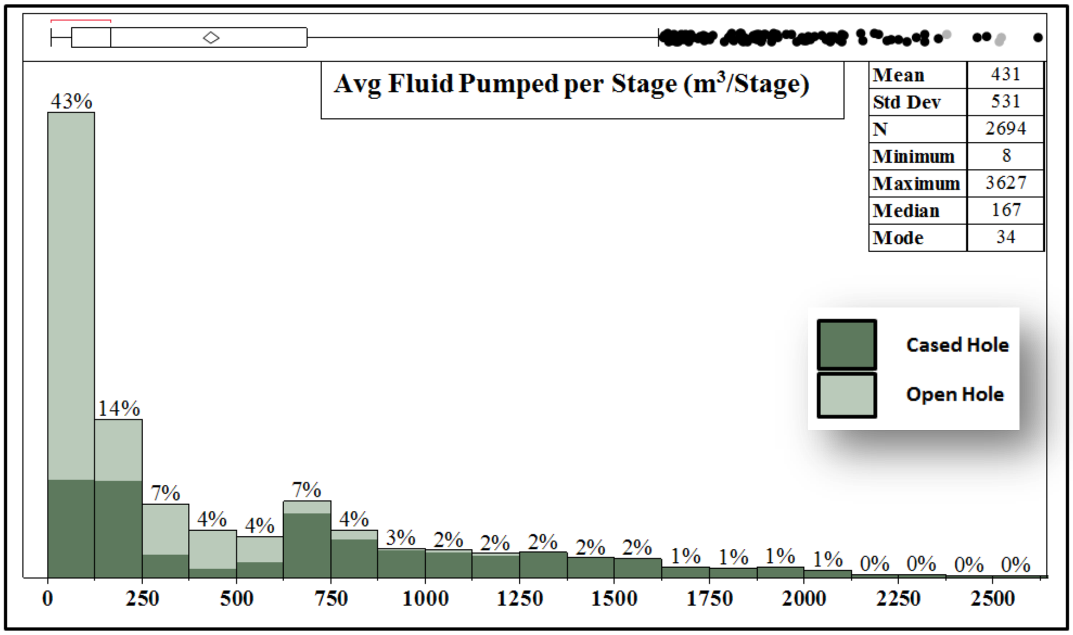

Figure 3.19 Avg. Total Pumped Fluid per Stage (m^3/stage) Data Distribution Histogram

3.4.2.12 Designed and pumped proppant (tonne). 88% of the wells in the database had data regarding the total amount of proppant placed, but no specific information about the types of proppant or proppant size were given. Figure 3.20 illustrates data distribution histograms of both designed and pumped proppant mass. It is clear that 65% of the wells in the Montney were treated with less than 1250 tonnes (2.7 million lbs.) the rest of the wells were treated with higher proppant masses, up to 2500 tonnes (5.5 million lbs.). It is also shown that higher proppant mass are associated with cased hole completion. Few wells had very high amount of proppant, which need a cross validation to confirm and detect possible outliers.

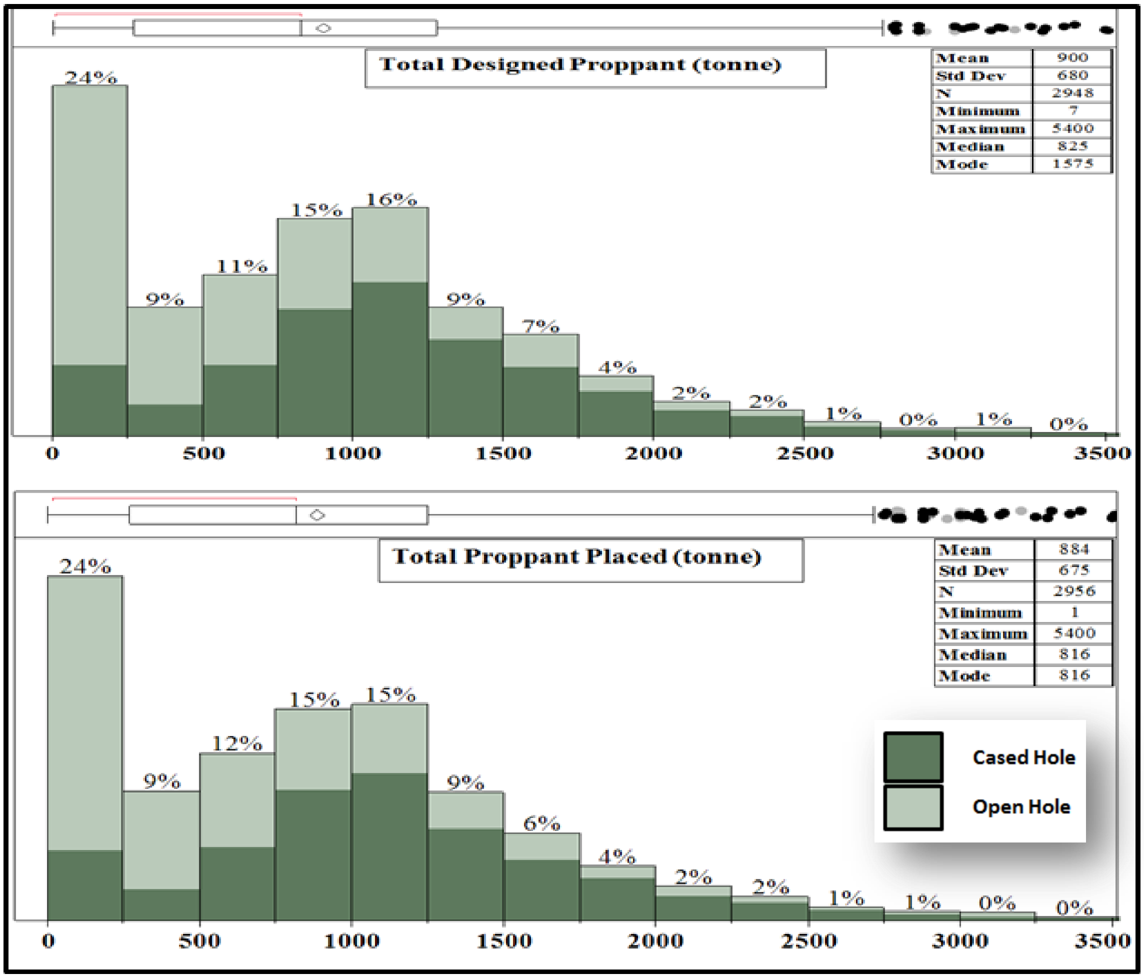

Figure 3.20 Total Designed and Placed Proppant Data Distribution Histogram

3.4.2.13 Normalized average placed proppant per stage (tonne/stage). The placed proppant was normalized by number of stages. Figure 3.21 shows the data distribution of this parameter. The figure shows that cased hole wells were treated with high amount of proppant per stage compared to the open hole completion. Most of the open hole wells were treated with less than 100 tonne/stage (220,000 lbs./stage) of proppant while many of the cased hole wells were treated with more than 100 tonnes/stage.

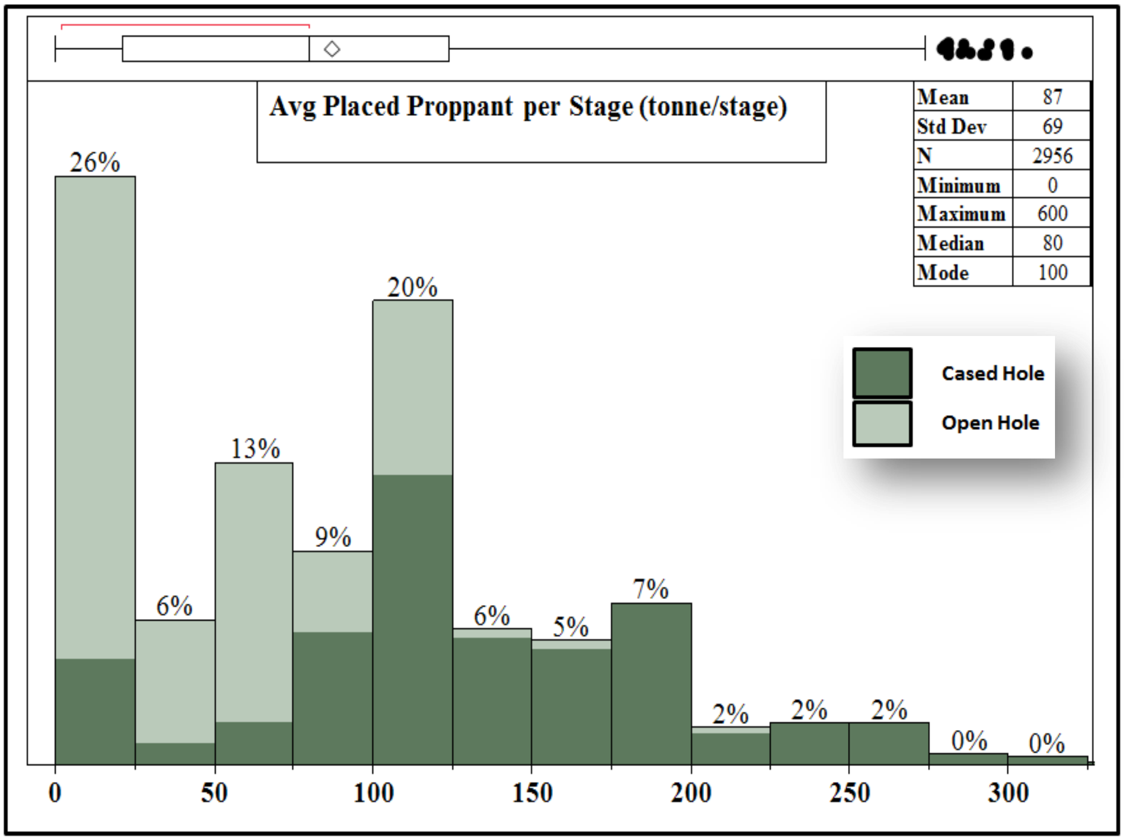

Figure 3.21 Normalized Avg. Placed Proppant per Stage Data Distribution Histogram

3.4.2.14 Normalized proppant pumped per length (tonne/m). This parameter was calculated and added to the dataset to be used in the analysis phase as a parameter that include the effects of amount of proppant pumped normalized by lateral length. Figure 3.22 shows the data distribution histogram for this normalized parameter. The diagram confirms that the cased hole wells, shown with dark shades, are commonly treated with more than 0.6 tonne/m while the open hole wells usually treated with less amount of proppant per meter length.

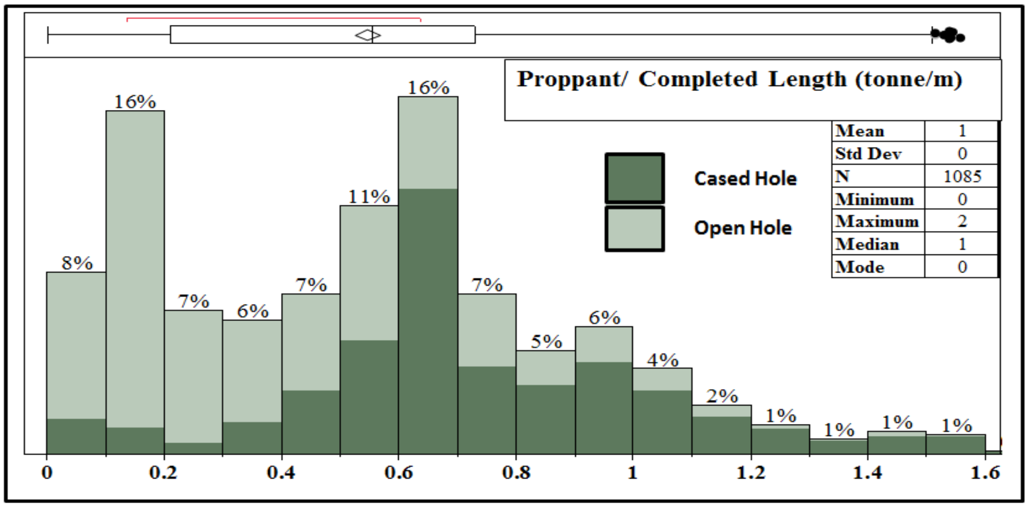

Figure 3.22 Normalized Total Proppant Pumped per Completed Lateral Length Data Distribution Histogram

3.4.2.15 Average fracturing spacing (m). This parameter represents the distance between the fracturing stages. Only 37% of the wells in the dataset included this information. Figure 3.23 displays the distribution of the data. Statistics for the well samples shows that the average spacing between stages in the Montney is 165 m and it is also evident that more than 70% of the wells stage's spacing fall in the range of 50 to 150 m. Open hole wells have smaller spacing between the stages. This is in consistent with the previous analysis of number of stages, where it was shown that higher number of stages were associated with open hole completion. It is important to recall the number of clusters per stage is not available for this study.

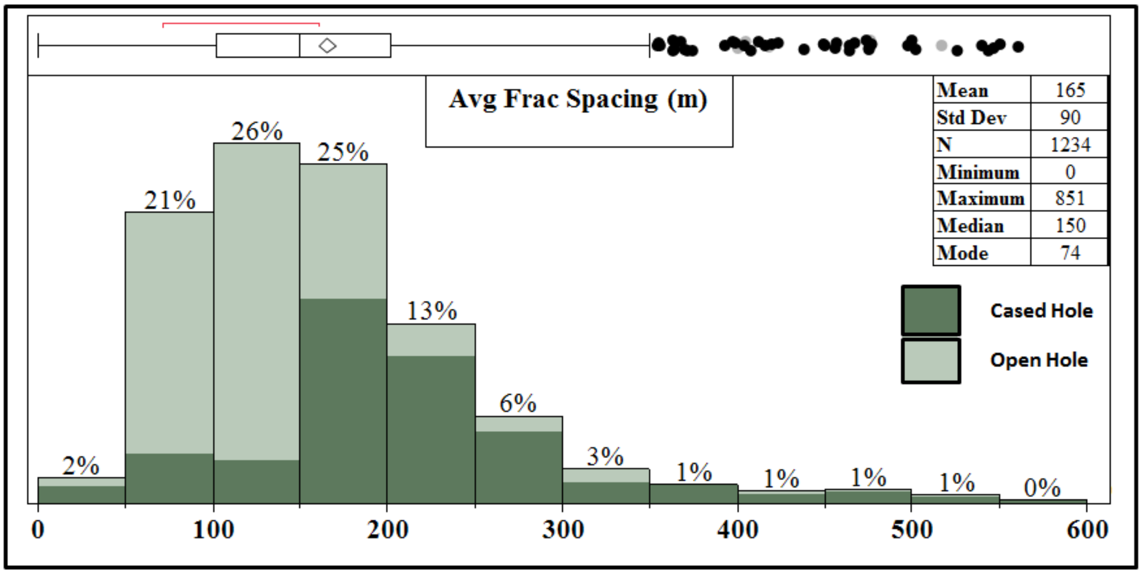

Figure 3.23 Average Fracture Spacing Data Distribution Histogram

3.4.2.16 Fracture closure gradient (Kpa/m). Closure stress is the minimum horizontal stress. Rocks with high closure stress take more horsepower to fracture than the same rocks with lower closure stress. The dataset has only 297 wells that report the closure pressure gradient.

The data distribution of this parameter is shown in Figure 3.24. It is observed that the cased hole wells are heavily distributed in the high closure gradient range while most of the open hole wells are concentrated in the low fracture closure gradient range. Table 3.8 converts the closure gradients from Canadian to the US units.

This observation needs more investigation and validation as the sample of data is relatively low and the methods of measuring the closure pressure is not reported in the dataset. It is not clear why cased hole completion would be associated with higher closure stress, unless operators simply prefers to use cased hole completions in the pre-identified regions of high closure gradient. Open hole completions required added pressure to activate the packers and it becomes more difficult to operate in high stress environments.

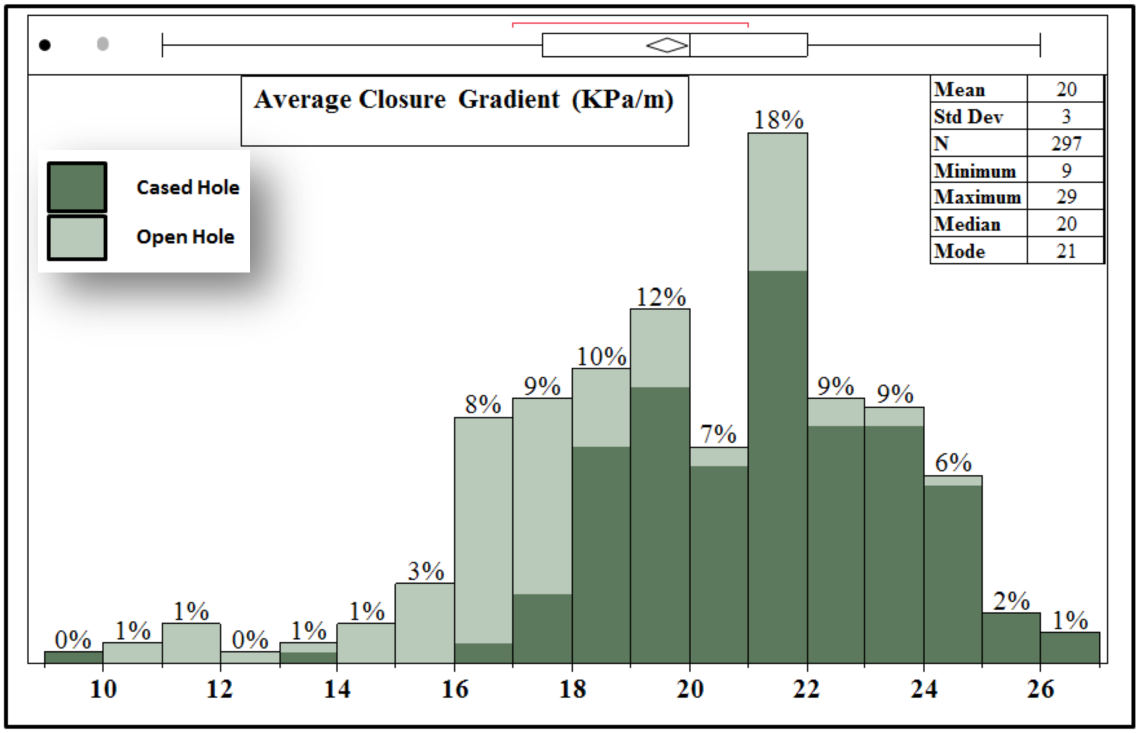

Figure 3.24 Avg. Fracture Closure Gradient Data Distribution Histogram

Table 3.8 Closure Pressure Gradient Unit Conversion

Closure Gradient in (Kpa/m)	Closure Gradient in (psi/ft.)
10	0.44
12	0.53
14	0.62
16	0.71
18	0.80
20	0.88
22	0.97
24	1.06
26	1.15

3.4.2.17 Avg. proppant concentration (lbs./gal). This parameter was calculated by dividing the total pumped proppant mass by the total pumped fluid volume, and applying an appropriate conversion factors to obtain the equivalent average proppant concentration in pounds per gallon units. This parameter combines the effects of total proppant and fluids pumped in the wells. Figure 3.25 depicts the distribution histogram for 80% of the wells in the data set. It is shown that 70% of the wells in the Montney were treated with a proppant concentration ranging between 0.5 to 2.5 (lbs./gal). Fewer wells were treated with higher proppant concentration, up to 8 lbs./gal, with the exceptions of very few removed outliers that exhibited extremely high and unreasonable concentrations value.

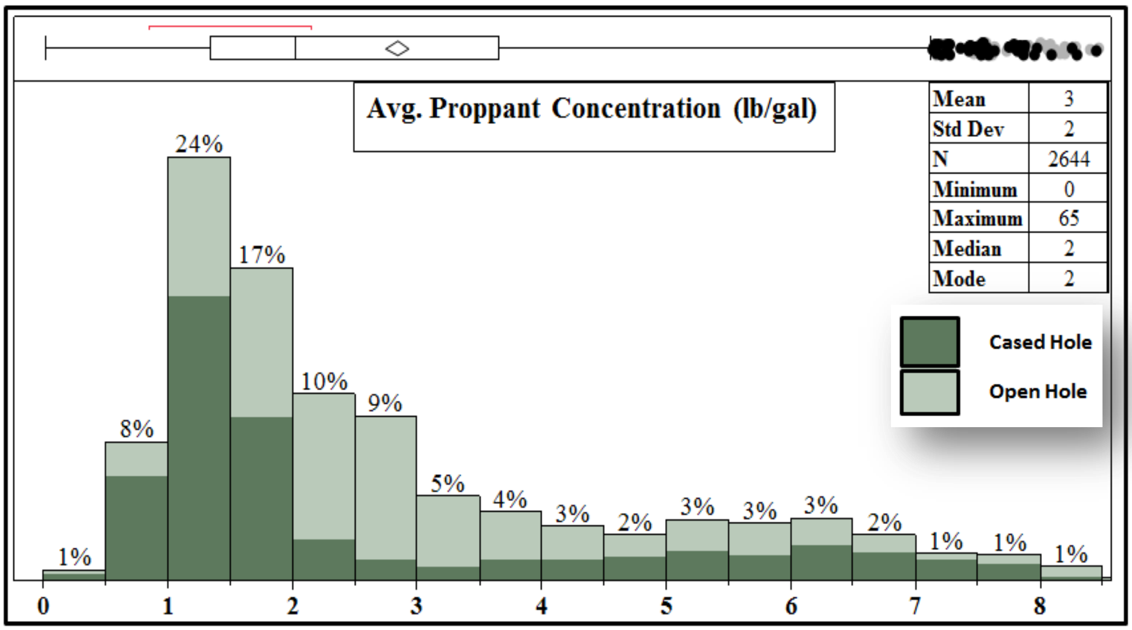

Figure 3.25 Average Proppant Concentration (lbs./gal) Data Distribution Histogram

3.4.2.18 Base fluid group. The dataset had two distinct basic groups of fracturing fluids, either water or oil base. 14 wells were reported to be treated with acid base fluid, two wells were reported to be stimulated with gas and 5 wells were reported to be treated with a mixture of oil and water. These wells were flagged and removed from the data set because of the insufficient number of wells to be included in a statistical comparison. Figure 3.26 and Figure 3.27 shows the data distribution of the original base fluid types in the dataset and the data distribution after eliminating the unwanted fluid groups, respectively.

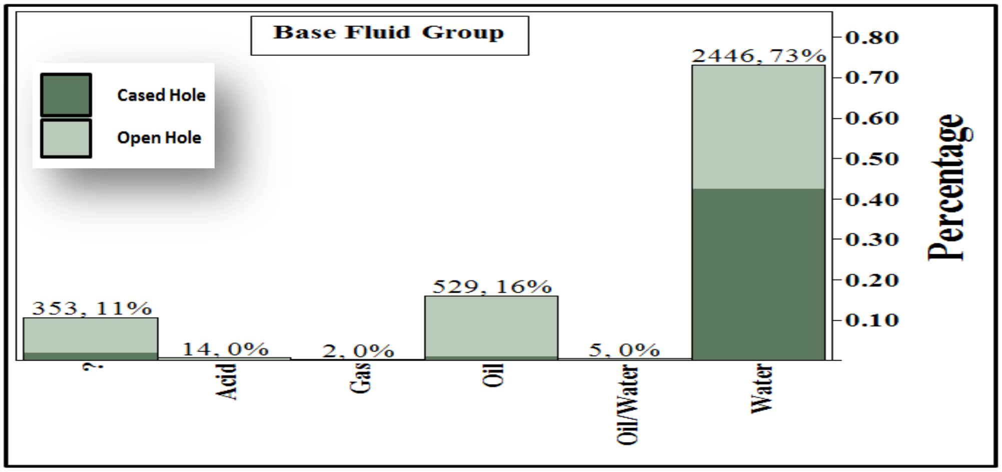

Figure 3.26 Fracturing Fluid Base Groups that were Originally in the Dataset

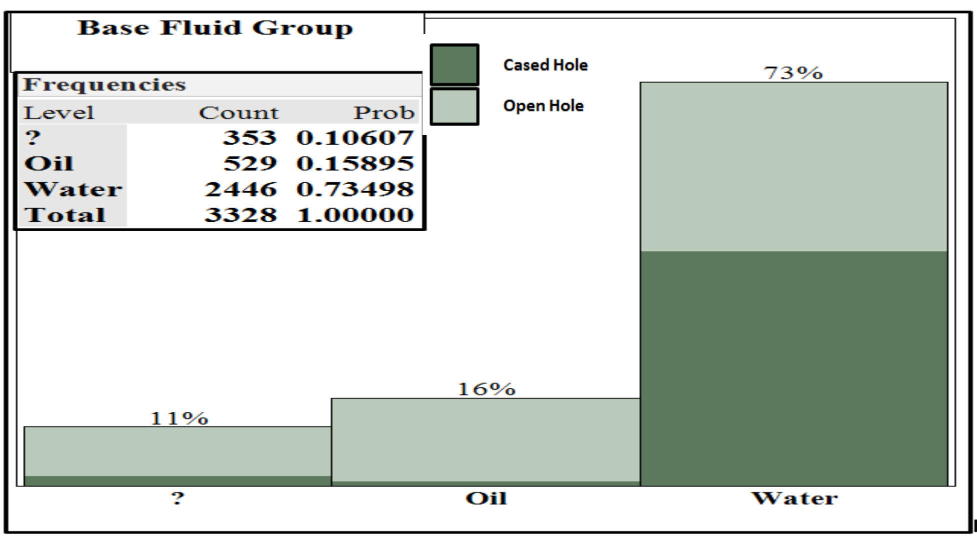

Figure 3.27 Fracturing Fluid Base Groups after Modifying the Dataset

The data histograms of the base fluid showed that the oil based treatment fluid was mainly carried out in the open hole completion.

3.4.2.19 Fracturing base fluid. In this categorical parameter, the water base fluid is more defined where it is sub-divided into water, surfactant and slick water. The difference between water and slick water is not defined in the dataset. They could be the same class, as some operators refer to the slick water and water. Figure 3.28 depicts the percentage of each fluid type in the dataset.

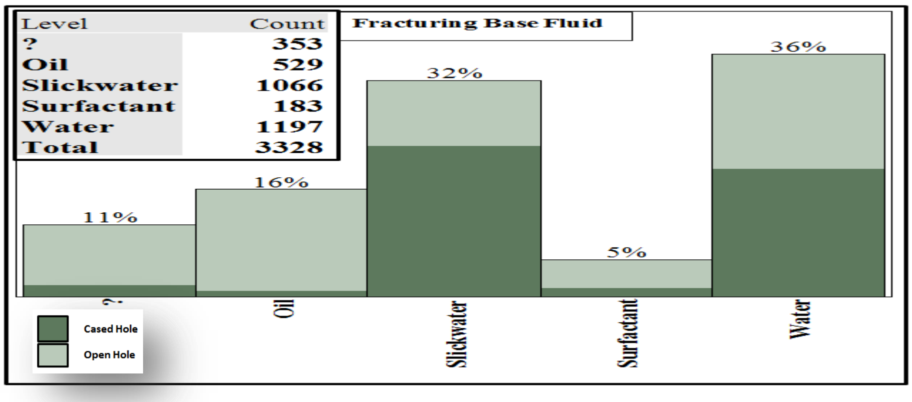

Figure 3.28 Hydraulic Fracturing Base Fluid Types Data Distribution Histogram

3.4.2.20 Fracturing fluid energizers. Approximately 20%–25% of all treatments contain an energizing gas (Economides 2000). In the study dataset, 47% of the wells were treated with gas energized fluids as it is shown in Figure 3.29.

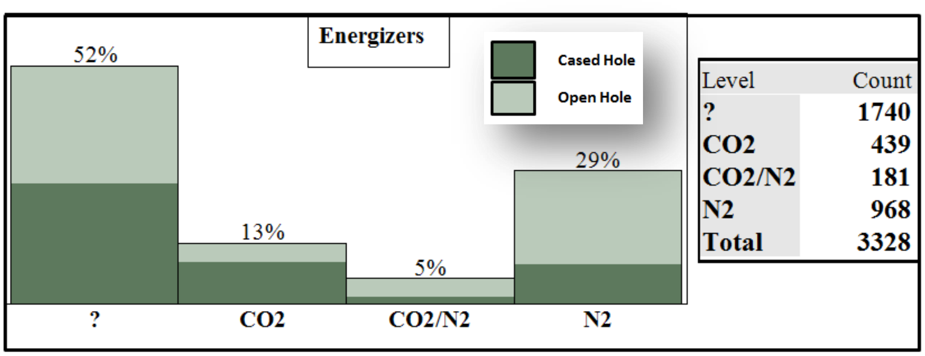

Figure 3.29 Fracturing Fluid Energizers Data Distribution Histogram

3.4.3. Dataset Preparation and Validation Phase. Identifying correlations and general trends between parameters such as well design, completion, stimulation, cost and production with a large dataset with more than 3300 wells, requires reliable data quality. Reliable data play a key role in the analysis of the parameters as bad quality data may result in misleading interpretation and may reduce the correlation coefficient.

It should be recognized that almost all datasets have some level of error and the dataset used in this study is likely no exception. Sources of possible errors may include general typographical errors, incorrect values assigned as a designed parameter rather than as an actual pumped parameter in addition to the possible errors during the process of entering the parameters and assigning a value for the wrong well entry point. However, every effort has been taken to ensure data quality in this work.

In order to minimize any error issues, data elements were subjected to different screening options in the interest of identifying and eliminating outliers and incorrect values. The main technique in this research employed box plot techniques and cross plots to identify outliers or misleading values, and exclude it from the dataset. In the statistical context, an outlier is simply viewed as an unusual extreme value for a variable, which is detected when a value is out of the range of certain statistical frequency. However, an extreme value does not definitely lead to a faulty value, as a statistically rare event could actually happen if it can be justified from an engineering standpoint. Therefore, a statistical approach is required to be coupled with the comparison of known limits and ratios of the parameters, and engineering judgment must be applied in each elimination process.

The objective of this stage is to ensure most of the data are validated and ready for any future analysis. The following sub-sections illustrates the parameters validation carried out during this study.

3.4.3.1 Validation of pumped fluid parameter. Fracturing fluid is a critical component of the hydraulic fracturing treatment. Its main functions are to open the fracture and to transport propping agent along the length of the fracture.

Figure 3.30 demonstrates the use of box plot techniques to statistically identify the practical range of the parameter. In this example, statistics shows that more than half of the wells in the dataset were treated with less than 2,500 M^3 and most of the rest of the wells treated with less than 20,000 M^3 of fracturing fluid. Extreme values of total fluid pumped are detected in a few wells with treatment fluid volume ranging between 20,000 to 50,000 M^3. Applying only a statistical approach results in eliminating 72 wells from the dataset. Figure 3.31 shows the final distribution histogram of total pumped fluid in the Montney wells, if once wells with extreme values were removed.

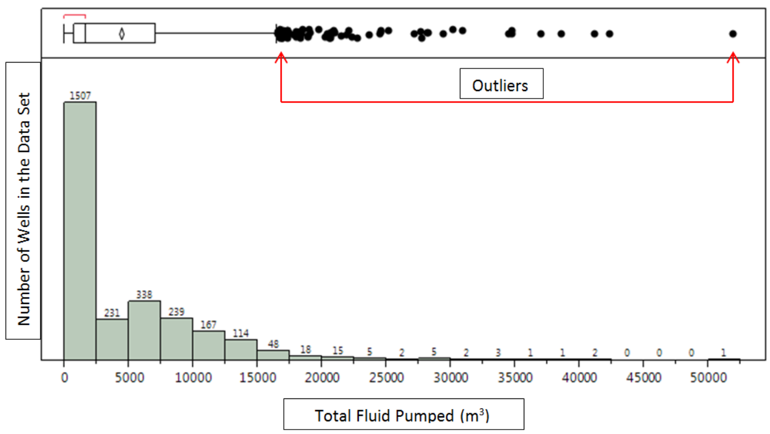

Figure 3.30 Using Box Plot Techniques to Identify Outliers in Total Fluid Pumped

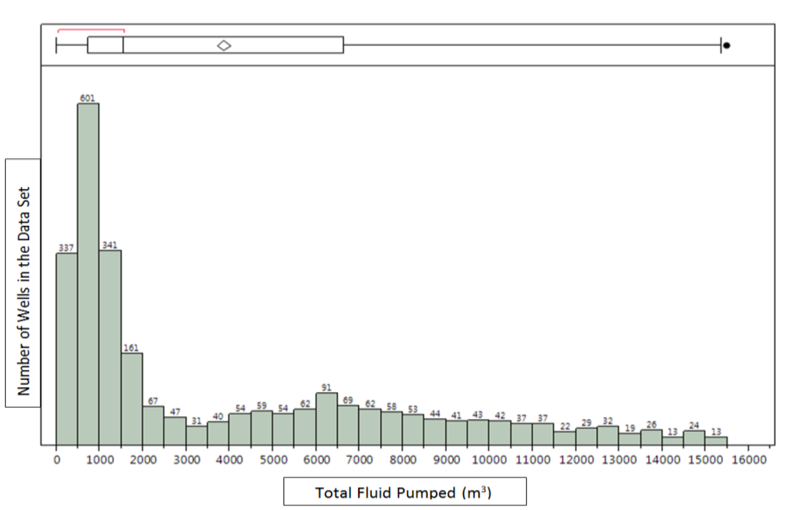

Figure 3.31 Distribution Histogram with Box Plot of Total Pumped Fluids in the Montney Wells after Removing Outliers

Figure 3.32 is a cross plot of load fluid versus total fluid pumped. This cross plot shows that most of the wells were treated with less than 20,000 M^3 of fluid and fewer wells that treated with higher volume of fluids. Most of these wells fall within the linear correlation of the two plotted parameters which means that these points should not be considered as an outliers in the dataset. This is because the high treatment fluid volume has been confirmed by other reported parameters from the dataset, and further grouping based on the well architecture and stimulation design parameters would interpret and justify the high fluid volume pumped in some of the wells.

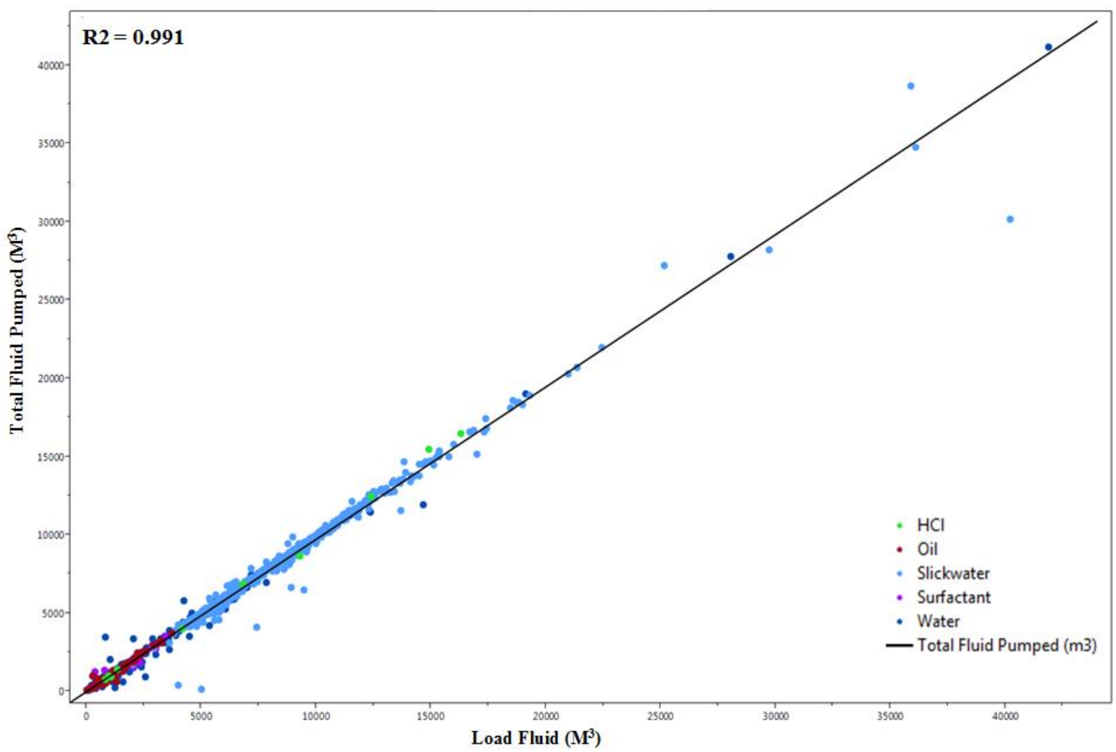

Figure 3.32 Cross Plot of Total Fluid Pumped (M^3) Versus. Load Fluid (M^3) for All Wells in the Montney Grouped in Colors Based on Treatment Fluid Type

Taking the analysis a step further and introducing the completed lateral length to the cross plot in Figure 3.32, demonstrates that the wells treated with very high values of fracturing fluid volume have the highest range of lateral length. This phenomena is also seen in Figure 3.33 which is a cross plot of total fluid pumped on the y-axis and load fluid on the x-axis, grouped horizontally in five ranges of lateral length. Wells are colored based on the fracturing base fluid groups. Figure 3.34 groups the wells in five brackets of calculated average proppant concentration. In this cross plot, it is shown that the wells with higher volume of treated volume fall in the low proppant concentration bracket. The plot also shows that the wells with high fluid volume were mainly treated with water base fracturing fluid type.

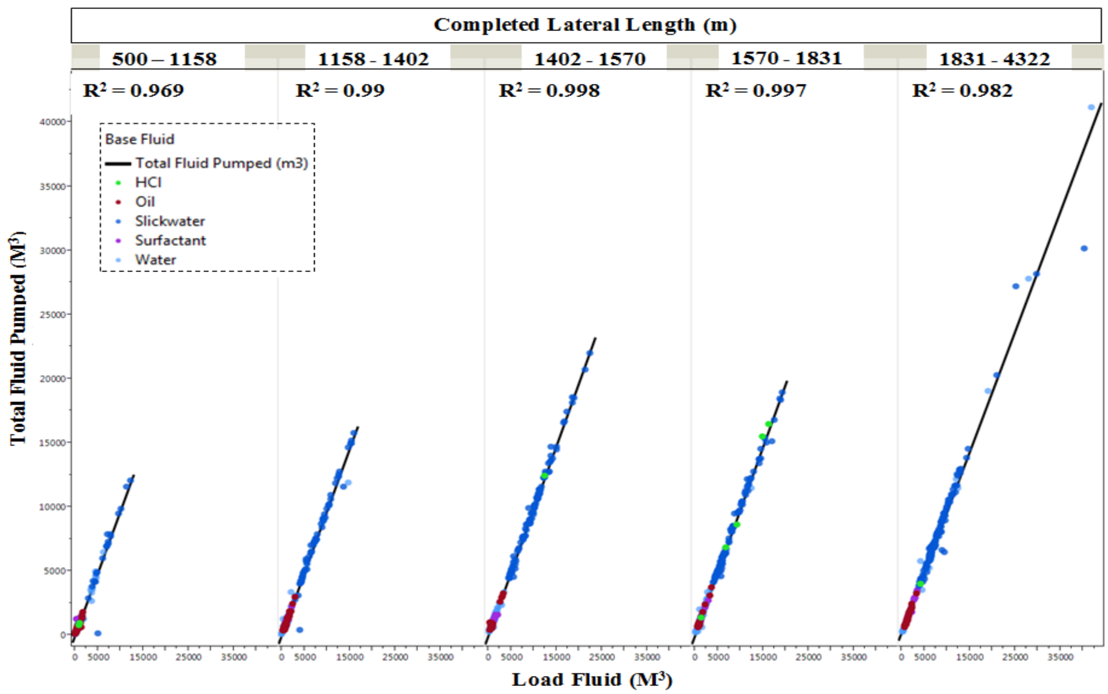

Figure 3.33 Cross Plot of Total Fluid Pumped (M^3) versus. Load Fluid (M^3) Grouped in Five Ranges of Completed Lateral Length (m) and Colored Basesd on Treatment Fluid Type

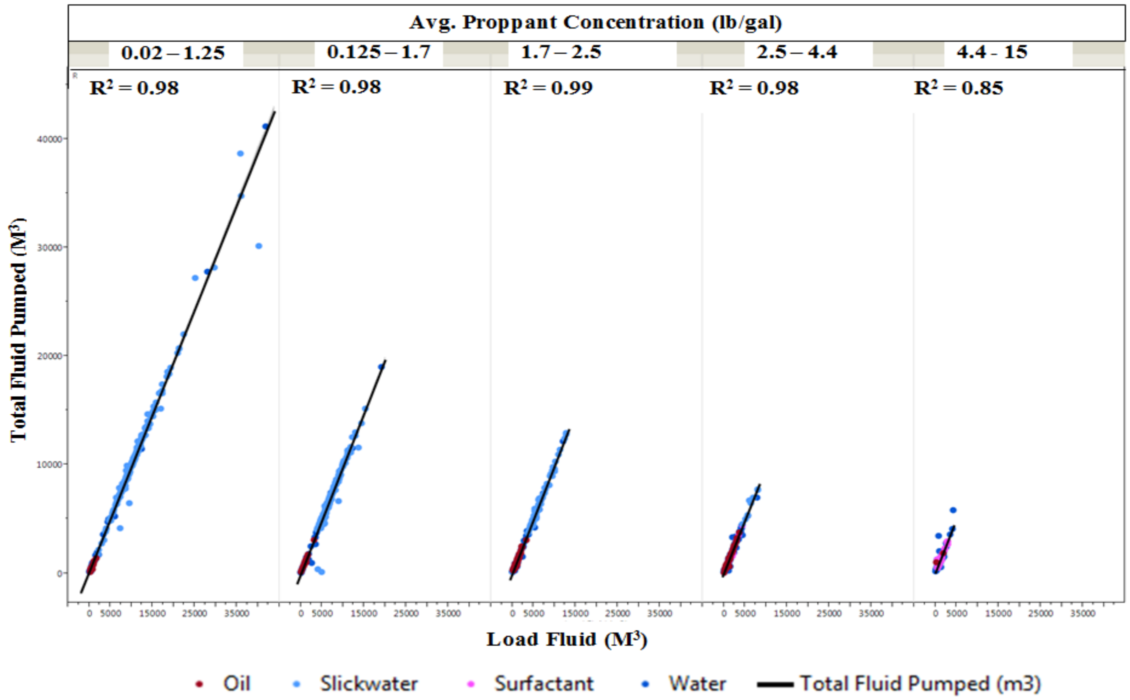

Figure 3.34 Cross Plot of Total Fluid Pumped (M3) versus Load Fluid (M3) for All Wells in the Montney Grouped in Five Ranges of Avg. Proppant Concentration (lbs./gal)

Based on the validation procedure followed above, the fluid parameter samples in the data set are believed to be accurate and reliable and there was no need for any wells to be eliminated as outliers.

3.4.3.2 Validation of completion cost data. Completion cost is reported in the dataset in two variables: Authorized for Expenditure (AFE) Completion Cost which is an estimated cost for project planning and total end of completion cost which is the final total cost of the completion after all operations are finished. It is worth pointing out that there is overlay between the completion and drilling cost as in some companies casing and some other services are charged to the drilling budget while in other companies the casing, perforation and stimulation are considered as part of the completion budget. Figure 3.35 is a cross plot of AFE completion cost versus the final completion cost. Comparing the two variables shows a linear correlation with 0.73 R^2. The wells was colored based on the different operators in the dataset.

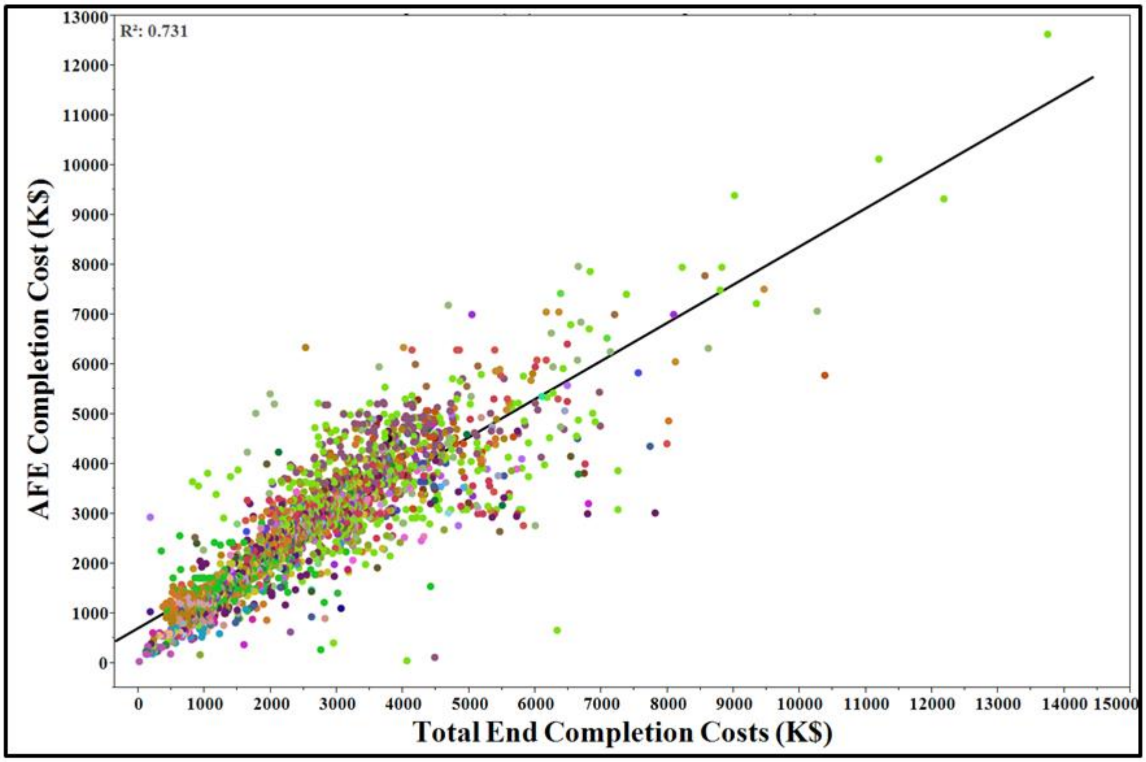

Figure 3.35 Cross Plot of AFE Completion Cost versus. Final Completion Cost with Wells Colored Based on The Operating Company

A few wells where scattered off the linear cost trend, these wells likely had low AFE value but higher end of completion cost, meaning the real final completion costs were higher than expected. Some other wells have high AFE but lower final cost, because no expected problems arose during the completion phase. The plot shows that the typical final completion cost of most of the wells in the Montney falls in the range between 1 to 5 million dollars. Since the database is for the Canadian resource, it is believed that all cost values are reported in Canadian dollars.

There is a strong linear correlation between the planned and actual costs. Figure 3.36 shows the same previous cross plot with further grouping of the wells based on ranges of the lateral length to confirm that there were no suspected outliers.

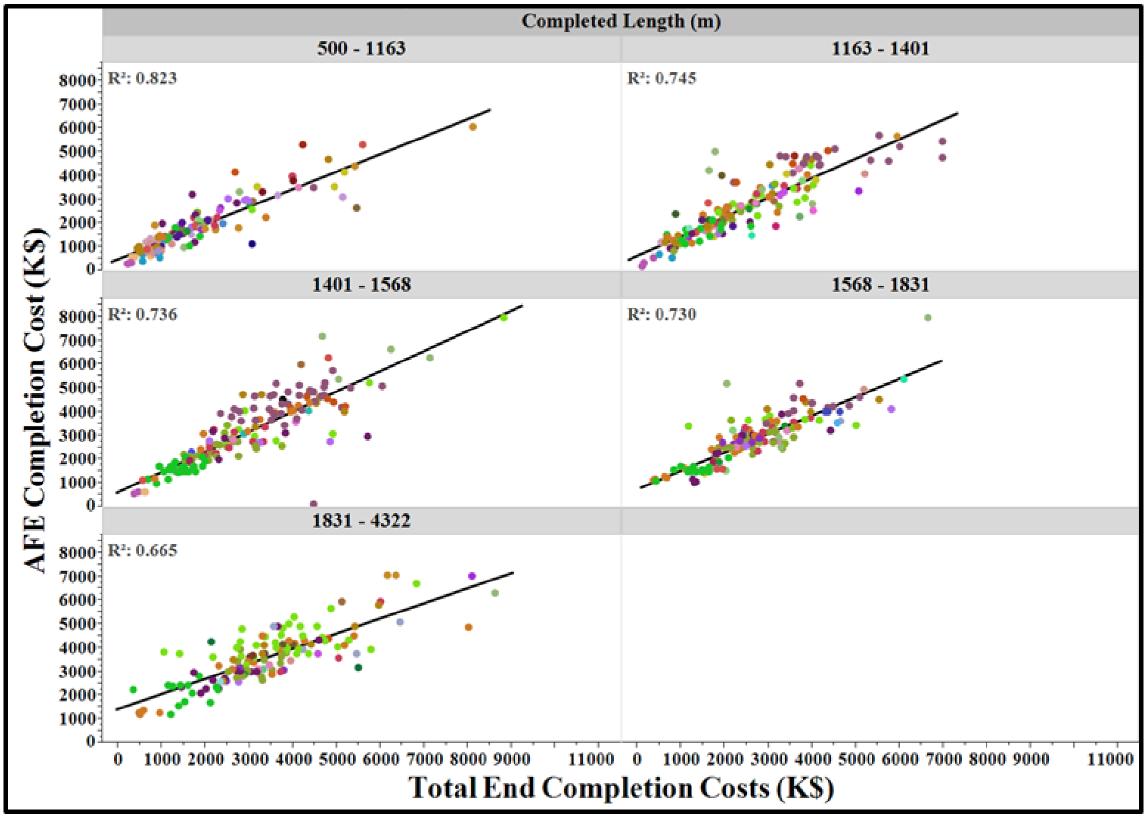

Figure 3.36 Cross Plot AFE completion Cost versus. Total End of Completion Cost Grouped based on Lateral length and Colored based on Operator Company

Figure 3.36 indicates there is little difference between planned and actual costs, which indicates service companies are completing Montney wells with little trouble time. This is particularly true for shorter laterals. Increasing lateral length decreases the R^2 value.

Figure 3.37 shows a similar cost comparison, but this time grouped by completion type. The plot confirms that cased hole completion cost is higher than the open hole completion. It was also evident that the cost of completing a cased hole well is more susceptible to lateral length increment, as it is determined from the changes in R^2 value for both types of completion as the lateral length increase.

Based on these analysis, it was determined that the completion costs in this dataset are clean and can be used for any future analysis.

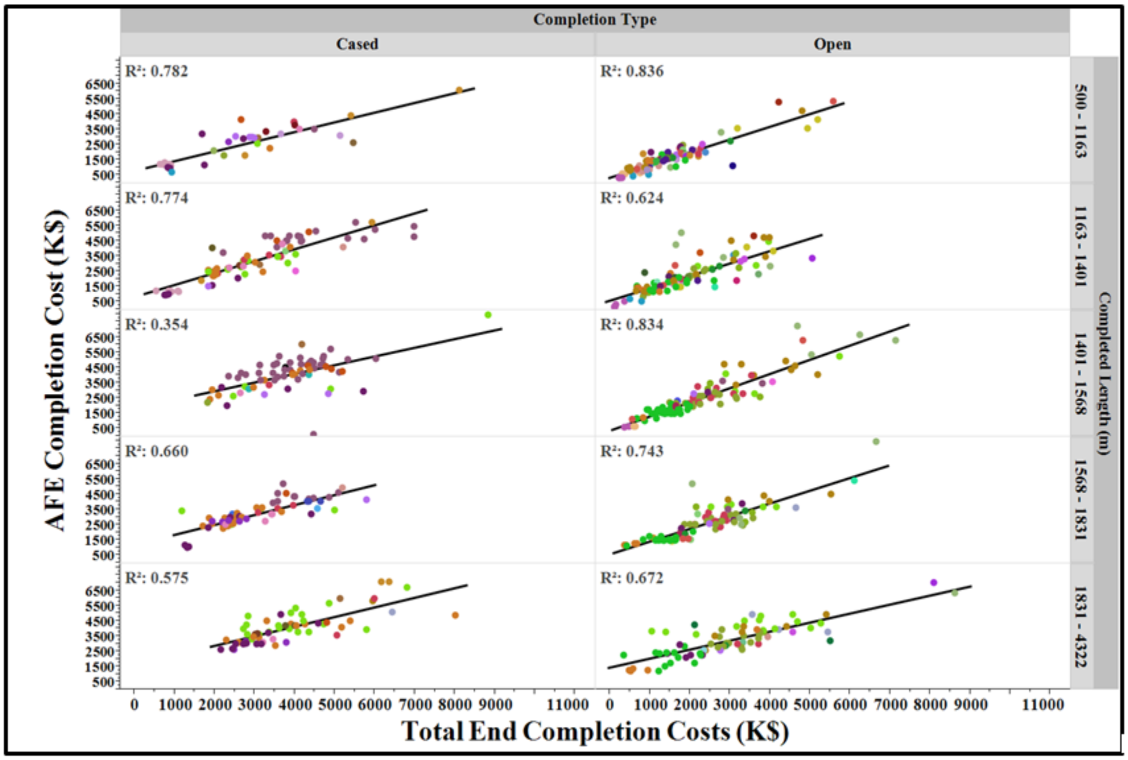

Figure 3.37 Cross Plot of AFE Completion Cost versus Final Completion Cost Grouped on Lateral length and Completion Type and Colored based on the Operating Company

3.4.3.3 Validation of number of stages data. A cross plot was generated to validate the number of stages reported in the dataset. The plot was grouped into two groups based on the completion type, cased and open hole. Each point on the plot represents a well from the dataset colored based on the stimulation company (Figure 3.38). This plot represents about one third of the dataset wells, because in the original dataset only 37% of the wells were incorporating the attempted number of stages.

The attempted number of stages and the actual number of stages showed a good correlation for the cased and open hole completion with R^2 of 0.958 and 0.911 respectively. This correlation implicates a slightly higher success rate in cased hole completions.

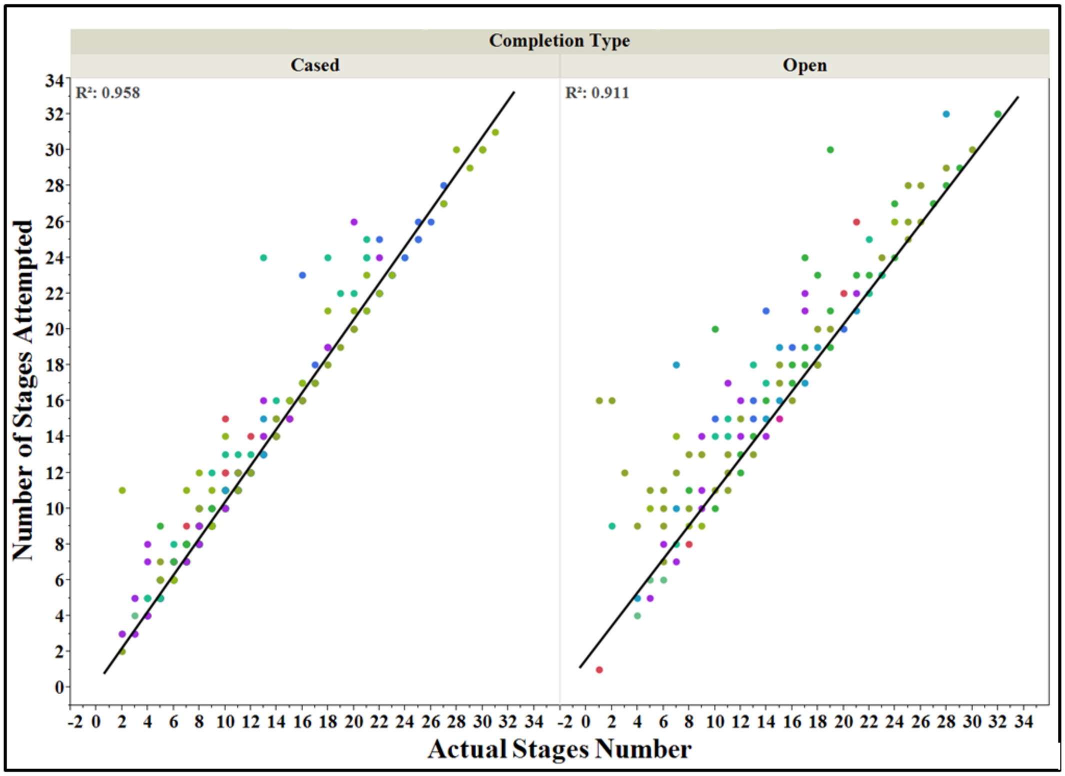

Figure 3.38 Cross Plot Number of Attempted Stages versus Number of Actual Stages Grouped Based on Completion Type and Colored Based on Stimulation Company

The number of stages was associated with the lateral length, as it is logical to expect an increase in the number of stages with increasing length of the horizontal lateral. Therefore, in Figure 3.39, lateral length was introduced to the previous plot to detect the effect of lateral length on the completion success rate of stages. In this plot most of the wells showed a good correlations. As it is shown below, the number of stages increase as the lateral length range increase in the open hole.

The reported average fracture spacing parameter in the dataset was also suspected to have an influence on the number of stages. Therefore; the lateral length in this plot was replaced with average spacing between stages as shown in Figure 3.40.

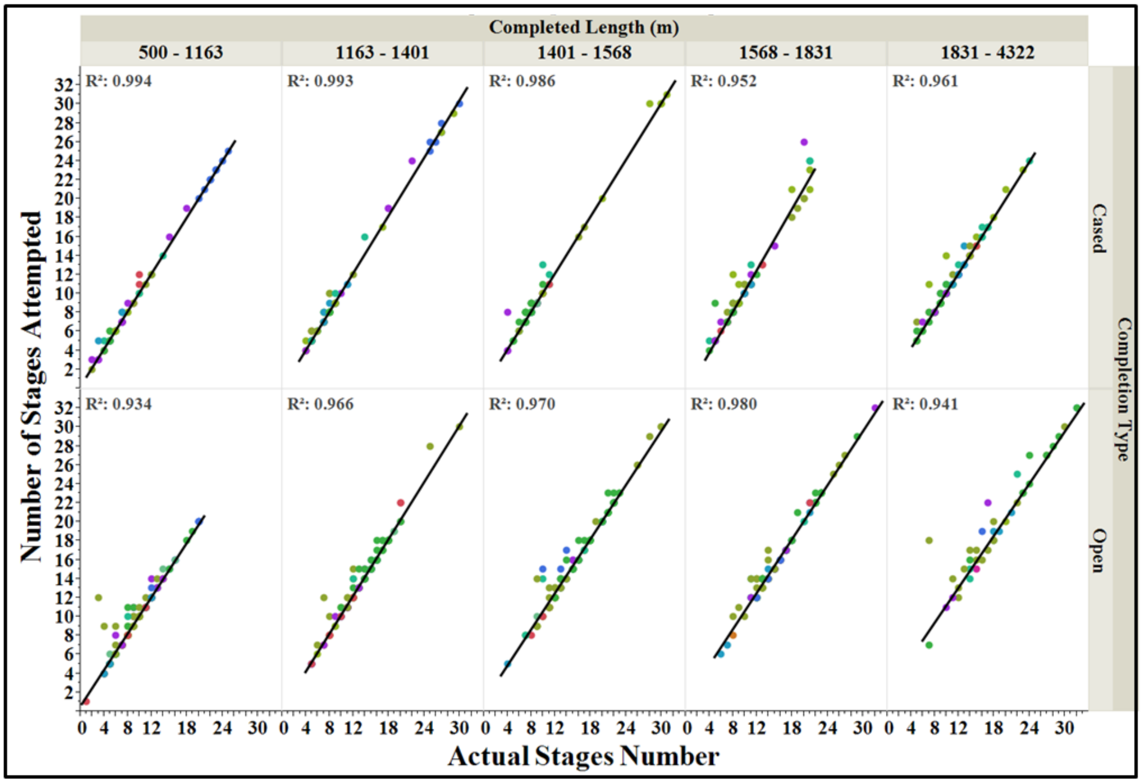

Figure 3.39 Cross Plotting Number of Attempted Stages versus Number of Actual Stages Grouped Based on Completion Type and Lateral Length and Colored Based on Stimulation Company

Including the average fracture spacing in the plot showes that the higher the spacing between stages, the greater the skew between number of attempted and actual stages. This effect is more obvious in the open hole wells, especially at the higher range of average fracture spacing. This support the idea that the very large spacing between stages in the Montney can negatively impact completion success.

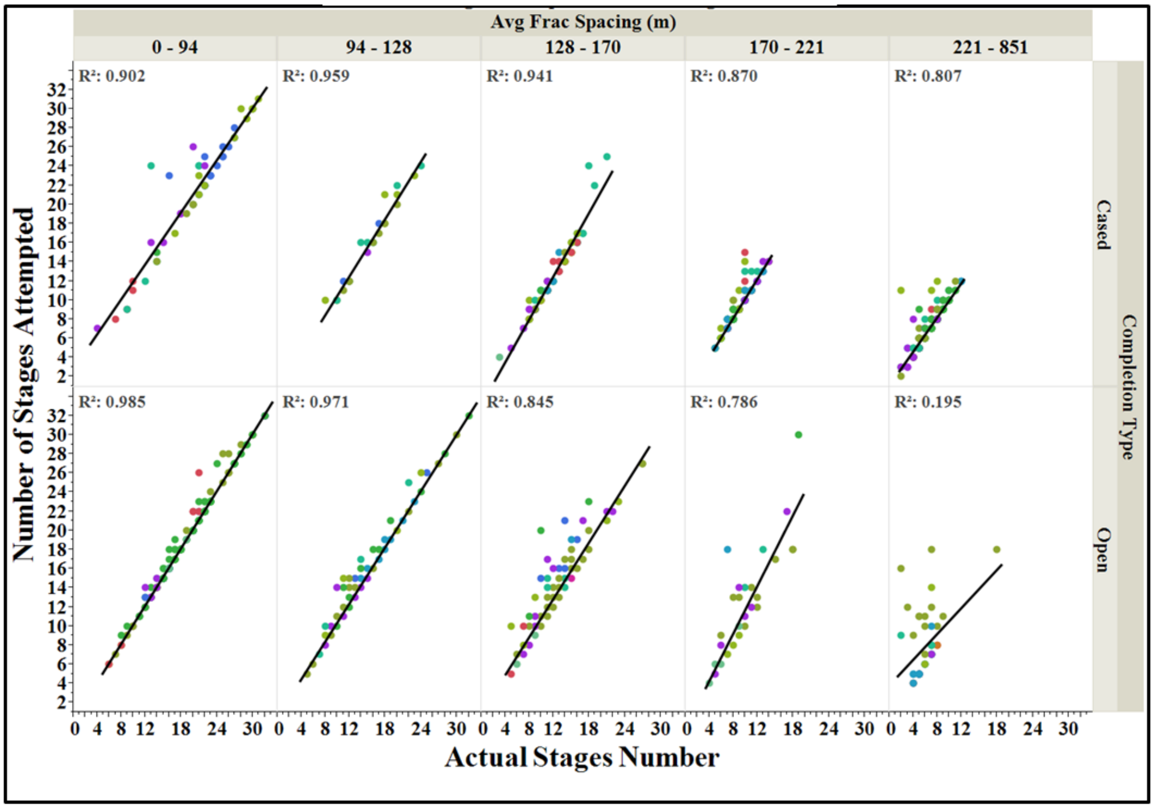

Figure 3.40 Cross Plotting Number of Attempted Stages versus Number of Actual Stages Grouped Based on Completion Type and Avg. Fracture Spacing and Colored Based on Stimulation Company

Based on the above validation cross plots, it is concluded that the number of stages reported in the dataset are accurate and clean from outliers.

3.4.3.4 Validation proppant data. The proppant distribution histogram shown earlier in section 3.4.2.12 indicated a possible outliers in the dataset based on the box plot. The questionable reported proppant mass data points were evaluated by cross plotting the total designed proppant mass versus the total actual pumped proppant to confirm the reliability of the reported values in the dataset and detect any possible off range parameters (Figure 3.41). The plot was grouped based on the type of completion and the points (wells) were colored based on the operating company. The plot reflected an excellent correlation between the two parameters for both completion types,. The plot justified the points with high value of proppant mass by depicting that they were falling exactly on the linear trend line of the correlation. Therefore it is concluded that there was no problem pumping the frac job as planned.

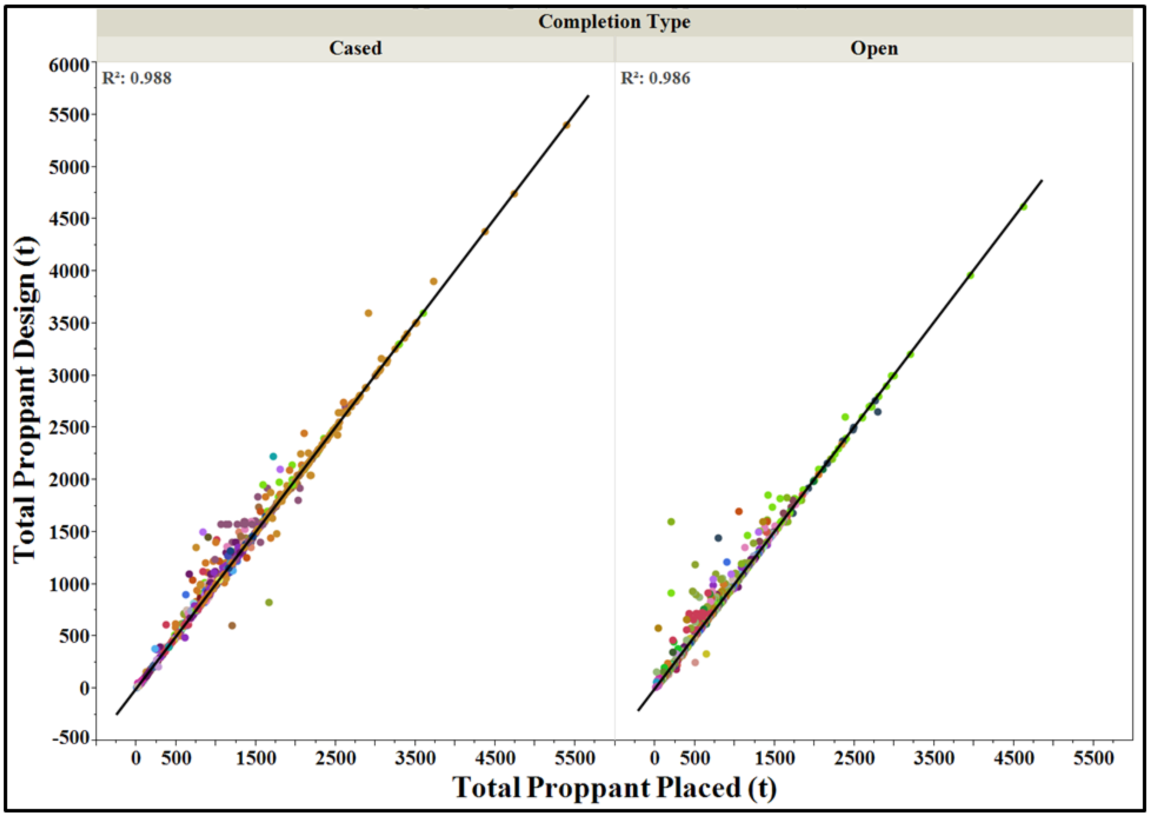

Figure 3.41 Cross Plotting Total Designed Proppant versus Total Actual Pumped Proppant Grouped Based on Completion Type and Colored Based on Operating Company

The lateral length was introduced to the latter cross plot, which is depicted in Figure 3.42. The plot showed that as the well length increase, the proppant mass pumped in the fracturing job also increase. Most of the proppant was pumped as designed regardless of lateral length.

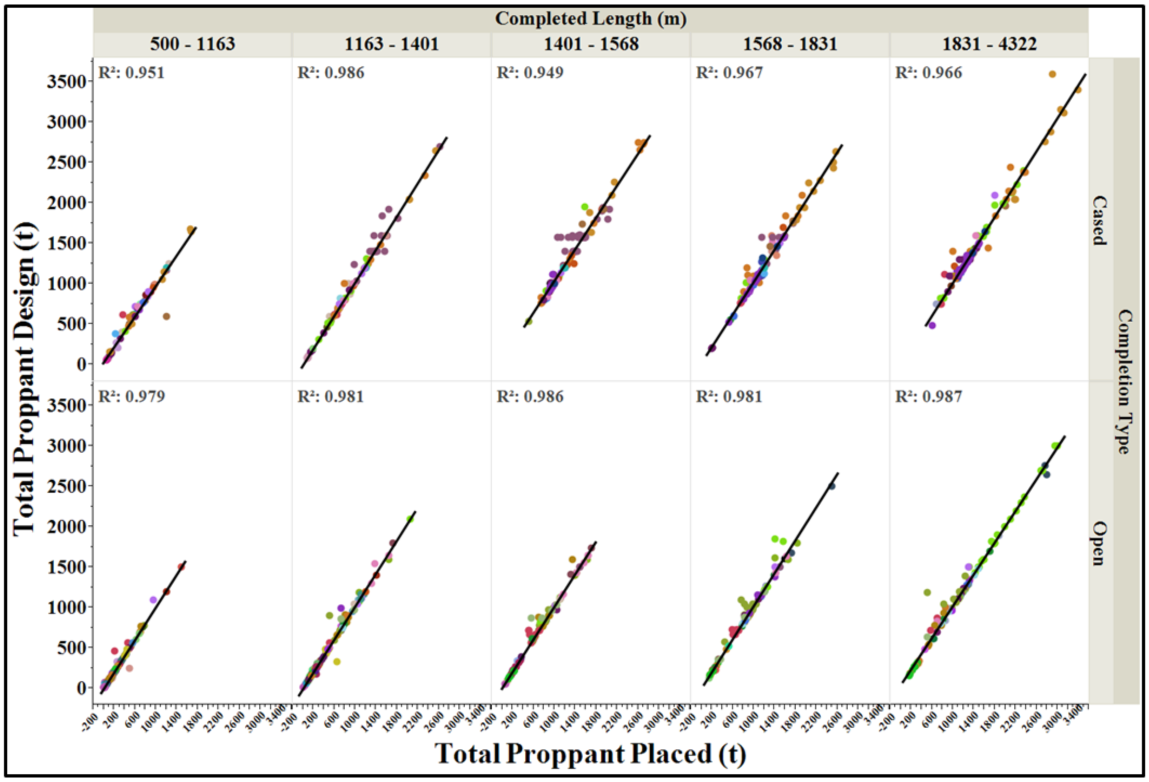

Figure 3.42 Cross Plotting Total Designed Proppant versus Total Actual Pumped Proppant Grouped Based on Lateral Length and Completion Type and Colored Based on Operating Company

Figure 3.43 depicts the fracturing fluid base group affiliation to proppant mass. It is evident that smaller amounts of proppant were used with oil base treatment fluid, and all the large amount of proppant were associated with Water base fluid.

Proppant type and quantities affected by the stress in the region and higher strength proppant usually required with higher anticipated closure pressure. In Figure 3.44 the closure gradient was introduced to the proppant data validation cross plot. It is shown that most of the wells with higher proppant pumped of more than 2000 tonne where located in the high closure stress bracket (0.97-1.28). However, this correlation is not certain because only 279 wells were available for this analysis.

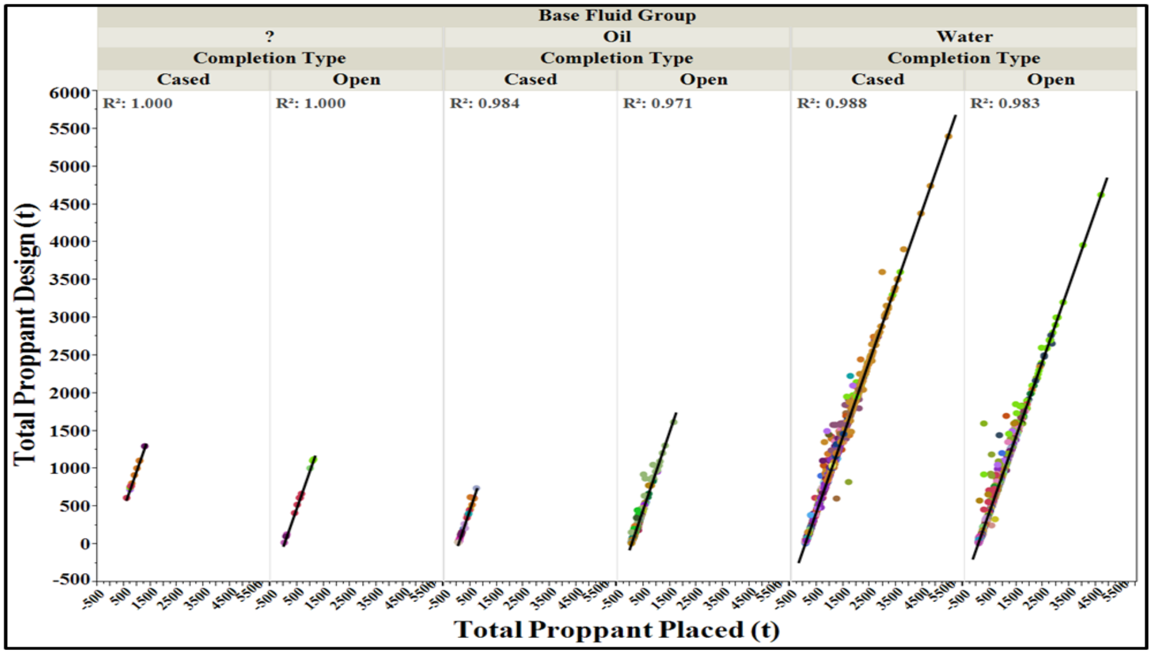

Figure 3.43 Cross Plotting Total Designed Proppant versus Total Actual Pumped Proppant Grouped Based on Treatment Fluid and Completion Type and Colored Based on Operating Company

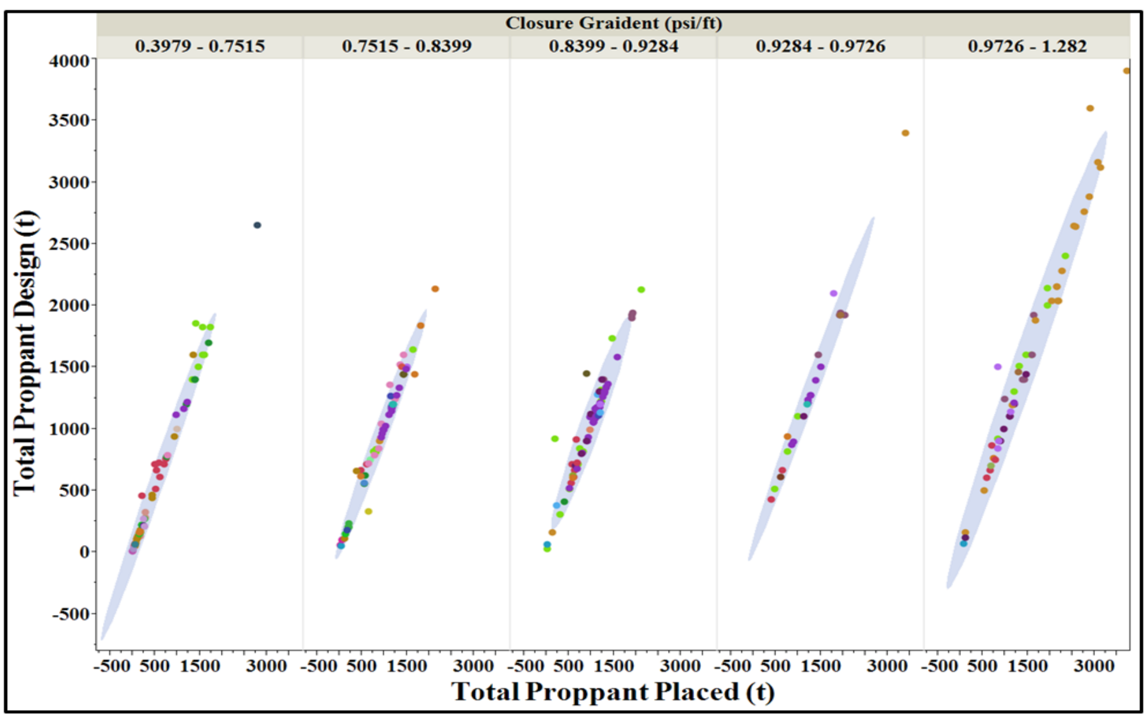

Figure 3.44 Cross Plot Shows the Relation of Closure Gradient and Proppant Mass

3.4.4. Final Parameters Selections and Normalizations. This is the stage that precedes the analysis and modeling. In this phase of the study, the final candidate parameters of the analysis section are selected and some normalizations are applied to the parameters.

The parameters that were applied mostly in the analysis section of this research can be summarized into three groups:

1. Normalized Parameters: these parameters were calculated and introduced to the original dataset parameters to involve the effect of more than one parameter per each analysis and to avoid bias conclusions. Table 3.9 illustrates the main normalized parameters that were calculated and applied to this study.

Table 3.9 List of Normalized Parameters Used in the Analysis

#	Normalized Parameters
1	Total Proppant /Lateral Length (tonne/m)
2	Total Fluid / Lateral Length (m3/m)
3	Total Proppant / Stage (tonne/stage)
4	Total Fluid / Stage (m3/stage)
5	Avg. Proppant Concentration (Total Proppant / Total Fluid) (lbs./gal)
6	Completion Cost /Lateral Length ($/m)
7	Recovery Fluid Pecentage (Recoverd Fluid / Total Fluid)
8	6 Months Cumulative Gas Production / Stage (mmcf/stage)
9	12 Months Cumulative Gas Production / Stage (mmcf/stage)
10	18 Months Cumulative Gas Production / Stage (mmcf/stage)
11	6 Months Cumulative Gas Production / Lateral Length (mmcf/m)
12	12 Months Cumulative Gas Production / Lateral Length (mmcf/m)
13	18 Months Cumulative Gas Production / Lateral Length (mmcf/m)
14	6 Months Cumulative Gas Production / Fracture Spacing (mmcf/m)
15	12 Months Cumulative Gas Production / Fracture Spacing (mmcf/m)
16	18 Months Cumulative Gas Production / Fracture Spacing (mmcf/m)
17	18 Months Cumulative Gas Production / Total Pumped Fluid (mmcf/m3)
18	18 Months Cumulative Gas Production / Total Pumped Proppant (mmcf/t)

2. Design Parameters: these parameters can be controlled by the operators and need to be optimized to achieve the best performance. Table 3.10 illustrates the main design parameters used in this study.

Table 3.10 List of Design Parameters Used in the Analysis

#	**Design Parameters**
1	Lateral Length (m)
2	Actual Number of Stages
3	Total Proppant Placed (tonne)
4	Total Fluid Pumped (m3)
5	Treatment Fluid Types

3. Uncontrolled Parameters: these parameters are either naturally exist and the operators cannot change them, such as principal stresses, or it is a response to a combination of variables, e.g. hydrocarbon production. Table 3.11 shown below illustrates the main uncontrolled parameters used in the analysis section of this study.

Table 3.11 List of Uncontrolled Parameters Used in the Analysis

#	**Uncontrolled Parameters**
1	Final Completion Cost ($)
2	Production of Gas (mmcf)
3	Production of oil (m bbl)
4	Production of Water (m bbl)
5	Recovery of Fluid (m3)
6	Avg. Closure Gradient (psi/ft)
7	IP of Gas (mcf/D)
8	IP of Oil (B/D)
9	IP of Water (B/D)

The modeling, evaluation and deployment phase of the data mining will be presented in the next chapters of this thesis.

4. ANALYSIS AND DISCUSSIONS

4.1. DEVELOPING GENERAL ANALYSIS TECHNIQUES

In this section the effect of well and hydraulic fracturing design parameters on the production performance of Montney well completions was investigated to understand the big picture of the different applications and techniques. As stated earlier, in section 3.4.2.2, there are 113 operators in the dataset, and each one of these companies operates with a different design and budget to meet their objectives. Trying to understand the effects of their different practices in an attempt to maximize initial production (IP) and ultimate cumulative production will help in the development of future wells.

After cleaning and evaluating the reliability of the parameters in the dataset, the analysis phase was started to identify trends and relationships between the parameters.

A statistical approach was employed to quantify the differences between cased and open hole completions by calculating the mean and median for the different cost, design and production parameters.

Another technique used extensively was to crossplot the average values of the variables and to group them with respect to other design parameters. The purpose of including more than one variable in each plot was to account for the interrelationships between the variables and as a comparison parameter to recognize the differences between their applications in the field. In many of the analyses conducted in this study, the completion type was set as a comparison parameter to give an idea about the differences between the open and cased hole wells.

As it was shown earlier in chapter three, cased hole and open hole wells were treated with different ranges in most of the parameters. The plots were grouped by the completion type to show the performance of each completion type.

The concept of the heat maps' technique was also used in this study to show the effect of a color coded single studied parameters on a pre-established trend to check the effect of increasing or decreasing the studied parameter on the performance trend by noticing the color changes.

Stimulated well performance in the unconventional resources depends on several parameters which could be controlled in the design phase, including completion lateral length, number of stages, total fluid volume pumped, total placed proppant mass, type of fluid and sand. Other parameters are uncontrolled and may change the well's performance for the same design parameters. Examples of these parameters include geology, porosity, permeability, and the principal stresses.

Simply cross plotting a parameter for all wells can only show a statistical correlation across the entire play, which may fail to identify that the parameters are profoundly related to each other. Hence, it is better that the parameters are grouped to represent cases that are more specific where other independent parameters are unified or their effects are normalized.

It must be stated that the trends presented in some of the plots are not necessarily intended to suggest a linear relationship with the studied variables, but rather clearly communicate whether a trend between the data is following an upward or downward direction.

4.1.1. Statistical Techniques. In this method, the wells in the dataset were split (based on the completion type) into two groups: open and cased hole wells. For each group a statistical value of the mean and median were calculated for several parameters of production, design and cost. These values were tabulated and each parameter was graphically represented by four bars that represent the mean and median for both cased and open hole wells in the dataset.

The parameters' data distribution histograms shown in chapter three indicated that not all of the parameters were normally distributed across the Montney play and depending only on the mean value in the comparison might include some bias in the decision. Therefore, the median was also calculated to represent the middle value for each parameter.

Table 4.1 and Table 4.2 list the mean and the median values respectively. In each table the completion type of 22 variables was compared based on the mean and median values. The number of samples (N) for each variable were also included in the tables to give an idea about the number of wells that the mean and median values were calculated from. The cased hole/open hole values of the mean were calculated and introduced to each table (yellow column on the right) to represent the percentage that the cased hole differs from the open hole. For example, if the cased/open was shown to be 1.5, this would imply that cased hole wells were 50% better than open hole wells.

Table 4.1 Parameters Mean Values in Cased and Open Hole Wells

	Parameter	N		Mean		Mean
		Cased Hole	Open Hole	Cased Hole	Open Hole	Cased / Open
Cost	Total End Completion Costs (K$)	1139	1361	3220	2152	1.50
	Total End Drilling Costs (K$)	325	960	2071	2653	0.78
Design	Proppant/Completed Length (t/m)	473	602	0.7454	0.3871	1.93
	Fluid/Completed Length (M3/m)	476	607	3.9921	1.5352	2.60
IP	IP Water (bwpd)	947	1081	72	63	1.14
	IP Oil (bopd)	126	455	120	185	0.65
	IP Gas (mcf/d)	1130	1189	3169	2522	1.26
Water	6 Mo Cum Prod Water (mbw)	984	1123	8	8	1.01
	12 Mo Cum Prod Water (mbw)	904	987	12	14	0.88
	18 Mo Cum Prod Water (mbw)	799	826	14	18	0.79
Oil	6 Mo Cum Prod Oil (mbo)	114	473	16	20	0.79
	12 Mo Cum Prod Oil (mbo)	73	382	27	30	0.91
	18 Mo Cum Prod Oil (mbo)	45	290	30	33	0.93
Gas	6 Mo Cum Prod Gas (mmcf)	1158	1240	398	303	1.31
	12 Mo Cum Prod Gas (mmcf)	1012	1042	760	534	1.42
	18 Mo Cum Prod Gas (mmcf)	883	859	1057	733	1.44
Gas / Length	6 Mo Cum Prod Gas / Completion Length (mmscf/m)	376	471	0.2798	0.2173	1.29
	12 Mo Cum Prod Gas / Completion Length (mmscf/m)	342	397	0.5277	0.3754	1.41
	18 Mo Cum Prod Gas / Completion Length (mmscf/m)	319	324	0.7320	0.5128	1.43
Gas / Stage	6 Mo Cum Prod Gas / Actual Stages Number (mmscf/stage)	1149	1231	50.09	26.57	1.89
	12 Mo Cum Prod Gas / Actual Stages Number (mmscf/stage)	1006	1036	95.78	48.73	1.97
	18 Mo Cum Prod Gas / Actual Stages Number (mmscf/stage)	879	854	135.11	69.36	1.95

Table 4.2 Parameters Median Values in Cased and Open Hole Wells

	Parameter	N		Median		Median
		Cased Hole	Open Hole	Cased Hole	Open Hole	Cased / Open
Cost	Total End Completion Costs (K$)	1139	1361	3064	1890	1.62
	Total End Drilling Costs (K$)	325	960	1921	2427	0.79
Design	Proppant/Completed Length (t/m)	473	602	0.6543	0.3056	2.14
	Fluid/Completed Length (M3/m)	476	607	3.3172	0.7118	4.66
IP	IP Water (bwpd)	947	1081	37	24	1.54
	IP Oil (bopd)	126	455	106	90	1.17
	IP Gas (mcf/d)	1130	1189	3125	2172	1.44
Water	6 Mo Cum Prod Water (mbw)	984	1123	4	3	1.33
	12 Mo Cum Prod Water (mbw)	904	987	7	5	1.40
	18 Mo Cum Prod Water (mbw)	799	826	8	6	1.33
Oil	6 Mo Cum Prod Oil (mbo)	114	473	14	11	1.27
	12 Mo Cum Prod Oil (mbo)	73	382	22	17	1.29
	18 Mo Cum Prod Oil (mbo)	45	290	18	21	0.86
Gas	6 Mo Cum Prod Gas (mmcf)	1158	1240	381	243	1.57
	12 Mo Cum Prod Gas (mmcf)	1012	1042	720	445	1.62
	18 Mo Cum Prod Gas (mmcf)	883	859	990	631	1.57
Gas / Length	6 Mo Cum Prod Gas / Completion Length (mmscf/m)	376	471	0.2559	0.1623	1.58
	12 Mo Cum Prod Gas / Completion Length (mmscf/m)	342	397	0.4852	0.3023	1.60
	18 Mo Cum Prod Gas / Completion Length (mmscf/m)	319	324	0.6476	0.4216	1.54
Gas / Stage	6 Mo Cum Prod Gas / Actual Stages Number (mmscf/stage)	1149	1231	43.70	20.86	2.10
	12 Mo Cum Prod Gas / Actual Stages Number (mmscf/stage)	1006	1036	83.90	40.25	2.08
	18 Mo Cum Prod Gas / Actual Stages Number (mmscf/stage)	879	854	117.57	58.67	2.00

The outcomes of the tables shown above were graphically represented in Figure 4.1 through Figure 4.22 to show the differences between open and cased hole wells based on the mean and median of the entire Montney play.

The main comparing parameters were the following:

- Initial Production: The IP of gas and water in cased hole wells showed higher production than the open hole wells, while the IP of oil was better in open hole than in cased hole wells. The median of open hole was lower than the cased hole, which means there are more wells in the open hole completion produced with less than the average.
- Cumulative Gas Production: The 6, 12 and 18 months' cumulative gas production in the cased hole was higher than the open hole. The median value of the open hole wells was less than the median of the cased hole wells.
- Cumulative Oil: The cumulative production of oil over 6, 12 and 18 months in the open hole completion is higher than the cased hole wells, after 18 months more than half of the wells completed with the open hole produced cumulative oil with a higher than average production of Montney open hole wells.
- Cumulative Water: Open hole completion produced more water than the cased hole completion.
- Cumulative Gas / Lateral Length: Normalized gas cumulative production per the completed length of the wells, showed that the cased hole wells performed better than the open hole wells in terms of the amount of gas produced for each completed meter.
- Cumulative Gas / Stage: Cased hole wells performed much better than open hole wells in terms of production per stimulated stages, which explains the higher use of proppant and fracturing fluid in cased hole wells. Although a lower number of stages were performed in the cased hole wells, the big fracturing treatment job and higher number of clusters per each stage doubled the effective production per stage in the cased hole wells, compared to the open hole wells.

- Completion Cost: The average completion cost of the cased hole wells showed a 50% increase, which is more than the open hole completion. This higher completion cost should be justified by the overall increase in production before making the decision on which completion methodology needs to be applied.
- Drilling Cost: The average and median drilling costs of the open hole completion showed an increase of 25-30% from the drilling cost of the cased hole wells. This may be associated with the problem that occurred during the drilling phase of the open hole section, e.g. the directional drill string was mechanically stuck down in the hole because of the instability or collapse of the well and was unable to be retrieved to the surface. Such problems might significantly increase the drilling cost of the well, because of the additional cost of the tools and sidetracking the lateral.
- Proppant / Lateral Length: More average proppant per unit length was placed in cased hole wells, compared to the open hole wells. Adding more proppant will increase the conductivity of the stimulated reservoir section and increase the production.
- Fluid / Lateral Length: Treatment fluid per unit length in cased hole wells is more than double the fluid volume per unit length in the open hole wells. Based on this normalized parameter, this parameter needs to be associated with the number of clusters for each stage. Pumping high volumes of fluid will generate longer fractures for the same number of fractures, or it will generate more fractures for a higher number of clusters.

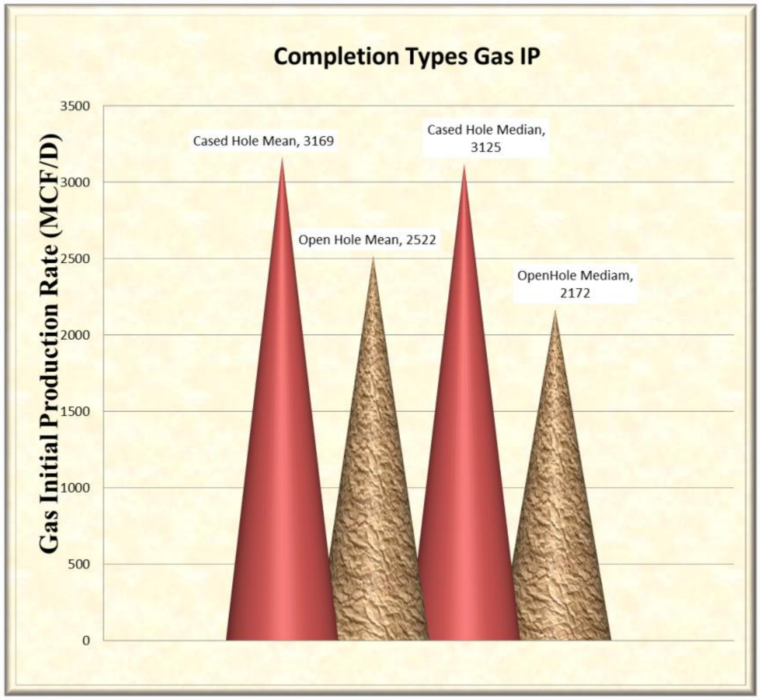

Figure 4.1 IP Gas of Completion Types

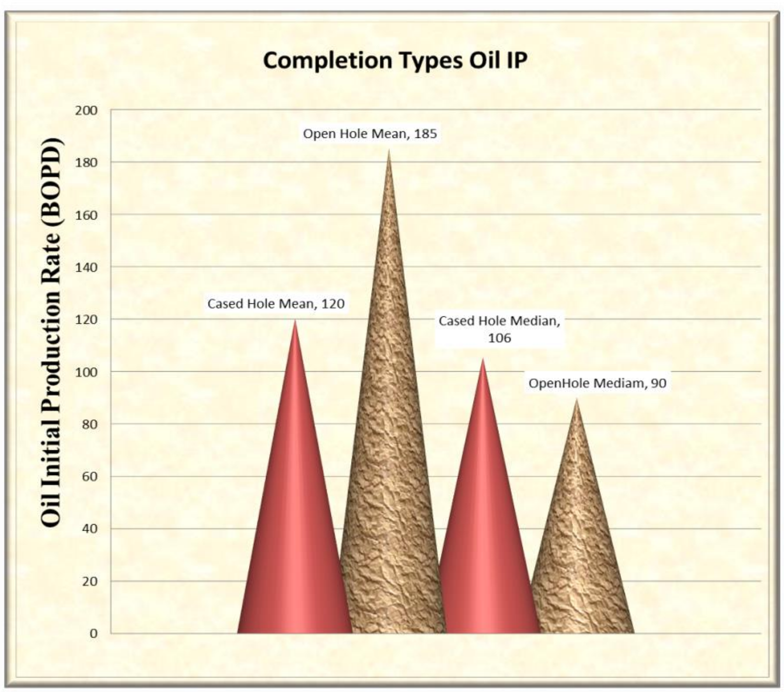

Figure 4.2 Completion Types Oil IP

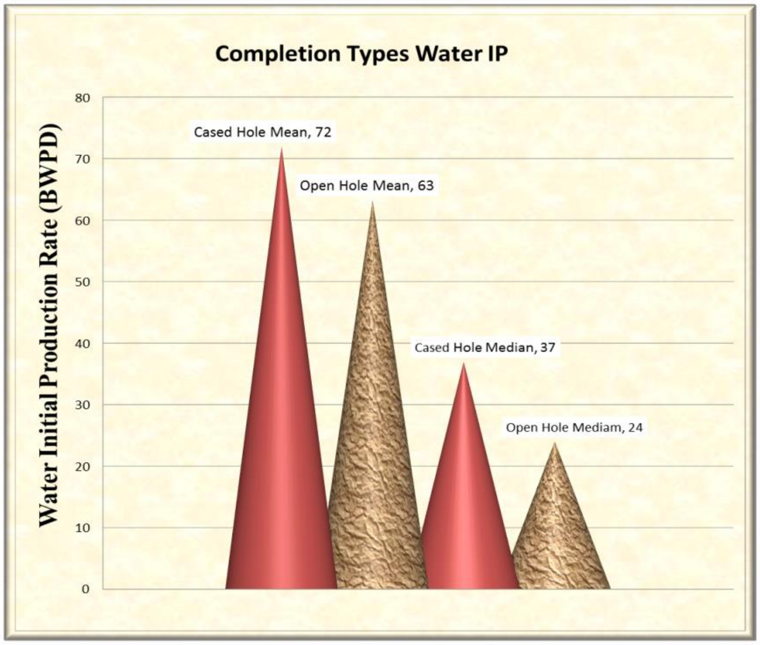

Figure 4.3 Completion Types Water IP

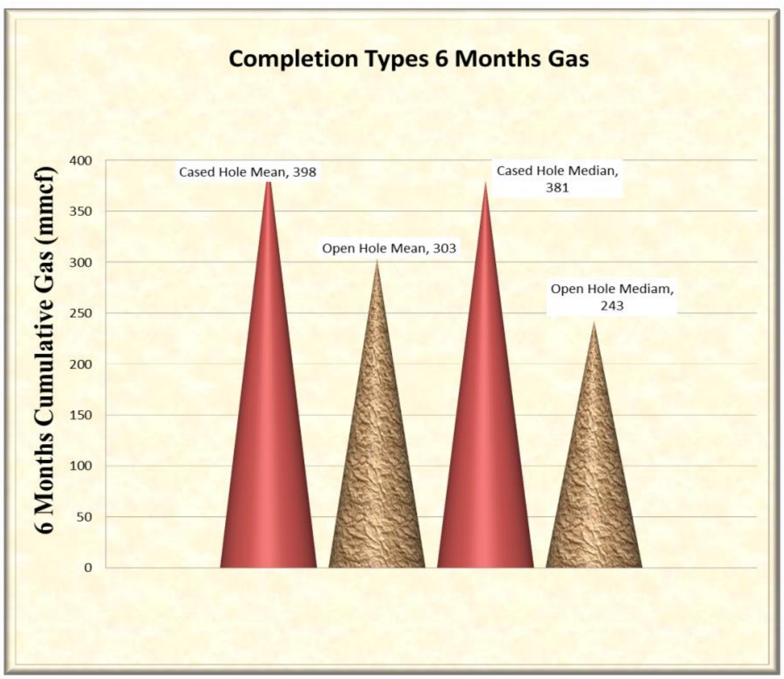

Figure 4.4 Completion Types 6 Months Gas Cumulative

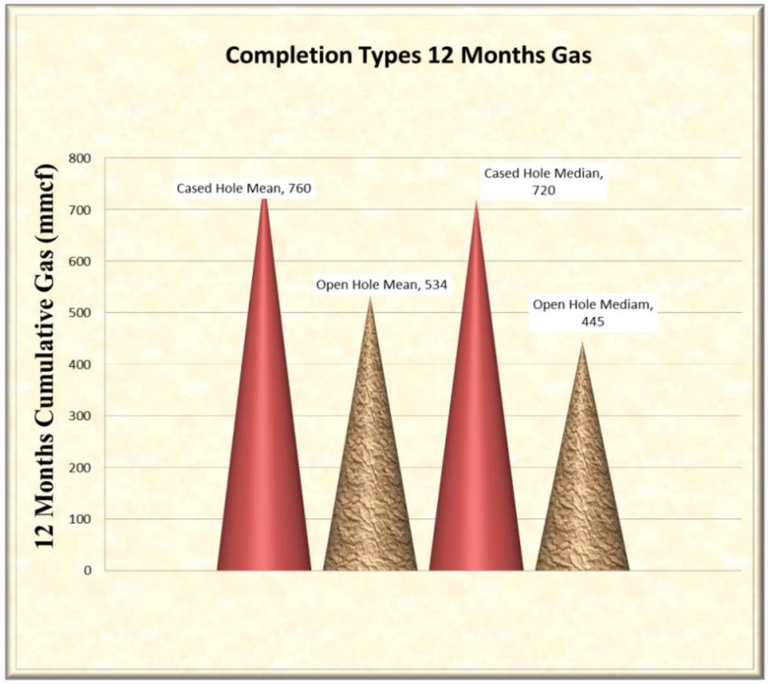

Figure 4.5 Completion Types 12 Months Gas Cumulative

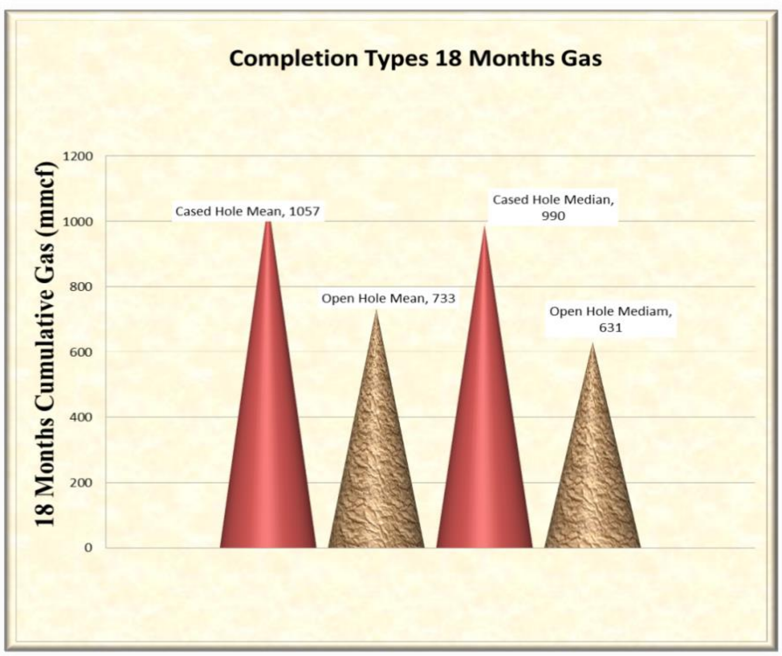

Figure 4.6 Completion Types 18 Months Gas Cumulative

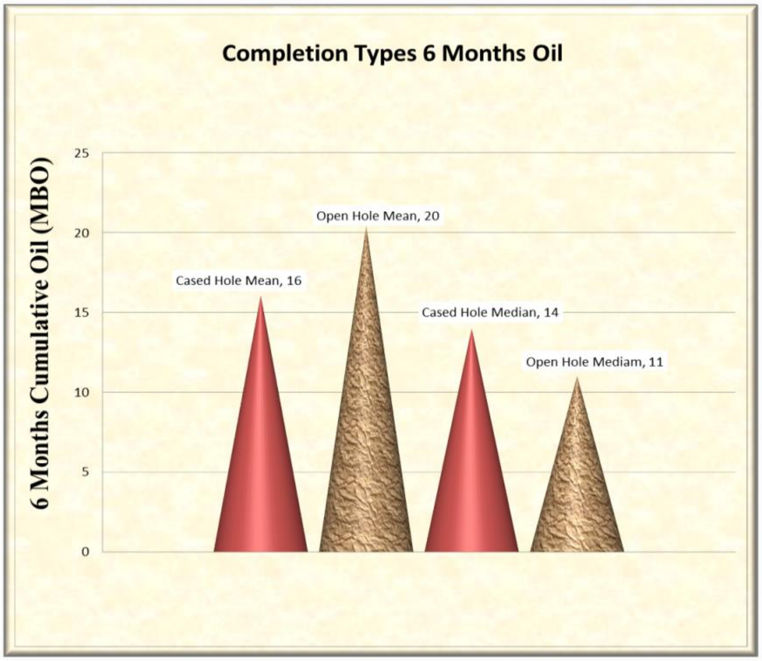

Figure 4.7 Completion Types 6 Months Oil Cumulative

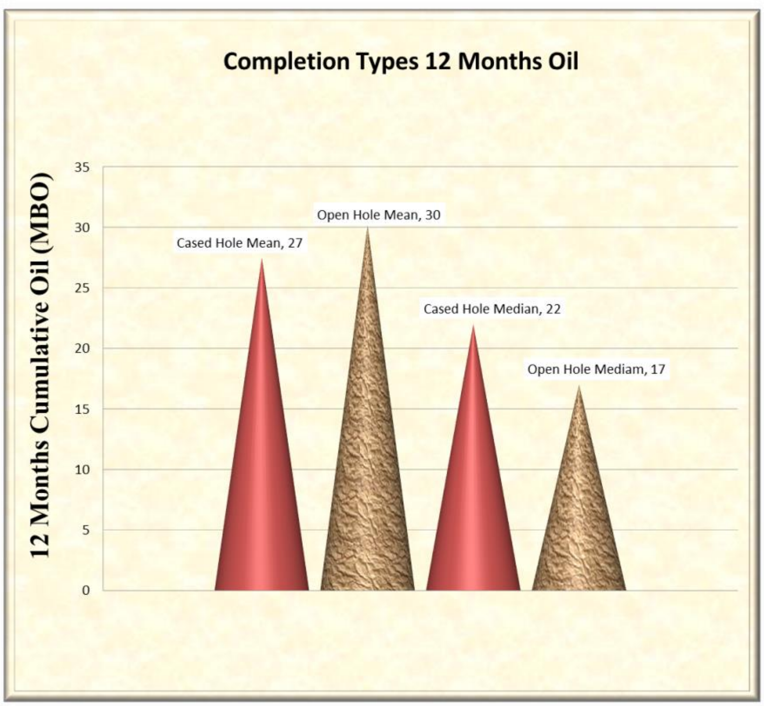

Figure 4.8 Completion Types 12 Months Oil Cumulative

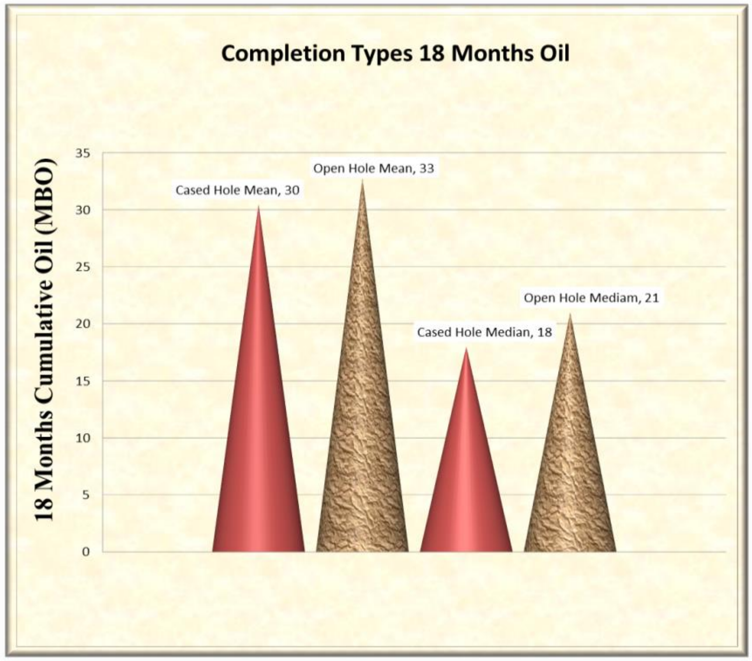

Figure 4.9 Completion Types 18 Months Oil Cumulative

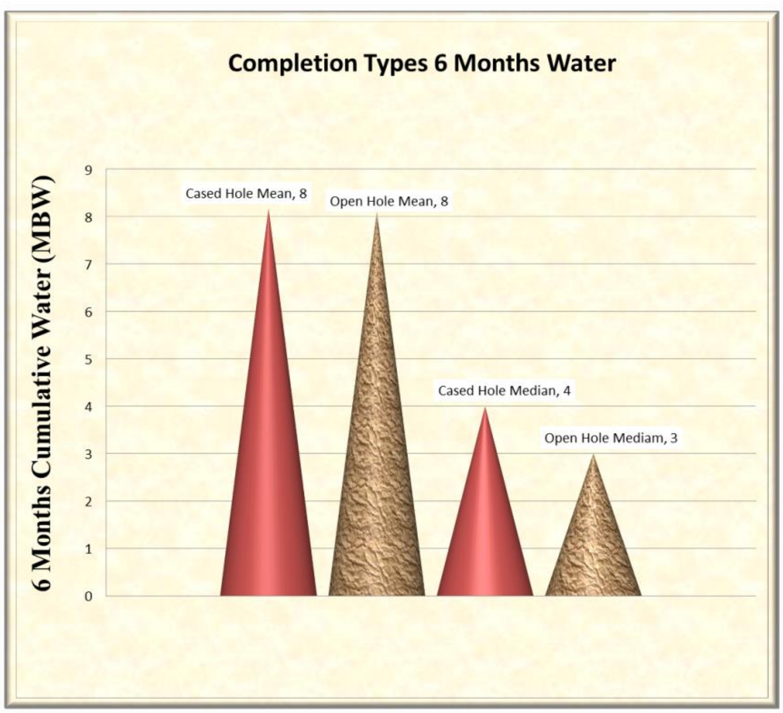

Figure 4.10 Completion Types 6 Months Water Cumulative

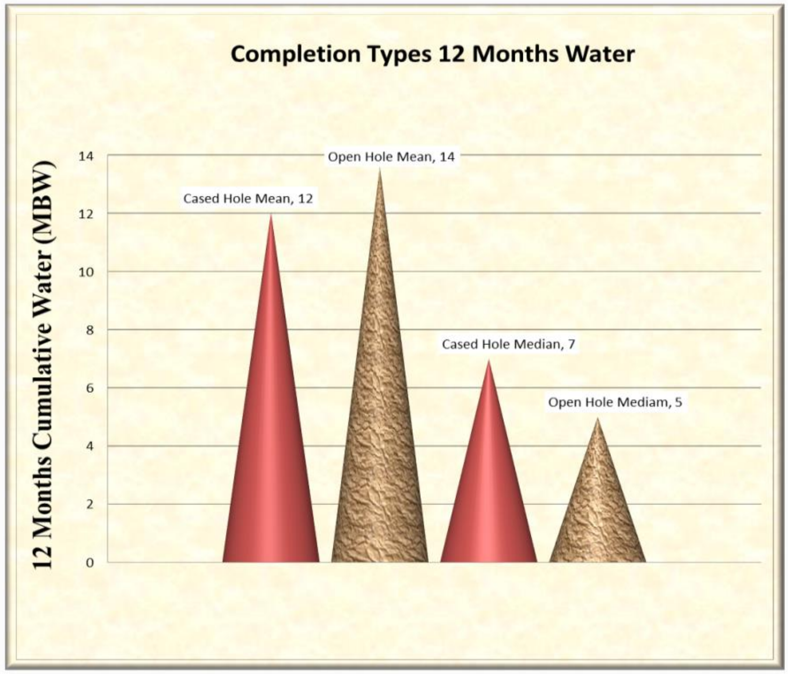

Figure 4.11 Completion Types 12 Months Water Cumulative

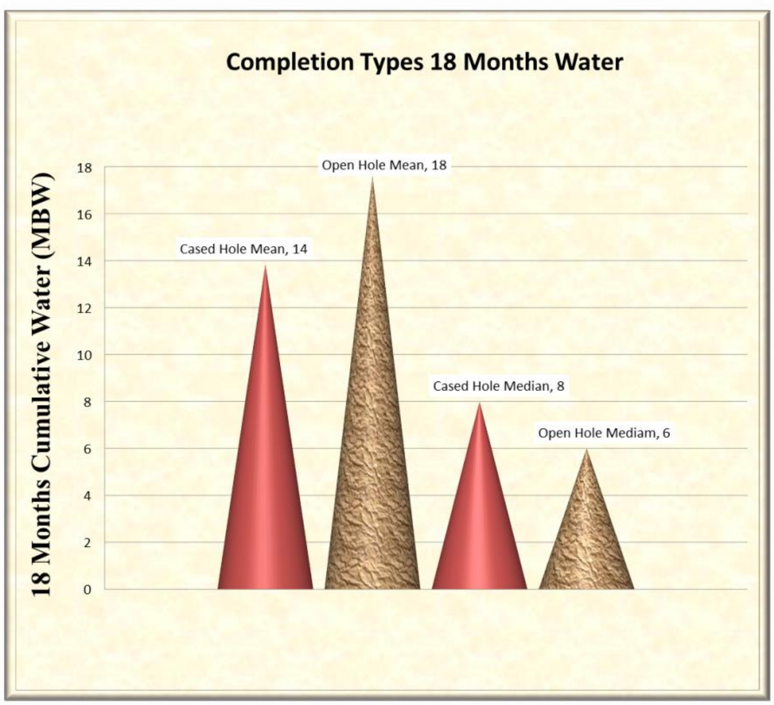

Figure 4.12 Completion Types 18 Months Water Cumulative

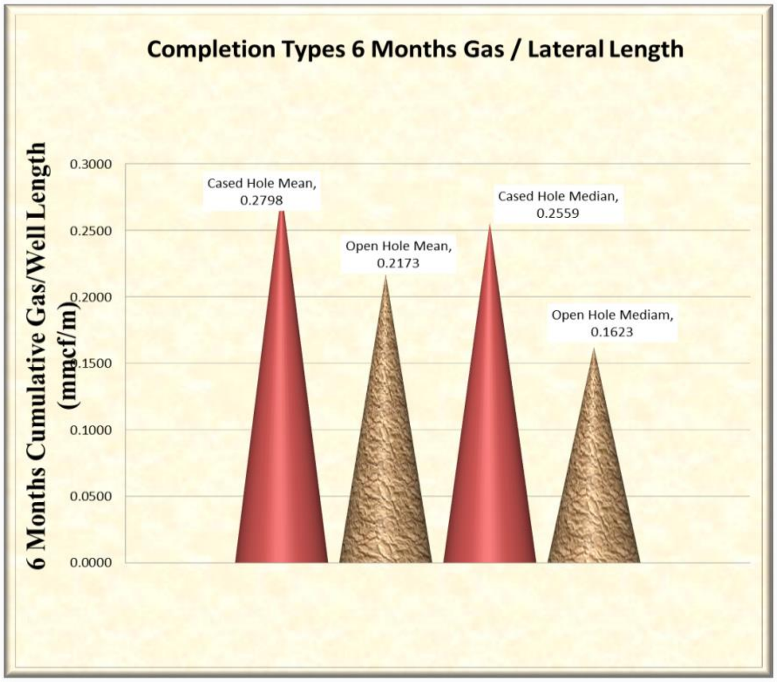

Figure 4.13 Completion Types 6 Months Gas Cumulative / Lateral Length

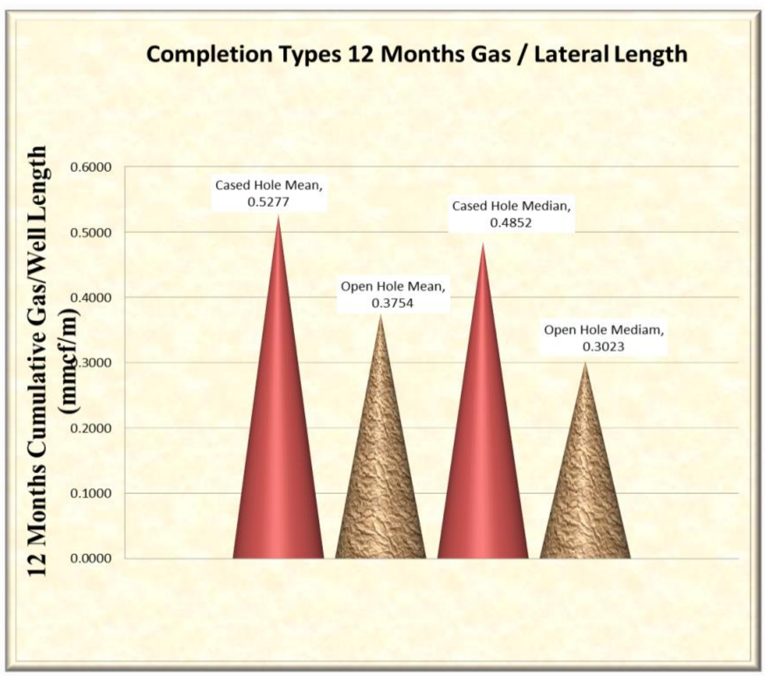

Figure 4.14 Completion Types 12 Months Gas Cumulative / Lateral Length

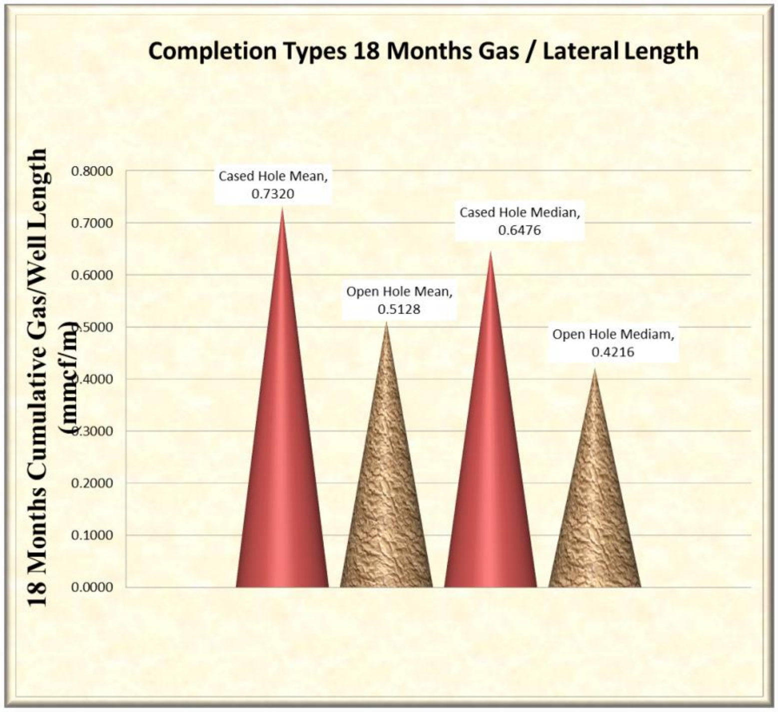

Figure 4.15 Completion Types 18 Months Gas Cumulative / Lateral Length

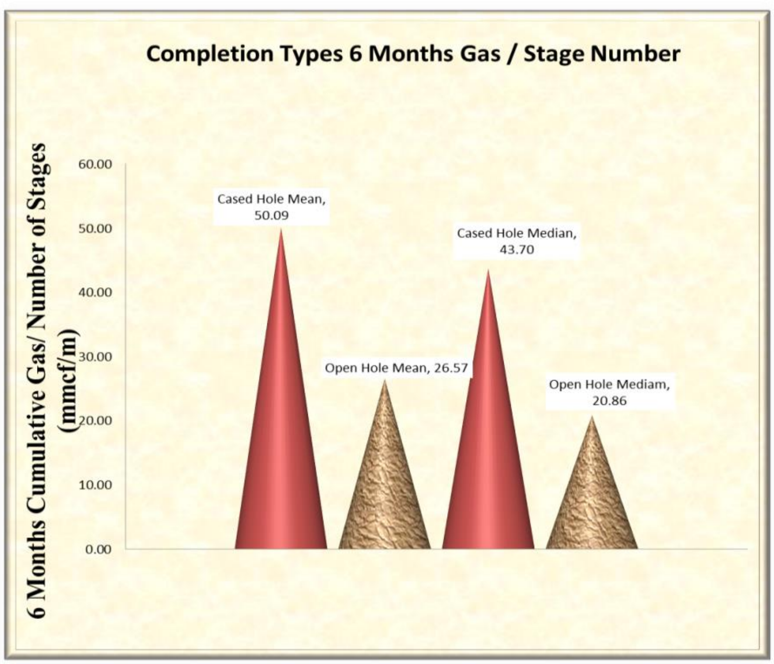

Figure 4.16 Completion Types 6 Months Gas Cumulative / Stage

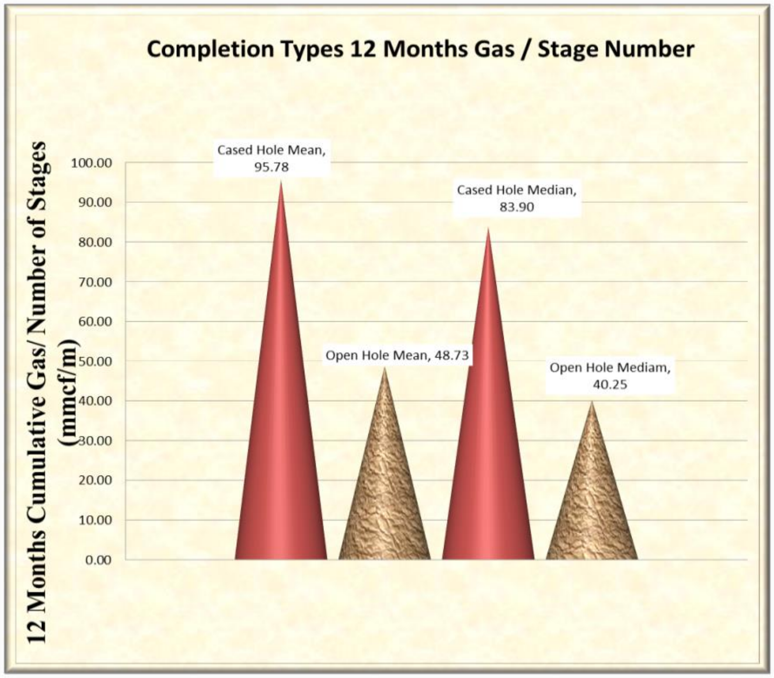

Figure 4.17 Completion Types 12 Months Gas Cumulative / Stage

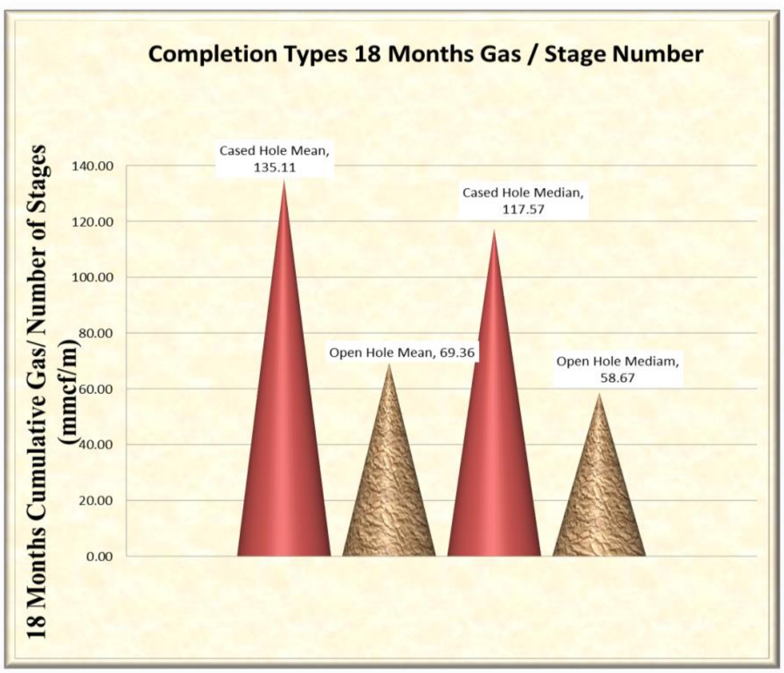

Figure 4.18 Completion Types 18 Months Gas Cumulative / Stage

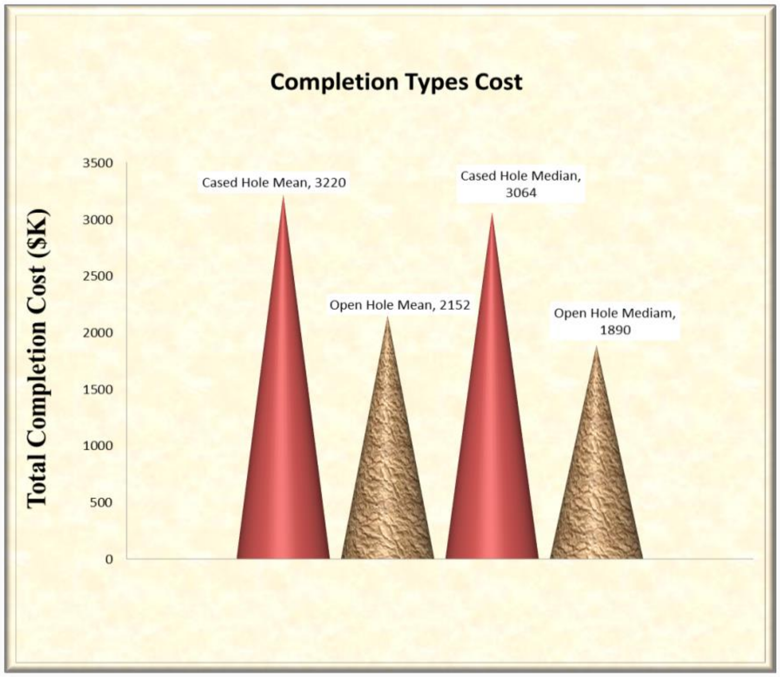

Figure 4.19 Completion Types Cost

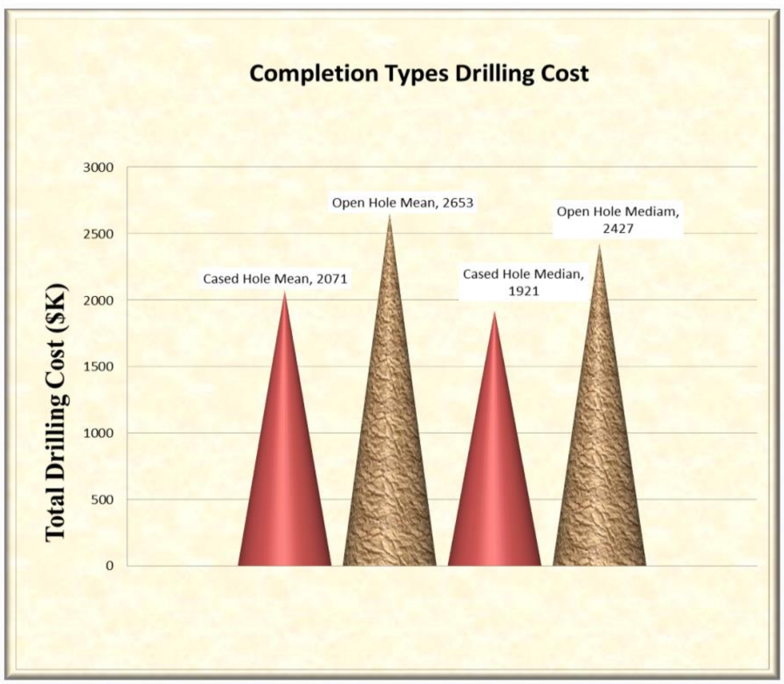

Figure 4.20 Completion Types Drilling Cost

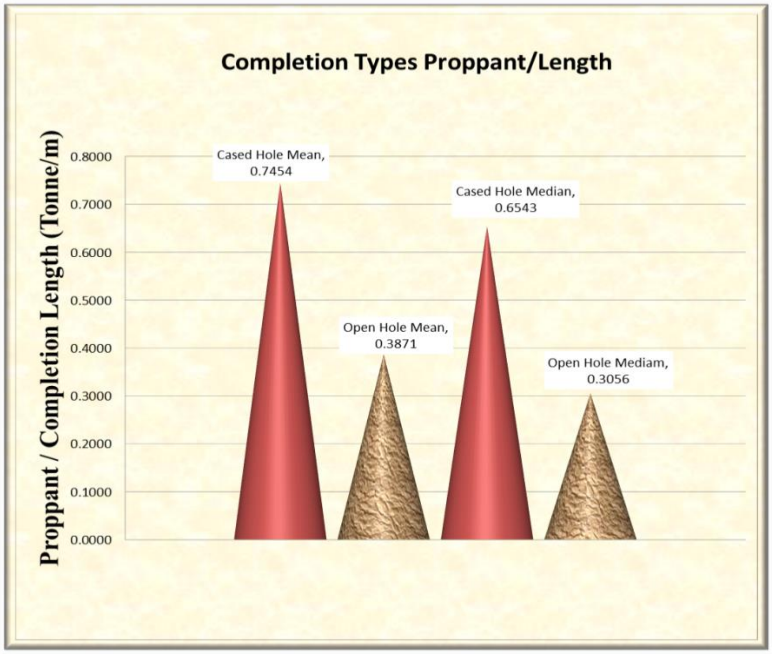

Figure 4.21 Completion Types Proppant per Length

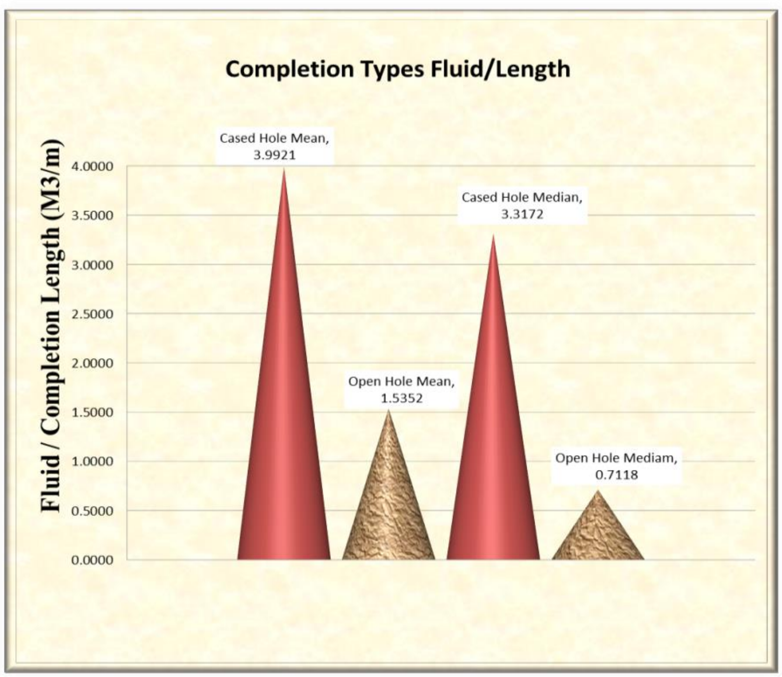

Figure 4.22 Completion Types Fluid per Length

4.1.2. Simple Cross Plots and Heat Map Techniques. In this approach, the parameters' interaction was tested throughout the entire dataset by plotting the average value of the response versus each of the variables. Grouping the plots based on different parameters was also introduced to some of the plots, which helped in better understanding the interrelated parameters in the dataset. Figure 4.23 depicts this technique, in which the data of completed lateral length versus the total completion cost were plotted for all of the wells in the Montney and grouped based on the completion type, cased versus open hole completion. A trend line was generated to connect the average total completion cost of the wells that were drilled with the same lateral length to represent the cost over length for each completion type. This plot example shows that the cased hole completion generally costs more than the open hole, and in both cases completion costs increase as the lateral length increases. In the previous section, the average completion cost per well in the Montney was calculated using the statistical technique and it was shown to be $3,220,000 for cased hole completion and $2,152,000 for open hole completion. The completion cost is different from well to well, because there are different parameters involved. The cost of the completion, and especially the cost of the fracture stimulation, depends on the type and volume of treatment fluid, type and mass of the pumped proppant, the fracturing fleets and many other factors including the hauling and disposal of waste and recovered liquids.

Several combinations of parameters were tested using the simple cross plot technique to identify the relationships between the design and performance elements in the Montney dataset. Some of the plots yield good correlations, which provides a generalized view of how that parameter varies over the entire play.

Figure 4.24 shows a cross plot of the recovered load fluid versus the total pumped load fluid, again the grouping was based on the completion type. It can be seen that there is a correlation between the pumped and the recovered load fluids. Based on the R^2 value, the cased hole completion shows a slightly better correlation than the open hole completion, and more fluid was recovered especially at the higher treatment volumes. Higher flowback fluid, particularly water-based, will increase the overall cost of the stimulation job because of the need for storing, hauling, recycling or disposal.

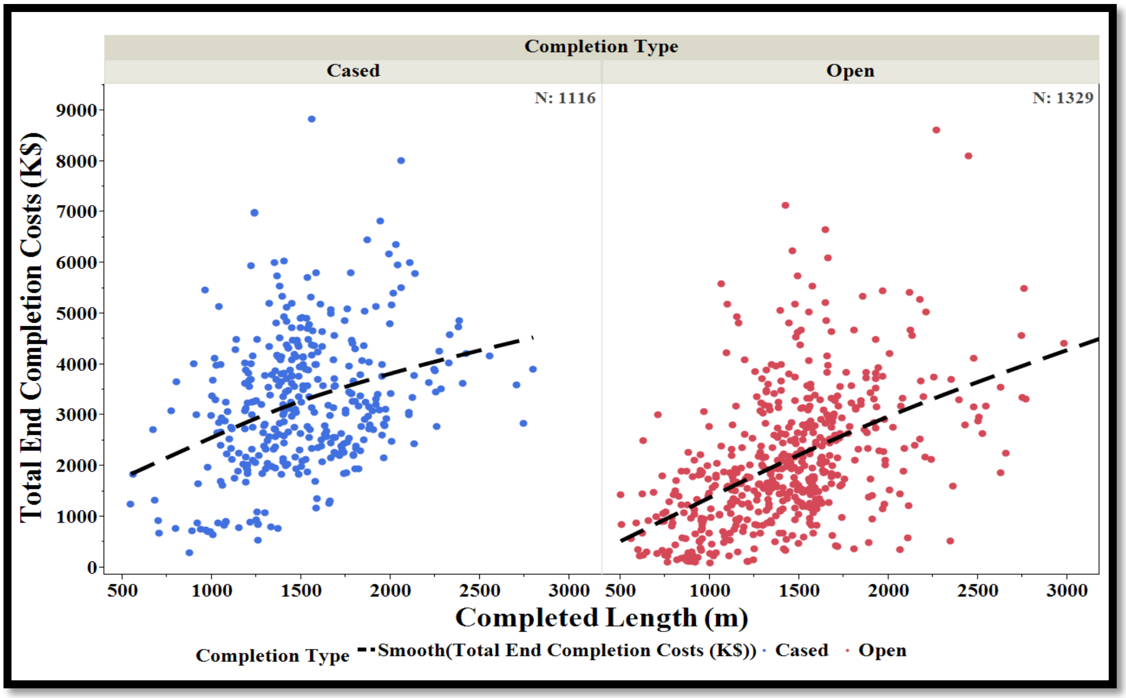

Figure 4.23 Completion Cost versus Lateral Length

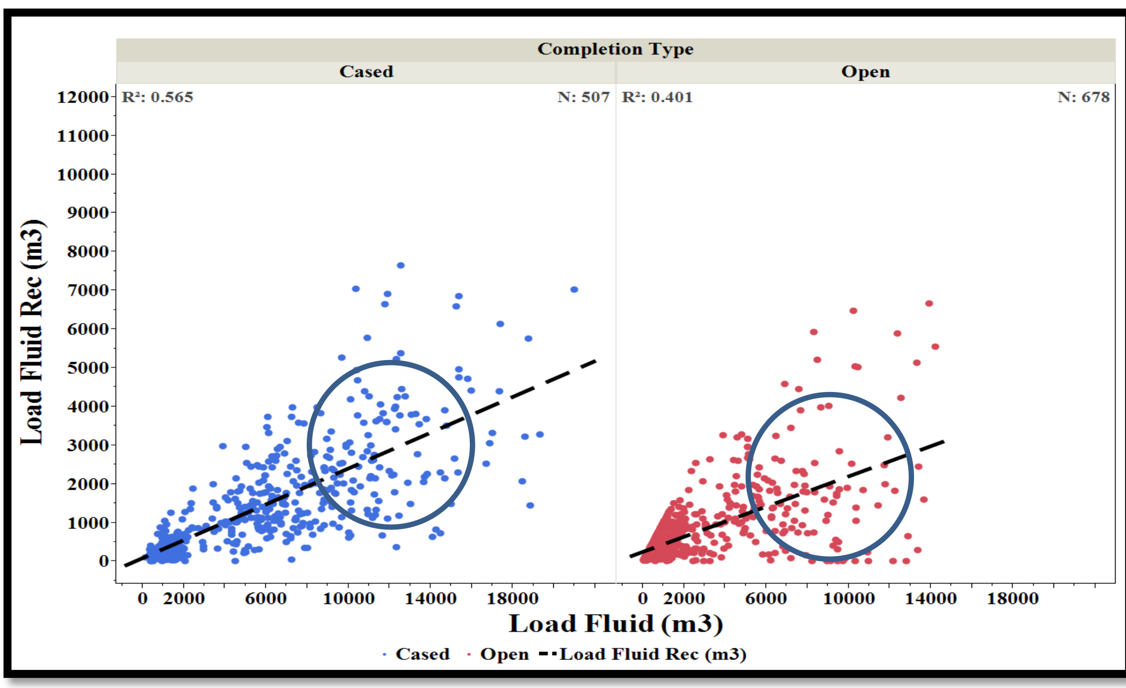

Figure 4.24 Recovered Load Fluid versus Pumped Load Fluid

The percentage of recovery was also cross plotted against the cumulative 6, 12 and 18 months to investigate the effects of load fluid recovery on the production performance of the entire play. Figure 4.25 plots the well's cumulative gas production over 6, 12 and 18 months against the percentage of fluid recovered after the fracturing treatment. The plot compares oil and water-based fracturing fluid. In both oil and water-based treatment fluid in the plot indicates a reduction in cumulative gas production as more treatment fluid percentage was recovered from the treated well. Although, this conclusion is not statistically supported by a high R^2 value, the same trend was confirmed by using heat maps techniques.

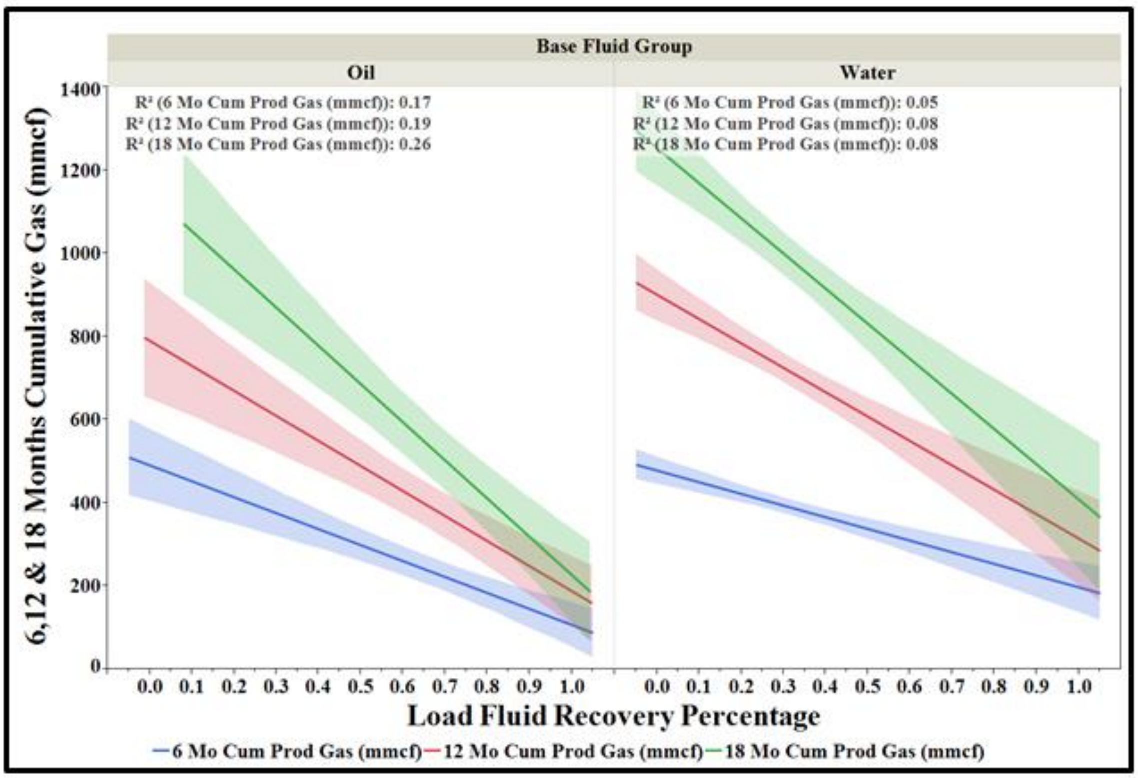

Figure 4.25 Cumulative Gas Production versus Load Fluid Recovery Percentage for 6, 12 and 18 Months

A heat map was used to further investigate this trend. In a heat map technique, a cross plot is generated and the wells are arranged and grouped in blocks that correspond to the x and y axes. These blocks are color coded in response to other tested variables to show the effect of the colored variable on the trend between the x and y variables.

The heat map interpretation becomes complicated as more variables and groups are introduced to the map. The simplest form of a heat map is to test only two variables to observe the direct effect of the colored variable on the other variable. To test only two variables using the heat map, the x and y axes will take the same variable to obtain a linear trend line and the other variable will be represented in colors with a designated scale. To identify the scale of the colored variable, a full understanding of the variable statistics and data distribution should be in place before setting the scale range values. For this research dataset all of the studied varibles were analyzed and distribution histograms were generated along with statistical tables to identify the variable mean, median, minimum and other significant statistical terms, as described in chapter three.

In order to set the color scale ranges accurately and to include the variables' distribution density in the analysis, the median value of the variable was chosen to represent the middle value of the color scale range and the maximum value was not chosen to be the highest value in the dataset. To determine the maximum value for the color scale, each variable distribution histogram was examined based on the highest value (with a reasonable count density) and set as the maximum. The minimum scale value was left as the statistical minimum.

For example, in Figure 4.26 to set a color scale of the load fluid recovery, the distribution histogram will be examined along with the summary statistic table. From the statistics table the median value will be set as the middle value for the scale which is 608, in the scale setting it will be entered as 600. The minimum value for the scale will be set to the minimum value from the statistics table, which in this example =0. To set the maximum value for the color scale, it is not recommended to select the statistics maximum directly; instead the distribution histogram should be examined. In this example the histogram shows that the value of 4000 m^3 is the highest value with 1% of the wells in the dataset, therefore; 4000 will be set as a miximum value for the color scale range and any value higher than 4000 will be colored the same as 4000.

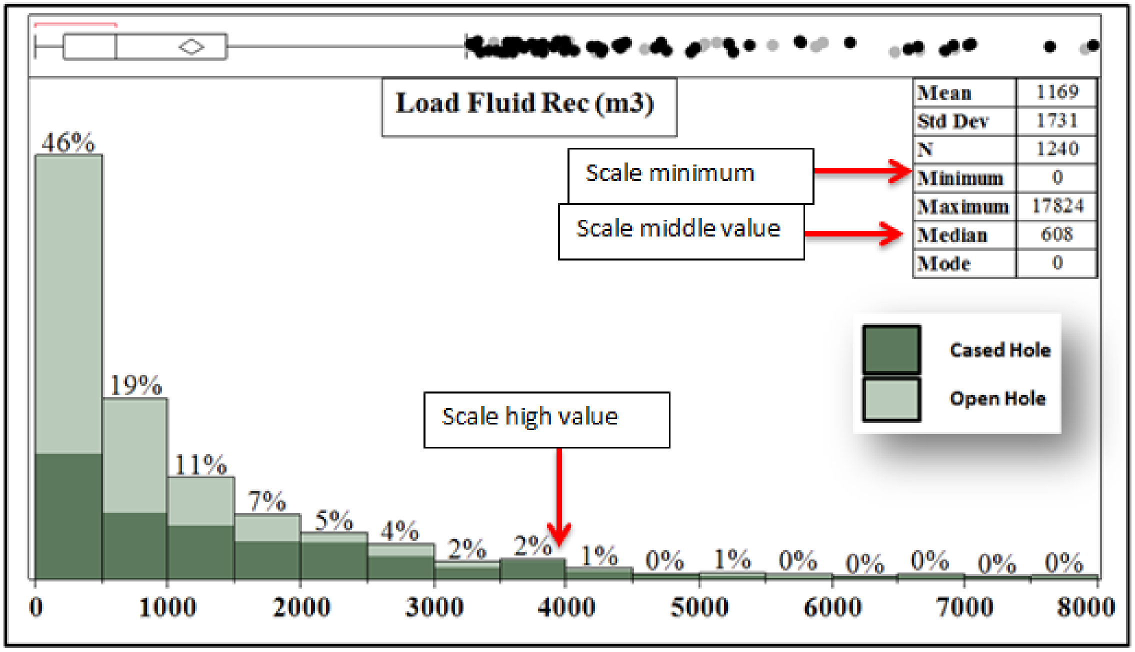

Figure 4.26 Example of Setting Color Scale for a Variable

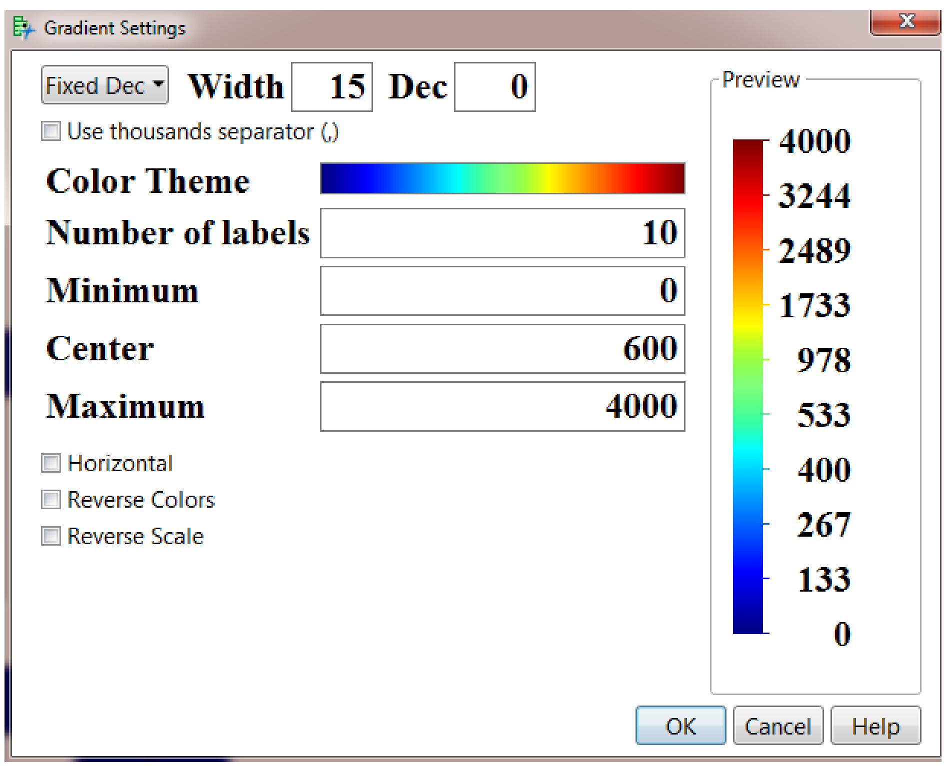

Figure 4.27 Example of Setting the Color Scale in JMP

Error! Reference source not found. combines two plotting techniques. On the eft side, the 6 months' cumulative production was plotted against the fracturing fluid recovery percentage, while on the right side of the figure a heat map was used to show the effects of fracturing the fluid recovery percentage on the 6 months' cumulative production, where the same variable is plotted on the x and y axes. The color scale was set based on the distribution histogram in Figure 3.18, as presented in chapter three. The minimum value was set as 0 and the middle value was set as 0.3, to match the median of the recovery fluid percentage across the entire dataset, and the maximum was set to 0.7.

Error! Reference source not found. clearly shows from the cross plot on the left ide that as more load fluid was recovered, the 6 months' cumulative gas production decreased. The heat map on the right side shows that the wells with a higher load fluid recovery, which are represented in red, produced the minimum cumulative 6 months of gas while the wells with lower fluid recovery, that are shown in blue, dominated the higher gas production range. The same trend was confirmed in Figure 4.29 and Figure 4.30 for the 12 and 18 months' cumulative gas recovery.

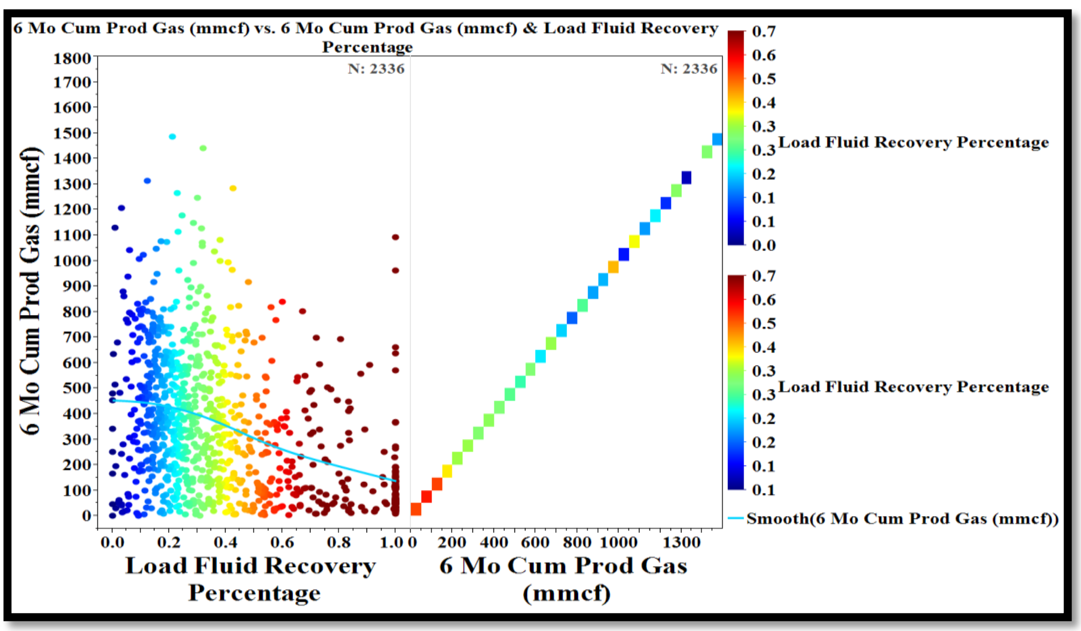

Figure 4.28 Cross Plot and Heat Map of 6 Months' Cumulative Production and Load Fluid Recovery Percentage

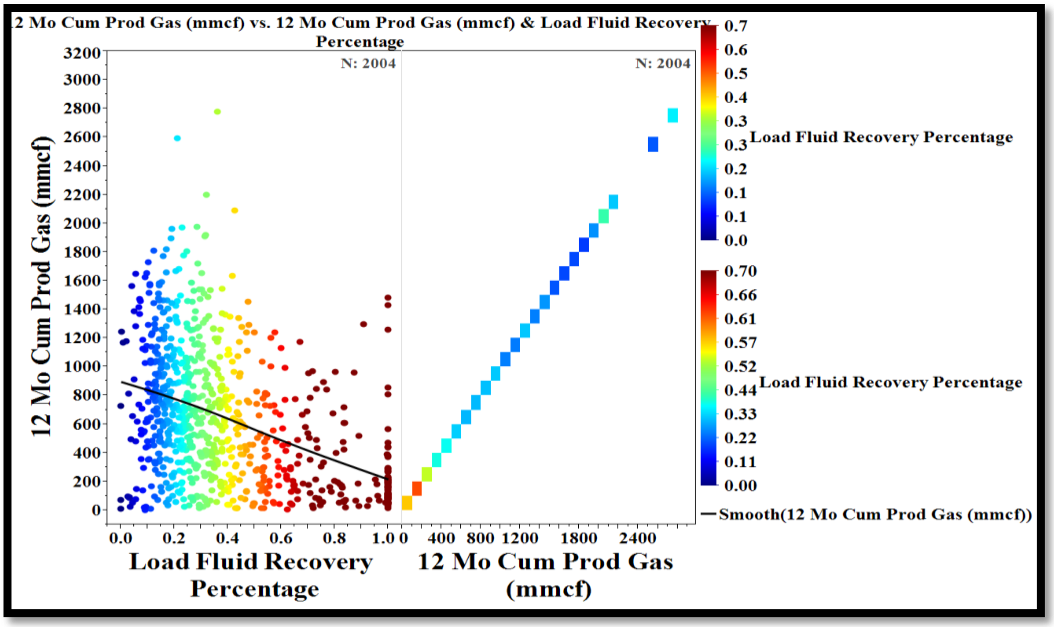

Figure 4.29 Cross Plot and Heat Map of 12 Months' Cumulative Production and Load Fluid Recovery Percentage

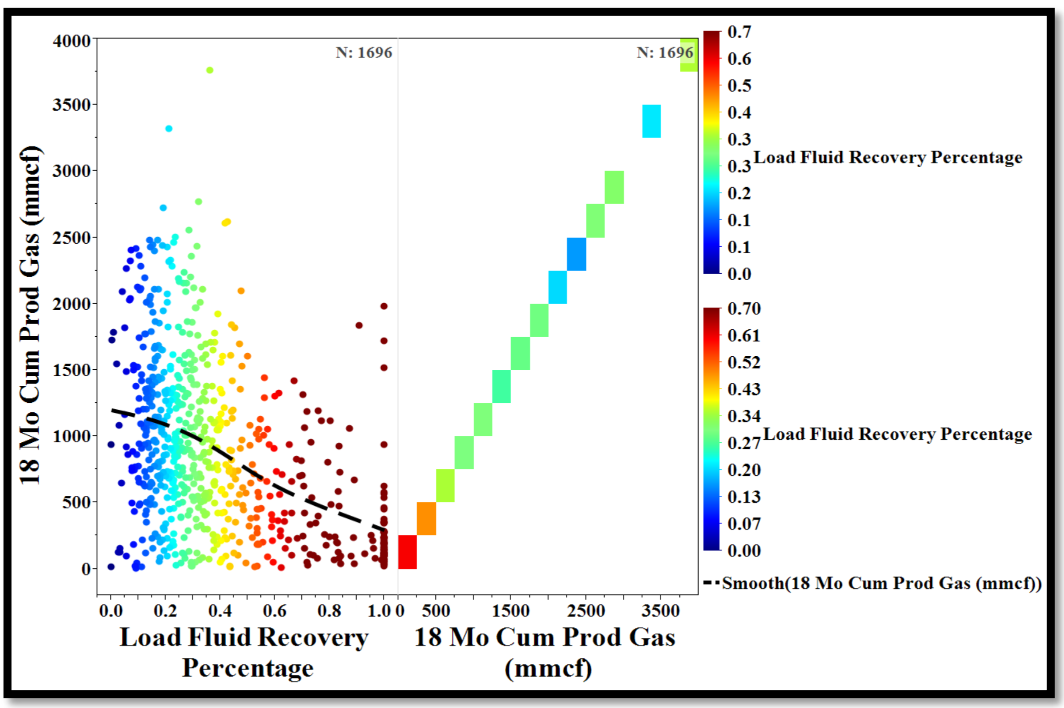

Figure 4.30 Cross Plot and Heat Map of 18 Months' Cumulative Production and Load Fluid Recovery Percentage

A heat map of the percentage of recovery fluid on both the 18 months of cumulative production and the IP is shown in Figure 4.31, The trends shown are similar to those seen in Figure 4.28 to Figure 4.30.

For the same cumulative production, the wells that produced with higher IP are less efficient than the wells that produced with lower IP, because the IP and cumulative production have a strong correlation, i.e higher IP should correspond to higher production, while in this case, the wells with a higher recovery percentage (colored in red and orange) had a high IP but produced the same cumulative production as the wells with the lower IP. This confirms the negative effect of high fluid recovery on the gas' cumulative production, because these wells produced at a faster decline rate compared to the other wells that produced the same cumulative production with a lower initial production rate.

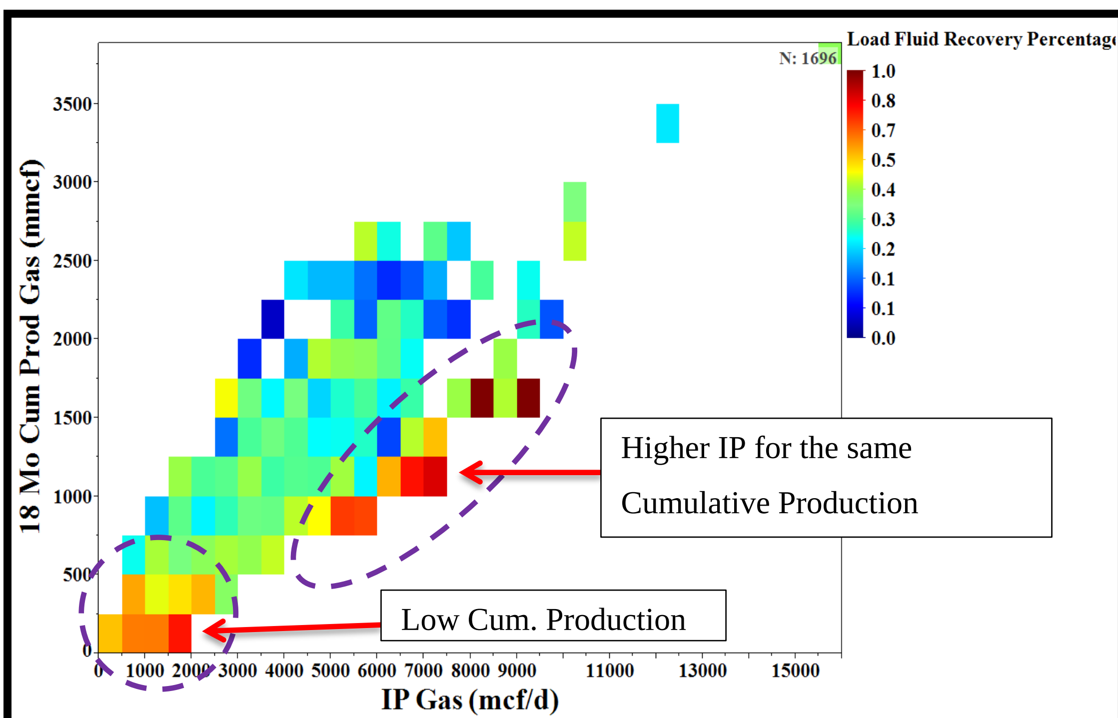

Figure 4.31 Heat Map of 18 Months Cum. Gas Production versus. Gas IP Colored Based on Fluid Recovery Percentage

A strong and expected trend was recognized in the dataset between the initial production rate and the cumulative production over time.

Plotting gas IP versus the 6 months of cumulative gas production results in a very strong correlation with an R^2 value of 0.867 for 1158 cased hole wells, and an R^2 of 0.936 for 1240 open hole wells. This correlation is presented in Figure 4.32.

The early production (6 months) and the IP showed a very good correlation, and based on that correlation the 12 and 18 months of cumulative production were plotted against the IP to verify the relationship. Figure 4.33 and Figure 4.34 show the correlation of the gas IP versus the 12 and 18 months' cumulative gas production, respectively. the relationship remains strong with a significant R^2 value. The R^2 value of correlation decreases slightly as the production time increases, but it continues to be significant. For instance, in the open hole wells the correlation R2 changed from 0.936 to 0.893, then 0.83 as the cumulative production time changed from 6 to 12 and then 18 months. Based on the above plots, wells with the highest IPs will be the best producers.

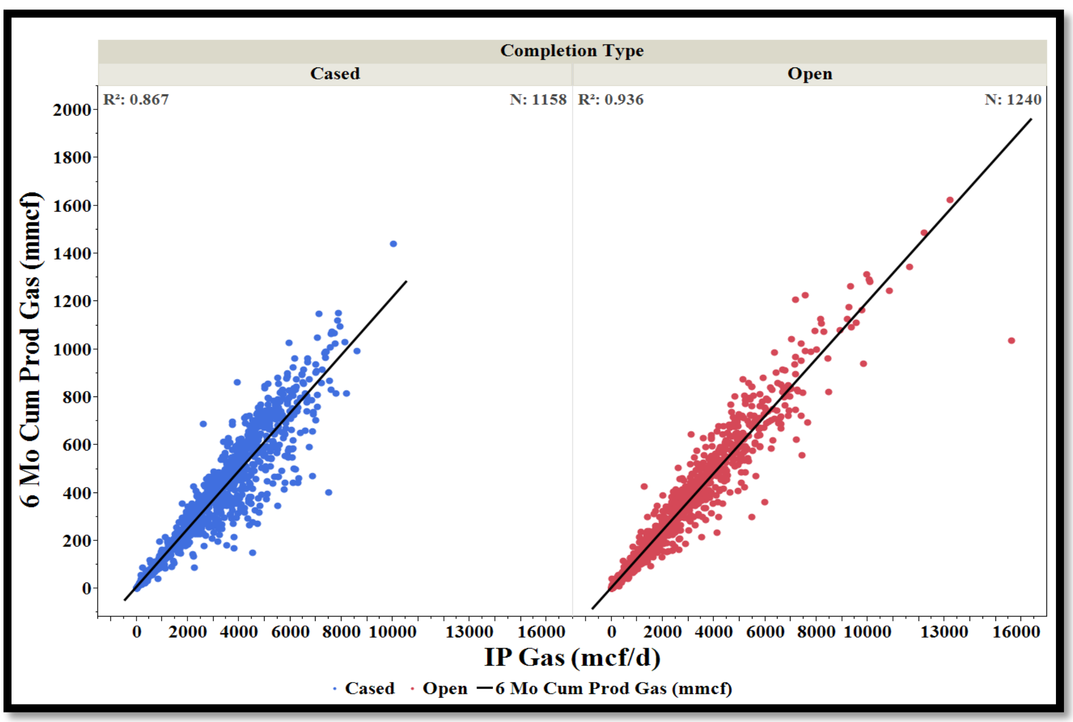

Figure 4.32 IP Gas versus. 6 Months of Cum. Gas Production

It is important to note that the dataset used in this study only includes the cumulative production up to 18 months. Hence, any correlation using cumulative production is limited up to this time. It is recommended to investigate the correlations for longer production periods, e.g. 5 years or more.

.

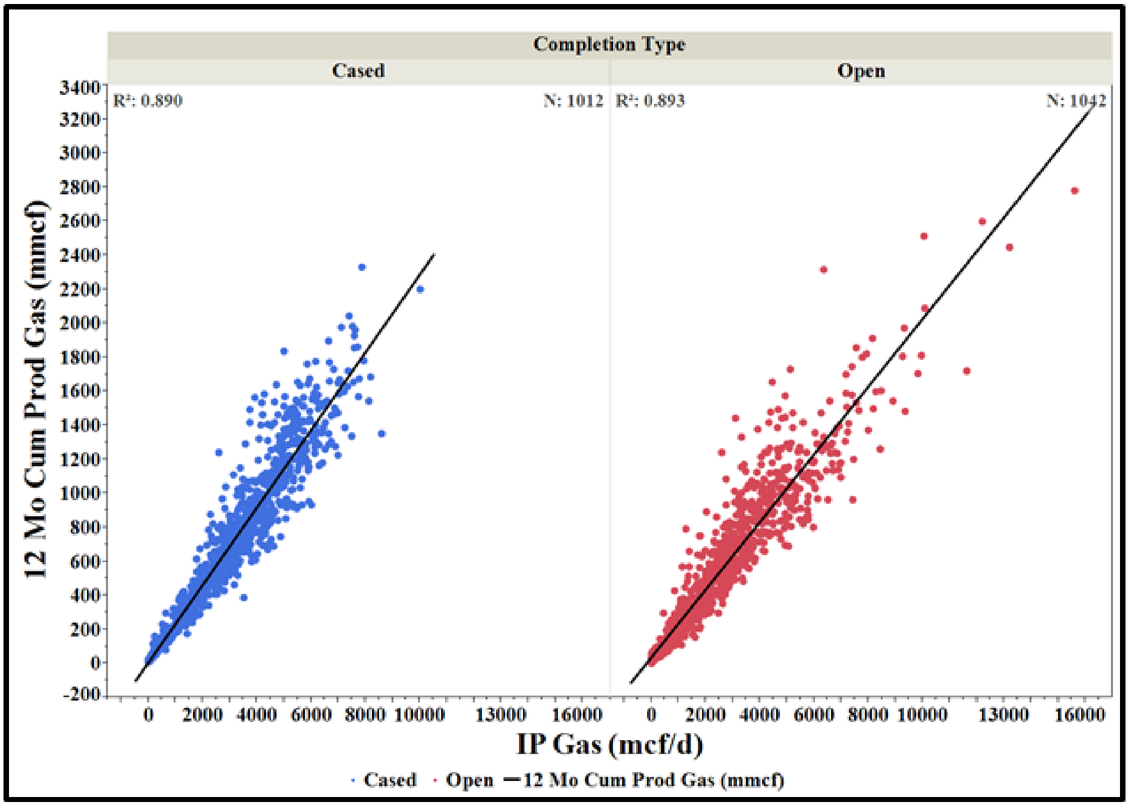

Figure 4.33 IP Gas versus 12 Months of Cum. Gas Production

Figure 4.35 through Figure 4.38 confirm the correlation IP and cumulative production for oil and water.

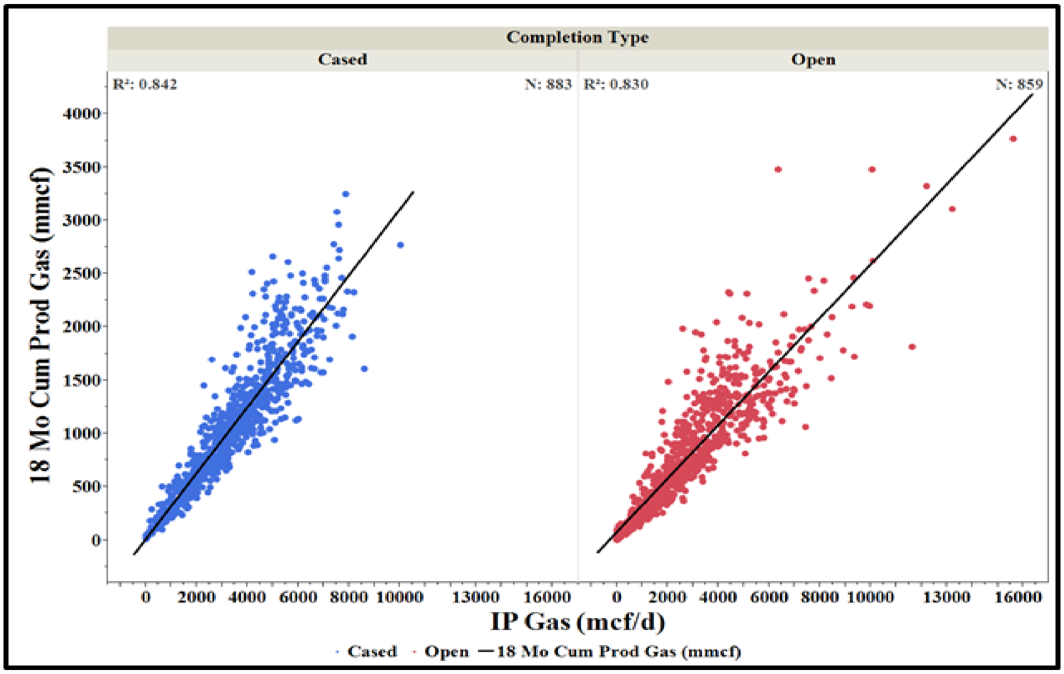

Figure 4.34 IP Gas versus 18 Months of Cum. Gas Production

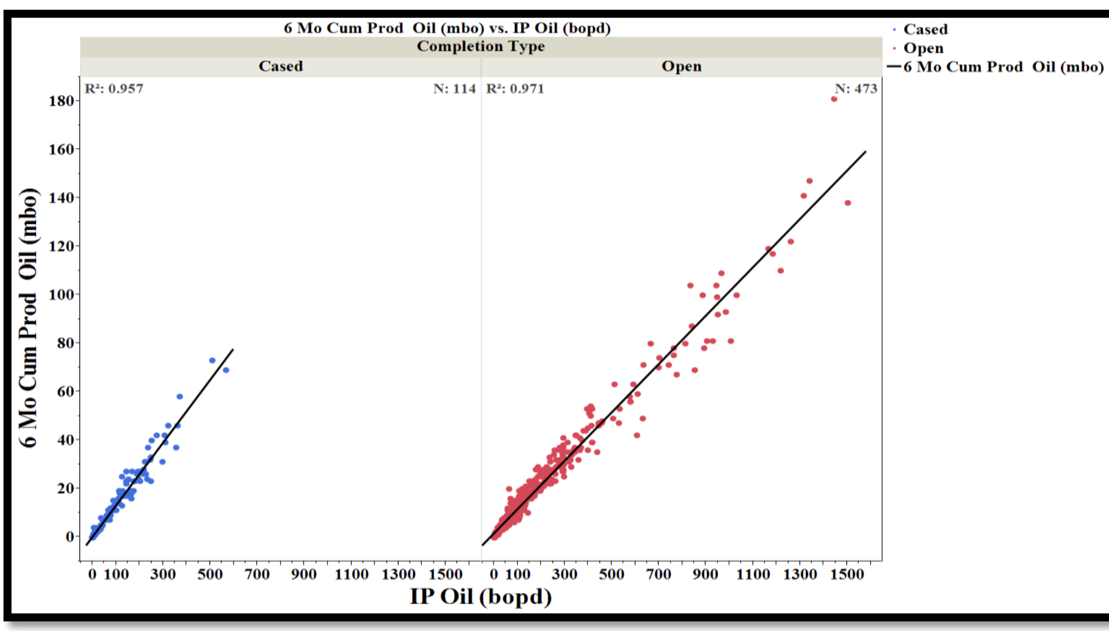

Figure 4.35 IP Oil versus 6 Months of Cum. Oil Production

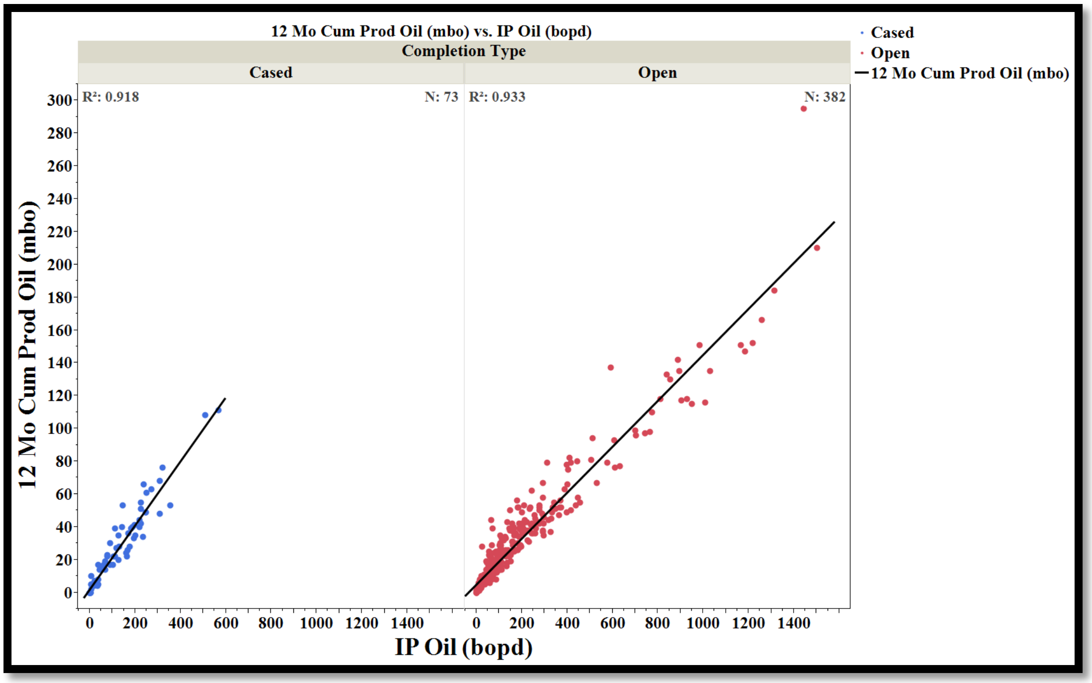

Figure 4.36 IP Oil versus 12 Months of Cum. Oil Production

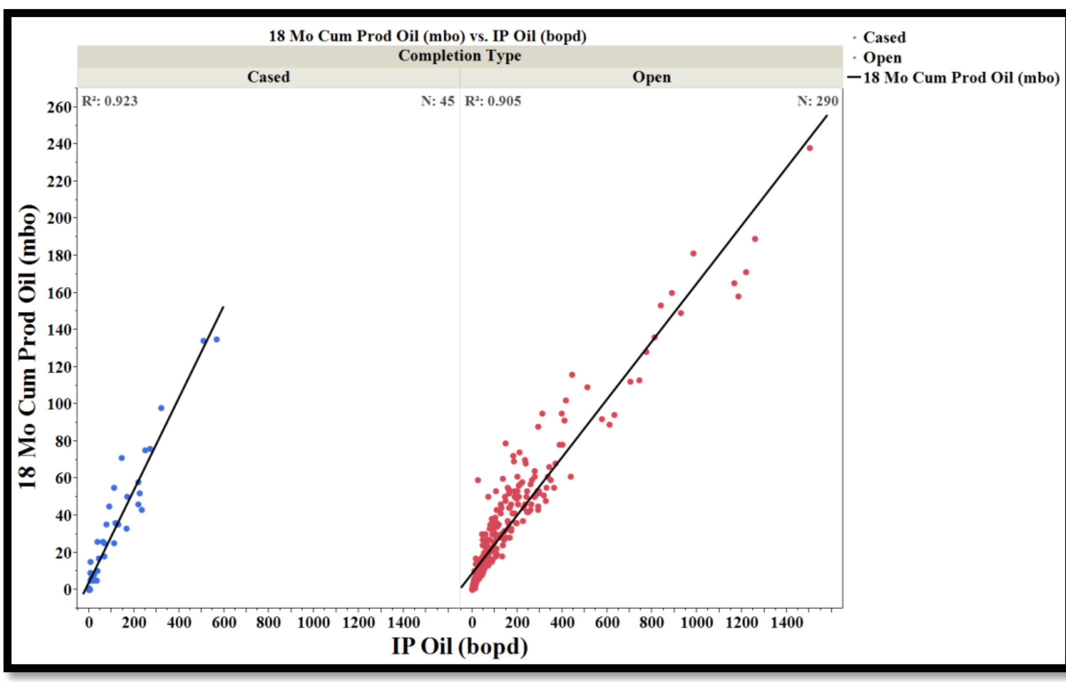

Figure 4.37 IP Oil versus 18 Months of Cum. Oil Production

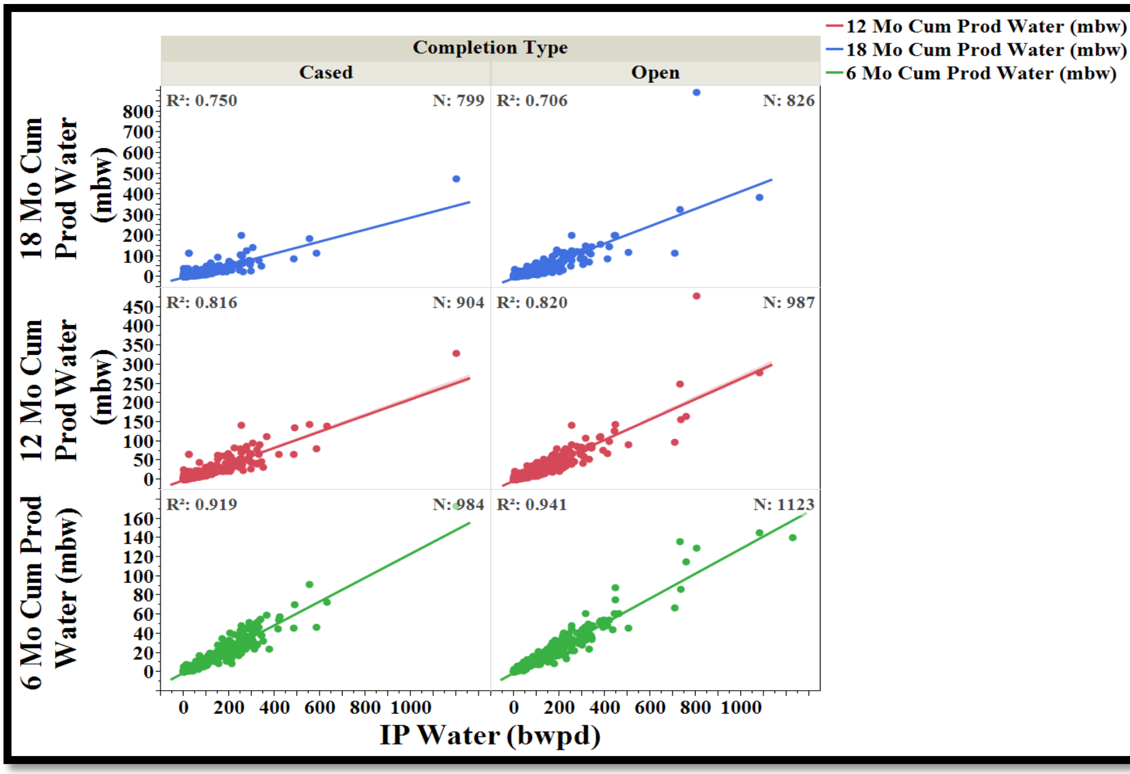

Figure 4.38 IP Water versus. 6,12 &18 Months of Cumulative Water Production

The dataset included information regarding both attempted and actual stages that were completed in the wells. This allowed for production to be normalized by stage. However, no detailed information about perforation clusters was given in the dataset.

The actual stages completed versus the normalized cumulative for 6 months of production per actual stage number (mmcf/stage) shows that the actual production per stage decreases as the number of treated stages increases. In other words, the effectiveness of the production per stage is decreased when comparing the production from one stage, which indicates that there is production interference between fracs in wells with high stage density.

This conclusion was also tested by the heat map techniques. Figure 4.39 shows the effect of increasing the number of stages on the effective production per stage. The plot is grouped by five ranges of proppant concentration brackets to show the effects of the concentration parameter on the effective production per stage and is also grouped based on the completion type.

This analysis shows that increasing the proppant concentration slightly increases the overall productivity of the well. Cased hole wells' production shows a better response to the proppant concentration than the open hole's completion. Figure 4.40 and Figure 4.41 confirm the same trends for the normalized 12 and 18 months of cumulative gas per stage.

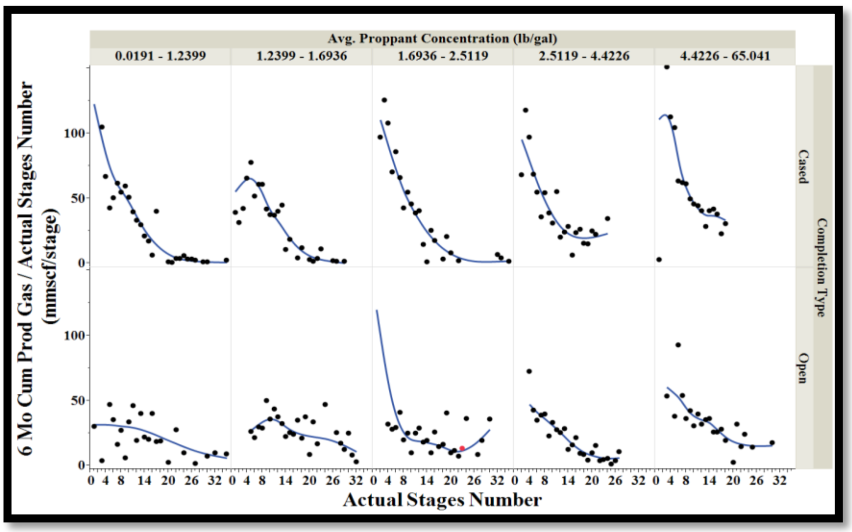

Figure 4.39 Normalized per stage 6 Months Cumulative Production versus Actual Stage Number

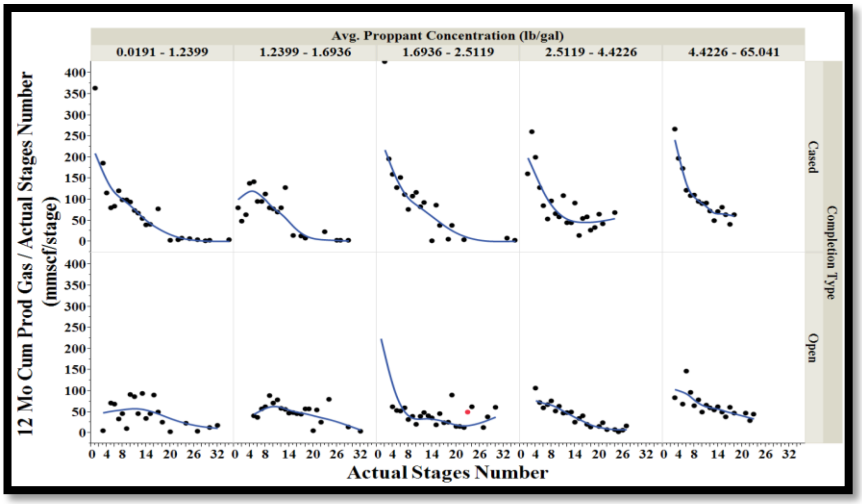

Figure 4.40 Normalized 12 Months Cumulative Production versus Actual Stage Number

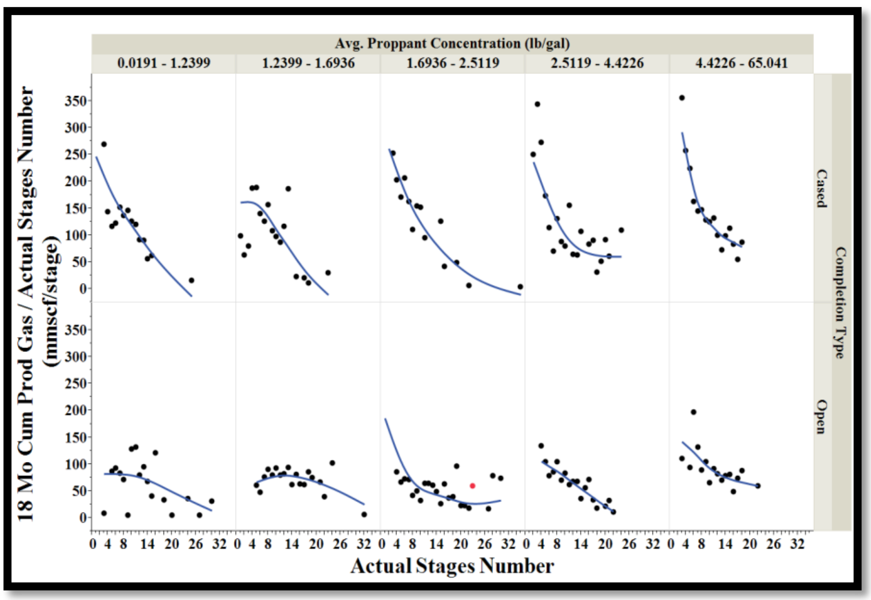

Figure 4.41 Normalized 18 Months Cumulative Production versus Actual Stage Number

Figure 4.42 shows the effective 18 months' cumulative production per stage versus the actual number of stages grouped into five brackets of gas IP, and colored based on the completion type into open hole red and cased hole blue. This plot also demonstrates that as the number of stages increase the effective production per stage decreases. This is an indication of frac interference.

The reduction in the effective production per stage as the stage number increases was validated using the heat map technique. In Figure 4.43 the actual number of stages were set on a color scale and the production variable was plotted as a single variable on the right side of each plot by assigning the same parameter on the x and y axes to test the effect of IP gas/stage, 6 months of gas production/stage, 12 months of gas production/stage and the 18 months of gas production/stage.

The heat map confirmed that the lowest effective cumulative production per stage was always associated with the highest number of stages. (Figure 4.44 - Figure 4.46)

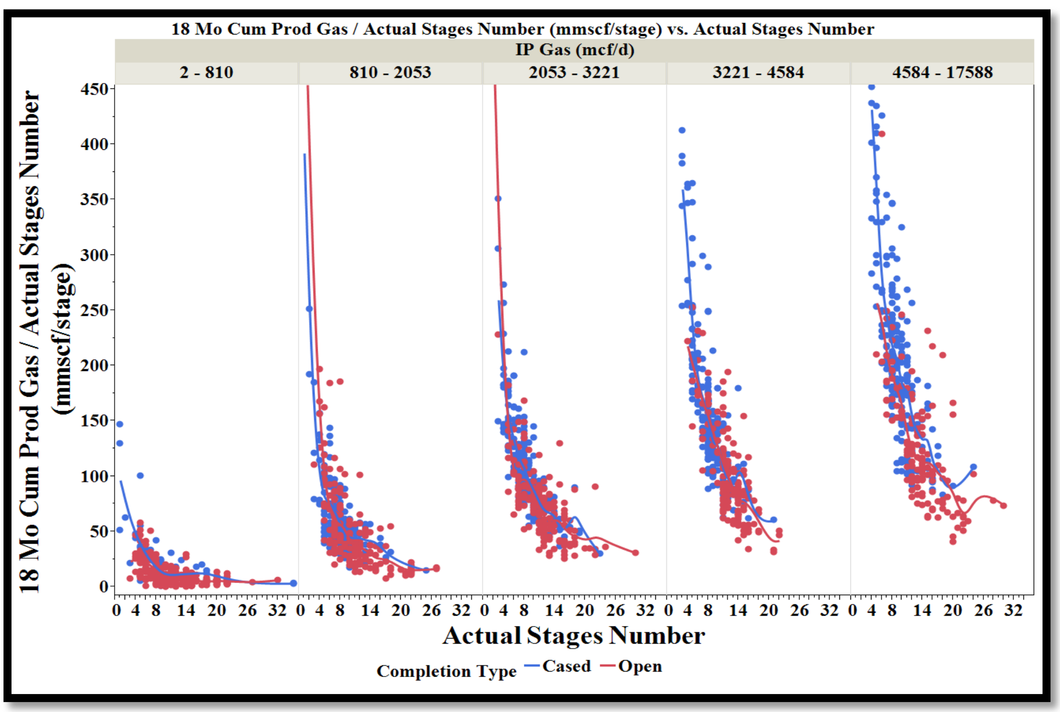

Figure 4.42 Actual Stage Number versus. 18 Months of Cum Gas / Actual Stage Number

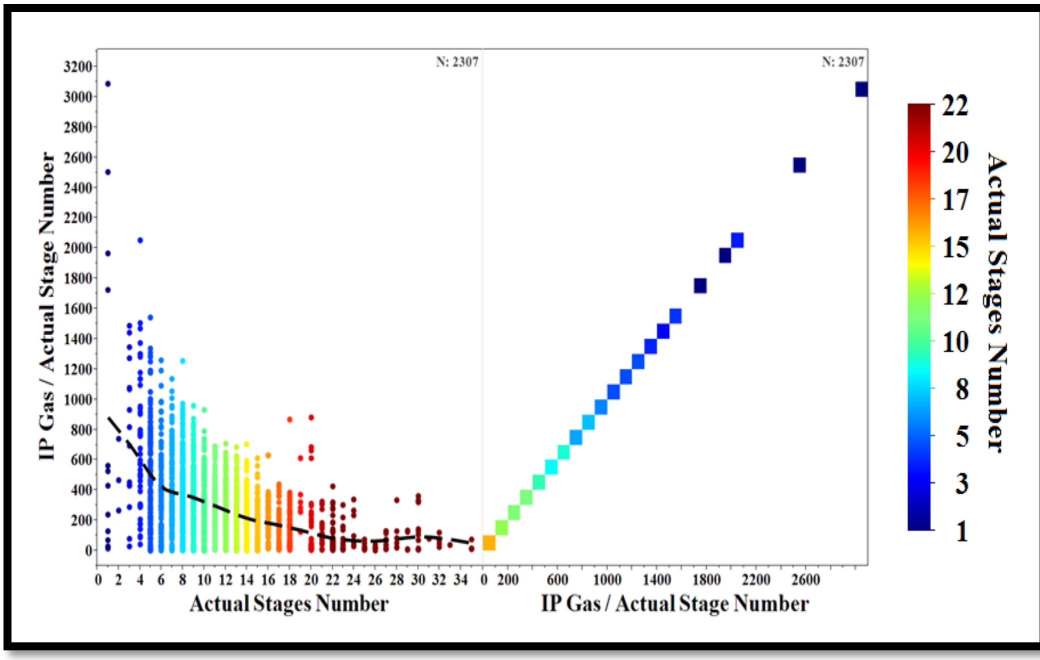

Figure 4.43 Heat Map of Normalized IP Gas/Stage versus Actual Number of Stages

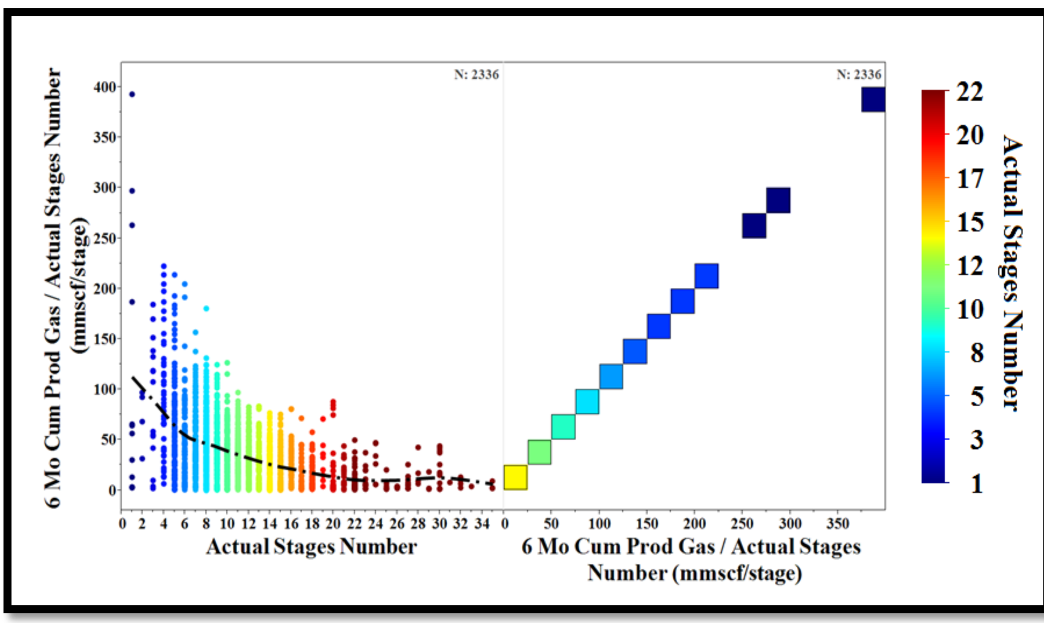

Figure 4.44 Heat Map of Normalized 6 Months of Cumulative Gas/Stage versus Actual Number of Stages

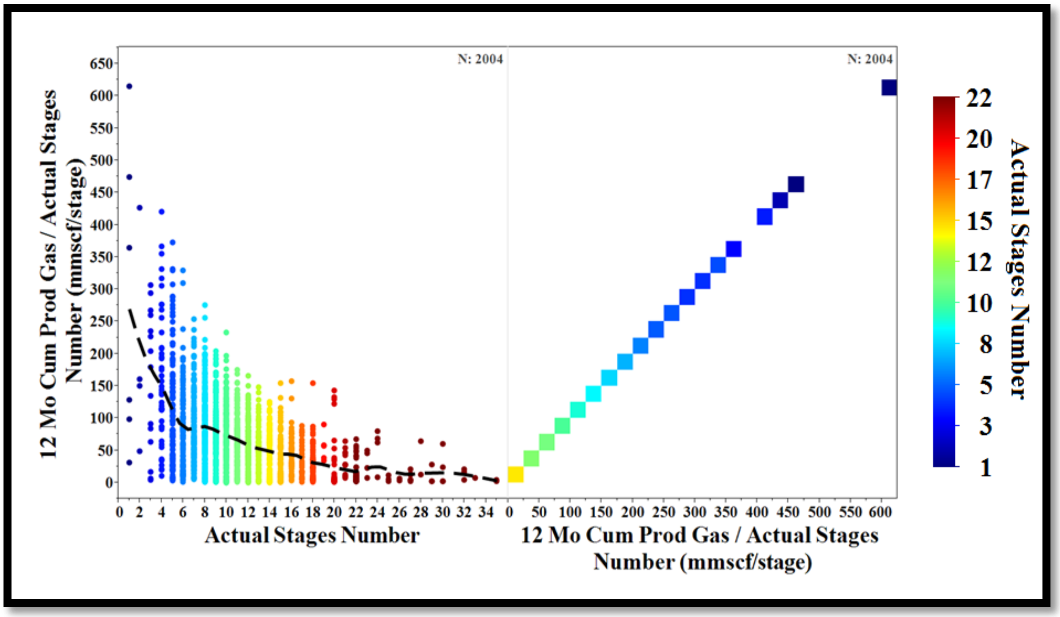

Figure 4.45 Heat Map of Normalized 12 Months of Cumulative Gas/Stage versus Actual Number of Stages

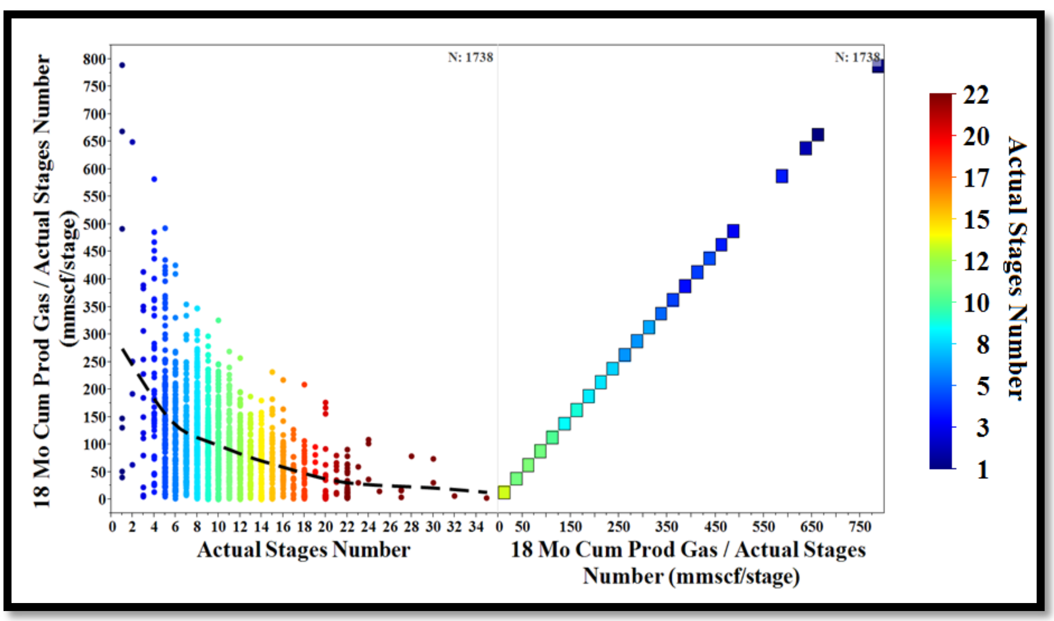

Figure 4.46 Heat Map of Normalized 18 Months of Cumulative Gas/Stage versus Actual Number of Stages

Figure 4.47 shows a simple cross plot of the cumulative 18 months of gas production normalized by the number of stages versus the actual number of stages for the two completion types in the entire dataset. The cased hole performs better than the open hole within the same trend.

Figure 4.48 introduces the effects of total proppant placed on the effective production per stage. It shows that increasing the proppant mass pumped increases the productivity per stage.

Figure 4.49 introduces the effects of the total fluid pumped on the effective production per stage. It shows that increasing the volume of the treatment fluid increases the productivity per stage.

Figure 4.50 introduces the effects of the completed lateral length on the effective production per stage. It shows that increasing the lateral length increases the productivity per stage. The plot also shows that the lateral length's effect on the productivity per stage becomes significant at 10 stages or higher.

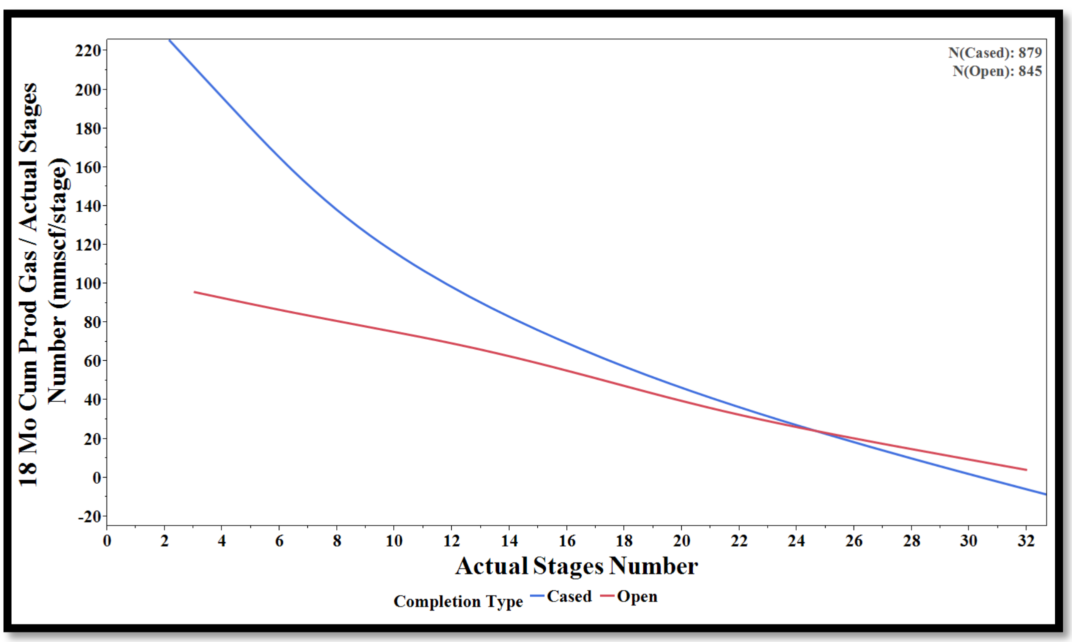

Figure 4.47 Cross Plots 18 Months Cum. Gas /Stage versus Number of Stages (Cased, Open) Hole

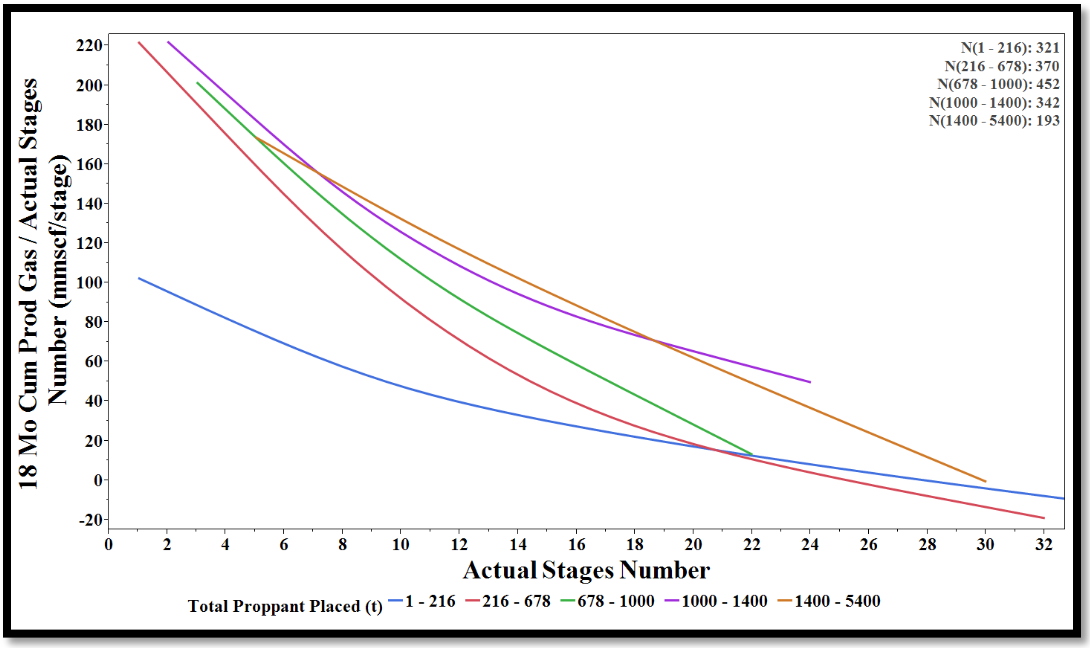

Figure 4.48 The Effects of Total Proppant Placed on the Cross Plot of 18 Months Cum. Gas /Stage versus Number of Stages

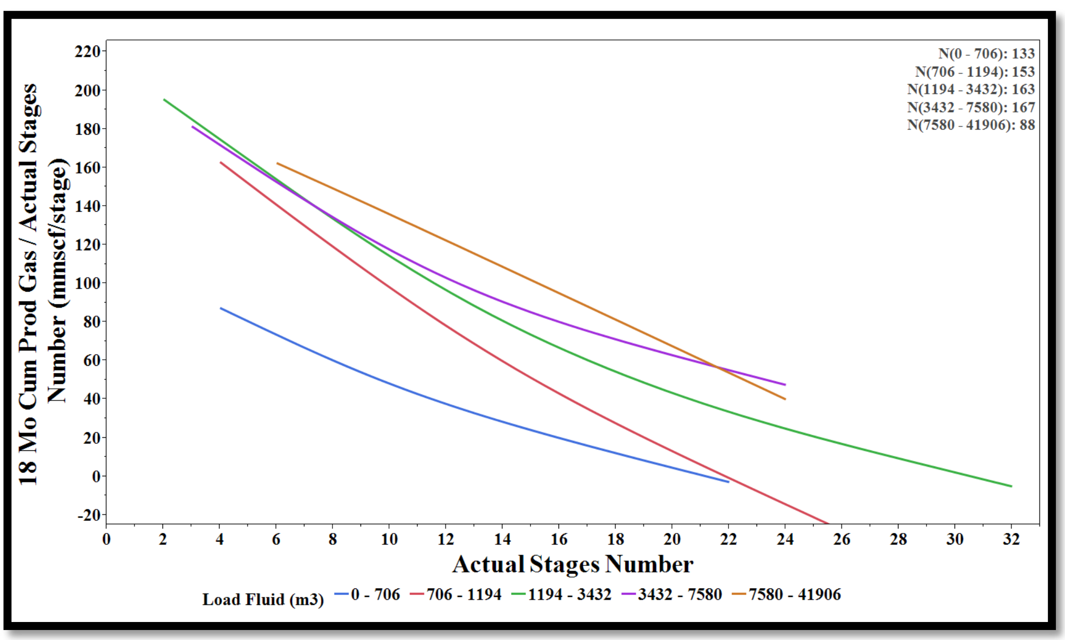

Figure 4.49 The Effects of Load Fluid on the Cross Plot of 18 Months Cum. Gas /Stage versus Number of Stages

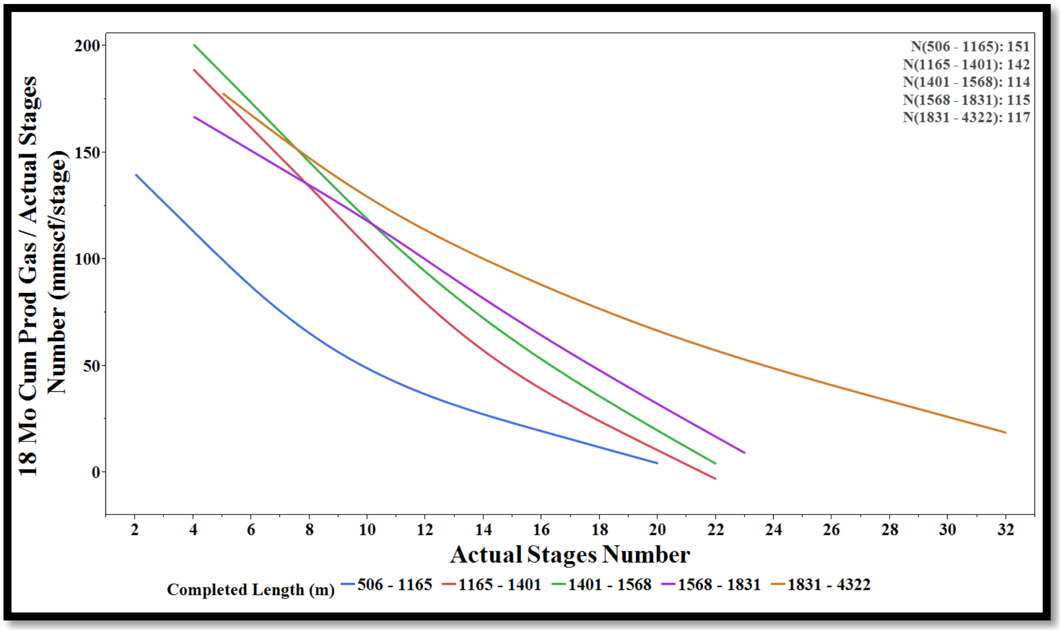

Figure 4.50 The Effects of Completed Lateral Length on the Cross Plot of 18 Months Cum. Gas /Stage versus Number of Stages

Figure 4.51 cross plots the 18 months of cumulative gas production per stage versus the average proppant per stage. The plot shows that the production per stage increases as the proppant per stage increases up to a point, after which the curve flattens. On average more than 200 tonnes/stage will not greatly improve the gas production per stage.

Figure 4.52 cross plots the 18 months of cumulative gas production per stage versus the average fluid pumped per stage. The plot shows that the production per stage increases as the fluid per stage increases up to a point, after which the curve flattens or drops. On average more than 1000 m^3/stage will not greatly improve the gas production per stage.

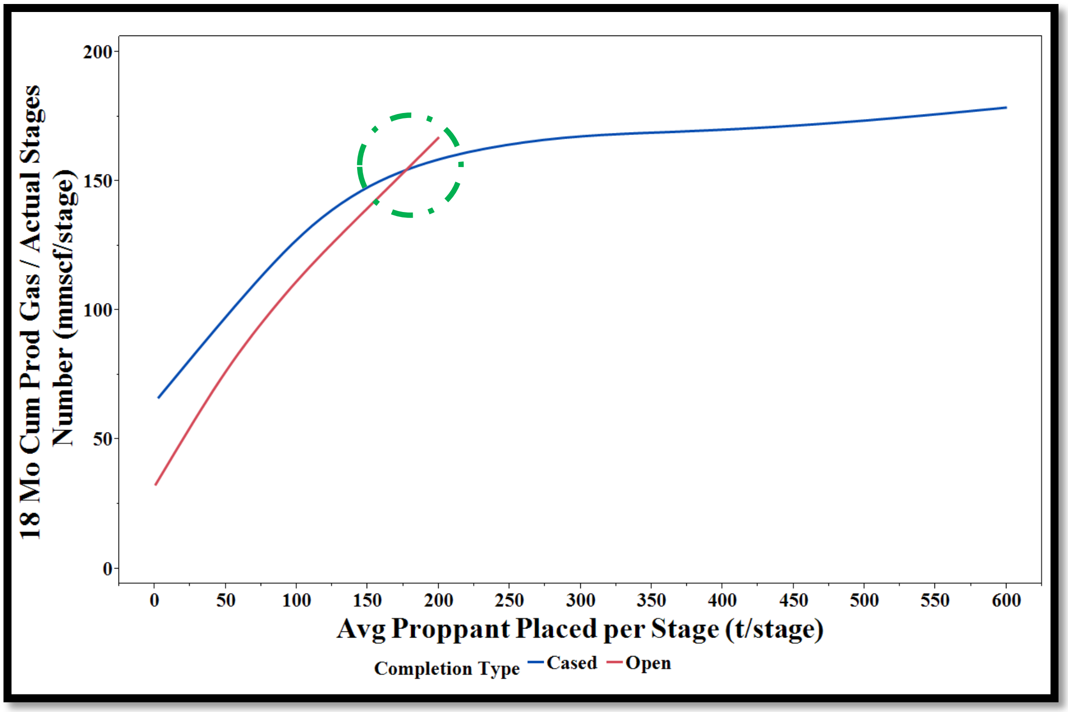

Figure 4.51 Cross Plot of 18 Months Cum. Gas/Stage versus Avg. Proppant/Stage

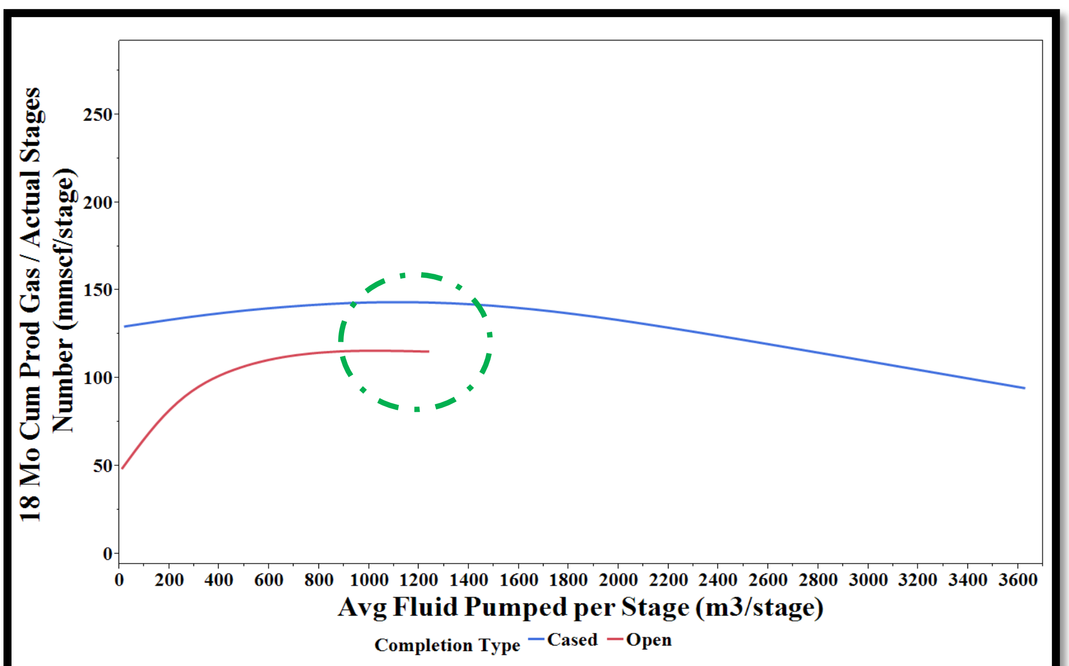

Figure 4.52 Cross Plot of 18 Months Cum. Gas/Stage versus Avg. Fluid / Stage

4.2. ANALYSIS OF PARAMETERS OVER TIME

To appreciate the trends in hydraulic fracturing designs that operators in the Montney adapted over time, the main parameters in the dataset were plotted as a snapshot over a timeframe of 10 years.

4.2.1. Completed Length (m). Over time, the operators in the Montney increased the wells' lateral length. Figure 4.53 shows the lateral length increment as a function of time. In this plot, it is noticeable that between 2005 and 2010 the wells' lateral length was ranging between 500-2000 m, then after 2010 the implemented lateral length started to increase further every year going up to more than 3500 m per lateral.

4.2.2. Number of Stages. The attempted and actual achieved number of stages over the time span between 2006 and 2014 was plotted in Figure 4.54 and Figure 4.55 respectively. The plot shows an increase in the number of stages over time. This trend of increasing the number of stages over time is in conformance with the lateral length increase.

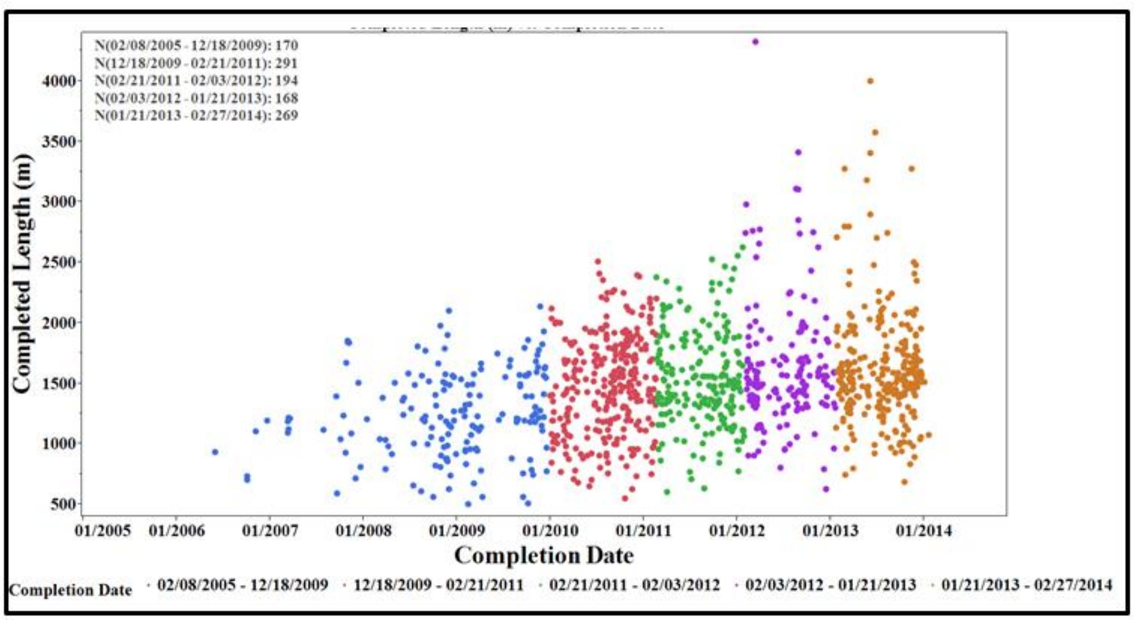

Figure 4.53 Completed Lateral Length Versus Time

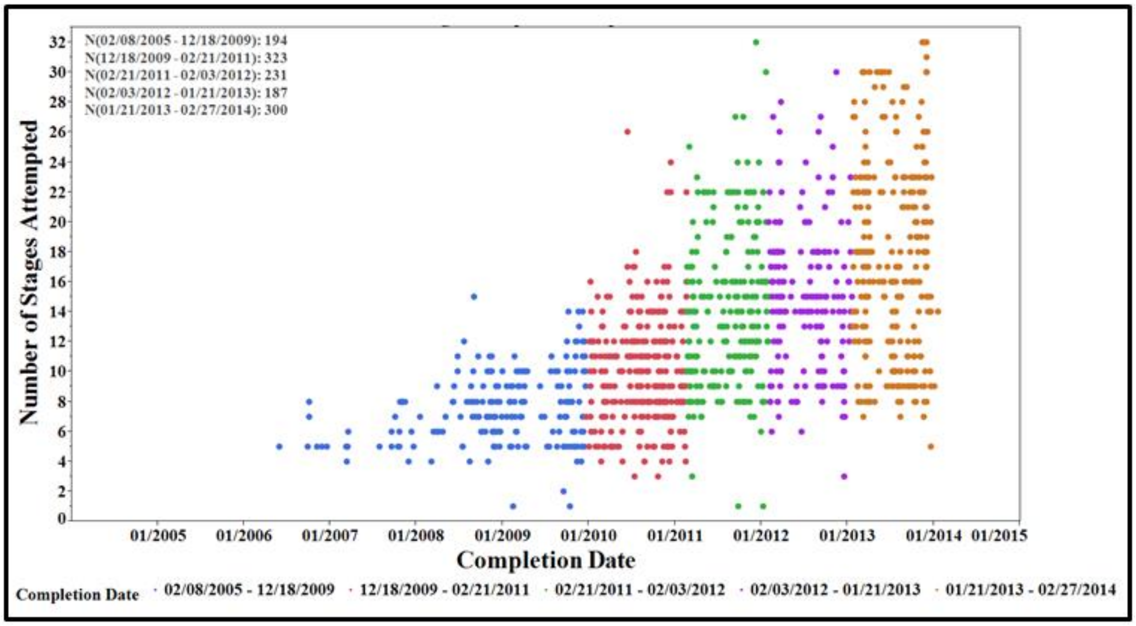

Figure 4.54 Attempted Number of Stages versus Time

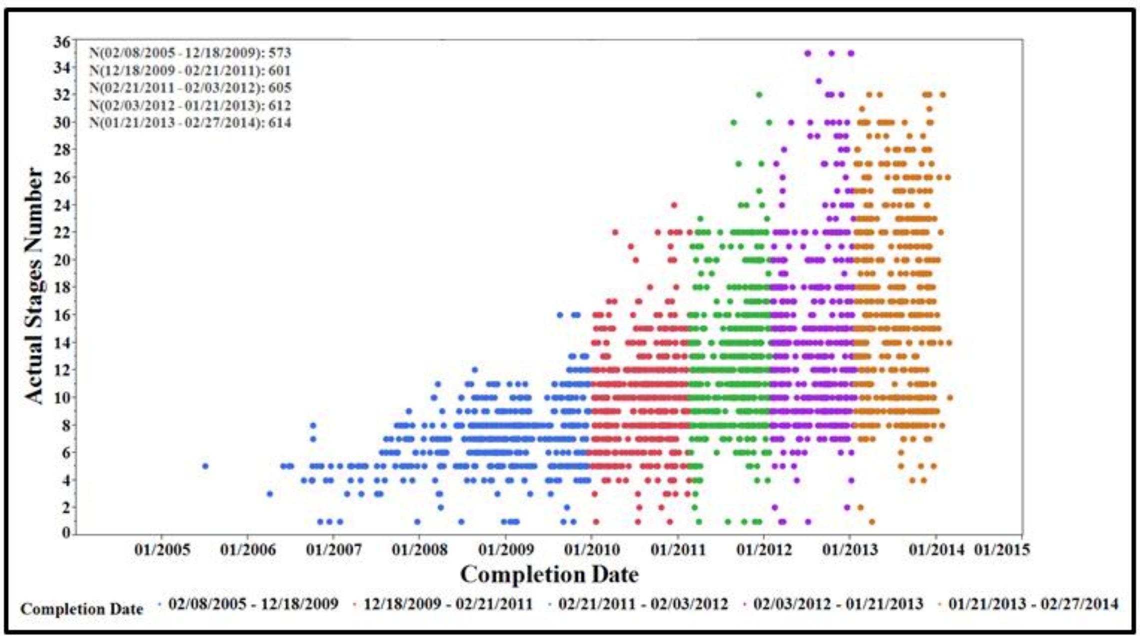

Figure 4.55 Actual Number of Stages versus Time

4.2.3. Total Proppant Placed (tonne). Proppant use in the Montney increased over time.

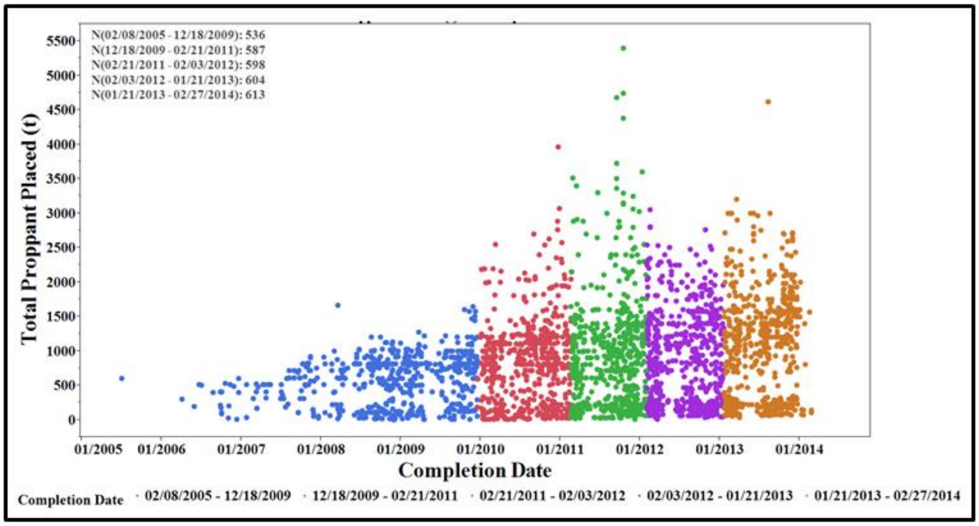

Figure 4.56 Total Proppant Placed versus. Time

4.2.4. Total Pumped Fluid (M^3). More fluid is pumped over time.

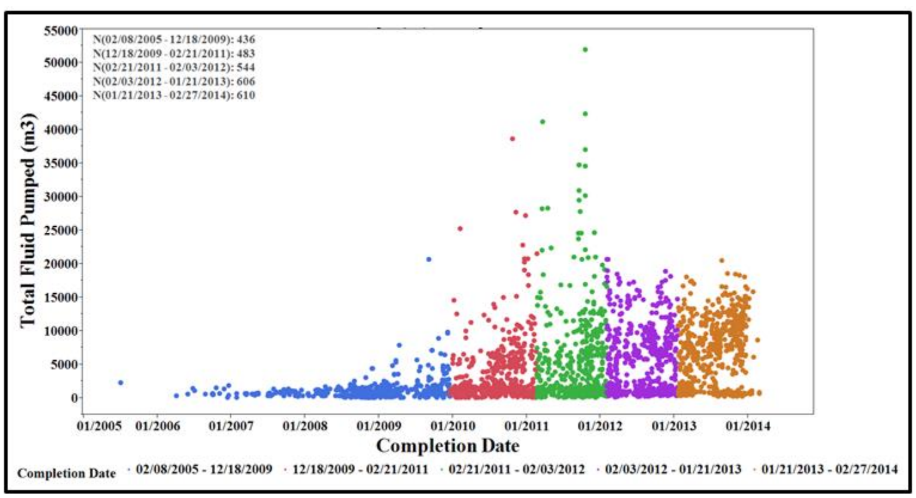

Figure 4.57 Total Pumped Fluid versus Time (M^3)

4.2.5. Type of Completion. Both completion types are used in the same ratio

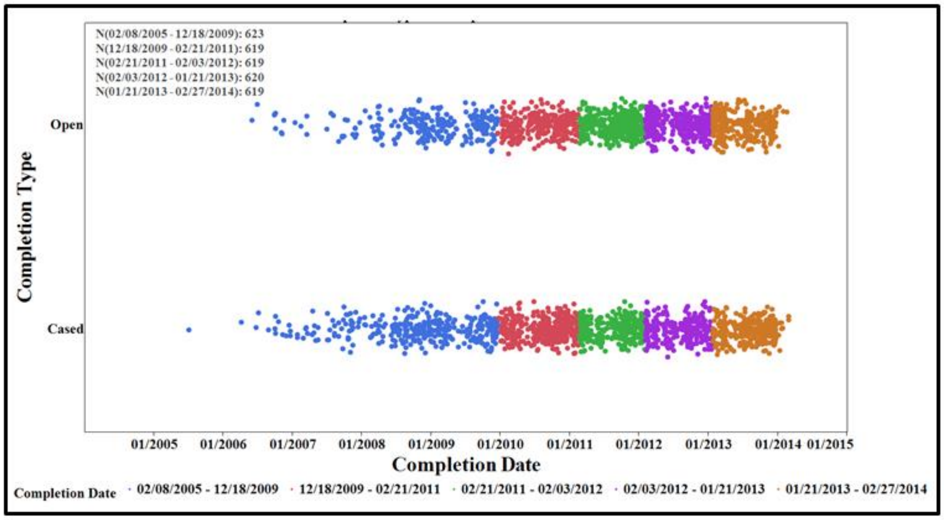

Figure 4.58 Completion Type versus Time

4.2.6. Proppant Concentration (lb/gal). The proppant concentration decreased over time.

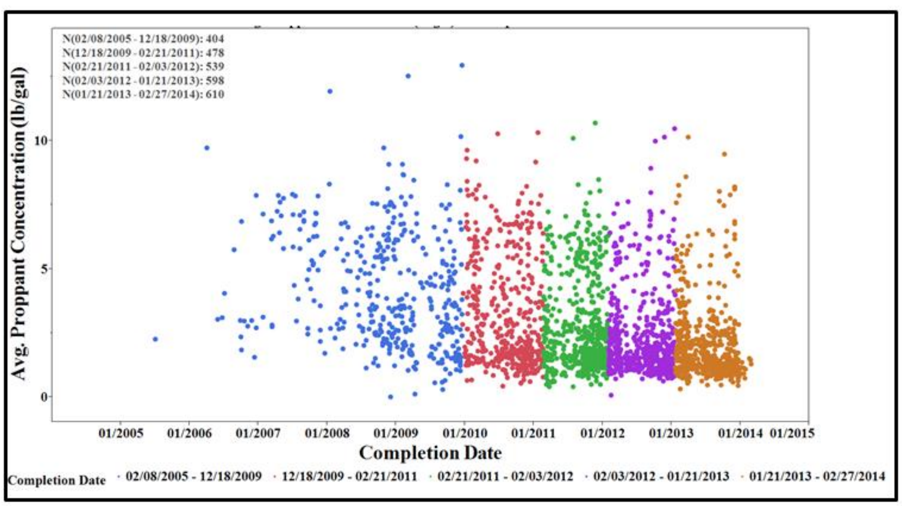

Figure 4.59 Proppant Concentration versus Time

4.2.7. 18 Months of Cumulative Gas Production (mmcf). Over time the gas production from the wells was improved.

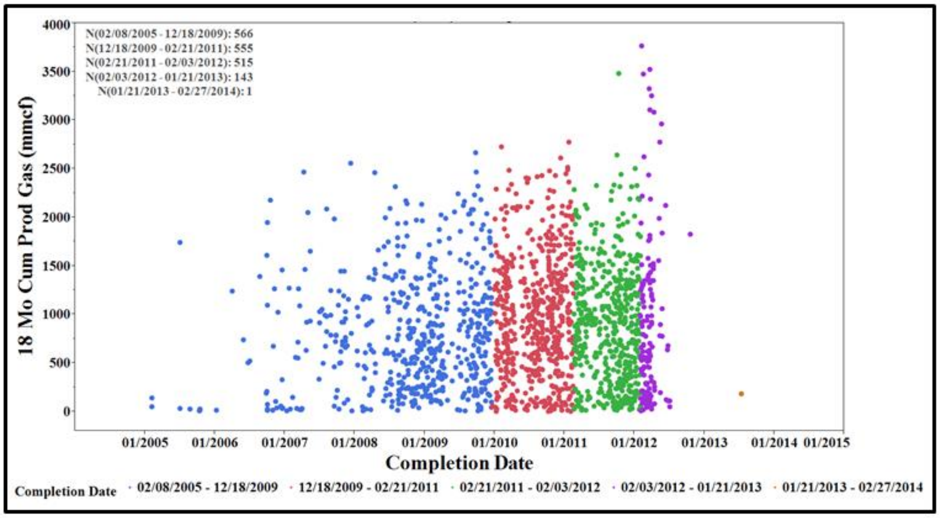

Figure 4.60 18 Months Cum. Gas versus. Time

4.2.8. IP Gas (mcf/D). The initial production of gas from the wells improved over time.

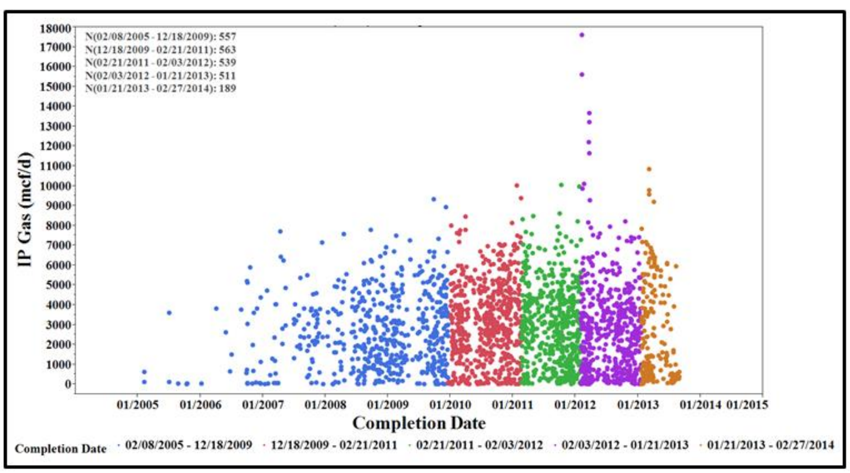

Figure 4.61 IP Gas versus. Time

4.2.9. Drilling and Completion AFE (K$). There was an increase in both the drilling and the completion budget over time.

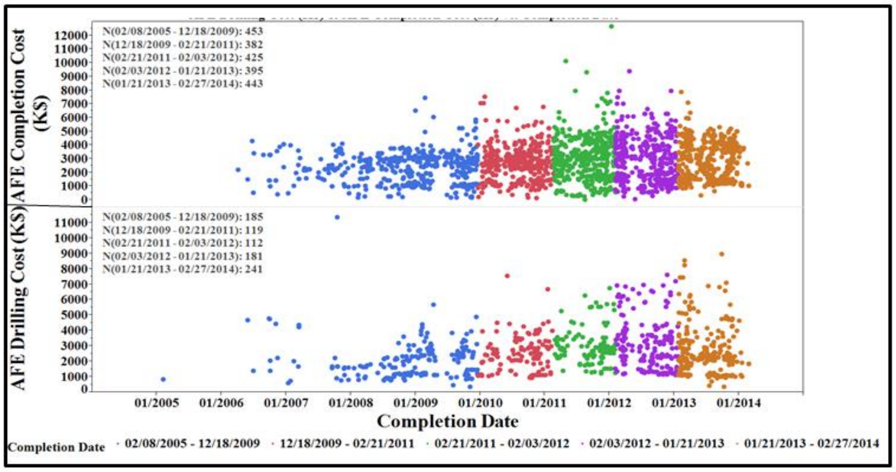

Figure 4.62 Drilling and Completion AFE versus. Time

4.2.10. Final Drilling and Completion Cost (K$). The drilling cost is increased over time.

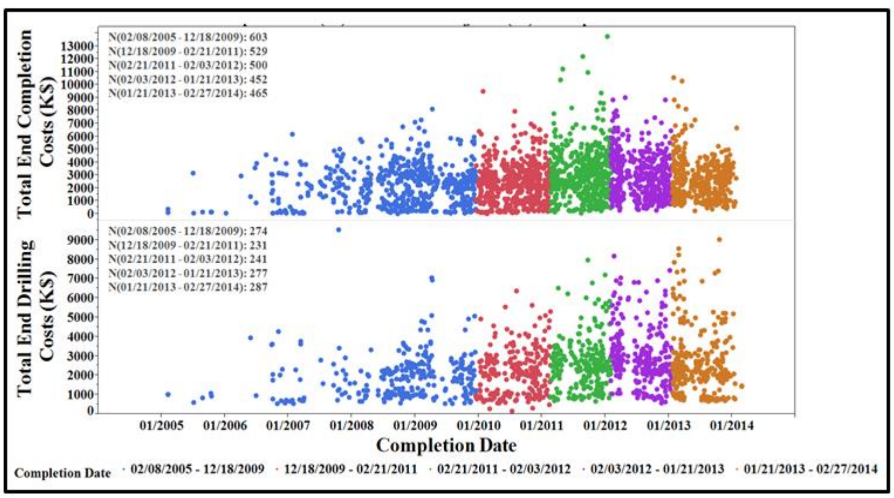

Figure 4.63 Final Drilling and Completion Cost versus. Time

4.2.11. Fracture Spacing (m). The fracture spacing decreased over the time as more stages were added to the wells.

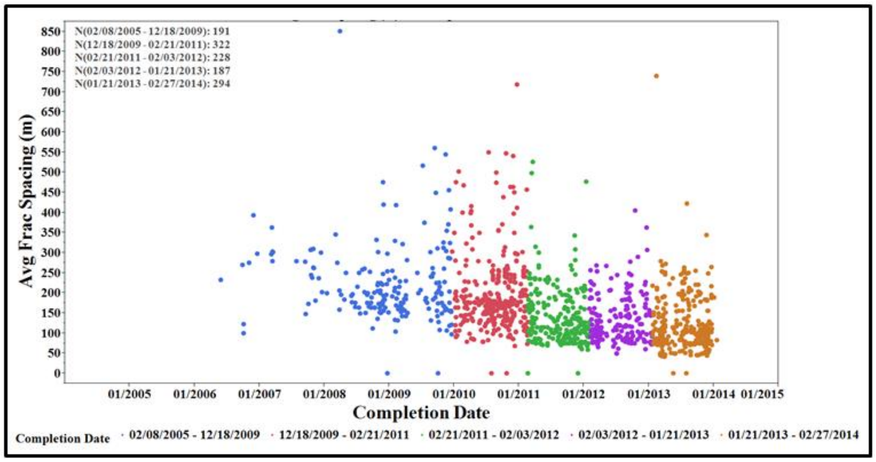

Figure 4.64 Fracture Spacing versus. Time

4.3. COST ANALYSIS

Different companies have taken very different approaches to well design using either plug and perf or ball and sleeve completions with a variety of fracture designs using slickwater, hybrid or cross-linked gel fluids and a variety of proppants from 100% natural sand to 100% ceramics. Consequently, it is not uncommon for different operators to have a difference of over 2 million dollars in their AFE’s solely because of the differences in their approach to the well’s completion and stimulation design. (Griffin et al. 2013)

To check the completion and stimulation cost effects on 18 months of cumulative production, a heat map was generated in Figure 4.65. The plot shows the completion cost effects on the cumulative 18 months of production. The 18 months of production were plotted on both the x and y axes, and the completion cost was set to be the coloring variable. The plot shows that the higher cumulative production was associated with the higher completion cost.

The drilling effects were also tested to check the production response. Figure 4.66 shows a heat map of 18 months of cumulative production on the y and x axes and the total drilling cost as the coloring parameter. The plot shows a good response in the cumulative 18 months of gas production with an increasing drilling cost of the well.

In conclusion, spending more on drilling and completion of the well yield a better cumulative production. Figure 4.67 is a cross plot between the final drilling cost on the y-axis and the final completion cost on the x-axis. The heat map technique was used in this plot to color the wells based on the value of the cumulative 18 months of gas production. The plot shows that the wells drilled and completed with higher costs seem to perform better than the lower cost wells.

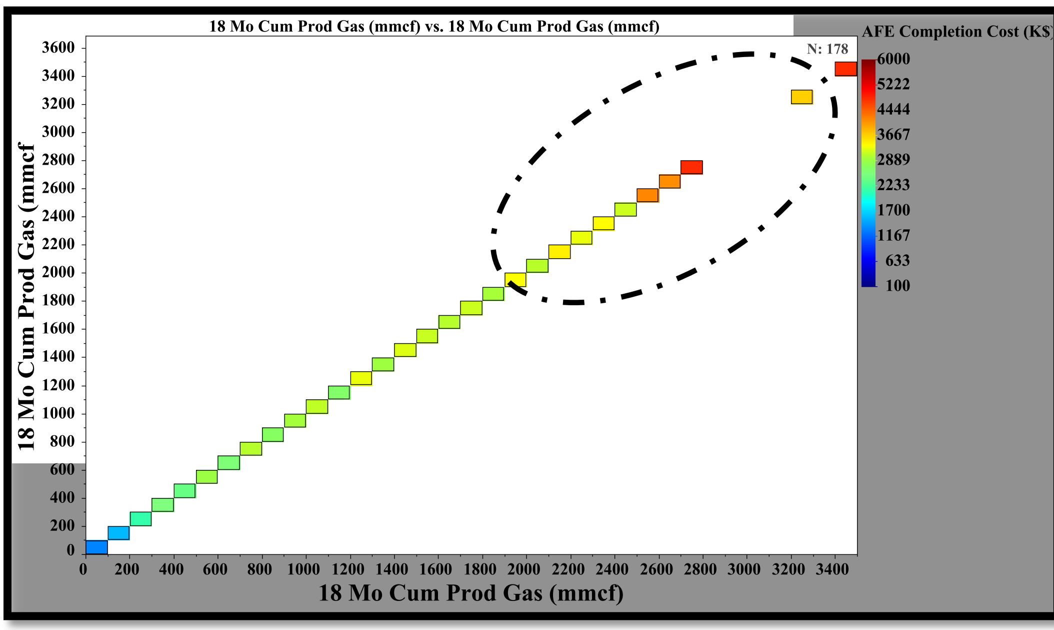

Figure 4.65 Heat Map Shows the Effects of Completion Cost on 18 Months Cum. Gas

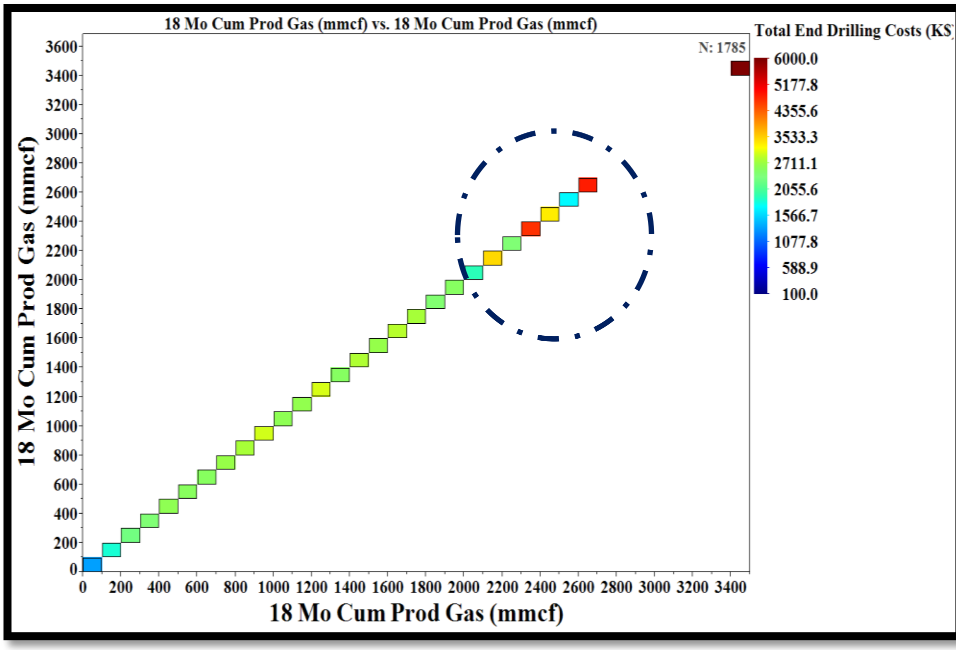

Figure 4.66 Heat Map Shows the Effects of Drilling Cost on 18 Months Cum. Gas

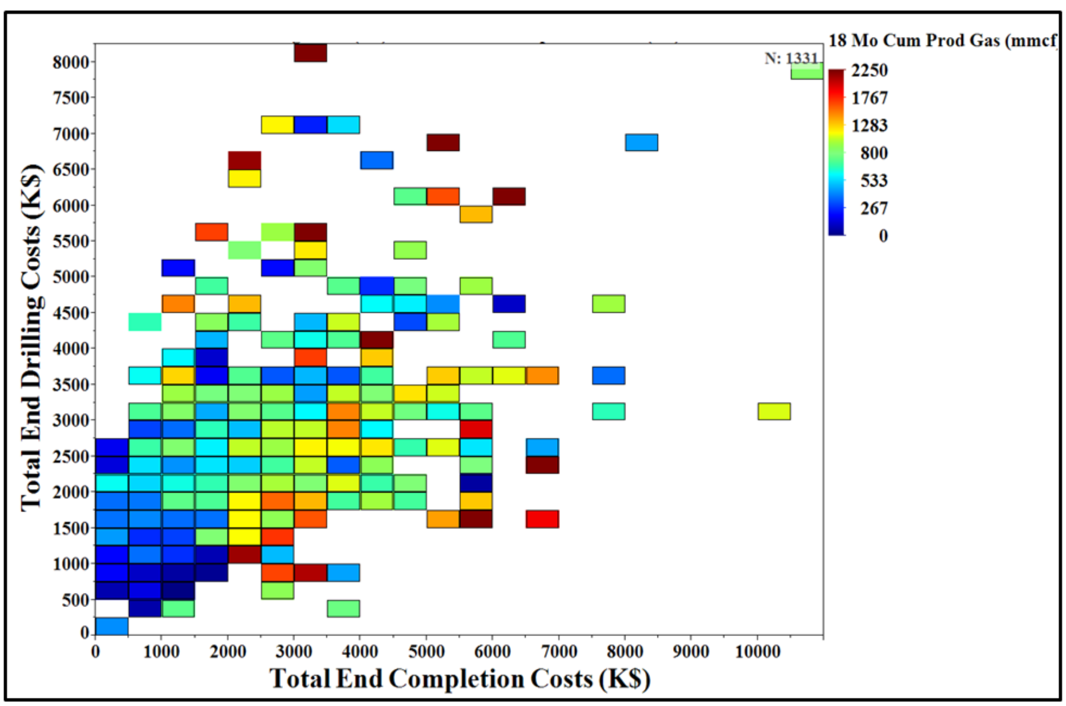

Figure 4.67 Heat Map Shows the Combination of Final Drilling and the Completion Cost Effects on 18 Months of Cum. Gas Production

5. CONCLUSIONS

The main findings and conclusions that are listed in this chapter are presented in two sections. The first section summarizes the main conclusions and observations that were obtained from analyzing the individual completion and stimulation parameters in the dataset. The second section summarizes the findings obtained from the analysis and from cross plotting of the parameters with each other.

5.1. PARAMETERS APPLICATIONS

This section summarizes the main conclusions related to understanding the completion and stimulation parameters along with their general applications in the Montney formation based on the data distribution histograms and data validation:

- Cased hole wells are stimulated with fewer stages than open hole wells.
- The typical range of fracturing fluid percentages in the Montney is 0.1% to 0.35%
- Fracturing with oil-based fluid yields a higher recovery percentage with some wells going up to 100%.
- Cased hole wells are treated with higher fluid volumes and proppant mass than open hole wells.
- Many of the cased hole wells were performed in the high closure stress regions.
- Depending only on the single outlier identification technique might result in an unnecessary elimination of the unique parameters. Combining the statistical methods with an expert opinion and engineering understanding of the parameters to justify the unique value helps in reducing the number of eliminating parameters and increases the trust level of the data quality.
- Based on the completion cost data validation, the operators in the Montney formation faced fewer troubles in performing the completions as planned. This is indicated by a good correlation between the AFE and the final completion costs

- Based on the number of stages' data validations, cased hole wells had a lower failure rate in implicating new stages than the open hole.
- Based on proppant placed data validation, in both open and cased hole wells, there were no difficulties in pumping the frac stage as all of the designed proppant were placed in the wells regardless of the lateral length.

5.2. FORMATION WIDE PERFORMANCE

This section lists the main findings related to the production performance and the main differences between the completion types based on statistical analysis and cross plots techniques:

- The average IP of Gas and water in the cased hole wells showed greater production than open hole wells.
- The average IP of oil in the open holes showed a greater production than cased hole wells.
- Open hole completion produced more water than the cased hole completion.
- The cumulative production of oil over 6, 12 and 18 months in the open hole completion is greater than the cased hole wells.
- Cased hole wells performed much better than open hole wells in terms of production per stimulated stage.
- The average completion cost of the cased hole wells showed a 50% increase not found in the open hole completion.
- The average drilling cost of the open hole wells is greater than the average drilling cost of the case hole wells by a factor of 25-30%.
- The treatment fluid per unit length in cased hole wells is more than double the fluid volume per unit length in open hole wells.
- In both oil and water based treatment fluid the cumulative gas production decreases as a result of a high percentage of treatment fluid recovered.
- There is a strong correlation between the IP of gas, oil and water. The cumulative production over time and wells with the highest IP's will be the best future producers.

- The actual production per stage decreases as the number of treated stages increases.
- Over time, from 2005 to 2014 a greater amount of stages, lateral length, proppant placed and fluid pumped were employed in the Montney formation every year.
- Spending more on the drilling and completion of the wells yields a greater cumulative production.

6. FUTURE WORK

Adding the geographical information represented by the longitude and latitude of each well will help to refine and classify the data more accurately by using the geographical information system (GIS) to identify the production's sweet spots across the Montney formation.

Geographically grouping the wells will remove some of the reservoir quality effects such as thermal maturity, layer thickness and pressure. This may possibly lead to more homogeneous groups of wells when attempting to define which parameters should be changed in order to increase well productivity or reduce the overall cost.

Having the well coordinates associated with the production layer thickness, or the true vertical depth (TVD), of the mid perforations can help in preparing a contour map of the thickness or lateral TVD for all of the wells in the Montney, as well as superimposing a bubble chart of the total cumulative production of the wells over the Montney's generated contour map. Applying these techniques, which are readily available within the JMP software package, will facilitate easier detection of trends and correlations of well productivity in response to different reservoir and design parameters. Simply plotting the top 10% of the producing wells in the Montney on the map will enable future investors and operators to identify the best locations in the area. Further comparisons and classifications can be applied to the well parameters in these particular areas to identify the best practices for future implementations.

Further details on proppant mesh size, type and concentration will also be good comparing factors that can be integrated for future studies.

APPENDIX

Table A.1. Units Conversion Factors Between the Canadian Units Used in the Dataset and The US Units

Unit Conversion		
Unit of Measure	**Reported Unit in the Canadian Dataset**	**Equivelent United States Field Units**
Length, Distance	1 meter (m)	3.28084 feet (ft)
Mass, Weight	1 metric ton (tonne)	2204.62262 pounds (lbs)
Volume	1 cubic meter (m3)	264.172 US Gallons (Gal)
Volume	1 cubic meter (m3)	6.2898 US bbl oil
Volume	1 cubic meter (m3)	35.3147 Cubic Foot (ft3)
Pressure	1 kilopascals (KPA)	0.145037738 psi
Pressure Gradient	1 kilopascal / meter (Kpa/m)	0.0442075025 psi / foot
Mass / Length	1 tonne / meter	671.968975 pounds / foot
Volume /Length	1 (cubic meter) / meter	80.5196416 gal / ft
Concentration	1 tonne / cubic meter	8.34540445 pounds / US gallon
Currency	1 $ Canadian	0.88 US

BIBLIOGRAPHY

Andrew MacMillan et al., 2006. Basic statistics. Available at: https://controls.engin.umich.edu/wiki/index.php/Basic_statistics:_mean,_median,_average,_standard_deviation,_z-scores,_and_p-value#Introduction. (Accessed on 07/21/2014)

Daniel T. Larose, 2005. *DISCOVERING KNOWLEDGE IN DATA An Introduction to Data Mining*

Devlin, T.F., 2009. *Moore, McCabe, and Craig's Introduction to the Practice of Statistics* Sixth Edit., Available at: http://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=8&cad=rja&uact=8&ved=0CFkQFjAH&url=http://people.uncw.edu/frierson/S215%5CStudent_JMP.pdf&ei=amI_VInSFY2dygTfkYHYCw&usg=AFQjCNHaMfbcSD_sx_erkdvhl0Xd013RpQ&sig2=j76AWJin-DokRbmRuQ0W4Q. (Accessed on 08/11/2014)

Economides, M.J., 2000. Reservoir Stimulation. In K. G. N. Michael J. Economides, ed. *Reservoir Stimulation*. J. Wiley, 2000.

Griffin, L.G. et al., 2013. The Value Proposition for Applying Advanced Completion and Stimulation Designs to the Bakken Central Basin.

Hand, D., Mannila, H. & Smyth, P., 2001. Principles of Data Mining Cambridge. *MIT Press*.

"Intelligent Solutions Inc." 2011. *OILFIELD DATA MINING*, Available at: http://www.intelligentsolutionsinc.com/PDFs/OilFieldDataMining.pdf (Accessed on 6/15/ 2014).

Kirkman, T.W., 1996. Statistics to Use. Available at: http://www.physics.csbsju.edu/stats/box2.html (Accessed on 10/18/2014).

Lafollette, R., Holcomb, W.D. & Aragon, J., 2012a. Impact of Completion System, Staging, and Hydraulic Fracturing Trends in the Bakken Formation of the Eastern Williston Basin. In Society of Petroleum Engineers, p. 12.

Lafollette, R., Holcomb, W.D. & Aragon, J., 2012b. Practical Data Mining: Analysis of Barnett Shale Production Results With Emphasis on Well Completion and Fracture Stimulation.

Lafollette, R., Holcomb, W.D. & Aragon, J., 2011. Practical Data Mining: Analysis of Barnett Shale Production Results With Emphasis on Well Completion and Fracture Stimulation. *SPE*, (152531).

Modeland, N., Buller, D. & Chong, K.K., 2011. Statistical Analysis of Completion Methodology.

National Energy Board, British Columbia Oil & Gas Commission, Alberta Energy Regulator, B.C.M. of N.G.D., 2013. Energy Briefing Note, The Ultimate Potential for Unconventional Petroleum from the Montney Formation of British Columbia and Alberta. Available at: https://www.neb-one.gc.ca/clf-nsi/rnrgynfmtn/nrgyrprt/ntrlgs/ltmtptntlmntnyfrmtn2013/ltmtptntlmntnyfrmtn2013-eng.pdf (Accessed on 09/13/2014)

Okouma Mangha, V. et al., 2012. Playwide Well Performance Analysis in Montney Siltstone. In Society of Petroleum Engineers, This paper was prepared for presentation at the SPE Canadian Unconventional Resources Conference held in Calgary, Alberta, Canada, 30 October–1 November 2012.

Petroleum Technology Alliance Canada & Science and Community Environmental Knowledge Fund, 2012. *The Modern Practices of Hydraulic Fracturing: A Focus on Canadian Resources*, Available at: http://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=2&cad=rja&uact=8&ved=0CCYQFjAB&url=http%3A%2F%2Fwww.capp.ca%2Fgetdoc.aspx%3FDocId%3D210903&ei=zsVWVOu7D4ecyQS2lYKoDg&usg=AFQjCNHN5pNy7FHJqm_xKyb6FNUGBBk6kA&sig2=MjhscGmgANYZWPZBVERSUScQ2A&bvm=bv.78677474,d.aWw (Accessed on 05/19/2014).

Roth, M. & Roth, M., 2013. An Analytic Approach to Optimizing Well Spacing and Completions in the Bakken/Three Forks Plays.

SAS Institute Inc., 2014. JMP® Pro. Available at: http://www.jmp.com/software/pro/.(Accessed on 09/10/2014)

The Gartner Group, 2013. The Gartner Group. Available at: www.gartner.com. (Accessed on 09/19/2014)

VISAGE & Jim Gouveia, 2014. Frac Analysis In VISAGE: How Adding More Stages Impacts Gas Production. Available at: http://www.visageinfo.com/2014/04/28/frac-analysis-in-visage-how-adding-more-stages-impacts-gas-production/ (Accessed 10/02/2014).

Wilson, B. et al., 2011. Comparative Study of Multistage Cemented Liner and Openhole System Completion Technologies in the Montney Resource Play.

VITA

Mustafa Adil Al-Alwani was born in September, 1987, in Baghdad, Iraq. He received his bachelor degree in Petroleum Engineering from Baghdad University, Baghdad, Iraq in 2010, he was the valedictorian of Petroleum Major of class 2006-2010 with an 87.3 % cumulative average, that is equivalent to a 4.0 GPA. After graduation, Mustafa joined Honeywell as an automation and control system engineer participating in junior engineers training program. In May 2011 he joined Rumaila Operating Organisation, a joint venture oil and gas operators consist of BP, CNPC and SOC, he worked as a drilling night companyman then promoted to be a day well site supervisor representing the operating company (BP) in the drilling well sites. In January 2013, he was awarded the Fulbright Scholarship Award to study Master's degree in Petroleum Engineering funded by the U.S. Department of States for the cycle of (2012-2014). He started at Missouri University of Science and Technology during the spring semester of 2013 to work under the supervision of Dr. Shari Dunn-Norman as a graduate research assistant and later to be the teaching assistant for the Artificial Lift Class of Fall 2014. His thesis topic is about applying statistics and data mining techniques to study the unconventional resources hydraulic fracturing stimulation. He received a Master of Science degree in Petroleum Engineering from Missouri University of Science and Technology in December 2014 with 4.0 GPA.