University of Calgary

PRISM Repository **https://prism.ucalgary.ca**

The Vault Open Theses and Dissertations

2018-03-19

Reservoir Characterization Using Well Logs and Digitized Core Images: A Case Study from the Montney Formation

Sun, Yu

Sun, Y. (2018). Reservoir Characterization Using Well Logs and Digitized Core Images: A Case Study from the Montney Formation (Master's thesis, University of Calgary, Calgary, Canada). Retrieved from https://prism.ucalgary.ca. doi:10.11575/PRISM/31753

http://hdl.handle.net/1880/106461

Downloaded from PRISM Repository, University of Calgary

UNIVERSITY OF CALGARY

Reservoir Characterization Using Well Logs and Digitized Core Images: A Case Study from the Montney Formation

by

Yu Sun

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE

DEGREE OF MASTER OF SCIENCE

GRADUATE PROGRAM IN CHEMICAL AND PETROLEUM ENGINEERING

CALGARY, ALBERTA

MARCH, 2018

© Yu Sun 2018

Abstract

Tight reservoirs typically have low permeability (<0.1 mD) and small scale (mms-cm) heterogeneity, which are difficult to characterize using conventional methods. The aim of this study is to develop a reliable permeability predictor and to characterize the small scale heterogeneity based on a limited petrophysical dataset from the finely laminated tight gas formation, the Montney Formation.

Three approaches, multiple variable regression, empirical model and artificial neural networks (ANNs), have been used to predict permeability based on well logs. The ANN developed from the workflow produces predictions for permeability with $R^2 = 0.99$ and mean square errors equal to 0.0069, which is more accurate than the other two methods.

A digital image analysis program is developed in Matlab, which can correct the effects of fractures, converts the image into gray value data and identifies the locations and thicknesses of laminations. By analyzing the digitized image data, lamination scale heterogeneity can be quantitatively characterized. By incorporating the gray values into the ANN, the network's ability to generalize is much improved, increasing the test set R^2 value from 0.26 to 0.89.

Acknowledgements

My great gratitude goes to my supervisor, Dr. Jerry Lee Jensen, who always guides and supports me with extreme patience. Without him, I could not even have an opportunity to study as a graduate student, let alone finish this work. I would like to thank Dr. Christopher Clarkson and Dr. Per Kent Pedersen for their valuable suggestions.

I would like to thank my colleagues, Mammad Mirzayev and Kushagra Kakar, for their help. I also thank Amin Ghanizadeh for sharing his data with me.

My thanks also go to my examination committee members, Dr. Jerry Lee Jensen, Dr. Christopher Clarkson and Dr. Hossein Hejazi, for reviewing my thesis.

I am very grateful to my friends including Liu Pei, Zhan Jie, Gao Tian, Yu Yanlong, VANCA and Shi Longyang for their support.

Funding for this project was provided by the Tight Oil Consortium (TOC). This support is gratefully acknowledged.

Finally, I would like to thank my parents for bringing me to this world.

Dedication

To my poor younger cousin. May you rest in peace.

Table of Contents

Abstract i
Acknowledgements ii
Dedication iii
Table of Contents iv
List of Tables vi
List of Figures viii

CHAPTER 1: Introduction 1
1.1 Background 1
1.2 Geological Context 2
1.3 Literature Review 5
1.3.1 The Montney Formation Characterization 5
1.3.2 Digital Image Analysis in Reservoir Characterization 7
1.3.3 Conventional Log-based Permeability Prediction 8
1.3.4 ANN-based Permeability Prediction 10
1.4 Methods Overview 11
1.5 Summary 12

CHAPTER 2: Data Description and Preparation 13
2.1 Data Available 13
2.2 Petrophysical Analysis 16
2.2.1 Core and Well Log Depth Matching 16
2.2.2 Data Screening 23
2.2.3 Core Petrophysics Analysis 28
2.3 Conclusions 36

CHAPTER 3: Reservoir Heterogeneity Characterization Using Digital Image Analysis 38
3.1 Introduction 38
3.2 Digitization Process 38
3.2.1 Image Preprocessing 38
3.2.2 Program Testing 41
3.2.2.1 Test 1 42
3.2.2.2 Test 2 45
3.2.2.3 Test 3 47
3.2.2.4 Test 4 49
3.2.2.5 Test 5 51
3.3 Heterogeneity Characterization 53
3.3.1 Comparison between Gray Value and Probe Permeability 53
3.3.2 Comparison of Gray Value and Mechanical Hardness 69
3.4 Discussion and Conclusions 73

CHAPTER 4: Permeability Prediction Using Artificial Neural Networks 75
4.1 Introduction.....75
4.2 Multiple Variable Regression Method.....76
4.3 Empirical Model Method.....79
4.4 Artificial Neural Network Method.....82
4.4.1 Introduction and Workflow.....82
4.4.2 Determining the Optimum Number of Hidden Neurons.....86
4.4.3 Determining the Optimum Number of Validation Checks.....88
4.4.4 Determining the Optimum Dataset Division.....91
4.4.5 Determining the Optimum Weights and Biases 92
4.4.6 Incorporating the Digitized Image Data into the ANN.....96
4.5 Investigation of Main Controls on Permeability.....102
4.6 Discussion and Conclusions 117

CHAPTER 5: Conclusions and Future Work.....119
5.1 Conclusions.....119
5.2 Future Work.....121

REFFERENCES 124

APPENDIX.....131

List of Tables

Table 1.1 Common sources for UGR reservoir, fluid and rock properties. Modified from Sondergeld et al. (2010a). 6

Table 1.2 Permeability determination empirical models derived from well log measurements (Modified from Di, 2015). 9

Table 2.1 Depth matching summary (- means shifting upwards; + means shifting downwards). 22

Table 3.1 Test 1 components. 43

Table 3.2 Test 2 components. 45

Table 3.3 Test components. 47

Table 3.4 Test 4 components 50

Table 4.1 Divide functions. 84

Table 4.2 Parameter setting. 86

Table 4.3 The recommended number of training times 87

Table 4.4 Parameter setting. 89

Table 4.5 The recommended number of training times 90

Table 4.6 Parameter setting. 92

Table 4.7 Parameter setting. 93

Table 4.8 Results summary. 101

Table 4.9 Parameter setting. 102

Table 4.10 Seven-factor ANOVA model obtained from MINITAB. 106

Table 4.11 Parameter setting. 108

Table 4.12 Six-factor ANOVA model obtained from MINITAB. 110

Table 4.13 Parameter settings. 112

Table 4.14 Five-factor ANOVA model obtained from MINITAB. 114

Table 4.15 The training algorthims..... 118

List of Figures

Figure 1.1 The study area and well location for the Montney Formation located in southwest Alberta, Canada (Modified from Ghanizadeh et al., 2015a)..... 3

Figure 1.2 Regional Triassic stratigraphic framework in the WCSB (From Davies, 1997)..... 4

Figure 1.3 Schematic representation of the proposed lithostratigraphic units in the Montney (From Dixon, 2000). The lithology changes are consistent with the decreasing depositional energy from east to west..... 4

Figure 2.1 The study area for the Montney Formation located in southwest Alberta, Canada (Modified from Ghanizadeh et al., 2015a)..... 15

Figure 2.2 An example of missing core. Figure 2.3 An example of unslabbed core..... 16

Figure 2.4 Depth shift of Interval 1 based on GR data. 17

Figure 2.5 Depth shift of Interval 2 based on GR data. 18

Figure 2.6 Depth shift of Interval 2 based on density data. 18

Figure 2.7 Depth shift of Interval 3 based on GR data. 19

Figure 2.8 Depth shift of Interval 3 based on porosity data..... 19

Figure 2.9 Depth shift of Interval 4 based on GR data. 20

Figure 2.10 Depth shift of Interval 5 based on GR data. 20

Figure 2.11 Depth shift of Interval 6 based on GR data. 21

Figure 2.12 Depth shift of Interval 7 based on GR data. 21

Figure 2.13 Depth shift of Interval 8 based on GR data. 22

Figure 2.14 Whole core permeability measurements. 24

Figure 2.15 Core plug permeability measurements. Core plugs with fracture issues are shown in red circles. 24

Figure 2.16 Profile permeability measurements. 25

Figure 2.17 Profile permeability measurements. 25

Figure 2.18 Profile permeability measurements. 26

Figure 2.19 Core plugs with fracture issues are shown in red circles. 27

Figure 2.20 The core is divided into 3 parts based on GR response. 28

Figure 2.21 Core slab images from Part 1; very fine grained well sorted clean siltstone and laminated shaly siltstone. 29

Figure 2.22 Core slab images from Part 2; very fine grained well sorted dispersed shaly siltstone. 29

Figure 2.23 Core slab images from Part 3; very fine grained well sorted highly fractured shaly siltstone. 29

Figure 2.24 Core plug permeability vs core plug porosity in the Montney. 31

Figure 2.25 Core plug permeability vs core plug porosity in the upper Montney. 31

Figure 2.26 Core plug permeability vs core plug porosity in the lower Montney. The core plugs with facture issues are identified with red circles. 32

Figure 2.27 Whole core permeability Kmax vs whole core porosity in the Montney. 32

Figure 2.28 Whole core permeability Kmax vs whole core porosity in the upper Montney. 33

Figure 2.29 Whole core permeability Kmax vs whole core porosity in the lower Montney. 33

Figure 2.30 Average permeability of the upper and lower Montney 34

Figure 2.31 Coefficient of Variation of permeability of the upper and lower Montney. 34

Figure 2.32 Number of censored permeability values of the upper and lower Montney. 35

Figure 2.33 Permeability measurements sample sufficiency assessment. No is the recommended number of measurements. N is the actual number of measurements. 35

Figure 2.34 Porosity measurements sample sufficiency assessment. No is the recommended number of measurements. N is the actual number of measurements. 36

Figure 3.1 Target area for illumination normalization; irregular boundaries caused by core plugs and breakages are shown in red boxes. The image is evenly divided into three parts whose boundaries are marked by red lines. 40

Figure 3.2 Core slab image having core plug, breakage and fracture effects. 42

Figure 3.3 Test 1 case. 42

Figure 3.4 Test 1 raw digitized results. 43

Figure 3.5 Test 1 corrected digitized results, showing the fractures have been removed. 44

Figure 3.6 Test 1 corrected digitized results, showing lamination thicknesses and center locations. 44

Figure 3.7 Test 2 case. 45

Figure 3.8 Test 2 raw digitized results. 46

Figure 3.9 Test 2 corrected digitized results showing the fractures have been remove. 46

Figure 3.10 10 Test 3 case. 47

Figure 3.11 Test 3 raw digitized results. 48

Figure 3.12 Test 3 corrected digitized results showing the fractures have been removed. 48

Figure 3.13 Test 3 corrected digitized results, showing lamination thicknesses and center locations. 49

Figure 3.14 Test 4 case. 49

Figure 3.15 Test 4 lamination thickness profile, showing lamination thicknesses and center locations. 50

Figure 3.16 Test 4 lamination thickness comparisons. 51

Figure 3.17 Test 5 case. 51

Figure 3.18 Test 5 raw digitized results. 52

Figure 3.19 Test 5 corrected digitized results showing the fractures have been removed. 52

Figure 3.20 Test 5 lamination thickness profile, showing lamination thicknesses and center locations. 53

Figure 3.21 Corrected gray value plot of 3040.6 to 3043.7 m interval. Two of the three intervals have been oversampled. 55

Figure 3.22 Lamination thickness profile of 3040.6 to 3043.7 m interval. 55

Figure 3.23 Lamination density and measurement sample density comparison. 56

Figure 3.24 Semivariogram of gray value data in the interval from 3040.6 to 3043.7 m. 56

Figure 3.25 Semivariogram of probe permeability data in the interval from 3040.6 to 3043.7 m. 57

Figure 3.26 Semivariogram of Log(Probe K) data in the interval from 3040.6 to 3043.7 m. 57

Figure 3.27 Semivariogram of de-trended probe permeability data in the interval from 3040.6 to 3043.7 m. 58

Figure 3.28 Core slab image from 3040.6 to 3043.7 m with 0.25 m cycles identified in red circles. 58

Figure 3.29 Semivariogram of gray value data in Interval 2. 58

Figure 3.30 Core slab image from 3080.7 to 3084.0 m. 59

Figure 3.31 Corrected gray value plot of 3080.7 to 3084.0 m interval..... 59

Figure 3.32 Semivariogram of gray value data in the interval from 3080.7 to 3084.0 m. 60

Figure 3.33 Semivariogram of probe permeability data in the interval from 3080.7 to 3084.0 m. 60

Figure 3.34 Semivariogram of Log(Probe K) in the interval from 3080.7 to 3084.0 m..... 61

Figure 3.35 Corrected gray value plot of 3135.2 to 3139.1 m interval..... 62

Figure 3.36 Core slab image from 3135.2 to 3139.1 m with 1 to 3 cm laminations. 62

Figure 3.37 Lamination thickness profile of 3135.2 to 3139.1 m interval. 62

Figure 3.38 Semivariogram of gray value data in the interval from 3135.2 to 3139.1 m. 63

Figure 3.39 Semivariogram of probe permeability data in the interval from 3135.2 to 3139.1 m. 63

Figure 3.40 Corrected gray value plot of 3140.15 to 3141.3 m interval..... 64

Figure 3.41 Core slab image from 3140.15 to 3141.3 m. 64

Figure 3.42 Lamination thickness profile of 3140.15 to 3141.3 m interval. 65

Figure 3.43 Semivariogram of gray value data in the interval from 3140.15 to 3141.3 m. 65

Figure 3.44 Semivariogram of probe permeability data in the interval from 3140.15 to 3141.3 m. 66

Figure 3.45 Corrected gray value plot of 3143.3 to 3145.9 m interval..... 67

Figure 3.46 Core slab image from 3143.3 to 3145.9 m. 67

Figure 3.47 Lamination thickness profile of 3143.3 to 3145.9 m interval. 67

Figure 3.48 Relationship of lamination number and lamination thickness. 68

Figure 3.49 Lamination density and sample density comparison. 68

Figure 3.50 No and N comparison. 69

Figure 3.51 Comparison of gray value, hardness and core plug of 3040.6 to 3043.8m interval.. 70

Figure 3.52 Gray value and hardness values compare favorably for this lower Montney interval, differing by about 2.5%. 71

Figure 3.53 Correlation between hardness and gray value data. 72

Figure 3.54 Hardness prediction results..... 72

Figure 3.55 Hardness prediction results..... 73

Figure 4.1 Multiple variables regression results. Figure 4.2 Multiple variables regression results. 77

Figure 4.3 2-interval model results. Figure 4.4 2-interval model results..... 78

Figure 4.5 Residuals compared to depth show an increase in variability with depth. 78

Figure 4.6 Residuals compared to Log(K) show a trend which the model does not include. 79

Figure 4.7 Empirical model results. Figure 4.8 Empirical model results. 80

Figure 4.9 2-interval model results. Figure 4.10 2-interval model results..... 81

Figure 4.11 Residuals compared to depth show an increase in variability with depth. 81

Figure 4.12 Residuals compared to Log(K) show a trend which the model does not include. 82

Figure 4.13 Neural network system. 83

Figure 4.14 Example of ANN architecture with one hidden layer. 84

Figure 4.15 Comparison of the averaged R^2 values from N groups of responses. 88

Figure 4.16 Comparison of the averaged R^2 values from N groups of responses. 91

Figure 4.17 ANN results. Figure 4.18 Test dataset results. 94

Figure 4.19 ANN results. 94

Figure 4.20 R^2 value of ANN, multiple variable regression and empirical model. 95

Figure 4.21 MSE of ANN, multiple variable regression and empirical model. 95

Figure 4.22 New ANN results. Figure 4.23 New ANN test results. 97

Figure 4.24 GR log data. 98

Figure 4.25 Comparison of Intervals 1 and 2. 98

Figure 4.26 Comparison of Intervals 2 and 3. 99

Figure 4.27 3-interval ANN results. Figure 4.28 3-interval ANN test dataset results. 99

Figure 4.29 3-interval ANN results. 100

Figure 4.30 2-interval ANN results. Figure 4.31 2-interval ANN test dataset results. 100

Figure 4.32 2-interval ANN results. 101

Figure 4.33 Normal probability plot obtained from MINITAB of the estimated effects. 104

Figure 4.34 Main effects plot obtained from MINITAB for the 7-factor case. 105

Figure 4.35 ANN results without RHOB. Figure 4.36 Original ANN results. 107

Figure 4.37 Normal probability plot obtained from MINITAB of the estimated effects. 109

Figure 4.38 Main effects plot obtained from MINITAB for 6 factors. 109

Figure 4.39 ANN results without RHOB and R. Figure 4.40 Original ANN results. 111

Figure 4.41 Normal probability plot obtained from MINITAB of the estimated effects. 113

Figure 4.42 Correlation between plug porosity and plug permeability in the upper Montney... 115

Figure 4.43 Correlation between plug porosity and plug permeability in the lower Montney... 115

Figure 4.44 Comparison between plug permeability, plug porosity and NPHS. 116

Figure 4.45 Correlation between NPHS and GR. 116

Figure 5.1 Depth shift example of Interval 1 122

CHAPTER 1: Introduction

1.1 Background

Unconventional resources such as shale gas, coal-bed methane, and tight gas constitute a significant percentage of the nature gas resource base and offer a significant potential for future reserve growth and production. However, complicated geological heterogeneities over different scales and very small permeability still remain challenges for an effective and detailed characterization of a tight gas reservoir, which is a key part in optimizing recovery.

The typical petrophysical tools, including core and wire line measurements, cannot fully characterize the heterogeneity at the lamination (mm) and laminaset (mms - cm) scales because of the limit of their investigation volume and sampling resolution. Meanwhile, the low permeability (<0.1 mD) and the weak permeability-porosity relationships of unconventional reservoirs also make the prediction of permeability difficult. These factors suggest that substantial amounts of core are required for permeability assessment. However, if permeability can be predicted accurately only from well log data and such predictors are general enough to be applied to other wells under similar conditions, the cost of core collecting and processing might be reduced. The major controls on permeability could also be determined from the predictor, which can inform well log measurement selection.

This study focuses on the characterization of a complex, finely laminated tight gas formation, the Montney Formation. It uses petrophysical and data analysis methods for the case where the reservoir sample data are limited, and includes two main objectives:

1. Characterize and analyze the degrees of heterogeneity, especially the lamination (mm) and laminaset (mms - cm) scale heterogeneity, of the Montney Formation.
2. Develop a correspondence between permeability and well log responses to predict the permeability distribution and identify the major controls on permeability.

1.2 Geological Context

The study area for the Montney Formation is located in southwest Alberta within the Western Canada Sedimentary Basin (WCSB) (Fig. 1.1). The Montney is one of the most active plays in Western Canada and it has been the target of hydrocarbon exploration since 1955 (Kendall, 1999).

The entire Triassic succession was interpreted to consist of three transgressive cycles (Gibson and Edwards, 1990). The first one includes the lower Triassic Grayling, Toad and Montney Formations. The second one is comprised of the middle to upper Triassic Liard, Doig and Charlie Lake Formations, and the last one consists of the upper Triassic Baldonnel, Pardonet and Bocock Formations (Fig. 1.2).

The lower Triassic Montney Formation was deposited in an open shelf marine environment along a passive continental margin prior to the development of a foreland basin during the Jurassic-Cretaceous (Rivard et al., 2014; Chalmers and Bustin, 2012; Zonneveld et al., 1997). The Montney unconformably overlies the Permian Belloy Formation and is overlain by the middle Triassic Doig Formation (Rivard et al., 2014; Chalmers and Bustin, 2012). Based on lithostratigraphy, the Montney can be subdivided into three informal members (Davies et al., 1997).

The lower member of the Montney is comprised of coarsening-upward, very fine grained sandstones and siltstone. The top of the lower member is marked by a coquina dolomitic member in the east, or by a sequence boundary below the upper member to the west.

The middle member consists of thick dolomitized coquina. It overlies the lower member in the eastern subsurface of west-central Alberta. The coquina bed was deposited in shallow water, an oxygenated environment dominated by bivalve mollusks (Davies and Sherwin, 1997).

The upper member is generally characterized of coarsening-upward siltstones with very fine grained sandstones. It overlies the middle member in the eastern shelf and a disconformity in the west.

A shale member was proposed in addition to the three members in the westward part because of the poorly defined coquina dolomite member (Dixon, 2000) (Fig. 1.3). The shale member is not present in Alberta, but becomes progressively thicker from the east to the west (Chalmers and Bustin, 2012; Dixon, 2000). This variation of lithology reflects the decrease of depositional environment energy from the east to the west. Thus the Coquinal Dolomite member is not present in the west.

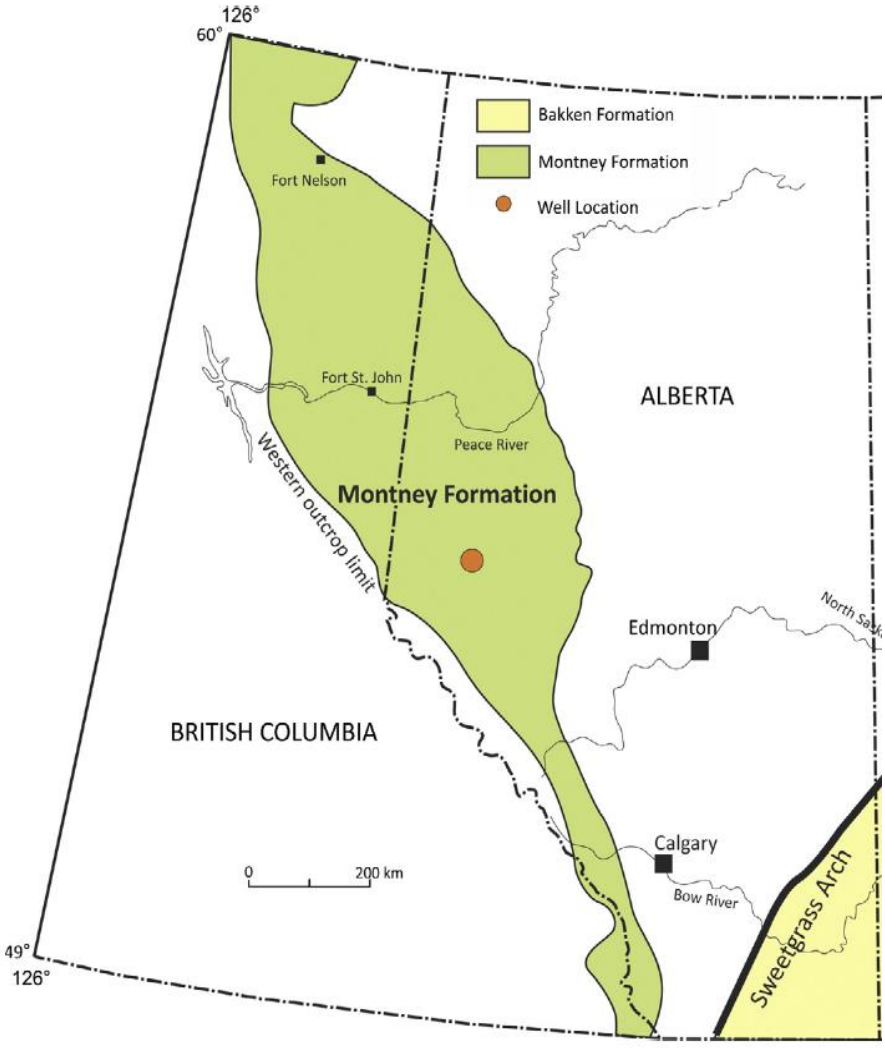

Figure 1.1 The study area and well location for the Montney Formation located in southwest Alberta, Canada (Modified from Ghanizadeh et al., 2015a).

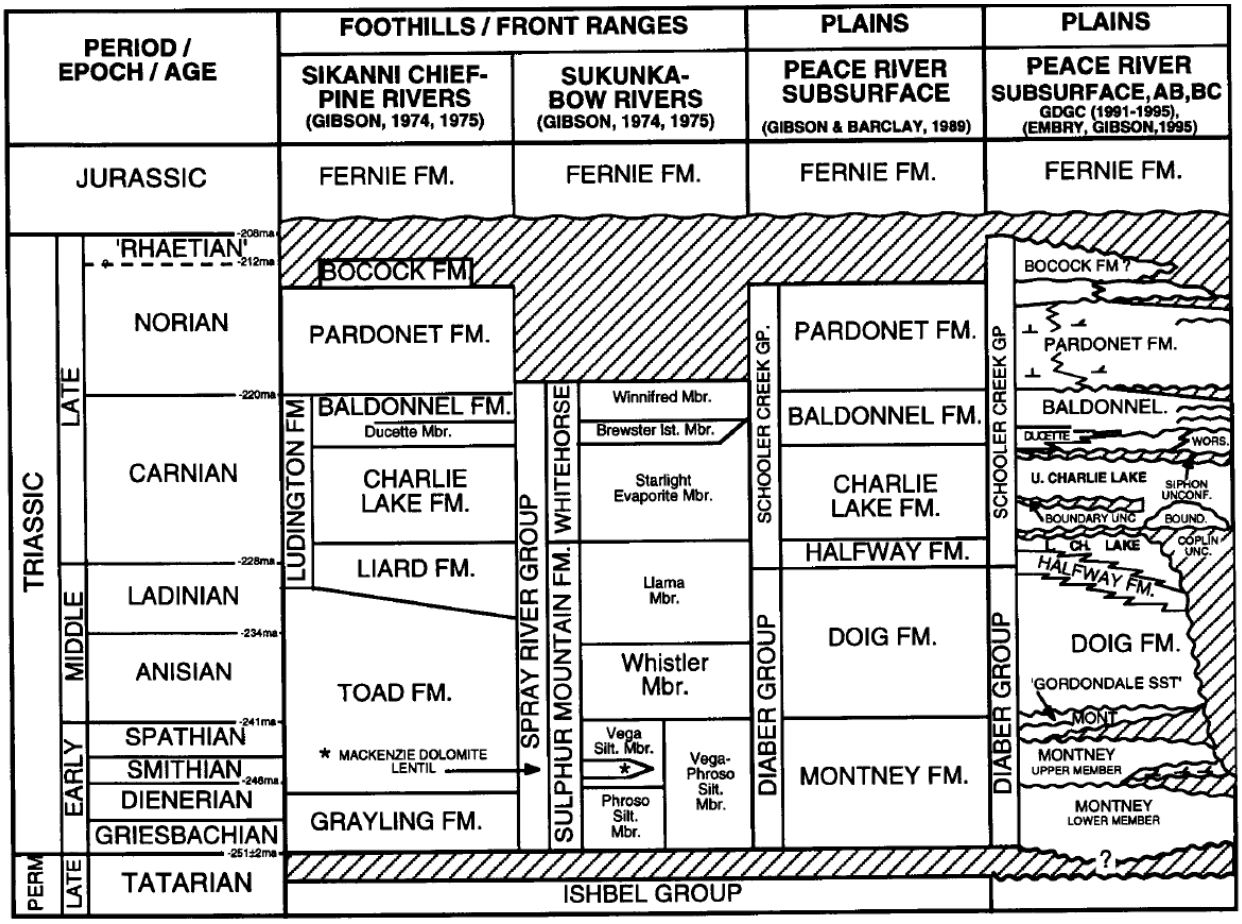

Figure 1.2 Regional Triassic stratigraphic framework in the WCSB (From Davies, 1997).

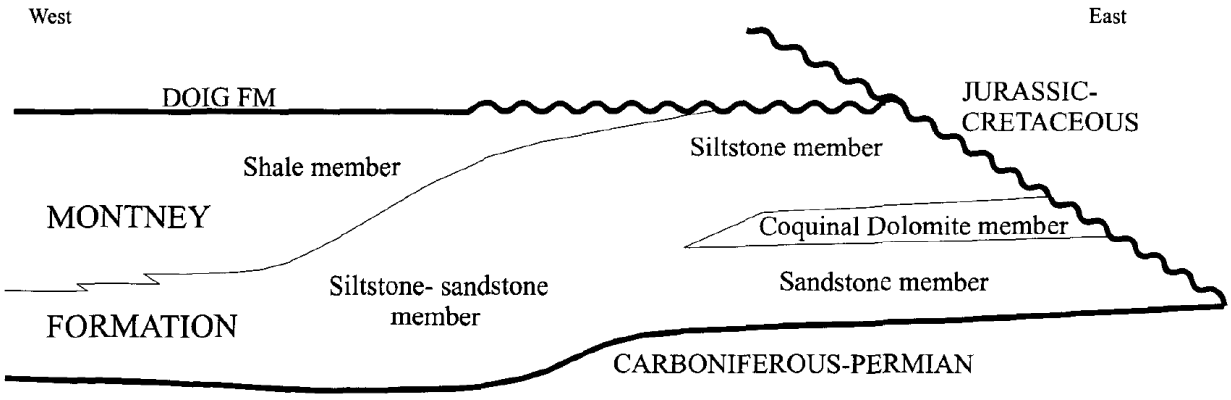

Figure 1.3 Schematic representation of the proposed lithostratigraphic units in the Montney (From Dixon, 2000). The lithology changes are consistent with the decreasing depositional energy from east to west.

1.3 Literature Review

In this study, we focus on reservoir characterization using digital image analysis and permeability prediction, especially using artificial neural networks (ANNs). The literature review consists of four parts:

1. The Montney Formation characterization
2. Digital image analysis in reservoir characterization
3. Conventional log-based permeability prediction
4. ANN-based permeability prediction.

1.3.1 The Montney Formation Characterization

Several studies have reported on petrophysical characterization of Montney cores. Clarkson et al. (2012a) thoroughly reviewed the techniques used for unconventional gas reservoir (UGR) characterization (Table 1.1). Clarkson et al. (2012b) conducted a comprehensive petrophysical analysis on a core from the Montney Formation including petrographic analysis, routine core analysis, profile permeability analysis, pulse-decay permeability analysis and Nitrogen adsorption isotherm analysis. They suggested that routine core analysis performed on full-diameter core is not useful for characterizing the subject tight gas reservoir due to the fine-scale heterogeneities. Derder (2012) characterized the Montney in the Pouce Coupe Field and discussed the degree of heterogeneity using an integrated dataset of core data, well log data and production data. He used the conventional porosity-permeability correlation to estimate the permeability, which is challenging in the highly heterogeneous Montney, so he recommended ANNs as a potential alternative method. Egbobawaye (2013) studied the sedimentology, ichnology, geochemistry and magnetostratigraphy of the Montney Formation in the Fort St. John area. Ghanizadeh et al. (2015a, 2015b) investigated petrophysical characteristics including porosity, pore size distribution, permeability, total organic carbon, grain size and mechanical hardness of the Montney based on a laboratory study. Their studied core samples and our studied

data come from the same well. Sanei et al. (2015) quantitatively analyzed various organic matter fractions within tight gas siltstone core samples from the Montney Formation of northeast British Columbia.

Table 1.1 Common sources for UGR reservoir, fluid and rock properties. Modified from Sondergeld et al. (2010a).

Reservoir Property	Data Source
Porosity	Helium gas expansion, mercury injection capillary pressure (MICP), nuclear magnetic resonance (NMR), log analysis (calibrated to core)
Permeability	Core Analysis: Steady-state and unsteady state (pressure- and pulse-decay), MICP Well-test Analysis: (pre- and post-fracture) Injection/falloff (IFOT), diagnostic fracture injection test (DFIT), post-fracture flow and buildup (FBU) Production Analysis: Rate-transient analysis (RTA), simulation history-matching
Pore Pressure	IFOT, static pressure/temperature gradient, perforation inflow diagnostic (PID), perforation inflow test analysis (PITA), various openhole tests
Water saturation	Core extraction (Dean Stark, Retort), capillary pressure, log analysis (using lab-based electrical property measurements)
Free and sorbed gas	Desorption canister testing & adsorption isotherms, calibrated log analysis
Total Organic Carbon	Leco TOC & RockEval (calculated)
Thermal Maturity	Vitrinite reflectance (Ro), RockEval (calc)

Rock composition	X-Ray Diffraction (XRD), Fourier-Transform Infrared (FTIR), visual point count (for optically-resolvable grains), EDAS (SEM), electron microprobe
Rock mechanical properties	Core measurements, log-derived (multi-component sonic, ex. DSI)
Fracture and closure stress	Mini-frac tests, DFIT, log-based (DSI) calibrated to core
Fluid properties	Mud log, produced gas and water analysis (PVT properties)
Temperature	Openhole logs, static pressure/temperature gradient, production logs

1.3.2 Digital Image Analysis in Reservoir Characterization

In terms of pore structure and geometry characterization, digital image analysis is a well-established method (Ehrlich et al., 1984, 1991a, b; McCreesh et al., 1991; Gerard et al., 1992). These studies used optical or electron microscopy to acquire images of polished thin sections of rock. To separate pore from solid phase, most optical images are converted into binary images by impregnating pore space with blue-dyed (Crabtree et al., 1984; Anselmetti et al., 1993) or fluorescent-dyed (Yanguas and Dravis, 1985; Ruzyla and Jezek, 1987) epoxy. The resulting images are binarized and analyzed by using standard image analysis software which can quantify the porosity evaluation over three orders of magnitude, from submicron to millimeter scales (Anselmetti et al., 1998). The characterization of nanometer-scale pore systems is also possible due to the development of the focused ion beam milling technique, higher magnification field emission scanning electron microscopy, and transmission electron microscopy (Keller et al., 2011; Chalmers et al., 2012).

Meanwhile, X-ray computer tomography (3D X-radiography or CT scanning) has become more commonplace in the petroleum industry and can produce multi-dimensional real time digital

images of core samples (Ketcham and Carlson, 2001; Rivers et al., 2004; Lesher et al., 2009). Medical X-ray CT and industrial CT systems are often used for their large core scanning capabilities (Baraka-Lokmane et al., 2009). For core samples, the surfaces as well as the internal features, including bedding features, cement distribution, small-scale grain size variation and density variation can be analyzed (Coles et al., 1991; Orsi et al., 1994; Coles et al., 1998).

At the core scale, digital images of slabbed whole core can replace traditional photography and offer quantitative analysis. Nielson (1995) discussed the use of digital core images in conjunction with image processing and GIS (geographic information system) software to analyze lithological and structural features. He used an ERDAS sequential clustering algorithm to create 60 classes from the digital images of core slabs and assigned those classes to the corresponding rock and mineral types to conduct the lithological characterization. Van Vliet and Verbeek (1995) focused on orientation and anisotropy measurements based on digital images of core slab. Prince et al. (2005) measured the grain size from the high-resolution (1200 dots per inch) digital image of the core slab, which is close to the grain size measured from thin sections and laser particle size analysis. He also used the measured grain size to estimate permeability, which was also confirmed by the permeability obtained from the NMR log and the core plug measurements. Nederbragt et al. (2006) discussed methods for extracting calibrated colour values from digital images of sediment core surfaces, which can be correlated with geochemical composition.

While digital image analysis is not widely used at the core scale, the previous investigations show that the information derived from core surfaces can be correlated with other petrophysical properties. Quantitative analysis can also improve the accuracy of collection, interpretation and archiving of reservoir data.

1.3.3 Conventional Log-based Permeability Prediction

By comparing the cost of permeability estimates from cores and well logs, Yao and Holditch (1993) found that the costs from core analysis were nearly 10 times greater than well logs. So many attempts have been made to find well log permeability predictors (Table 1.2). These predictors were reviewed by Di (2015).

Most of the listed predictors are applicable for conventional formations. But the more complicated petrophysical and geological characteristics of the Montney call their applicability into question. Di (2015) also noted this. Among these studies, only the geological conditions from Yao and Holditch (1993) are similar to the Montney. Di (2015) also assessed the Yao and Holditch (1993) model for the Cardium Formation and found the model works best in the Cardium A shaly siltstone interval, with $R^2 = 0.94$ but shale and conglomerate show relatively weak correlations between the model and core measurements, with $R^2 < 0.5$. We also test their model for suitability for the Montney Formation (Chap. 4).

Table 1.2 Permeability determination empirical models derived from well log measurements (Modified from Di, 2015).

Authors	Permeability Correlations	Explanatory variables
Tixier, 1949	$K = C\left(\frac{2.3}{R_o(\rho_w - \rho_o)}\frac{\Delta R}{\Delta h}\right)$	R_o:sample resistivity; ρ: density; $\frac{\Delta R}{\Delta h}$: resitivity gradient
Coates and Dumanoir, 1973	$K^{1/2} = \frac{c}{w^4}\left(\frac{\emptyset^{2w}}{R_w/R_t}\right)$	w: tortuosity factor; R_w, R_t: water and true resistivity
Yao and Holditch, 1993	$K = \frac{U\emptyset^{e1}(1 - I_{GR})^{e2}R_{ild}^{e3}}{\left(R_{ild}/R_{sfl}\right)^{e4}}$	I_{GR}: gamma log index; R_{ild}, R_{sfl}: deep, shallow resistivity
Saner et al., 1997	$K = \frac{10^{7.04}}{F^{4.19}}$	F: formation factor
Mohaghegh et al., 1997	$K = 126.5 + 0.001R_t - 50.3\rho_D + 0.06I_{GR}$	R_t: deep resistivity; ρ_D: bulk density; I_{GR}: gamma log index
Xue et al., 1997	$log(K) = 0.15\Delta t - 0.019\emptyset_d - 0.039GR + 0.022RR - 7.73$	Δt: sonic travel time; $\emptyset_d$: density porosity; GR: gamma ray; RR: ratio of deep resistivity to shallow resistivity

1.3.4 ANN-based Permeability Prediction

ANNs have been increasingly applied to predict reservoir properties using well log data (Mohaghegh et al., 1995; Saemi et al., 2007; Irani and Nasimi, 2011; Ali Ahmadi et al., 2013). Previous investigations (Wong et al., 2000; Aminian et al., 2000) have indicated that ANNs can predict formation permeability even in highly heterogeneous reservoirs using well log data with good accuracy.

An ANN architecture consisting of a three layered feed-forward network trained using the back propagation algorithm can map any non-linear relationship with a desired degree of accuracy (Hornik et al., 1989), so this type of neural network is used in this study. Division of data is also an issue for ANNs. ANNs perform best when they do not extrapolate beyond the range of the data used for training (Flood and Kartam 1994; Minns and Hall 1996; Tokar and Johnson 1999). Consequently, a separate validation set is needed to ensure that the model can generalize within the range of the data used for training. A modification of the above data division approach is cross-validation (Stone 1974) in which the data are divided into three sets: training, testing and validation. For example, the dataset may be divided by proportions so that 70% are allocated to the training subset and 15% is selected for each of the validation and testing subsets. In the majority of ANN applications in permeability predictions, the data are divided into the subsets on an arbitrary basis. Tokar and Johnson (1999) found that the way the data are divided can have a significant impact on the results. Some studies have used ad-hoc methods to ensure that the data used for training and validation subsets have the same statistical properties (Braddock et al. 1998; Campolo et al. 1999; Tokar and Johnson 1999). Bowden et al. (2002) used a genetic algorithm to minimize the difference between the means and standard deviations of the data in the training, testing and validation sets. Thus this study, we will propose a new way to divide data to improve the ANN performance.

1.4 Methods Overview

This study can be divided into two major parts: core digital image analysis and permeability prediction from ANN. The applied methods are summarized as follows.

- Core digital image analysis:
    1. Matlab is used to convert the core slab image into a 2 D matrix with elements presented at the gray value scale. Each gray value is adjusted based on the light intensity to ensure that each image is under the same illumination level.
    2. The digital image analysis programs are tested with 4 synthetic simplified cases and 1 real case before they are applied to analyze the core slab images. The aims of these tests are to confirm that the image analysis program correctly removes the effects of fractures and that the lamination thicknesses are correctly calculated.
    3. The results calculated from the programs are compared with profile permeability data and the mechanical hardness data.

- Permeability prediction from ANN:
    1. Apply multiple variable regression method and Yao and Holditch (1993) permeability model to predict permeability that will be compared with ANN results.
    2. Determine the optimum ANN model by optimizing the number of hidden neurons, the number of validation checks, dataset division and weights and biases. The optimum number of hidden neurons and validation checks are determined by comparing the averaged R^2 (coefficient of determination) of 500 trials for each value. The optimum dataset divisions are determined based on the frequency of each data point in the best 500 results of 10000 trials. The optimum weights and biases are determined by comparing the R^2 of 10000 trials.

3. The digitized image data from Chapter 3 are incorporated into the ANN model optimization.
4. The ANOVA model is designed to investigate the importance of well log inputs on the prediction results. Probability plots are used to detect the factors and interaction effects that are due to random errors.

1.5 Summary

Low permeability, scale problems and heterogeneity are main challenges for permeability prediction in tight formations (Di, 2015). So the results from conventional log-based permeability prediction are questionable. According to the literature review, ANN is an applicable approach for log-based permeability prediction in highly heterogeneous reservoirs. But due to relatively large the sample spacing and investigation volume, it is difficult for well log and core measurements to fully characterize the cm-scale heterogeneity in tight formations (Solano et al., 2012). The literature review suggests core digital image analysis can provide corresponding petrophysical information at a high resolution. Thus, this study tries to solve these challenges based on the Montney Formations by using ANN and digital image analysis.

The results indicate that ANN is accurate in predicting permeability but its degree of overfitting remains to be assessed. The digital image analysis can reveal the lamination scale heterogeneity which might be overlooked by core measurements.

CHAPTER 2: Data Description and Preparation

2.1 Data Available

The data available, both well logs and core measurements, for this study come from well 13-28-064-05 in the fine-grained low-permeability intervals of the Montney Formation in southwest Alberta (Fig. 2.1). These data include:

1. Full diameter core measurements, including core spectral gamma ray measurements, core bulk density log measurements, whole core permeability and porosity measurements. Aluminum sleeve coring equipment and oil base mud were used to core the subject well. Core spectral gamma ray activities and bulk density index are measured prior to other measurements. One hundred four full diameter intervals were selected and cut from the whole core using dry air as the lubricant. After the Dean Stark analysis, the core samples were cleaned using toluene extraction equipment and dried in a FrictionAire oven. Permeability to air was measured in the horizontal (K_{max}. K_{90}) and vertical (K_v) directions. Porosity was determined by Boyle's Law using helium as the gaseous medium.
2. Core plug measurements including ambient porosity and overburden-stressed air permeability measurements. A total of ninety-eight 25.4 mm (1 inch) diameter core plugs were selected and retrieved from the cores. Ambient porosity was measured at zero confining pressure. Overburden air permeability was measured at 5520 kPa hydrostatic net confining pressure by the non-steady-state pressure-decay permeameter CMS-300, but only 93 measurements were completed due to core fractures and sample breakages.
3. Conventional well logs, including neutron porosity, sonic, PE, density, resistivity, GR and caliper logs. All these logs were measured by Halliburton in 2013.
4. Gray-scale core slab images. Photographs were taken by Core Laboratories. There are 68 core slab images covering the Montney Formation from 3040.6 m to 3242.4 m, but 52.4 m of the cores are missing (Fig. 2.2) or not slabbed (Fig. 2.3). Thus 149.3 m of core, broken into 49 short intervals, can be digitized.

5. Profile (probe) permeability provided by Ghanizadeh et al. (2015a). Profile permeability measurements were acquired on the surfaces of core slabs and core plugs with a PDPK-400™ (CoreLab) probe permeameter using a mean gas (N_2) pressure of 170.3 kPa (24.7 psia). Profile permeability values are measured by flowing gas through a small injection tip pressed and sealed against the surface of an unconfined slab or plug. The theory, design, and operation of the nonsteady-state (pressure-decay) profile permeameter (PDPK-400™) are described by Jones (1994). The gas permeability values were corrected for slip-flow (Klinkenberg effect). The procedure for correcting profile permeability values for slip-flow effect is explained in detail in Jones (1994). The uncertainty of the profile permeability measurements was about ± 5% (Jones, 1994).
6. Mechanical hardness data provided by Ghanizadeh et al. (2015b). Mechanical hardness tests were performed using a portable Equotip® Piccolo 2 hardness tool (Proceq SA, Switzerland) at the same locations as probe permeability measurements were acquired.

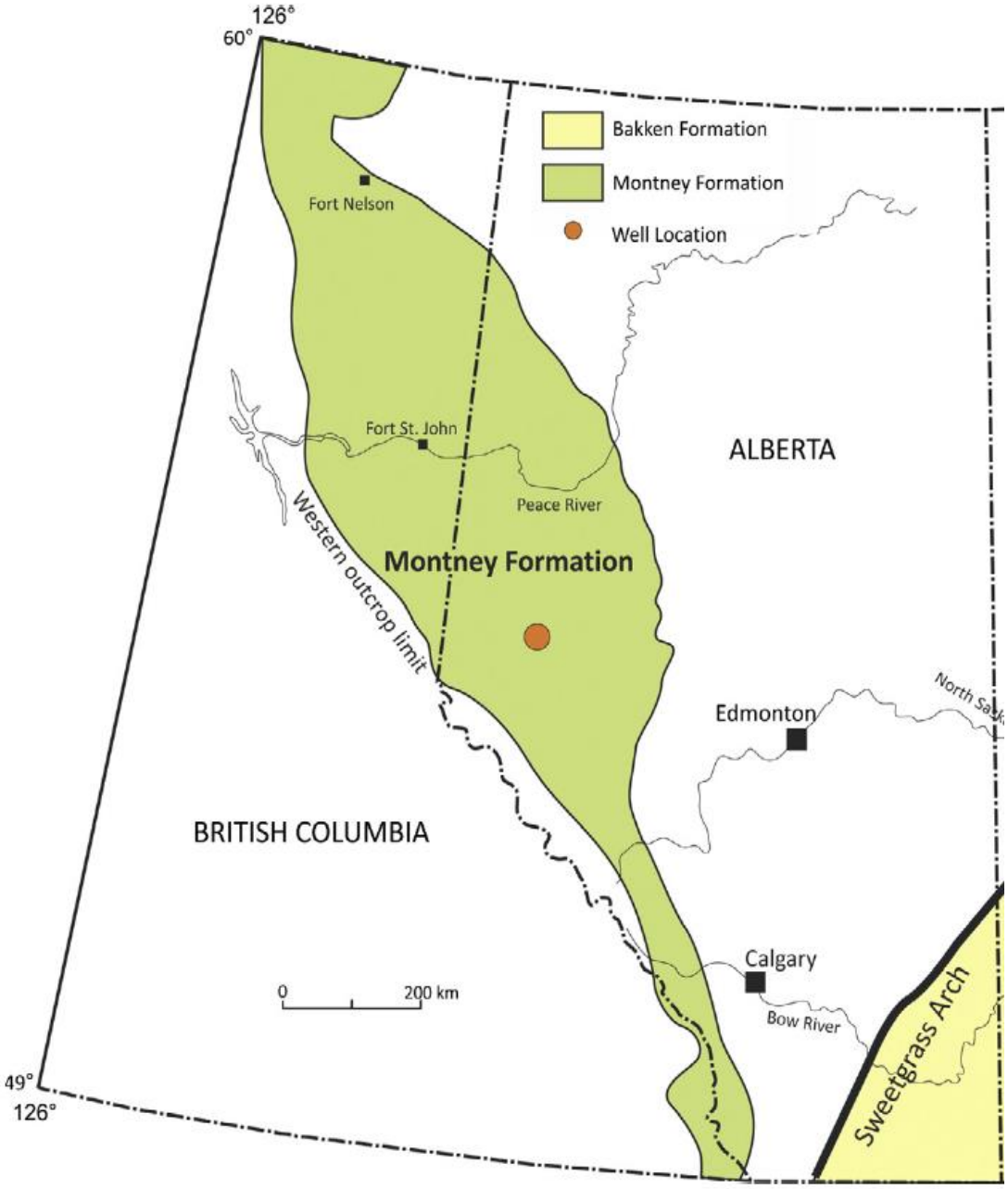

Figure 2.1 The study area for the Montney Formation located in southwest Alberta, Canada (Modified from Ghanizadeh et al., 2015a).

Figure 2.2 An example of missing core.

Figure 2.3 An example of unslabbed core.

2.2 Petrophysical Analysis

2.2.1 Core and Well Log Depth Matching

Depth is the single measurement that can integrate all other measurements made in a wellbore and cores. Recorded core depths are normally inconsistent with well logging depths because of incomplete core recovery and differences between coring and logging tools. Thus, the depths of core and well log measurements should be matched before any further analyses are made.

In this study, the spectral GR log data measured on the cores and GR log data measured in the wellbore are used to conduct the depth shift. The depths of core are shifted to optimize the agreement of spectral GR log data and GR log data. Due to incomplete core recovery, the whole core is divided into 8 intervals and each interval should be shifted separately (Figs. 2.4-2.13).

Care also had to be taken so that the core depth shift of a deeper interval did not cause an overlap with a shallower interval. The depth matching results are summarized in Table 2.1.

For Interval 1, 4 and 5, their GR data match core GR data well after the shift (Figs. 2.4, 2.8 and 2.9). For Interval 6, 7 and 8, their GR data also match core GR well (Figs. 2.10, 2.11 and 2.12). Their matches can be improved but then overlaps will be caused between each interval. For Interval 2, the shift distance cannot be identified solely based on GR data, so wellbore density log data and core density log data are also used to help match the depths (Fig. 2.6). For Interval 3, whole core porosity data, core plug porosity data and wellbore density porosity log data are also used to help match the depths (Fig. 2.8).

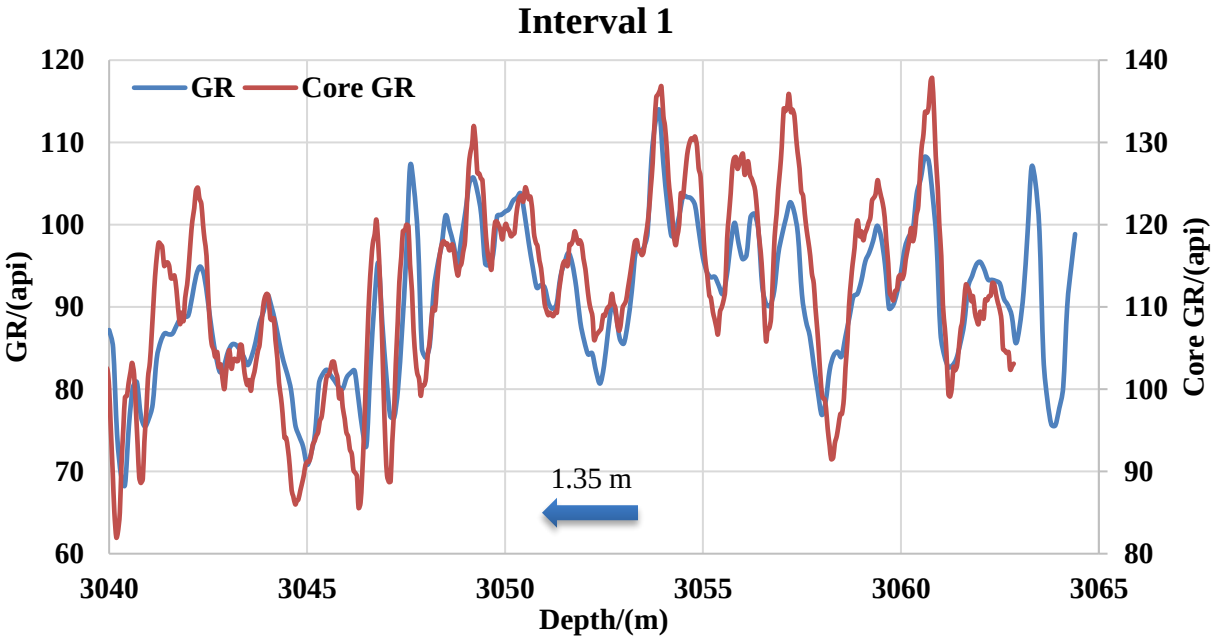

Figure 2.4 Depth shift of Interval 1 based on GR data.

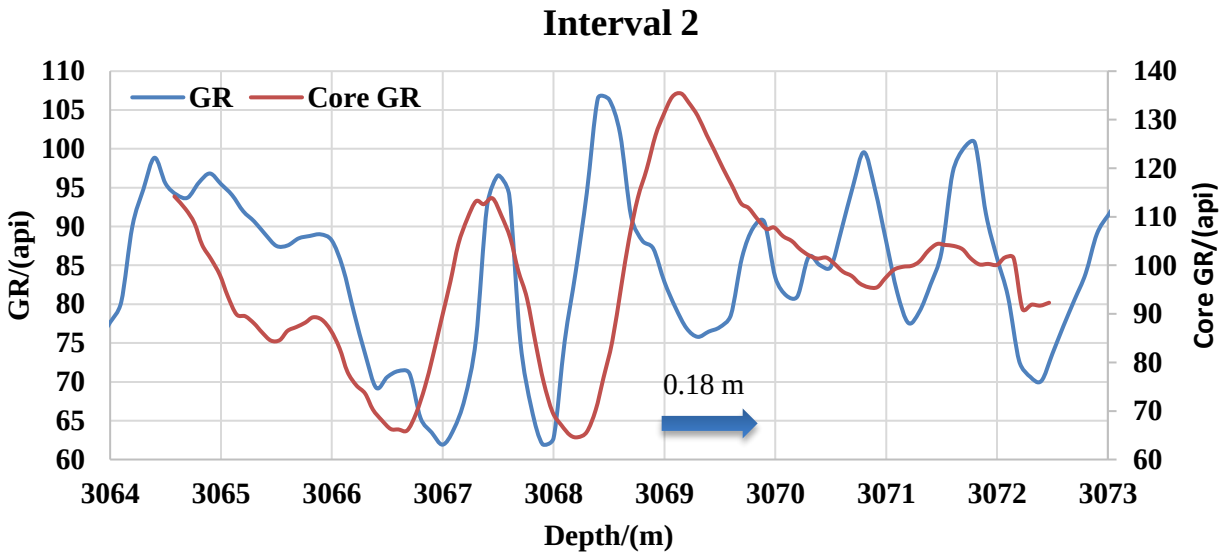

Figure 2.5 Depth shift of Interval 2 based on GR data.

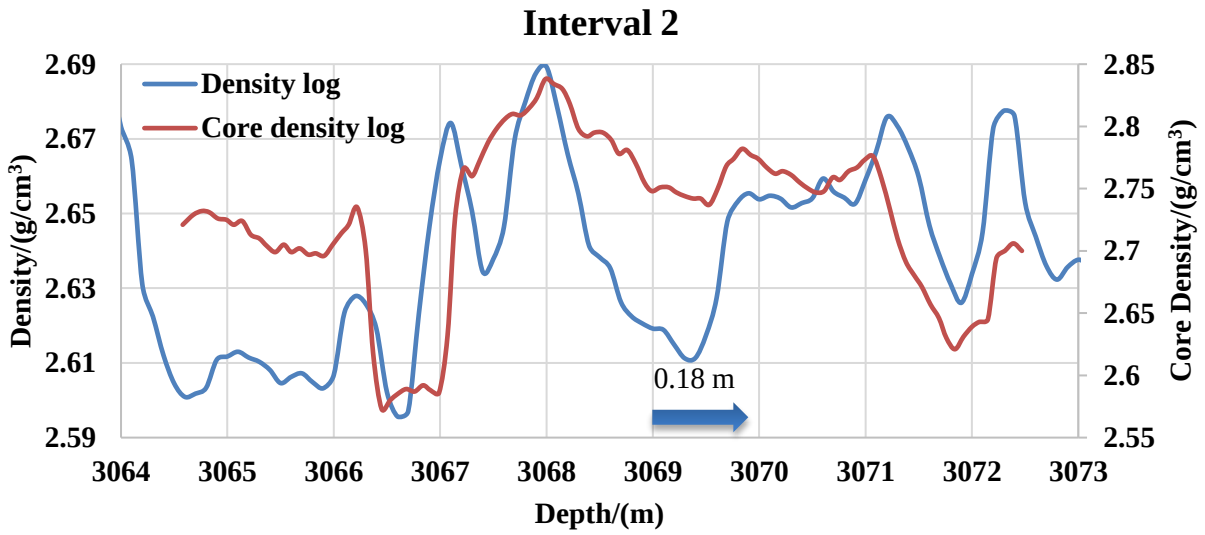

Figure 2.6 Depth shift of Interval 2 based on density data.

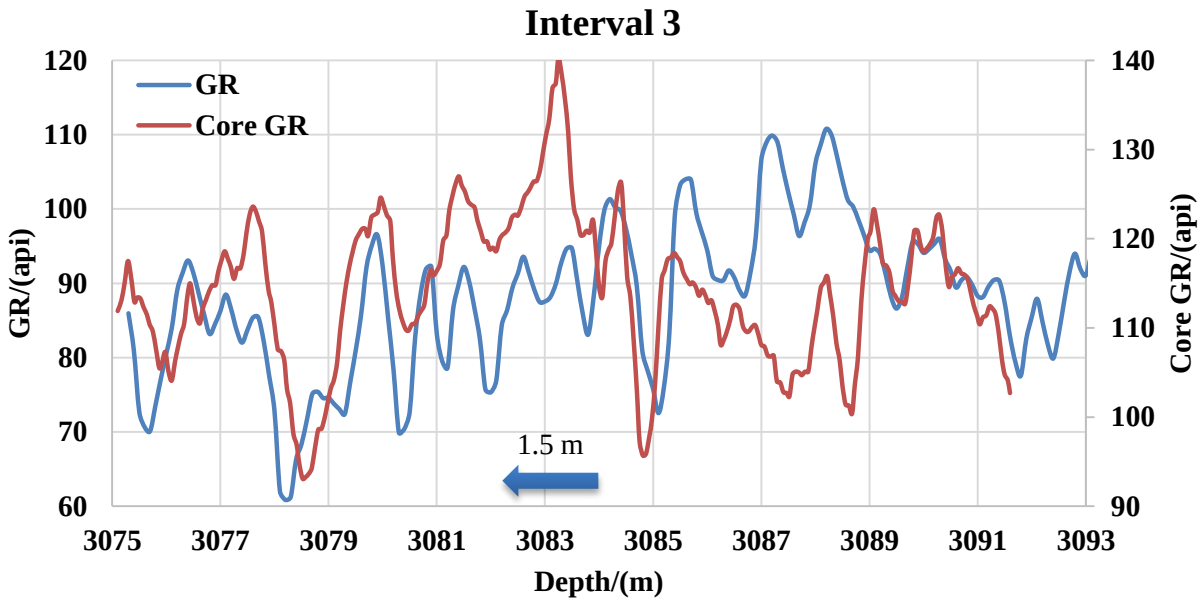

Figure 2.7 Depth shift of Interval 3 based on GR data.

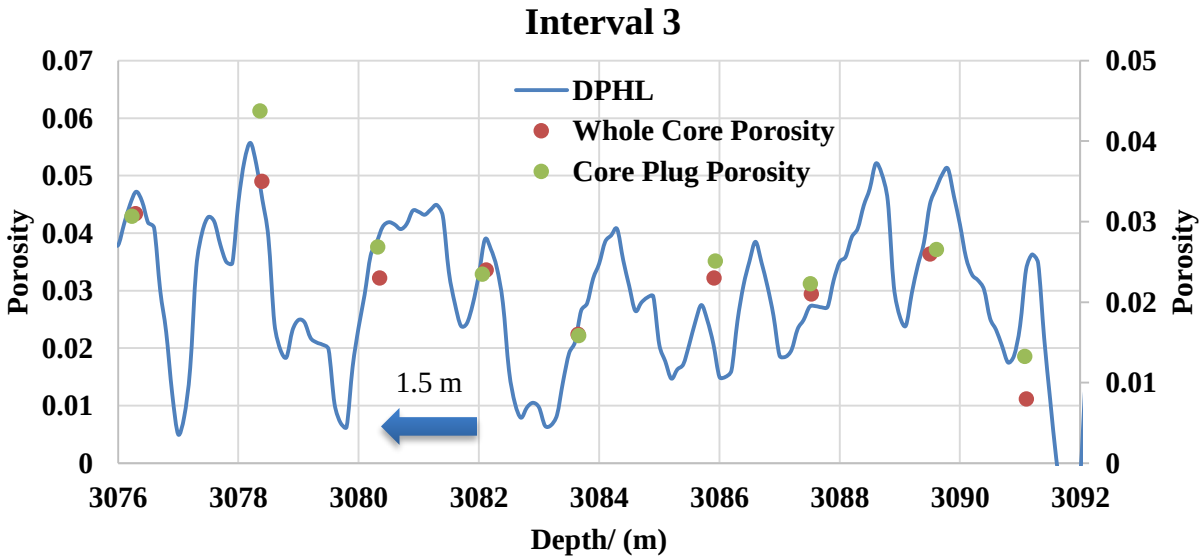

Figure 2.8 Depth shift of Interval 3 based on porosity data.

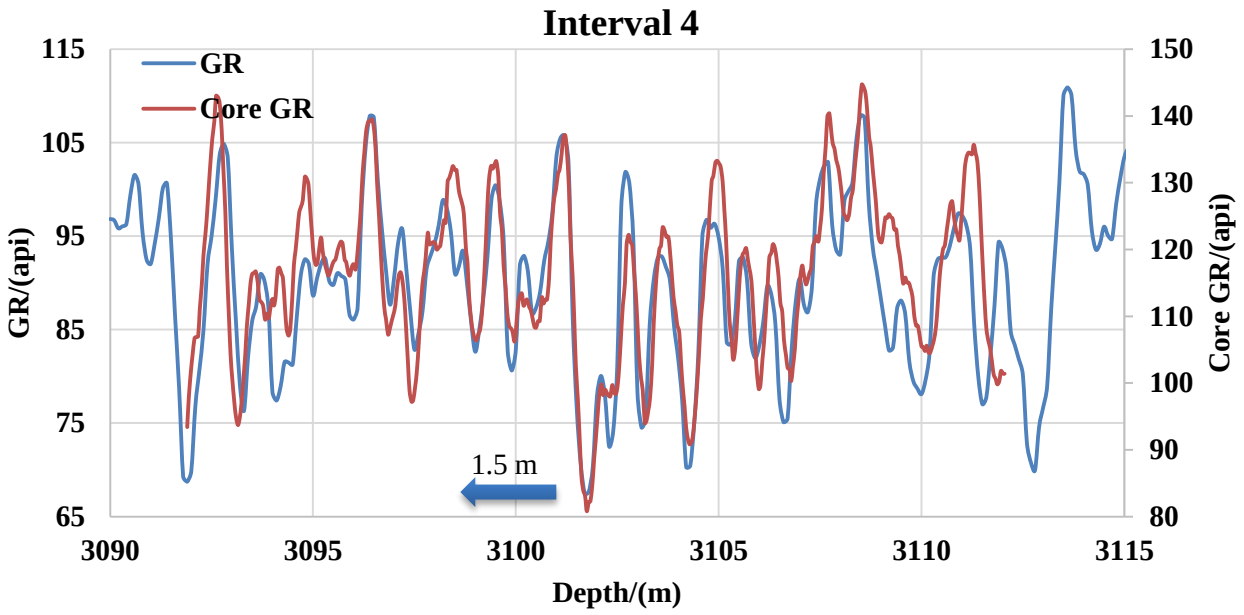

Figure 2.9 Depth shift of Interval 4 based on GR data.

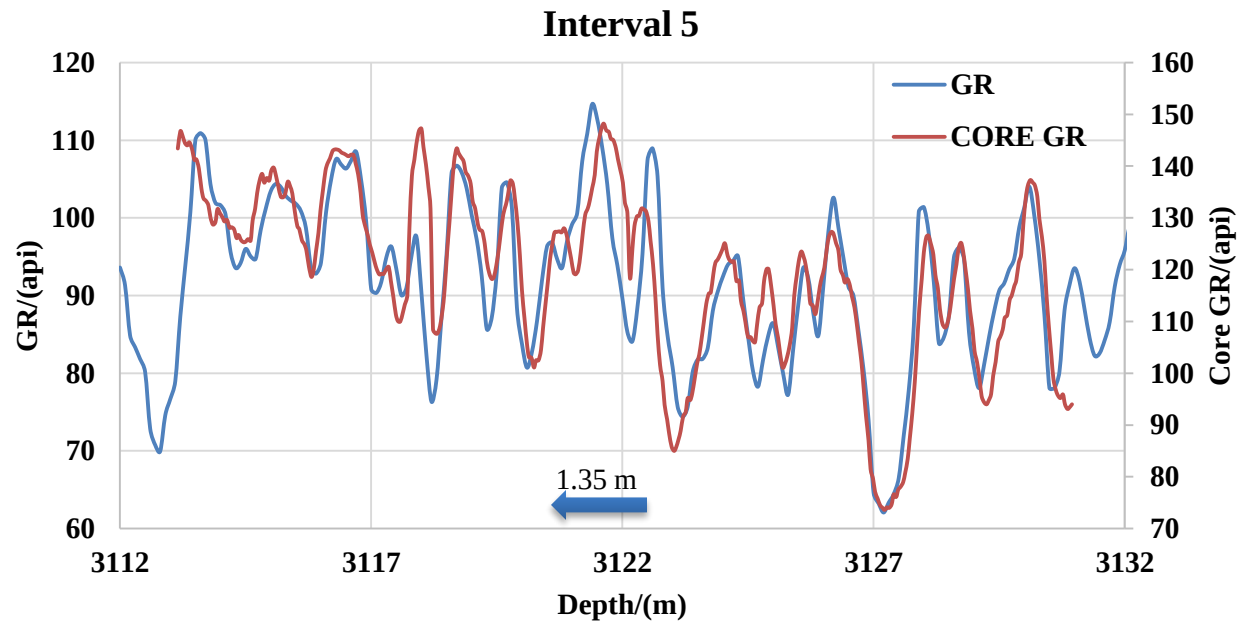

Figure 2.10 Depth shift of Interval 5 based on GR data.

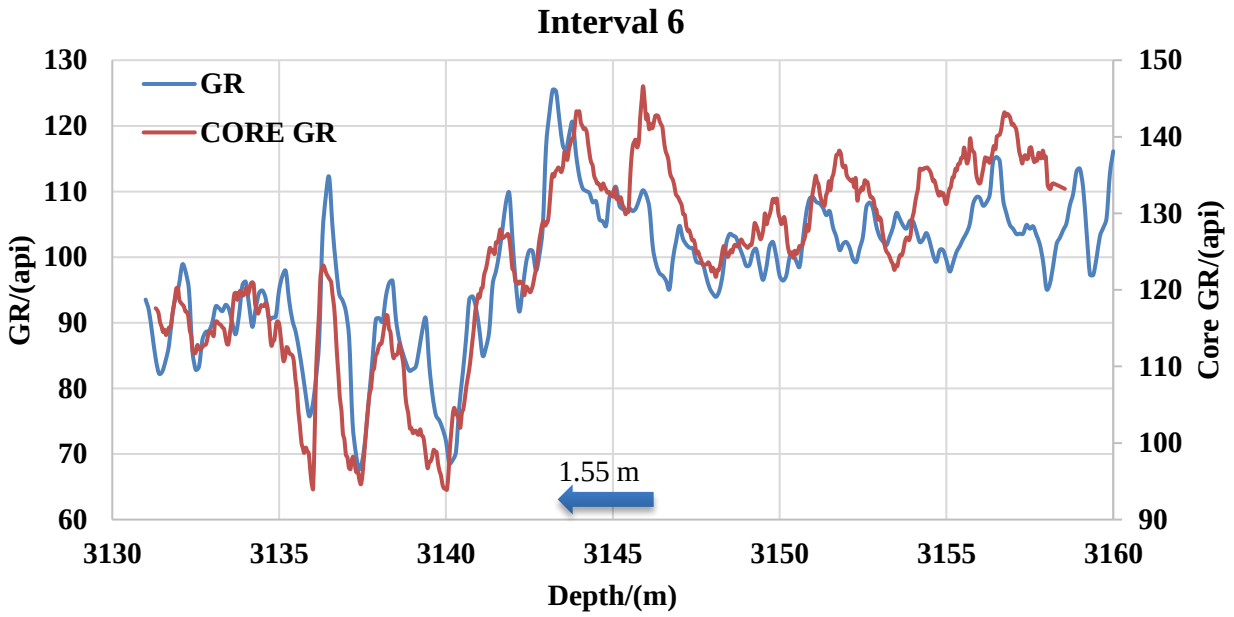

Figure 2.11 Depth shift of Interval 6 based on GR data.

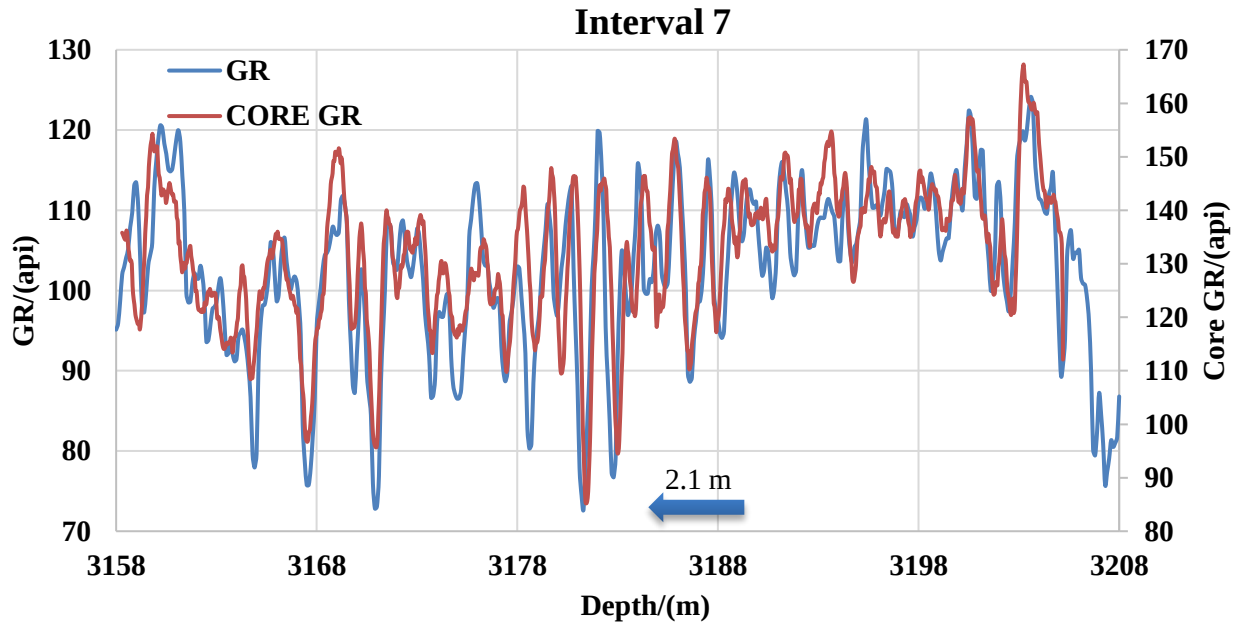

Figure 2.12 Depth shift of Interval 7 based on GR data.

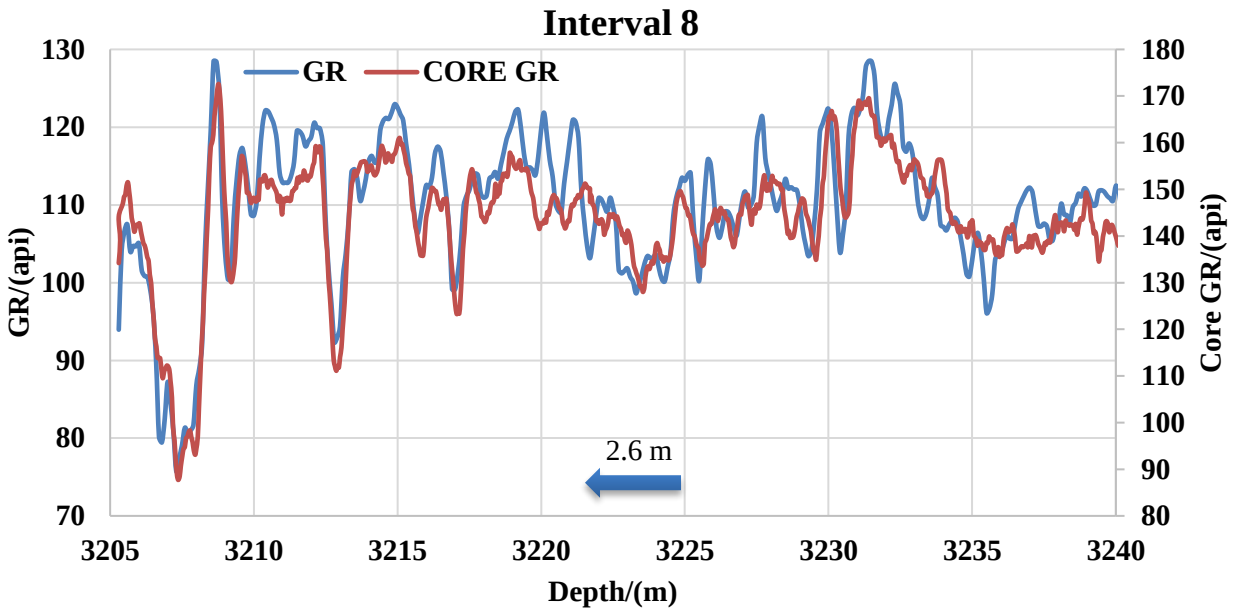

Figure 2.13 Depth shift of Interval 8 based on GR data.

Table 2.1 Depth matching summary (- means shifting upwards; + means shifting downwards)

Before Shift	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Upper Depth/(m)	3039.96	3064.40	3076.60	3093.40	3114.50	3132.60	3160.40	3207.80
Bottom Depth/(m)	3064.20	3071.97	3093.10	3113.55	3132.30	3159.85	3207.30	3251.85
Shifted Depth/(m)	-1.35	+0.18	-1.50	-1.50	-1.35	-1.55	-2.10	-2.60
After Shift								
Upper Depth/(m)	3038.61	3064.58	3075.10	3091.90	3113.15	3131.05	3158.30	3205.20
Bottom Depth/(m)	3062.85	3072.15	3091.60	3112.05	3130.95	3158.30	3205.20	3249.25

2.2.2 Data Screening

The resolution of the whole core permeability measurements is 0.01 mD and the measured values are also censored at 0.01 mD (Fig. 2.14). But the permeability in the Montney Formation may be as small as 4×10^{-4} mD based on the core plug permeability measurements (Fig. 2.15) and profile permeability measurements (Figs. 2.16 – 2.18). The whole core permeability measurements cannot characterize the dominant permeability variations that are lower than 0.01 mD so they will not be used as the target data in the permeability prediction analysis. In addition, the profile permeability values are only measured over a combined 12.27 m long interval of the cores (Figs. 2.16 – 2.18) so that they also cannot represent the permeability distribution of the entire formation. Thus, they will not be used as the target in predicting permeability but they will be used to compare with the image digitization results. Thus, the core plug permeability measurements are the only available data that are consistently measured on the entire Montney Formation and can characterize the dominant 10^{-3} to 10^{-4} mD level permeability variations at the same time, which will be the target data for permeability prediction.

But there are five core plugs that cannot produce effective permeability measurements due to core fractures at 3177.51 m, 3178.59 m, 3180.78 m, 3217.96 m and 3222.71 m respectively (Figs. 2.15 and 2.19). These core plugs all come from the shalier portions in the lower Montney and the permeability measurements near them range from 8.4×10^{-4} mD to 3.6×10^{-3} mD, so the small permeability might be underrepresented.

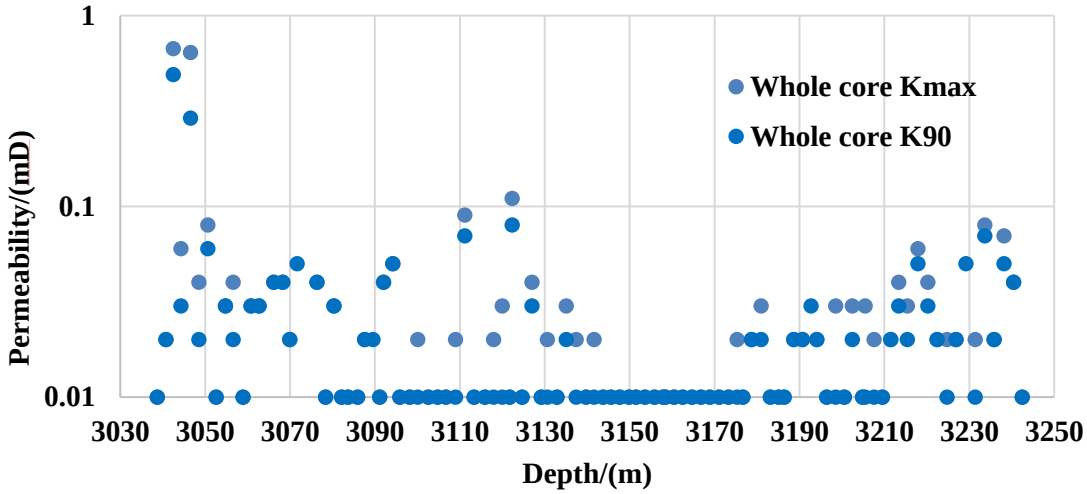

Figure 2.14 Whole core permeability measurements.

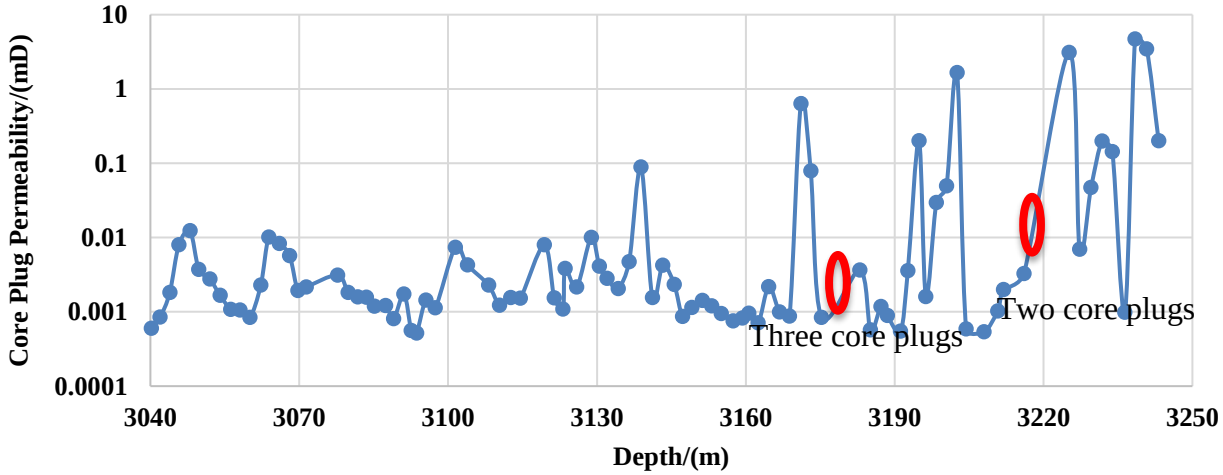

Figure 2.15 Core plug permeability measurements. Core plugs with fracture issues are shown in red circles.

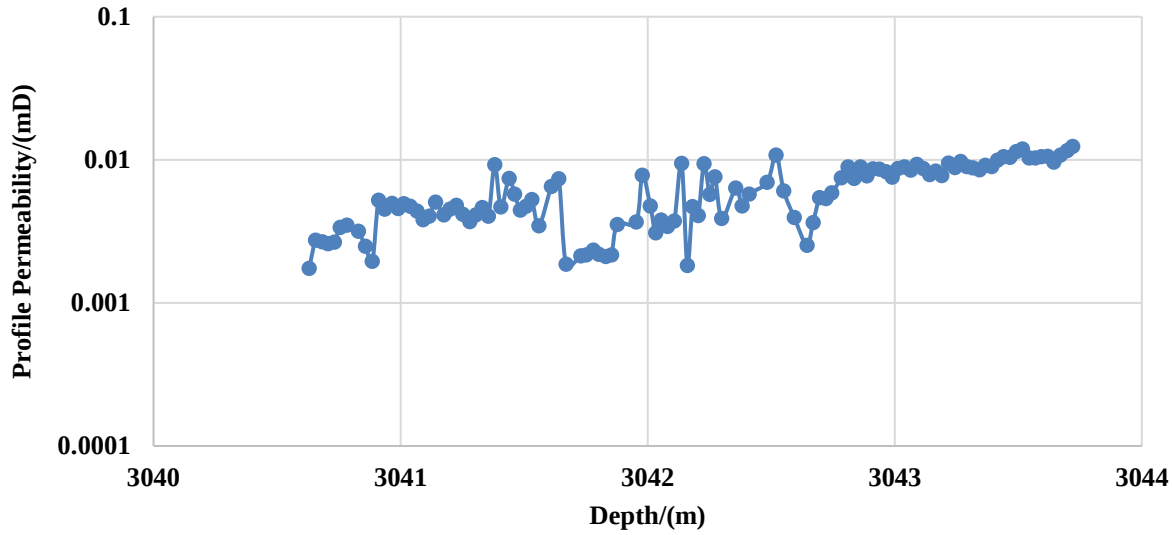

Figure 2.16 Profile permeability measurements.

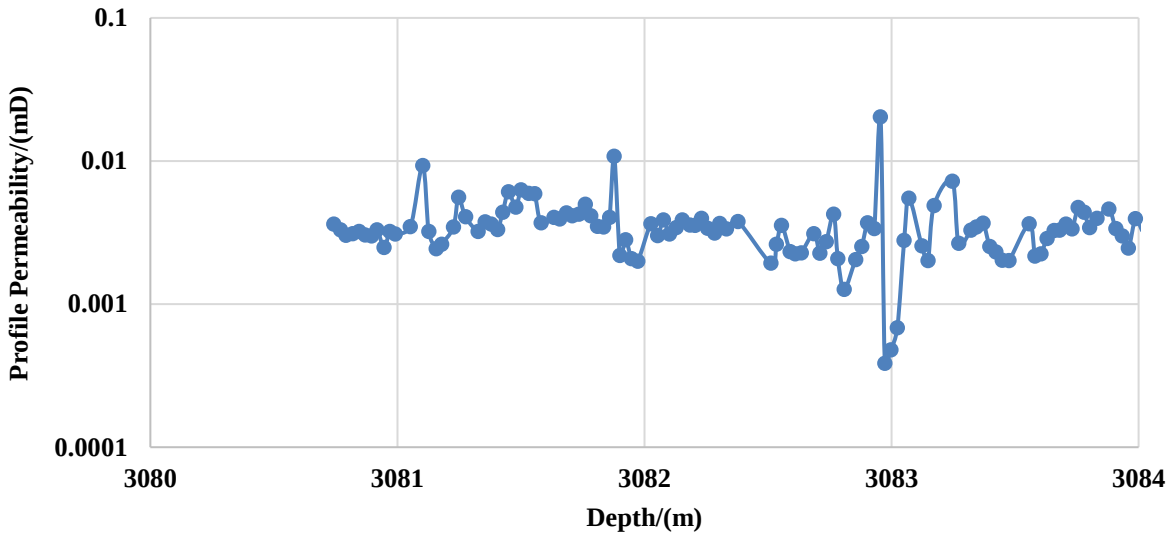

Figure 2.17 Profile permeability measurements.

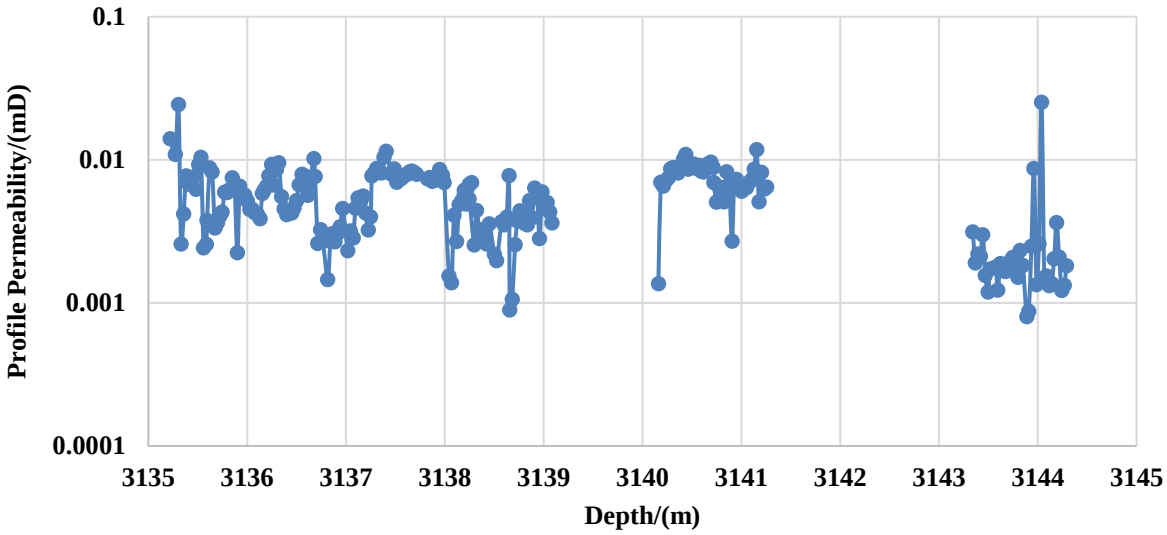

Figure 2.18 Profile permeability measurements.

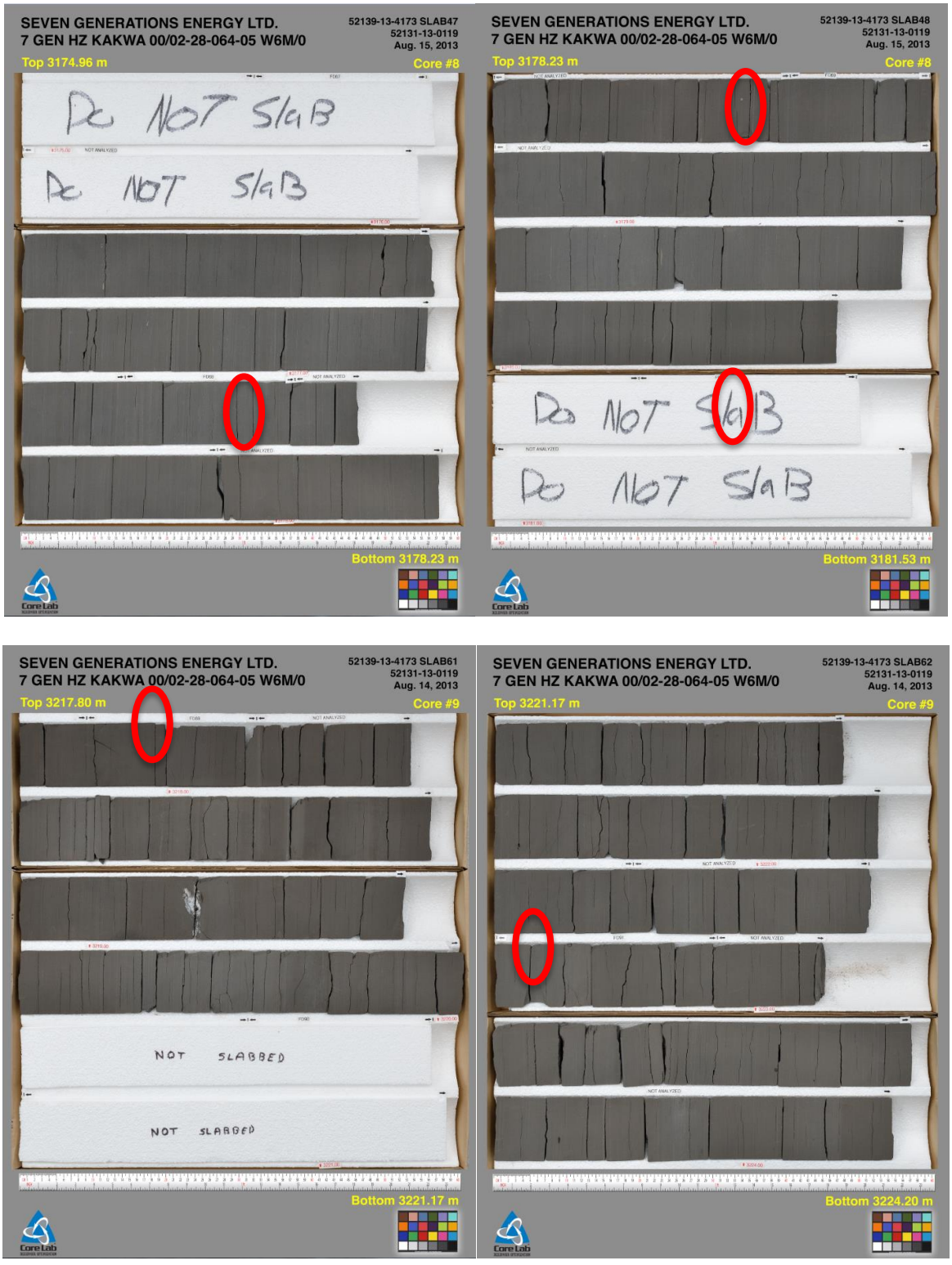

Figure 2.19 Core plugs with fracture issues are shown in red circles.

2.2.3 Core Petrophysics Analysis

The Montney succession is divided into three informal members by Davies et al. (1997): the lower member, the middle member and the upper member. The Lower member of the Montney consists of coarsening-upward, very fine grained sandstones and siltstones. Thick dolomitized coquina of the middle member is underlain by the Lower member of the Montney through a north-to-southwest-trending belt in the eastern subsurface. The upper Montney member is generally characterized by coarsening-up siltstones with very fine grained sandstones and dolomitized coquina facies occurring locally.

Based on the GR data, the core can be divided into 3 parts (Fig. 2.20). Part 1 is mainly comprised of clean siltstone and laminated shaly siltstone (Fig. 2.21). Parts 2 and 3 are composed of dispersed shaly siltstone (Figs 2.22 and 2.23). The middle member, composed of thick dolomitized coquina, is missing in this well according to the core slab images and descriptions. So interval 1 is identified as the upper Montney and interval 2 and 3 are identified as the lower Montney. This division is also supported by the ANN and image results in Section 4.4.6 Incorporating the digitized image data into the ANN, Chapter 4.

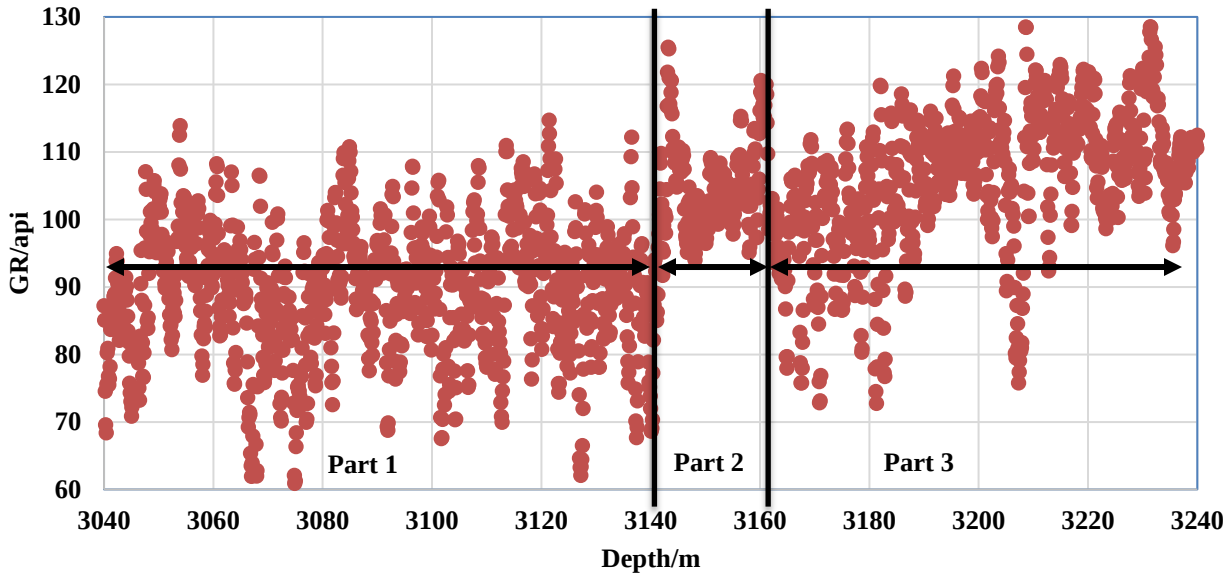

Figure 2.20 The core is divided into 3 parts based on GR response.

Figure 2.21 Core slab images from Part 1; very fine grained well sorted clean siltstone and laminated shaly siltstone.

Figure 2.22 Core slab images from Part 2; very fine grained well sorted dispersed shaly siltstone.

Figure 2.23 Core slab images from Part 3; very fine grained well sorted highly fractured shaly siltstone.

The core plug porosity shows a very weak correlation with the core plug permeability in the Montney (Fig. 2.24). But, the correlation between the core plug permeability and porosity is much stronger in the upper Montney than in the lower Montney (Figs. 2.25 and 2.26). According to the lithological descriptions of the core plugs, a total of 12 core plugs have identified fracture issues, 5 of which cannot produce an effective permeability measurement. The other 7 fractured cores are all located in the lower Montney and have relatively high permeability which may be caused by fractures (Fig. 2.26). The GR log data shows the lower Montney has higher shale content than the upper Montney (Fig. 2.20), which can also be observed from the core slab images (Figs. 2.21 - 2.23). This trend may also be affected by organic matter content which also contributes to gamma radioactivity. Ghanizadeh et al. (2015a) measured the total organic carbon content in our core samples. Their results show that the Montney samples are organic-lean (<0.7 wt.% TOC content) which agrees with other Montney sample results from west-central Alberta (Beaton et al., 2010). This increasing shale effect may also weaken the relationship of permeability and porosity in the lower Montney.

The whole core porosity shows a weak correlation with the whole core permeability K_{max} in the Montney (Fig. 2.27). Similar to the core plugs, this correlation is stronger in the upper Montney than in the lower Montney (Figs. 2.28 and 2.29). But in both members, the correlation is weak, which may result from the censorship at 0.01 mD and the resolution limit of the whole core permeability measurements. Meanwhile, the upper Montney has higher average permeability, a higher coefficient of variation (C_V) of permeability and fewer censored permeability values than the lower Montney (Figs. 2.30 – 2.32), which may be caused by the increasing shale effect in the lower Montney that also weakened the relationship of permeability and porosity.

The recommended number of measurements, No, is used to assess the sample sufficiency of core plug and whole core measurements. No is based on the C_V of data, where $No = (10C_V)^2$ (

Jensen et al., 2000). No assumes one wants to estimate the arithmetic mean permeability of the interval to within ±20% with 95% probability. N is the actual number of measurements. Both core plug and whole core permeability are largely undersampled at the 20% tolerance (Fig. 2.33) but are adequately sampled for a 91% tolerance and a 51% tolerance respectively. For porosity,

both core plug and whole core measurements are adequately sampled at the 20% tolerance (Fig. 2.34).

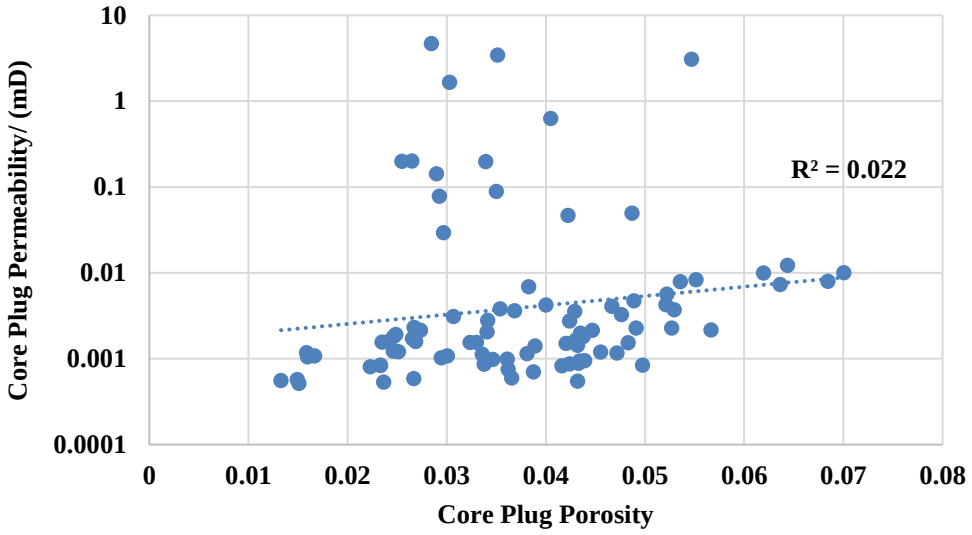

Figure 2.24 Core plug permeability vs core plug porosity in the Montney.

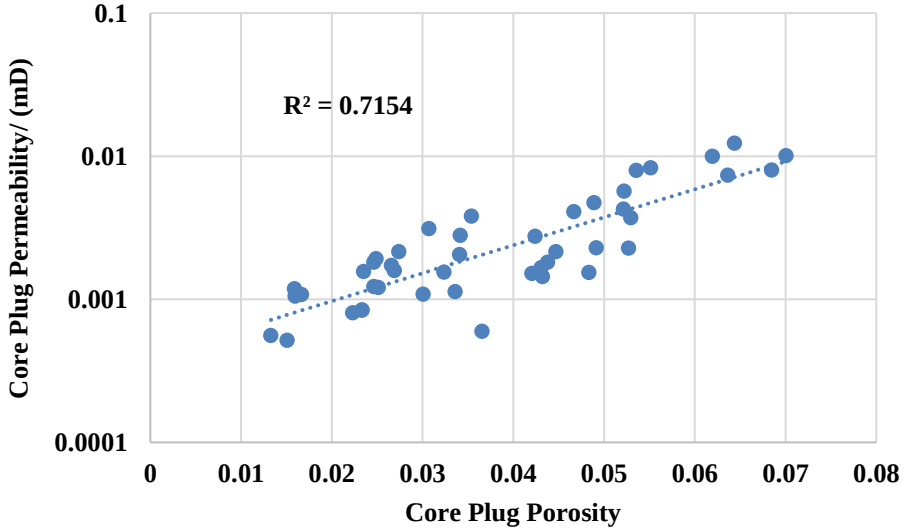

Figure 2.25 Core plug permeability vs core plug porosity in the upper Montney.

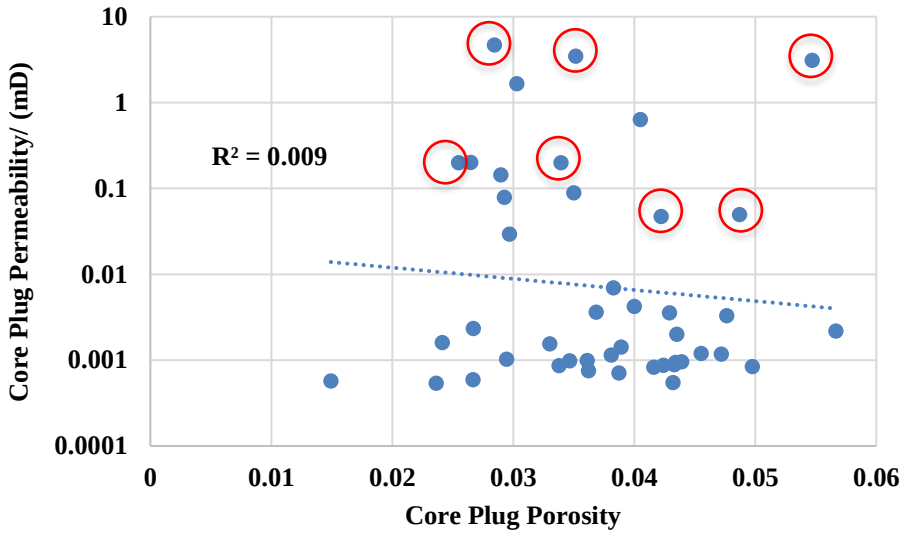

Figure 2.26 Core plug permeability vs core plug porosity in the lower Montney. The core plugs with facture issues are identified with red circles.

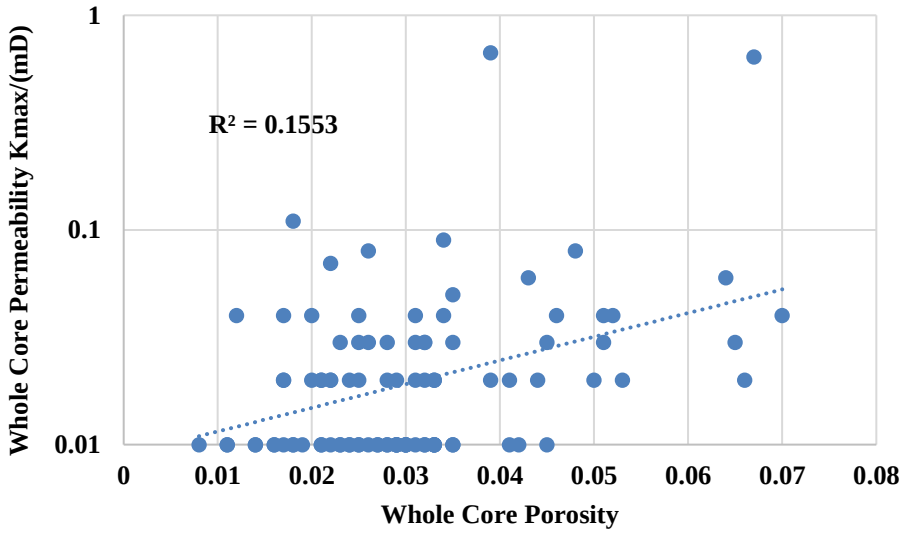

Figure 2.27 Whole core permeability Kmax vs whole core porosity in the Montney.

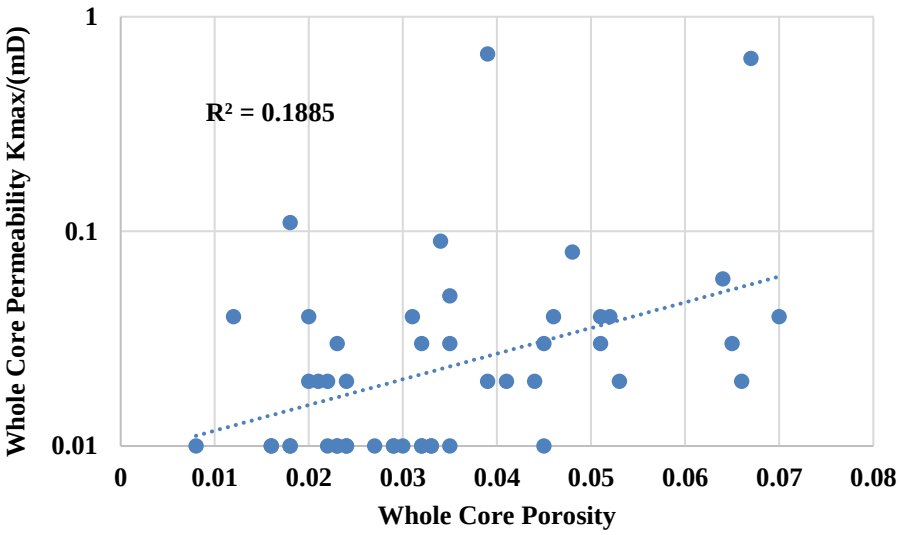

Figure 2.28 Whole core permeability Kmax vs whole core porosity in the upper Montney.

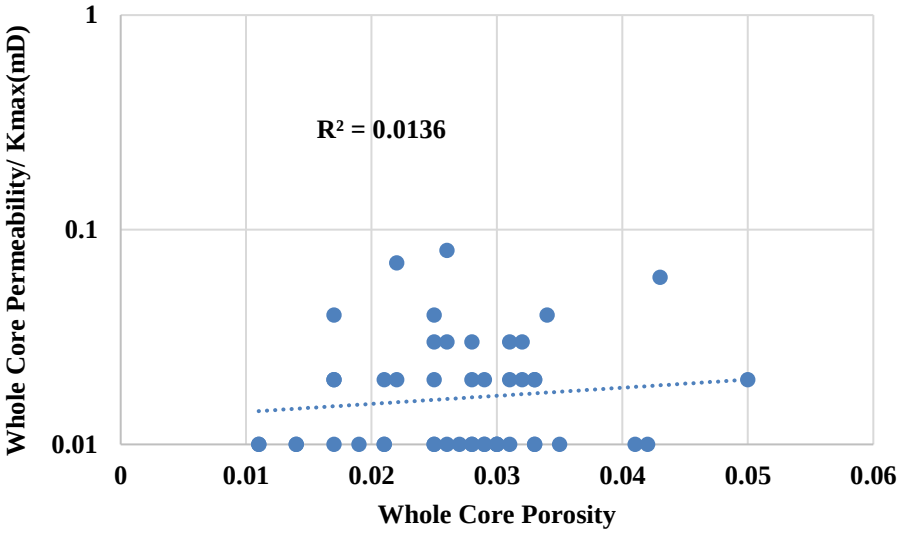

Figure 2.29 Whole core permeability Kmax vs whole core porosity in the lower Montney.

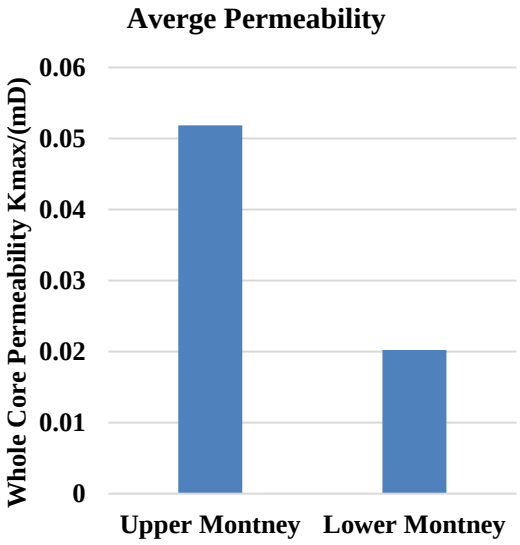

Figure 2.30 Average permeability of the upper and lower Montney

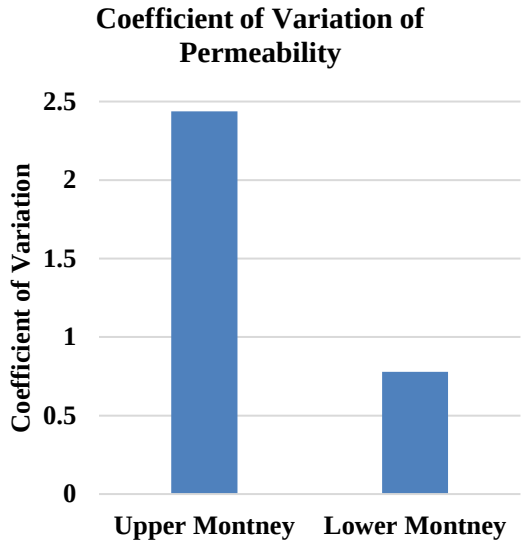

Figure 2.31 Coefficient of Variation of permeability of the upper and lower Montney.

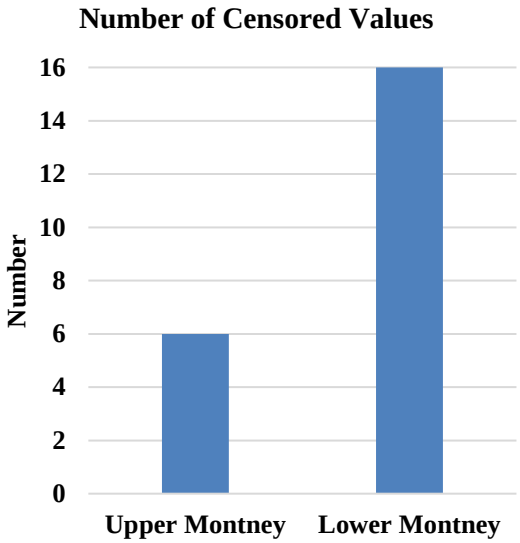

Figure 2.32 Number of censored permeability values of the upper and lower Montney.

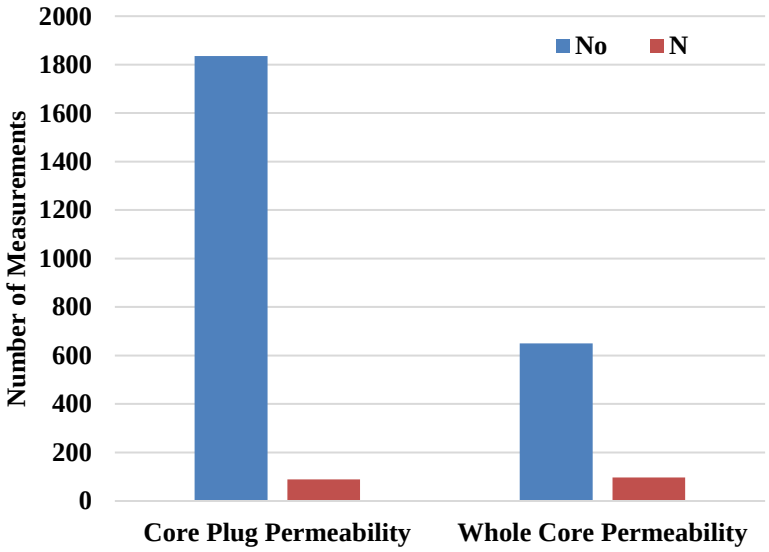

Figure 2.33 Permeability measurements sample sufficiency assessment. No is the recommended number of measurements. N is the actual number of measurements.

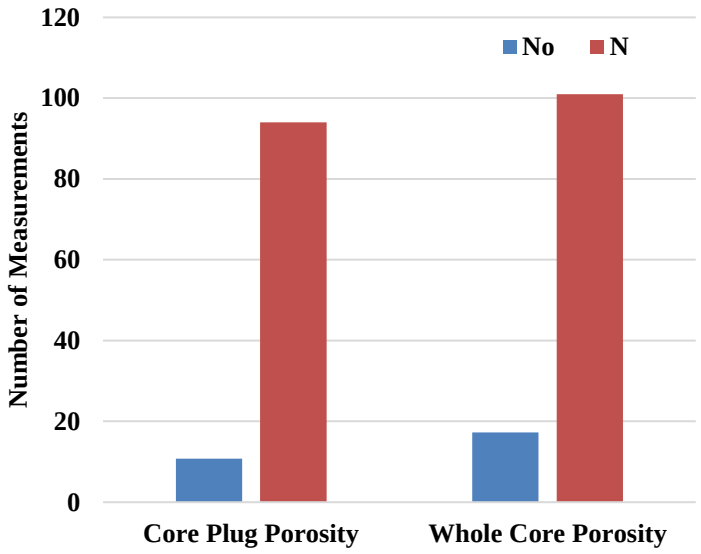

Figure 2.34 Porosity measurements sample sufficiency assessment. No is the recommended number of measurements. N is the actual number of measurements.

2.3 Conclusions

This chapter focus on the screening, assessment and analysis of the petrophysical data, which provides the basis for further investigations. The main results are listed below:

1. The middle Montney is missing in the studied well. The core descriptions and images show that the Montney in the studied well is dominantly composed of variable amounts of interbedded fine-grained shale, sandstone and siltstone.
2. The whole core permeability measurements cannot characterize the dominant small permeability variations (10^{-3} to 10^{-4} mD) because of their resolution limit and censorship. The vertical permeability measurements are also censored, so the anisotropy (K_v/K_h ratio) also cannot be characterized.
3. The core plug permeability measurements are the only available data that are consistently measured on the entire Montney Formation and can characterize the dominant 10^{-3} to 10^{-4} mD level permeability variations at the same time, which will be the target data for permeability prediction.

4. Both core plug and whole core permeability are largely under-sampled at the 20% tolerance but are adequately sampled for a 91% tolerance and a 51% tolerance respectively. The five missing core plug permeability measurements might make the small permeability underrepresented.
5. The porosity-permeability relationship is weak in the Montney, suggesting that porosity (i.e., sorting) is not the unique permeability controlling factor. But the relationship is much stronger in the upper Montney than in the lower Montney, which may result from the increasing shale effect in the lower Montney.

CHAPTER 3: Reservoir Heterogeneity Characterization Using Digital Image Analysis

3.1 Introduction

For effective reservoir simulation models, it is important to estimate permeability and other petrophysical properties accurately over different scales of geological heterogeneity. Due to the limited volume of investigation and sampling resolution, many petrophysical tools cannot fully characterize the heterogeneity at the lamination (mm) and laminaset (mms – cm) scale, which may be important in tight gas reservoirs. However, digital imaging can improve the efficiency and accuracy of collection, interpretation and archiving of reservoir data in a cost effective manner. Images can be digitized to identify important features and characterize small scale heterogeneity because of the high resolution of digital images.

This chapter presents a method to convert core slab images into digitized gray value data with necessary adjustments and corrections for light intensity and core breakages. The gray value data are used to assist the artificial neural network to predict permeability accurately over different scales, which is discussed in Chapter 4. The gray value data are also compared and correlated with the probe permeability data and mechanical hardness data measured from laboratory.

3.2 Digitization Process

Each image is composed of a large number of pixels which reflect a specific intensity of gray. This intensity can be measured quantitatively as a gray scale value. Based on this principle, the core slab images are digitized into gray scale values using Matlab.

3.2.1 Image Preprocessing

The analysis of the digitized image is based on the assumption that each image is taken under the same illumination environment so that the same mineral type will present the same gray scale value. Thus the illumination normalization should be conducted before any further analysis.

The core rests in Styrofoam trays, giving a white background in the core photos, which can show the degree of illumination. The average gray scale value of the white background is considered as the target value to conduct the illumination correction (Fig. 3.1). The target value of the first slab image is used as the benchmark value for the normalization. For each slab image, the ratio of target value to benchmark value is calculated and applied to correct the whole image.

$$Ratio = \frac{target\ value}{benchmark}$$

Results show that the deviation of all images is less than 5%, and is therefore not an important influence.

Also, core plugs and breakages will cause irregular boundaries of core slabs (Fig. 3.1). Thus each image is evenly divided into three parts (Fig. 3.1). The middle (central) part of the image is used for further digitization.

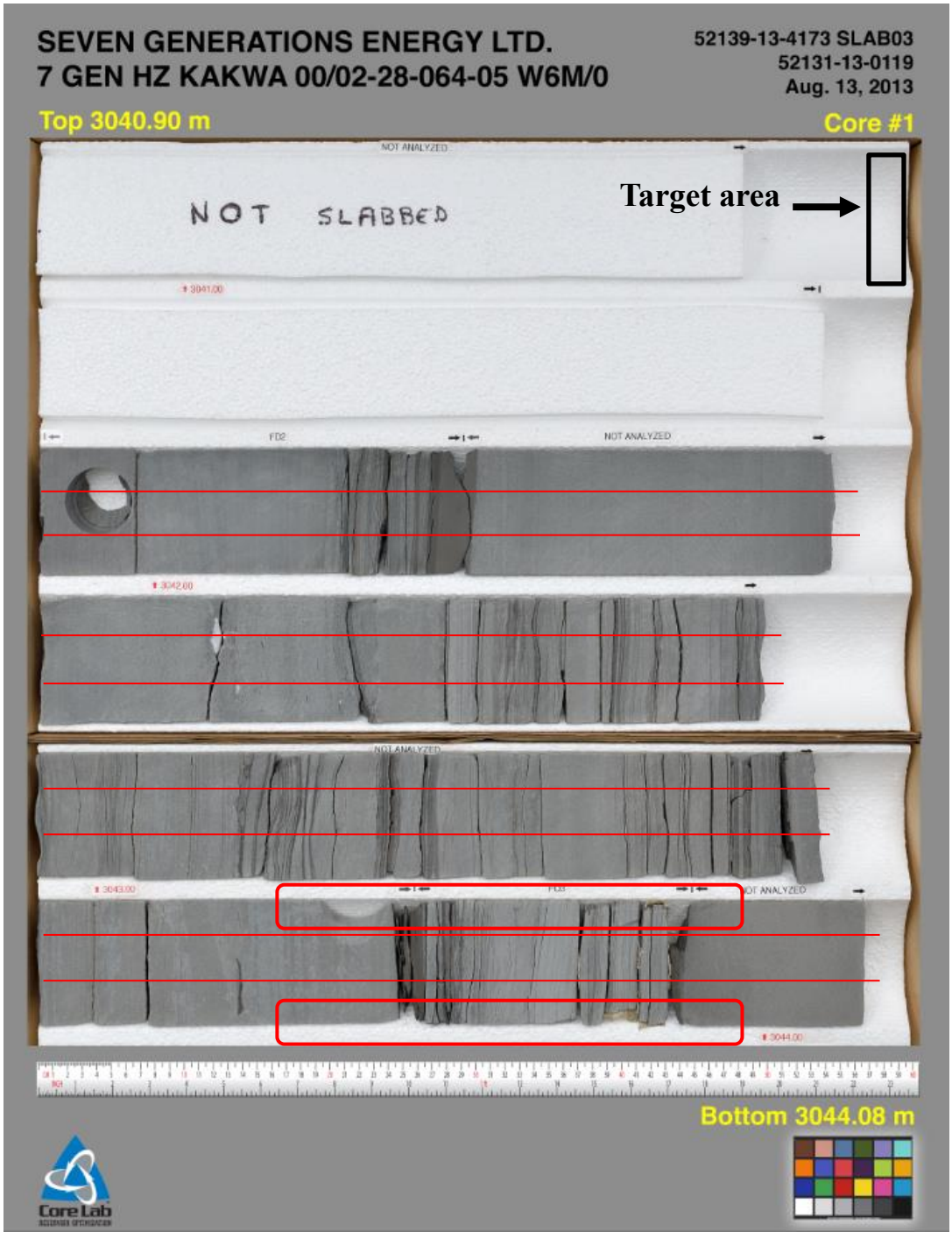

Figure 3.1 Target area for illumination normalization; irregular boundaries caused by core plugs and breakages are shown in red boxes. The image is evenly divided into three parts whose boundaries are marked by red lines.

3.2.2 Program Testing

After digitization and intensity adjustment, a core slab image will be transformed into a 2 D matrix with elements presented at the gray value scale that defines 0 as the 100% black and 255 as the 100% white. The lamination orientations are generally horizontal, so the average value of each column is used to represent that column. An example digitization result is shown in Fig. 3.2. But core breakages or fractures will show extremely small gray values due to the dark color of shadows which do not reflect lithology changes. These low-value spikes may highly influence the investigation of heterogeneities and core breakages and take 2.6% of the whole core length in this study, so these events should be corrected.

The space fractures and breakages occupy were once filled with the matrix before they occurred. So the fracture and breakage effects can be corrected by replacing them with the surrounding matrix. The missing areas caused by core plug voids are also corrected in the same way (Fig. 3.2). The detailed logic of this adjustment is described in APPENDIX. On the other hand, dark features such as laminations are important for the investigation of heterogeneities. Thus a program is developed in Matlab to correct the effects of core breakages and plugs and identify and extract the locations and thicknesses of laminations.

The digitized photograph analysis programs are tested with 4 synthetic simplified cases and 1 real case before they are applied to analyze the core slab images. The aim of these tests is to confirm that the image analysis program correctly removes the effects of fractures and that the lamination thicknesses are correctly calculated. The first case is relatively simple and each subsequent case tests the program with more complicated geometries.

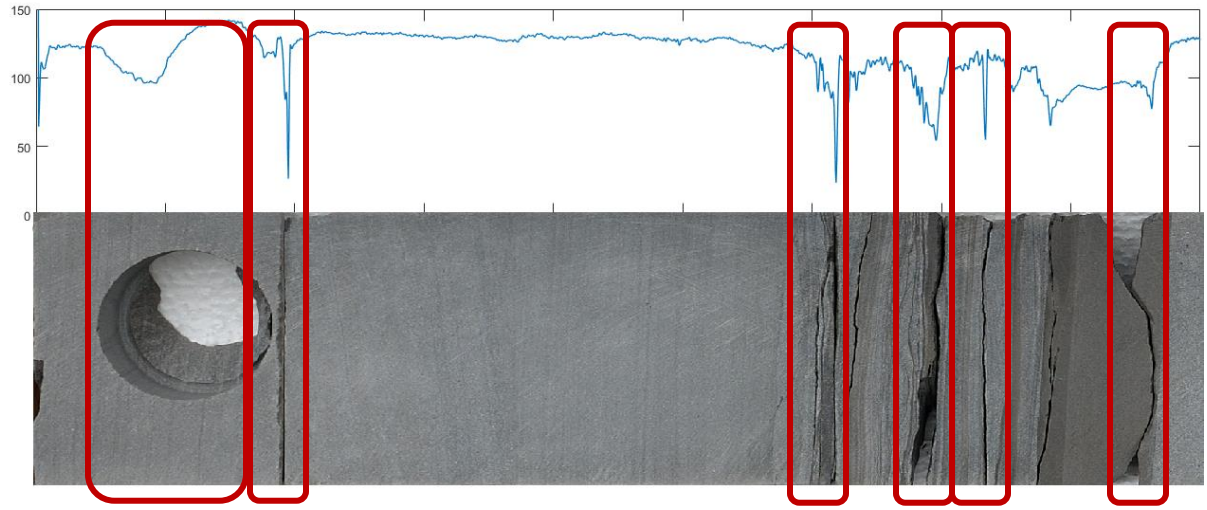

Figure 3.2 Core slab image having core plug, breakage and fracture effects.

3.2.2.1 Test 1

The case is composed of three components (Fig. 3.3): matrix, fractures and laminations (Table 3.1). By comparing the raw digitized results (Fig. 3.4) with the corrected digitized results (Fig. 3.5), the fractures are corrected and replaced by the surrounding components. The lamination thickness profile (Fig. 3.6) shows that the lamination locations and thicknesses are correctly captured and presented by the program. This test indicates that the programs are effective to capture and characterize the straight and sharp boundary fractures and laminations.

Figure 3.3 Test 1 case.

Table 3.1 Test 1 components.

	Fig. 3.6 color	Gray Value	Thickness	Boundary
Matrix		150	1300 pixels	
Fracture		10	10 pixels each	Straight Sharp
Lamination		100	10, 15, 30 and 100 pixels respectively	Straight Sharp

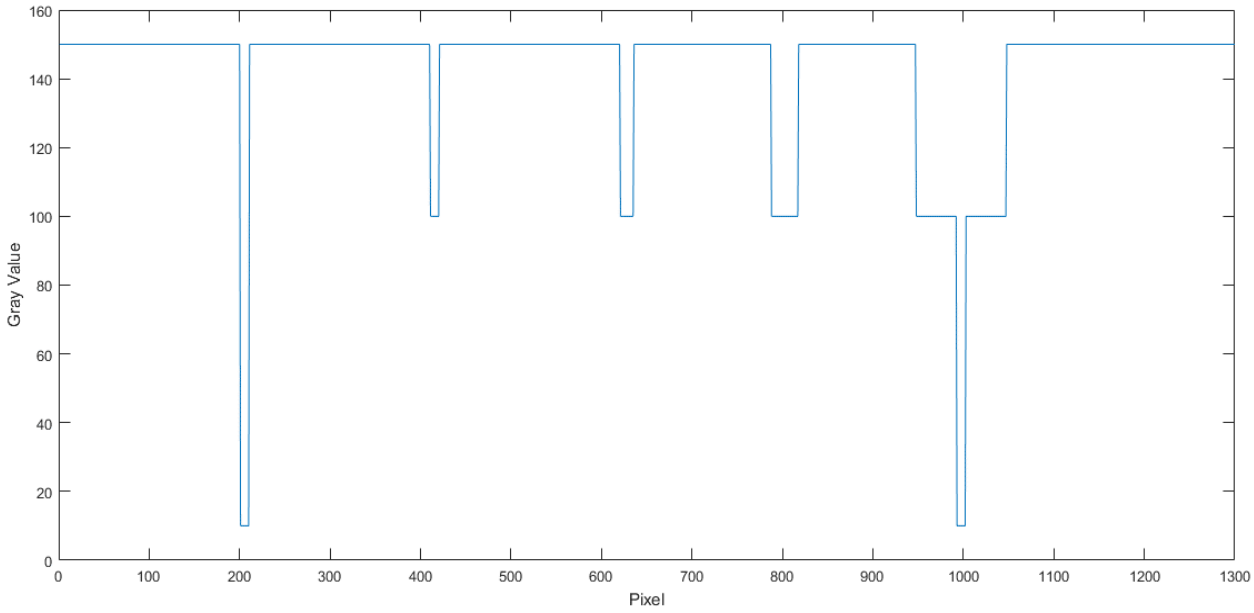

Figure 3.4 Test 1 raw digitized results.

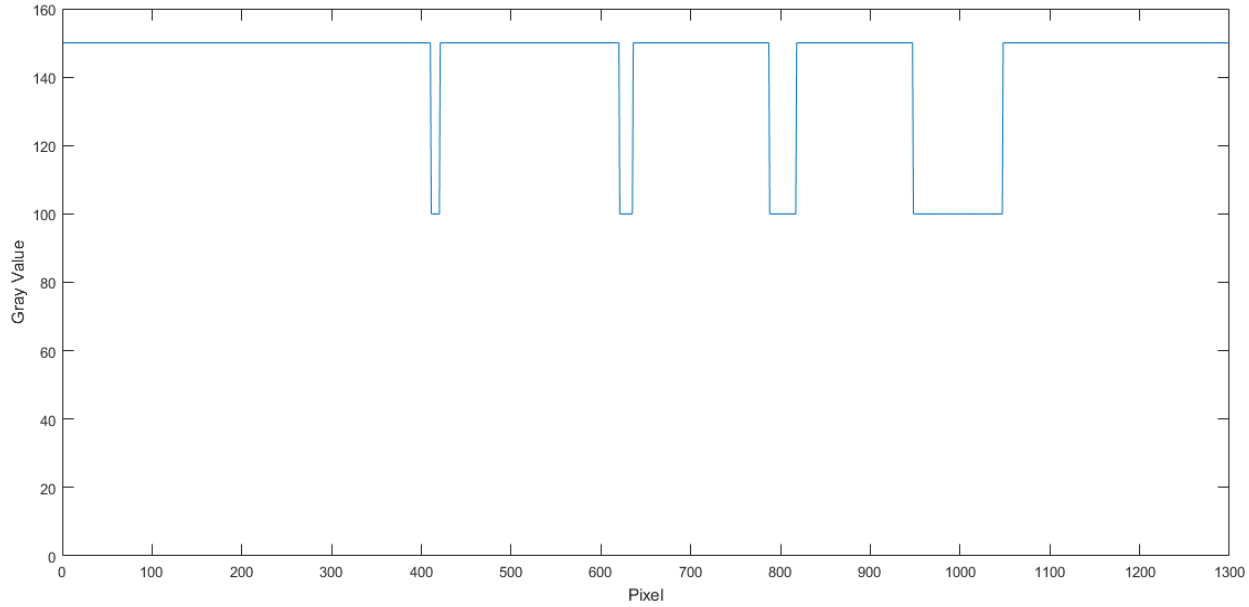

Figure 3.5 Test 1 corrected digitized results, showing the fractures have been removed.

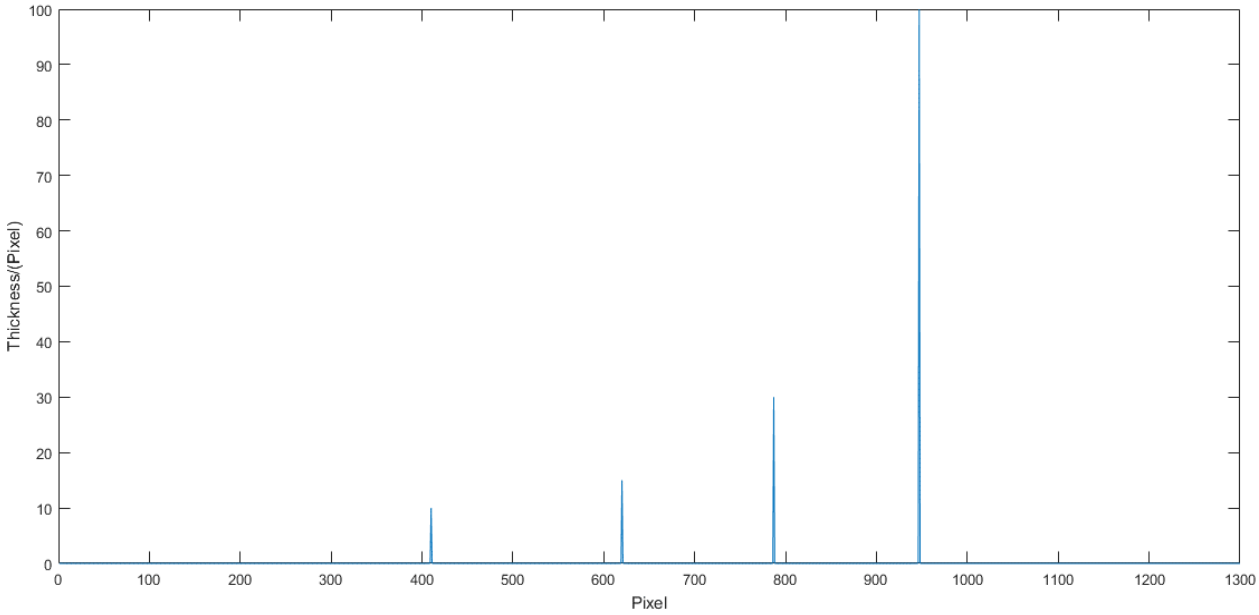

Figure 3.6 Test 1 corrected digitized results, showing lamination thicknesses and center locations.

3.2.2.2 Test 2

The case is composed of two components (Fig. 3.7): matrix and fracture (Table 3.2). The raw digitized result (Fig. 3.8) depicts the curved boundary and gradual boundary of fractures by the irregular gray value change. The corrected digitized results (Fig. 3.9) suggests that the program is effectively correcting the curved and gradual boundary fractures in the matrix.

Figure 3.7 Test 2 case.

Table 3.2 Test 2 components.

	Fig. 3.7 color	Gray Value	Thickness	Boundary
Matrix		150	500 pixels	
Fracture 1		10	39 pixels (Averaged)	Curved Sharp
Fracture 2			18 pixels	Straight Gradual
Fracture 3			27 pixels (Averaged)	Curved Gradual

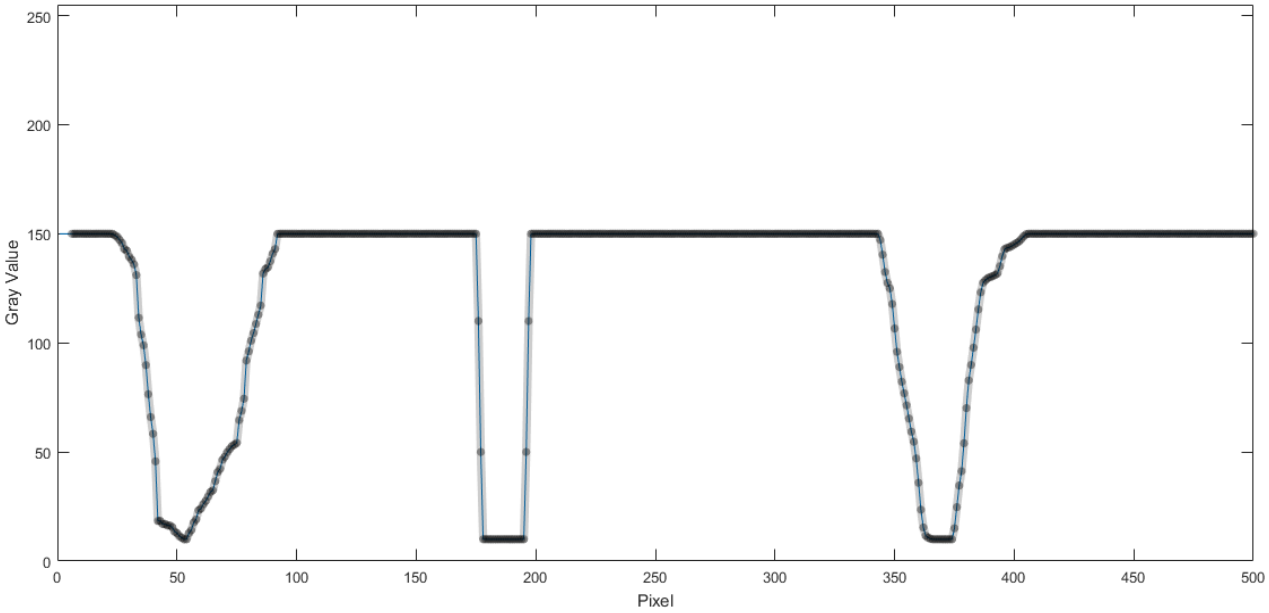

Figure 3.8 Test 2 raw digitized results.

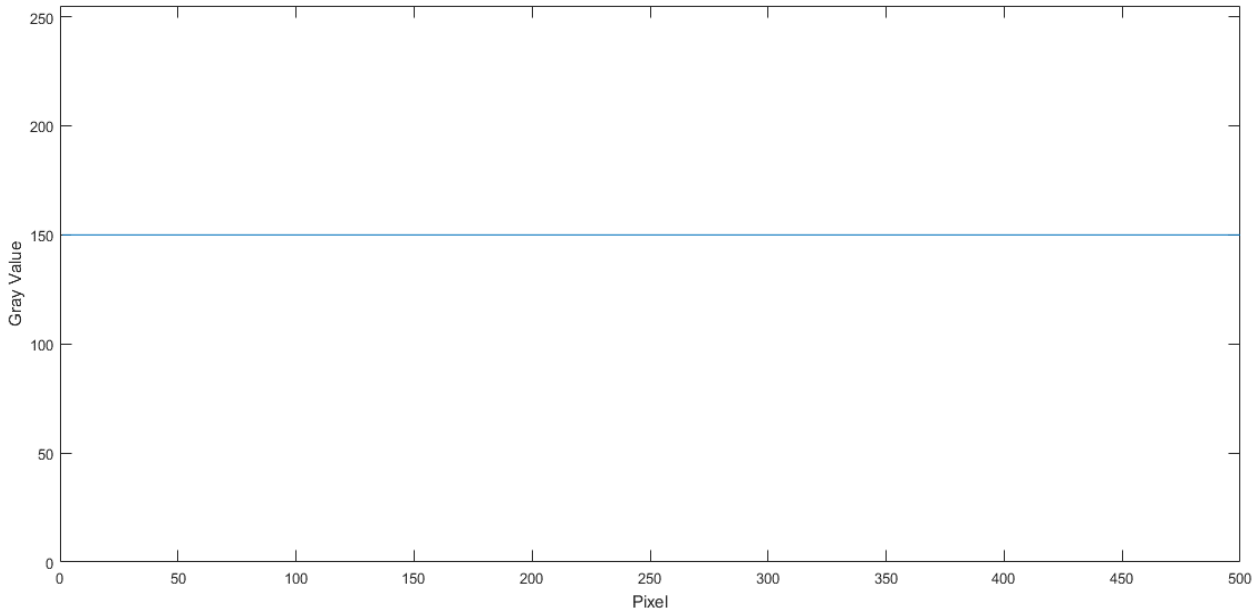

Figure 3.9 Test 2 corrected digitized results showing the fractures have been remove.

3.2.2.3 Test 3

The case is composed of three components (Fig. 3.10): matrix, laminations and fractures (Table 3.3). The results (Figs. 3.11 – 3.13) indicate that the programs are effective to correct the curved and gradual boundary fractures in laminations.

Figure 3.10 10 Test 3 case.

Table 3.3 Test components.

	Fig. 3.10 color	Gray Value	Thickness	Boundary
Matrix		150	500 pixels	
Lamination		100	80, 40 and 70 pixels respectively	Straight Sharp
Fracture 1		10	39 pixels (Averaged)	Curved Sharp
Fracture 2			18 pixels	Straight Gradual
Fracture 3			27 pixels (Averaged)	Curved Gradual

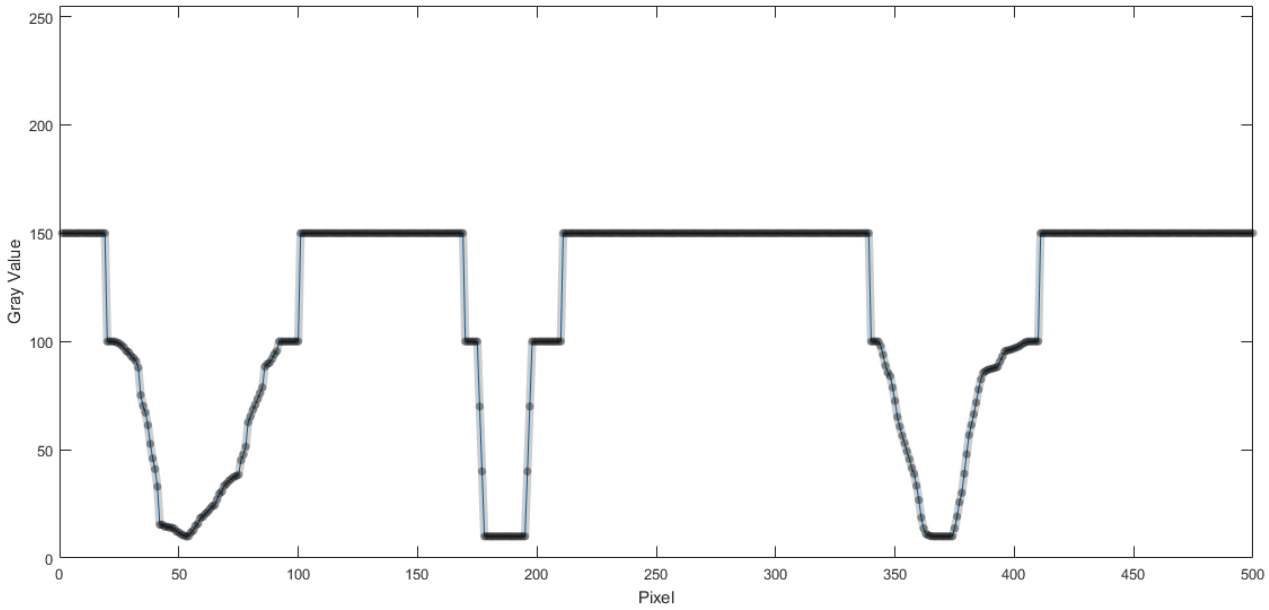

Figure 3.11 Test 3 raw digitized results.

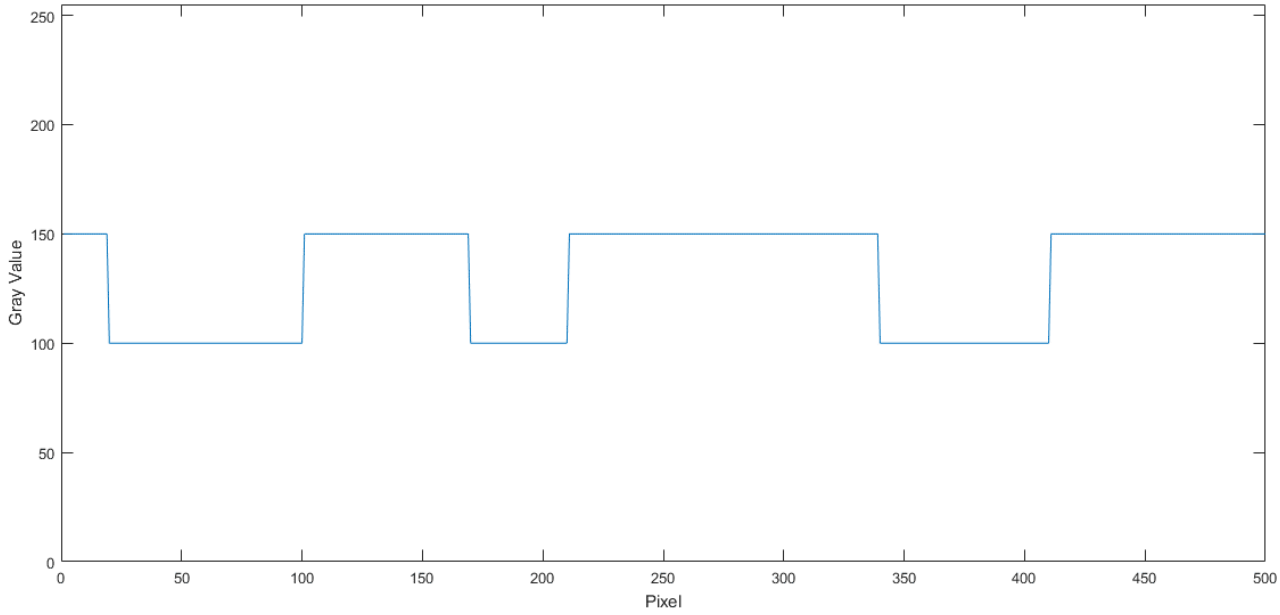

Figure 3.12 Test 3 corrected digitized results showing the fractures have been removed.

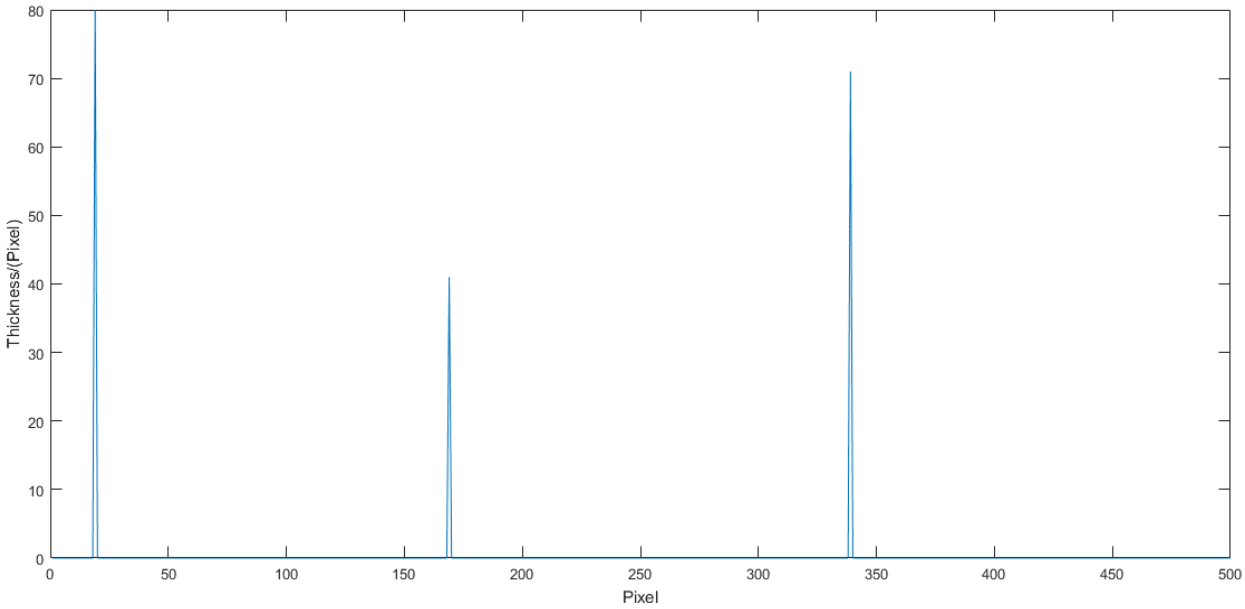

Figure 3.13 Test 3 corrected digitized results, showing lamination thicknesses and center locations.

3.2.2.4 Test 4

The case is composed of two components (Fig. 3.14): matrix and lamination (Table 3.4). The lamination thickness profile (Fig. 3.15) shows that the program can locate the curved and gradual boundary laminations correctly. The program can also approximate the thickness to within 15% (Fig. 3.16).

Figure 3.14 Test 4 case.

Table 3.4 Test 4 components

	Fig. 3.14 color	Gray Value	Thickness	Boundary
Matrix		150	500 pixels	
Lamination 1		10	39 pixels (Averaged)	Curved Sharp
Lamination 2			18 pixels	Straight Gradual
Lamination 3			27 pixels (Averaged)	Curved Gradual

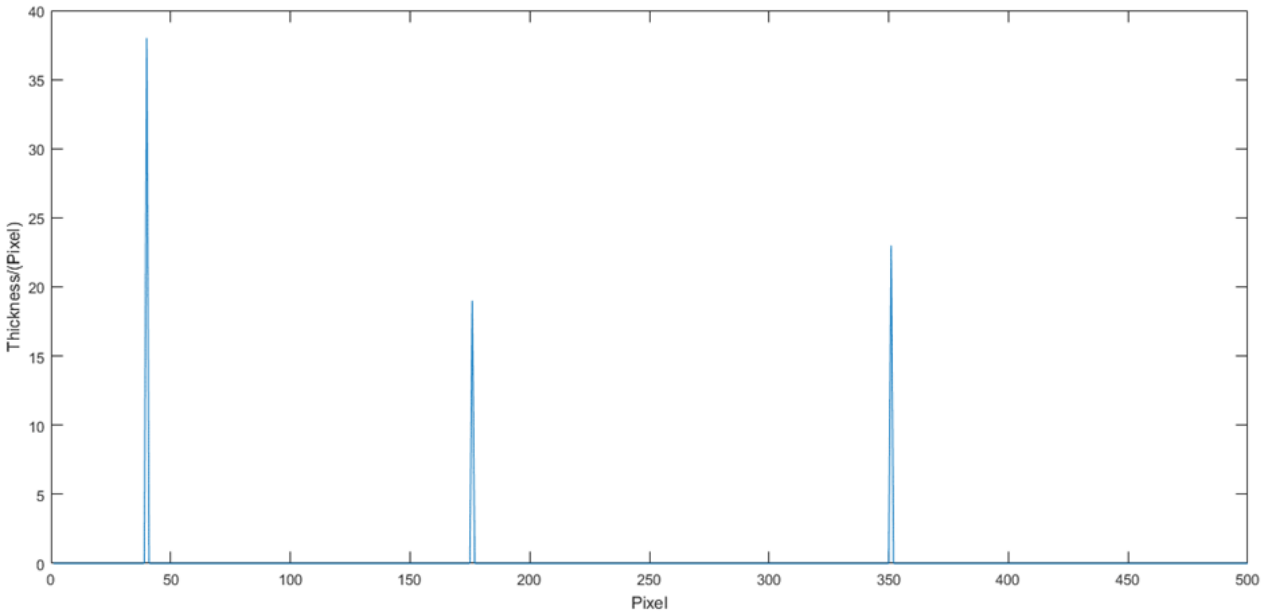

Figure 3.15 Test 4 lamination thickness profile, showing lamination thicknesses and center locations.

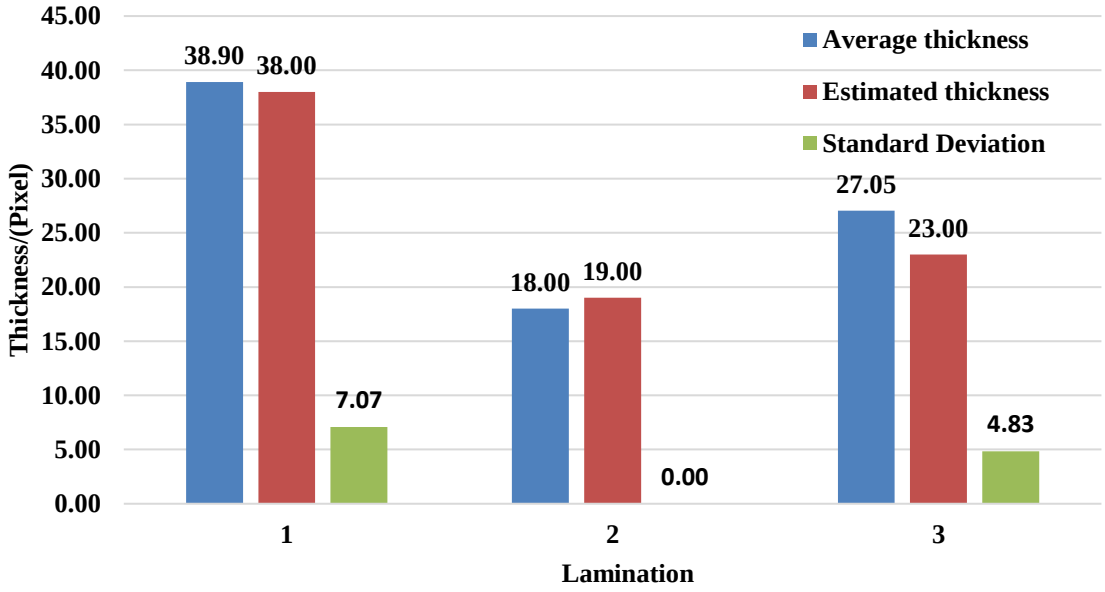

Figure 3.16 Test 4 lamination thickness comparisons.

3.2.2.5 Test 5

The case is a real core slab image composed of matrix, laminations and fractures. The results (Figs. 3.17 – 3.20) indicate that the programs are effective to correct the real case. The fracture effects in the raw results (Fig. 3.18) are removed (Fig. 3.19) and the lamination thickness and location are calculated (Fig. 3.20). But the corrected results also show more variations in the non-fractured zones compared to the raw results (Fig. 3.18 and 3.19) because only the central third of the image is used for the program.

Figure 3.17 Test 5 case.

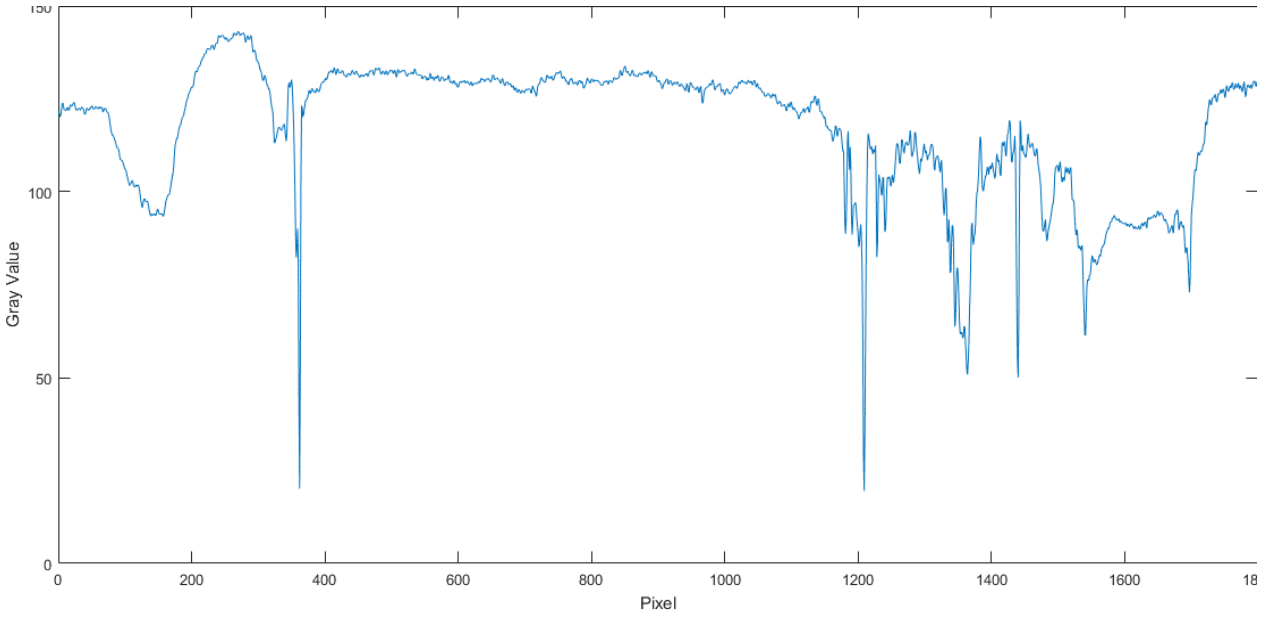

Figure 3.18 Test 5 raw digitized results.

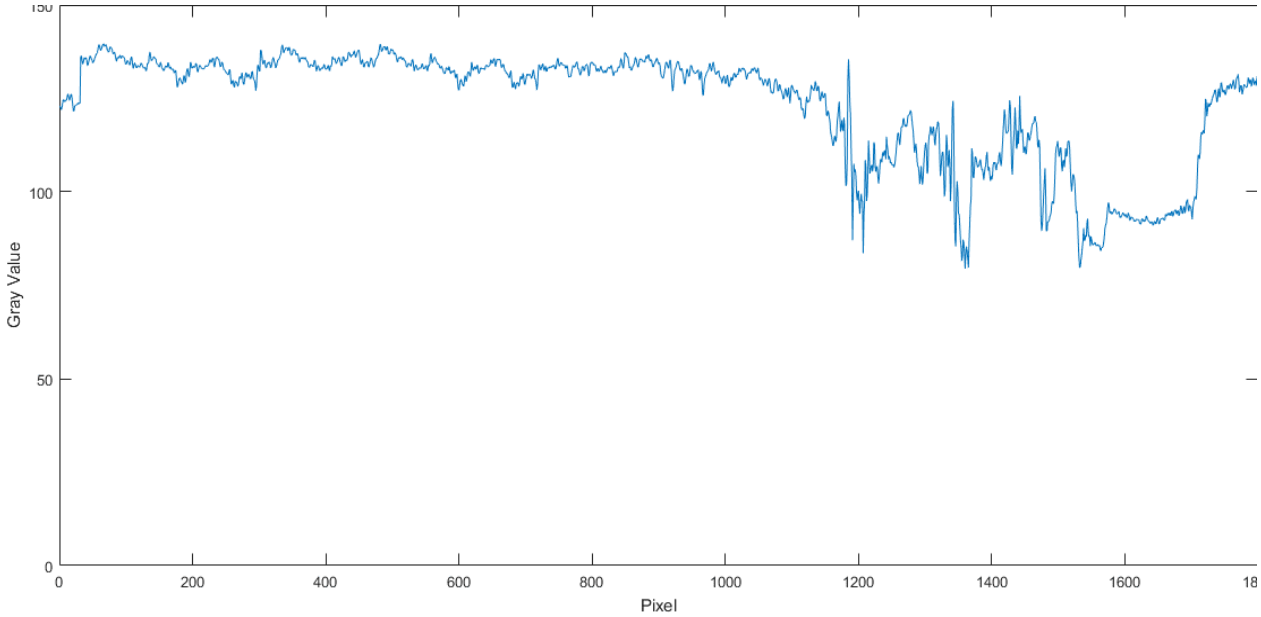

Figure 3.19 Test 5 corrected digitized results showing the fractures have been removed.

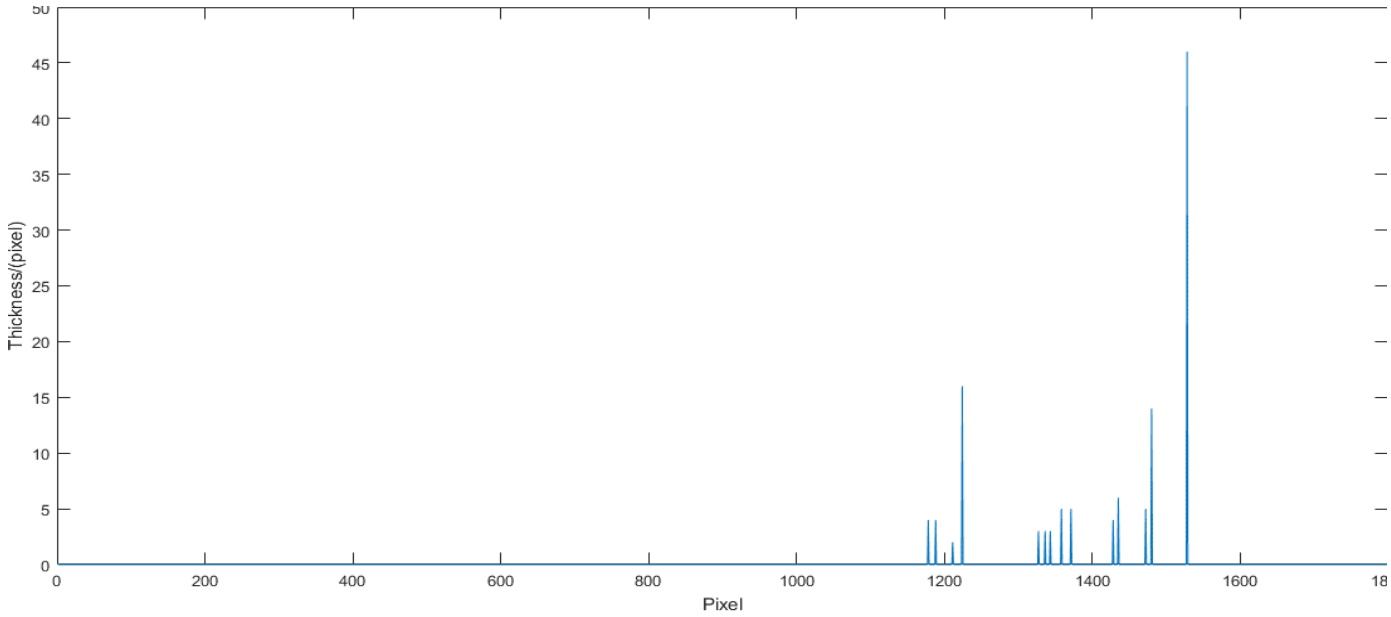

Figure 3.20 Test 5 lamination thickness profile, showing lamination thicknesses and center locations.

3.3 Heterogeneity Characterization

From the last section, the gray value data and lamination data can be accurately derived from the core slab images. In this section, they will be used to compare with the probe permeability data and the mechanical hardness data to investigate the capacity of these measurements on characterizing heterogeneity at different scales.

3.3.1 Comparison between Gray Value and Probe Permeability

For this core, only five intervals, 3040.6 – 3043.7 m, 3080.7 – 3084.0 m, 3135.2 – 3139.1 m, 3140.15 to 3141.3 m and 3143.3 to 3145.9 m, are measured for probe permeability.

3.3.1.1 Interval 3040.6 – 3043.7 m

The gray value and probe permeability for the interval, 3040.6 – 3043.7 m, are plotted (Fig. 3.21). This interval can be divided into 3 sub-intervals: Interval 1 (relatively homogeneous), Interval 2 (relatively heterogeneous) and Interval 3 (homogeneous). The lamination profile and the lamination density plot show that Interval 2 has a dense distribution of laminations while Intervals 1 and 3 have much fewer laminations (Figs. 3.22 and 3.23).

No is the recommended number of measurements based on the C_V of probe permeability data, where $No = (10C_V)^2$ (Jensen et al., 2000). No assumes one wants to estimate the arithmetic mean permeability of the interval to within ±20% with 95% probability. N is the actual number of measurements. For the homogeneous intervals, probe permeability and gray value have similar distributions, and No < N, which means these intervals are oversampled. For the heterogeneous interval, probe permeability does not fully characterize the heterogeneity, and No > N, which means the interval is under-sampled at the 20% tolerance but is adequately sampled for a 24% tolerance. This might result from the dominant lamination scale heterogeneities in Interval 2, where the lamination density is much larger than the measurement sample density (Fig. 3.23).

The semivariograms of the gray value and probe permeability data are also plotted (Figs. 3.24 - 3.27). For the gray value data, the correlation length is 0.12m with approximately 0.25 m cycles which can be identified on the core images (Figs. 3.24 and 3.28). These cycles mostly come from Interval 2, so the semivariogram of the gray value data in Interval 2 is separately plotted in Fig. 3.29, which also shows similar correlation length and cycles as the previous. For the probe permeability data, the correlation length is over 1.6 m and those cycles are missed (Figs. 3.25 and 3.26). There is a clear trend for probe permeability data, so the de-trended data are also calculated, which still misses the cycles (Fig. 3.27). From the comparison, the probe permeability data miss the cycles and cannot capture the lamination scale heterogeneities.

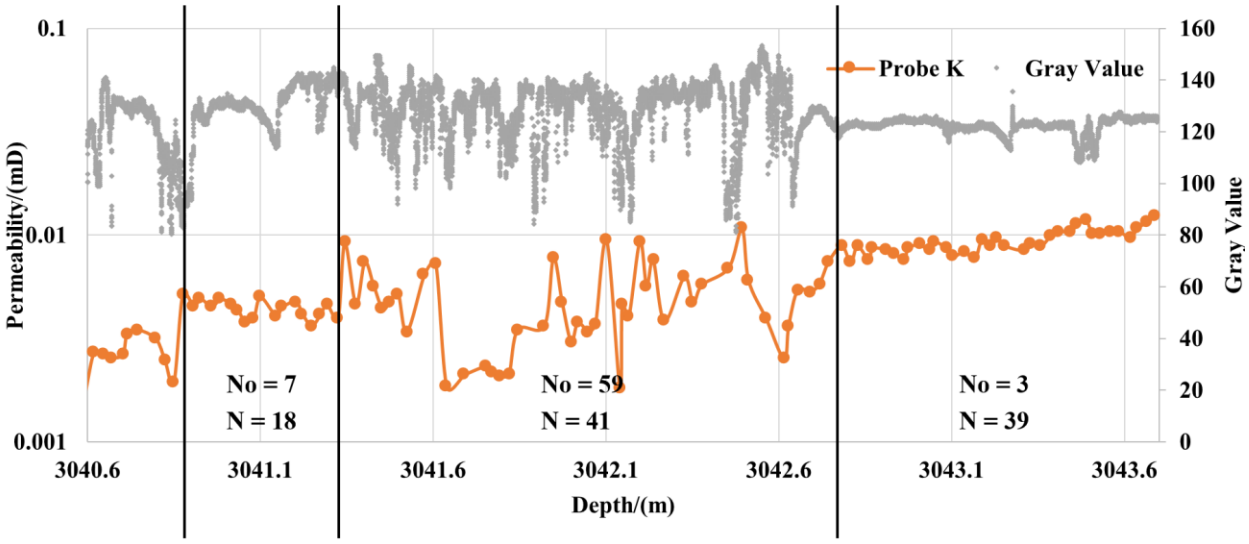

Figure 3.21 Corrected gray value plot of 3040.6 to 3043.7 m interval. Two of the three intervals have been oversampled.

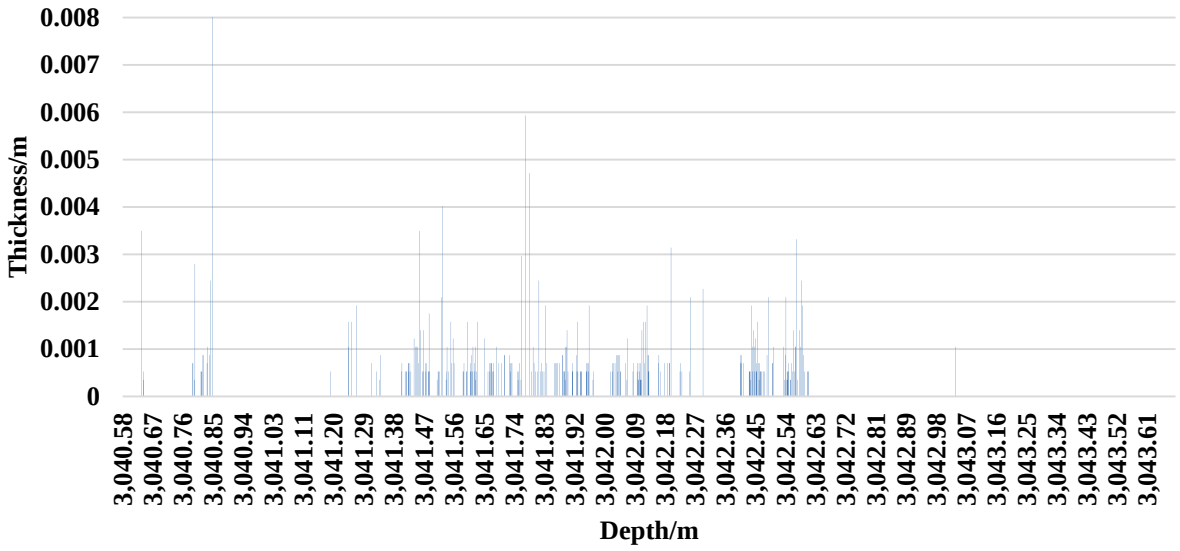

Figure 3.22 Lamination thickness profile of 3040.6 to 3043.7 m interval.

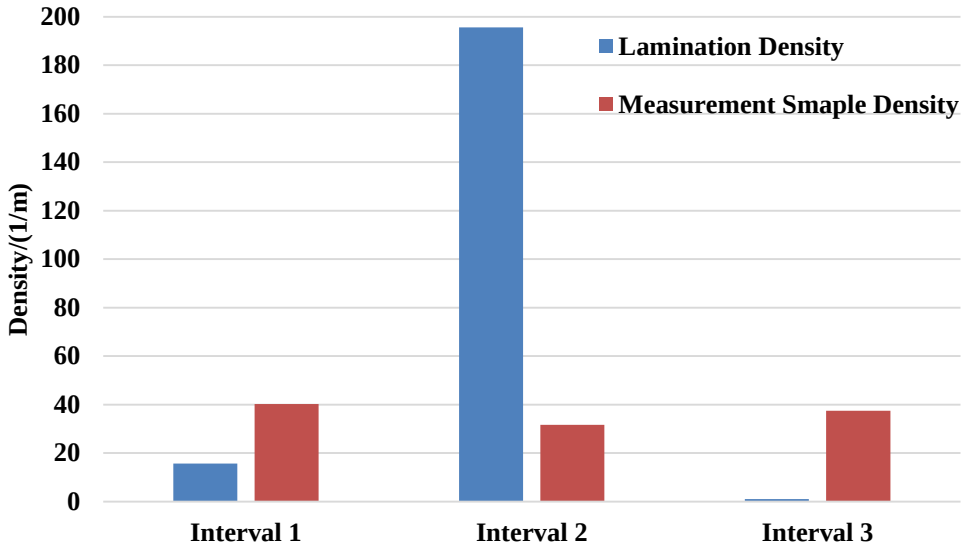

Figure 3.23 Lamination density and measurement sample density comparison.

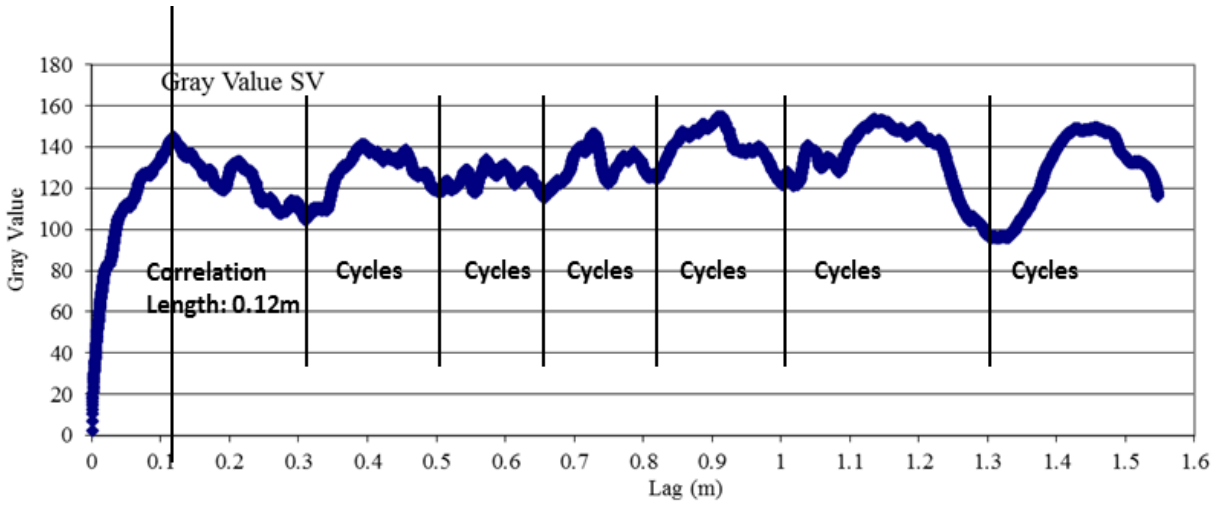

Figure 3.24 Semivariogram of gray value data in the interval from 3040.6 to 3043.7 m.

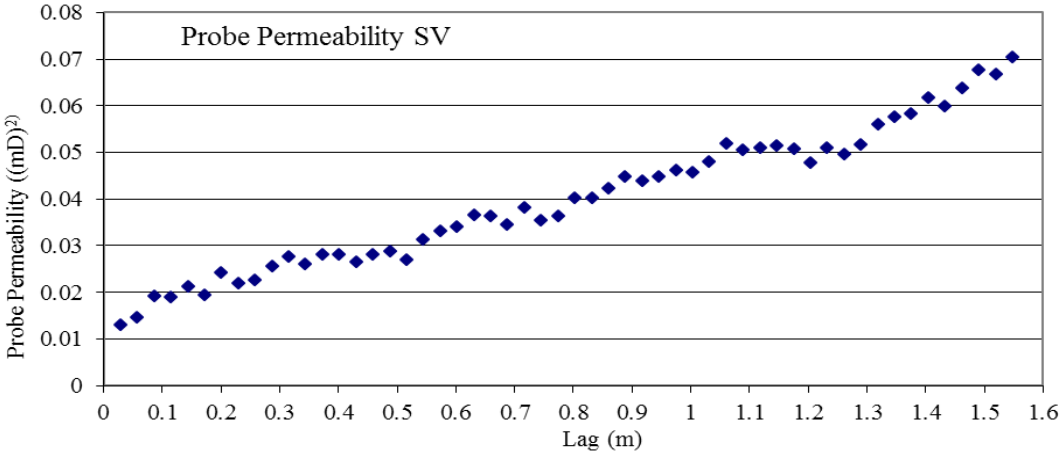

Figure 3.25 Semivariogram of probe permeability data in the interval from 3040.6 to 3043.7 m.

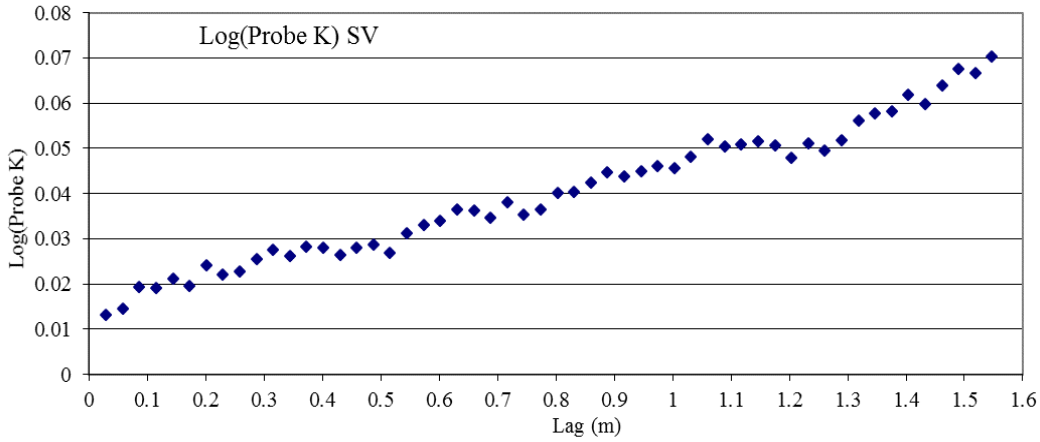

Figure 3.26 Semivariogram of Log(Probe K) data in the interval from 3040.6 to 3043.7 m.

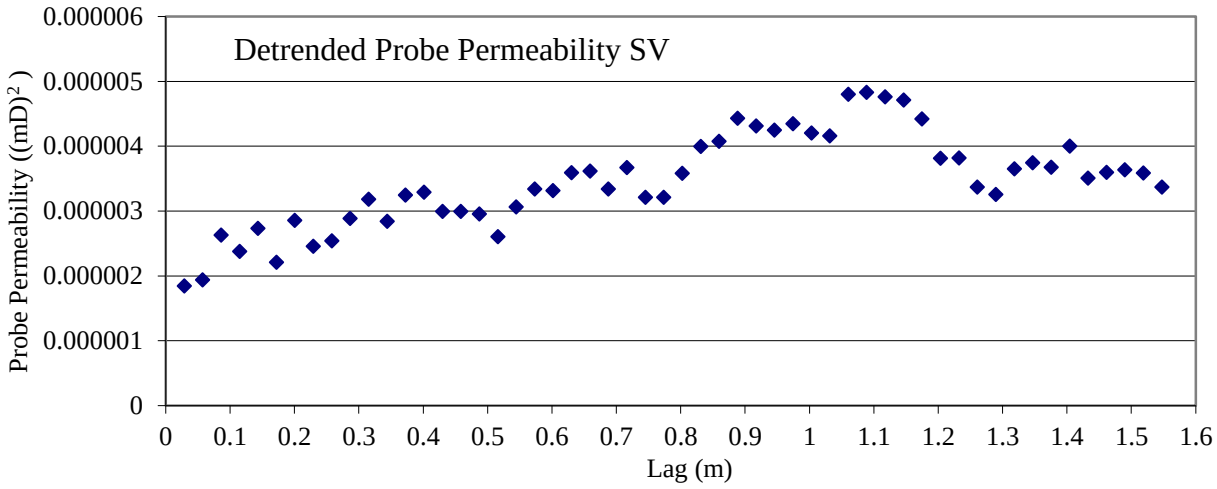

Figure 3.27 Semivariogram of de-trended probe permeability data in the interval from 3040.6 to 3043.7 m.

Figure 3.28 Core slab image from 3040.6 to 3043.7 m with 0.25 m cycles identified in red circles.

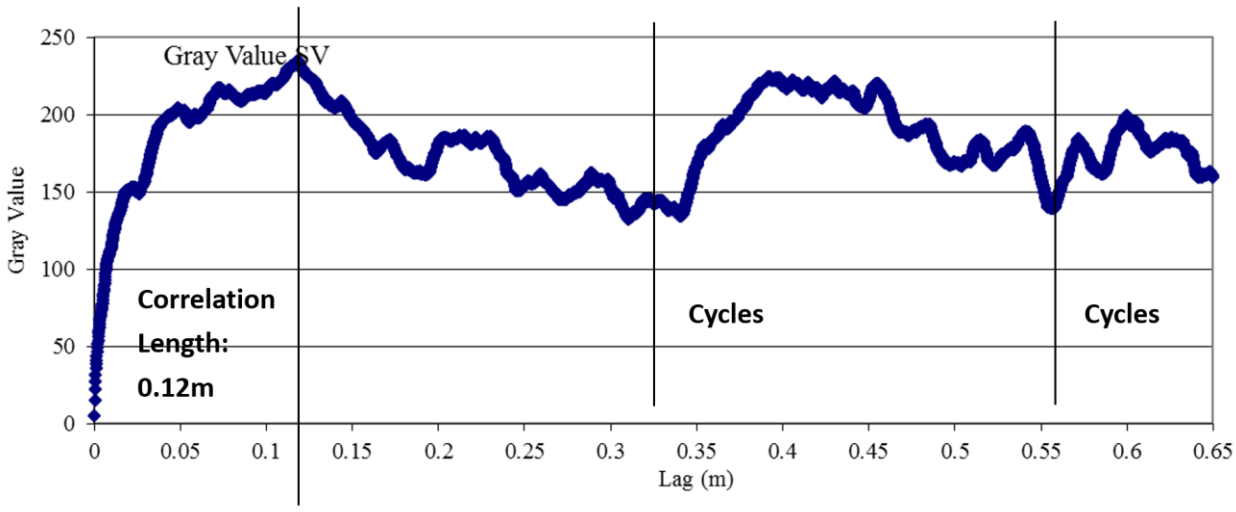

Figure 3.29 Semivariogram of gray value data in Interval 2.

3.3.1.2 Interval 3080.7 – 3084.0 m

The interval, 3080.7 – 3084.0 m, is characterized as the very fine grain, laminated shaly siltstone (Figs. 3.30 and 3.31). Semivariograms of the gray value data and probe permeability data show similar patterns with small scale cycles (Figs. 3.32 - 3.34). These small scale cycles result from the dense distribution of the laminaset scale heterogeneities (Fig. 3.30).

Figure 3.30 Core slab image from 3080.7 to 3084.0 m.

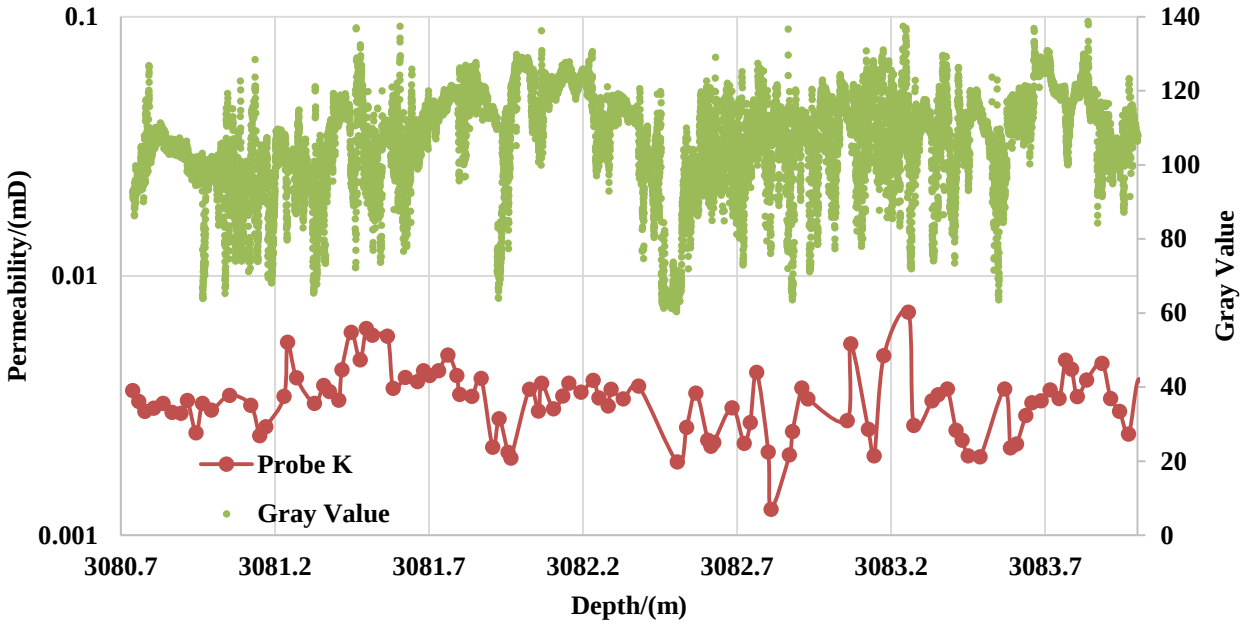

Figure 3.31 Corrected gray value plot of 3080.7 to 3084.0 m interval.

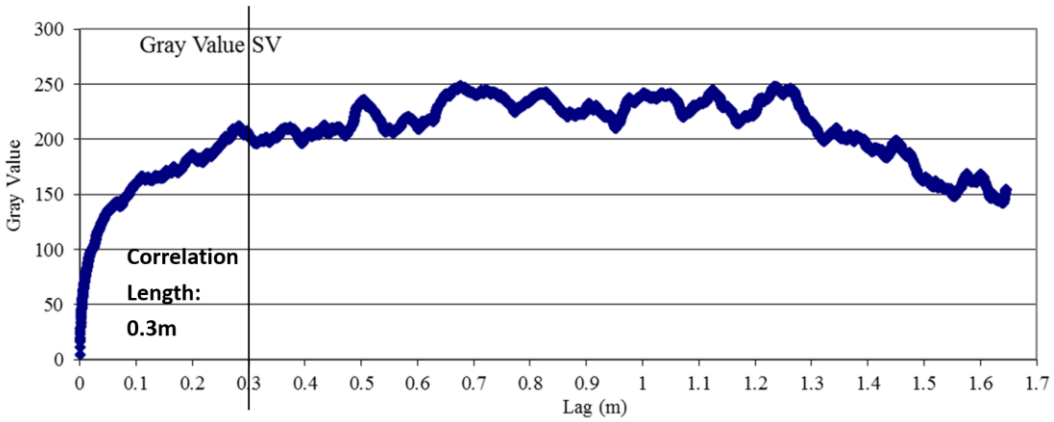

Figure 3.32 Semivariogram of gray value data in the interval from 3080.7 to 3084.0 m.

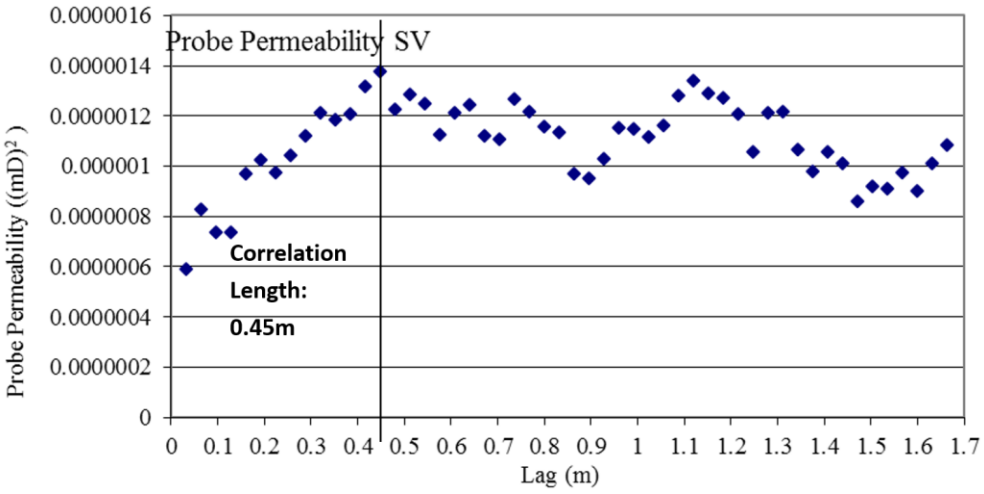

Figure 3.33 Semivariogram of probe permeability data in the interval from 3080.7 to 3084.0 m.

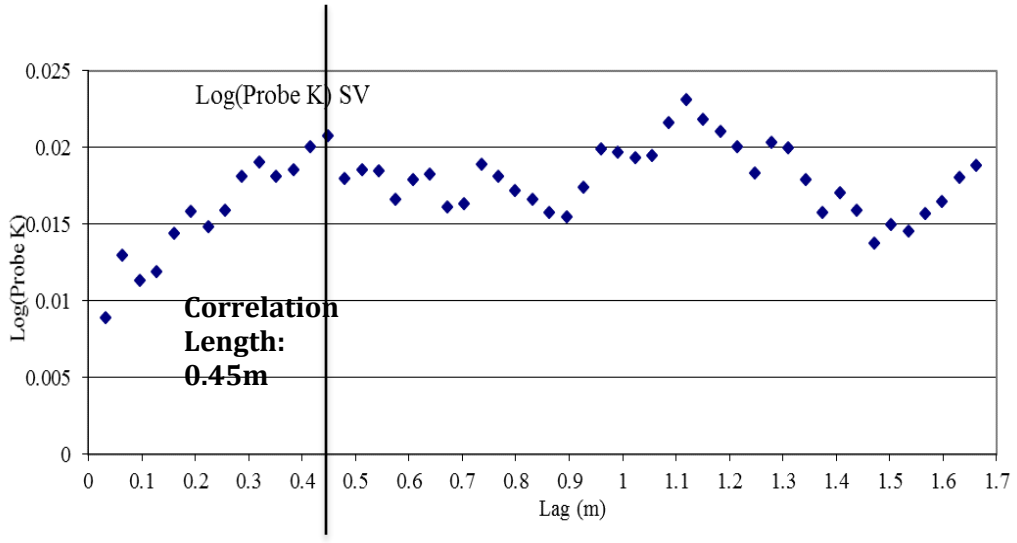

Figure 3.34 Semivariogram of Log(Probe K) in the interval from 3080.7 to 3084.0 m.

3.3.1.3 Interval 3135.2 to 3139.1 m

The interval from 3135.2 to 3139.1 m located at the bottom of the upper Montney is digitized (Fig. 3.35). It consists of clean siltstone with 1 to 3 cm laminations (Figs. 3.36 and 3.37). From the semivariogram of the gray value data, approximate 0.2 m cycles can be identified with a 0.06 m correlation length (Fig. 3.38). These cycles are caused by the laminations in the core (Fig. 3.36). The semivariogram of the probe permeability data can also capture these cycles with a 0.5 m correlation length (Fig. 3.39).

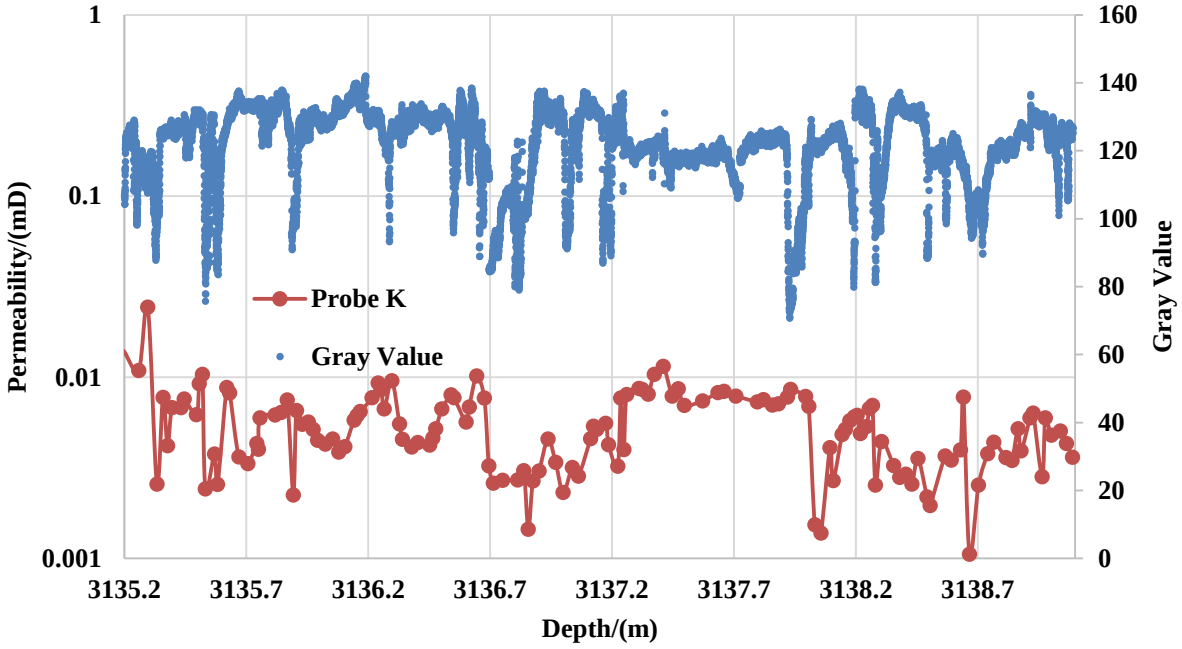

Figure 3.35 Corrected gray value plot of 3135.2 to 3139.1 m interval.

Figure 3.36 Core slab image from 3135.2 to 3139.1 m with 1 to 3 cm laminations.

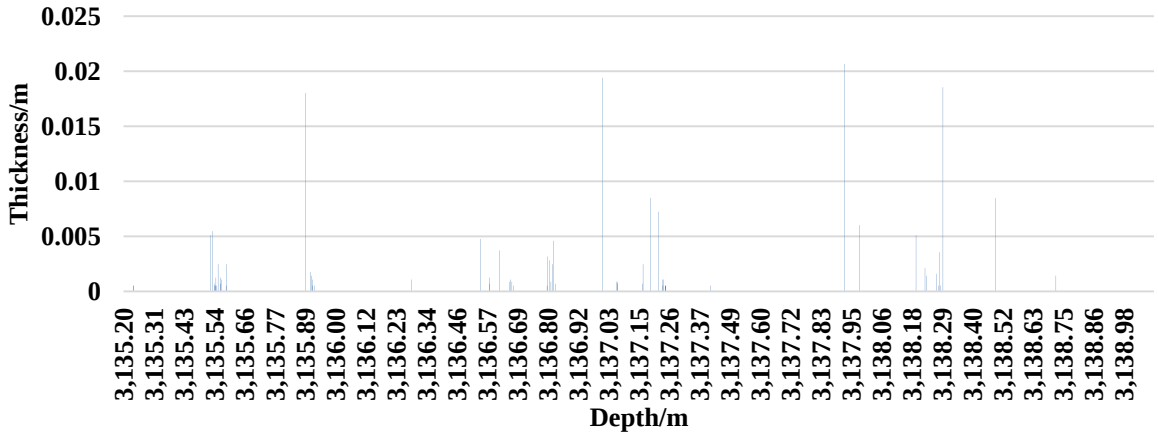

Figure 3.37 Lamination thickness profile of 3135.2 to 3139.1 m interval.

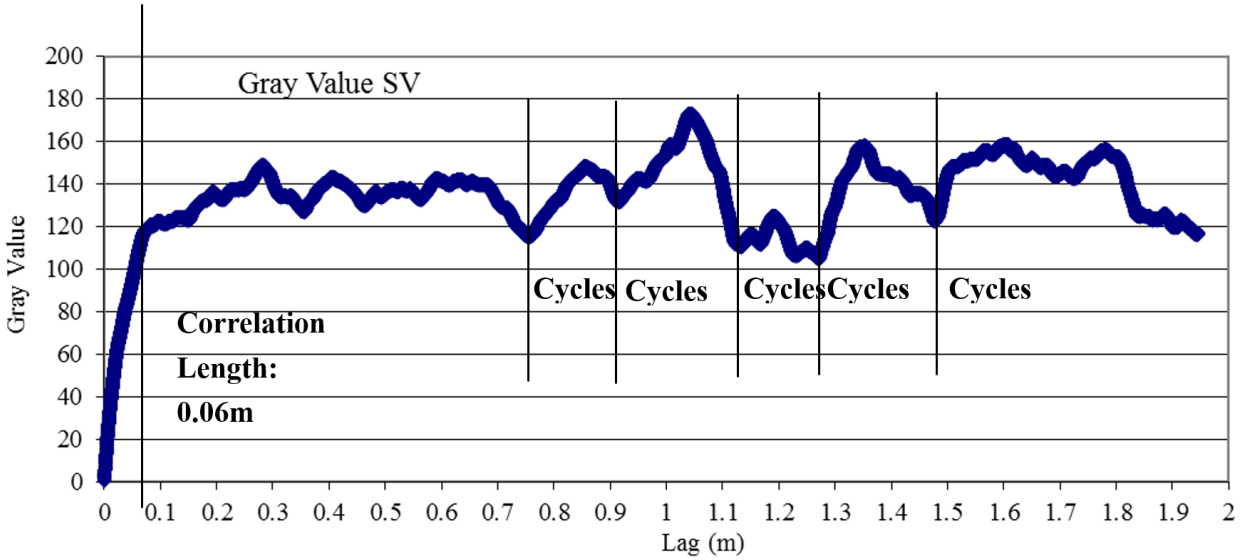

Figure 3.38 Semivariogram of gray value data in the interval from 3135.2 to 3139.1 m.

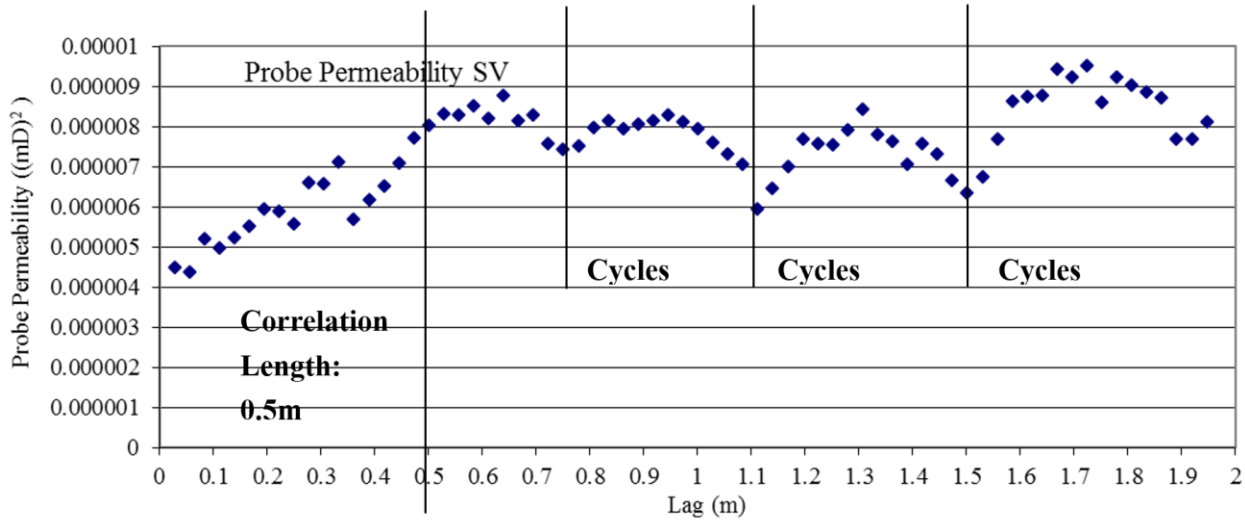

Figure 3.39 Semivariogram of probe permeability data in the interval from 3135.2 to 3139.1 m.

3.3.1.4 Interval 3140.15 to 3141.3 m

The interval from 3140.15 to 3141.3 m located at the bottom of the upper Montney is digitized (Fig. 3.40). It is very clean siltstone (Figs. 3.41 and 3.42). The distributions of gray value data and probe permeability data are very similar (Fig. 3.40), and the semivariograms also agree with each other (Figs. 3.43 and 3.44).

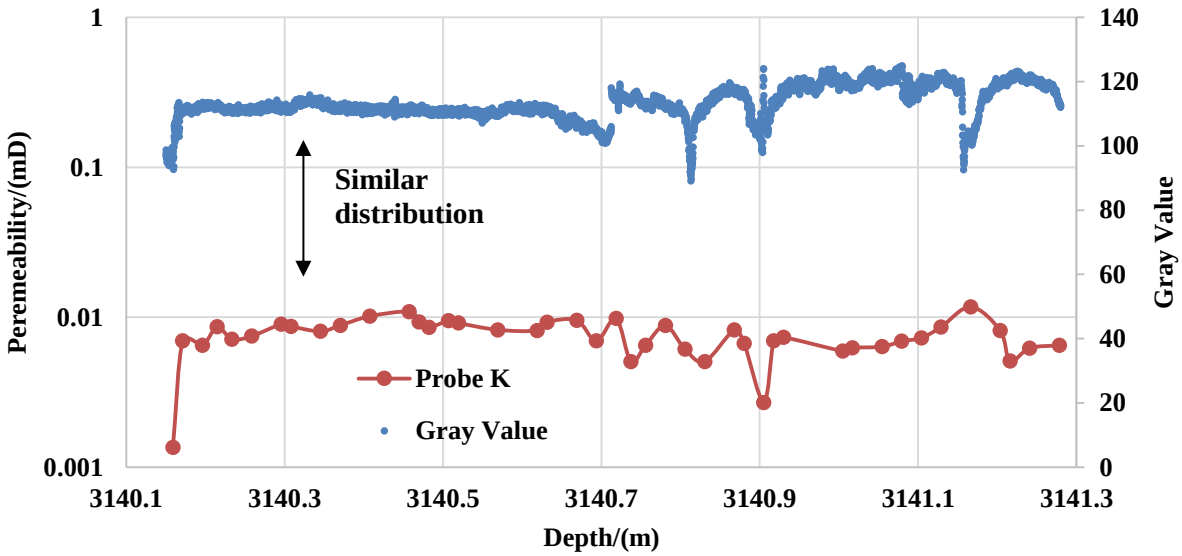

Figure 3.40 Corrected gray value plot of 3140.15 to 3141.3 m interval.

Figure 3.41 Core slab image from 3140.15 to 3141.3 m.

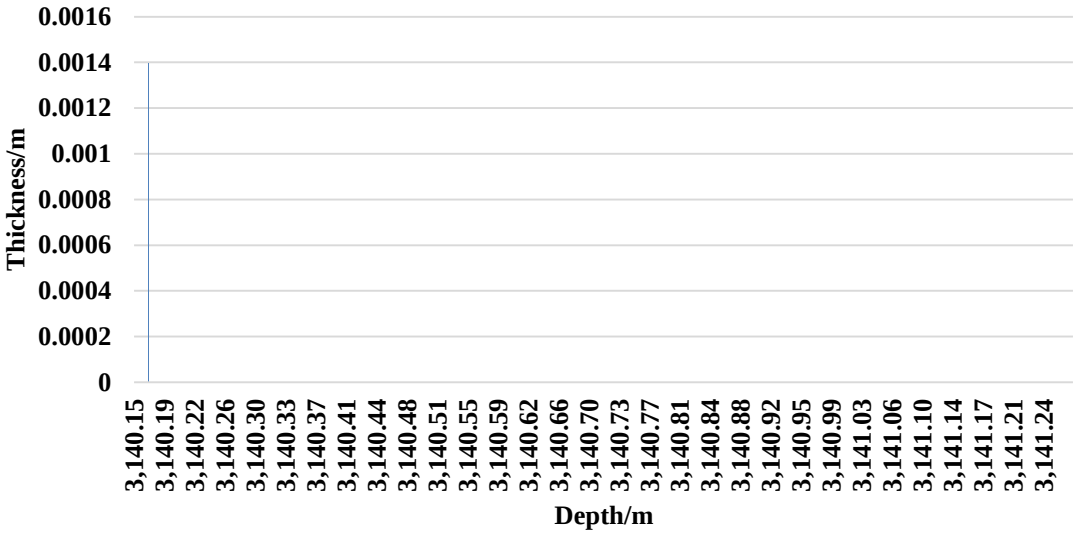

Figure 3.42 Lamination thickness profile of 3140.15 to 3141.3 m interval.

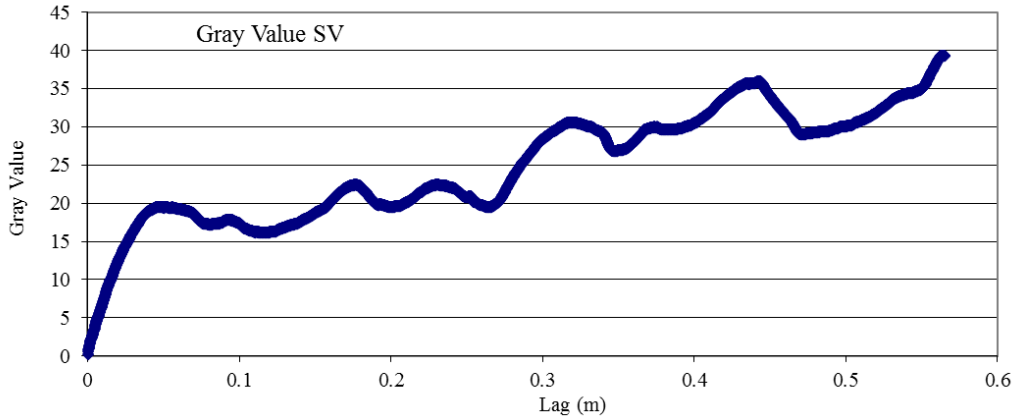

Figure 3.43 Semivariogram of gray value data in the interval from 3140.15 to 3141.3 m.

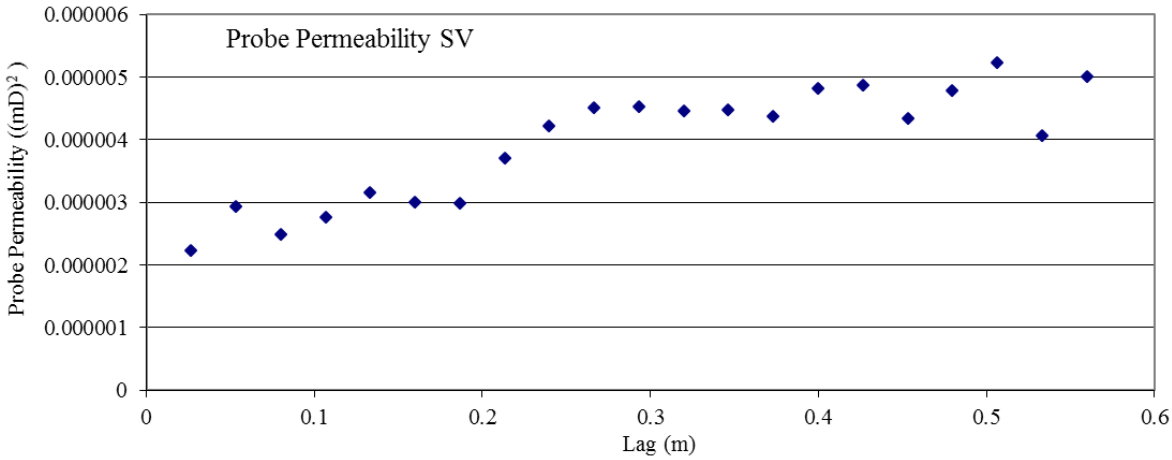

Figure 3.44 Semivariogram of probe permeability data in the interval from 3140.15 to 3141.3 m.

3.3.1.5 Interval 3143.3 to 3145.9 m

The interval from 3143.3 to 3145.9 m located at the bottom of the upper Montney is digitized (Fig. 3.45). It is finely laminated siltstone with dense distribution of thin laminations (Figs. 3.46 and 3.47).

The relationship of lamination number and lamination thickness is investigated based on the gray value data (Fig. 3.48). Most of the laminations are in mm-scale which are thinner than 2.5 cm, the averaged sampling spacing of probe permeability (Fig. 3.48), so they cannot be fully characterized by probe permeability measurement. Large lamination density will affect the characterization capacity of probe permeability. Take Interval 3135.2 to 3139.1 m, Interval 3140.15 to 3141.3 m and Interval 3143.3 to 3145.9 m for example. The lamination density of Interval 3143.3 to 3145.9 m is much larger than its sample density, so probe permeability measurements may miss a large amount of heterogeneity (Fig. 3.49). The No is also much larger than N for this interval, which means it is highly under-sampled (Fig. 3.50). The other two intervals have relatively small lamination density compared to the sample density and they are both adequately sampled at the 20% tolerance (Figs. 3.49 and 3.50).

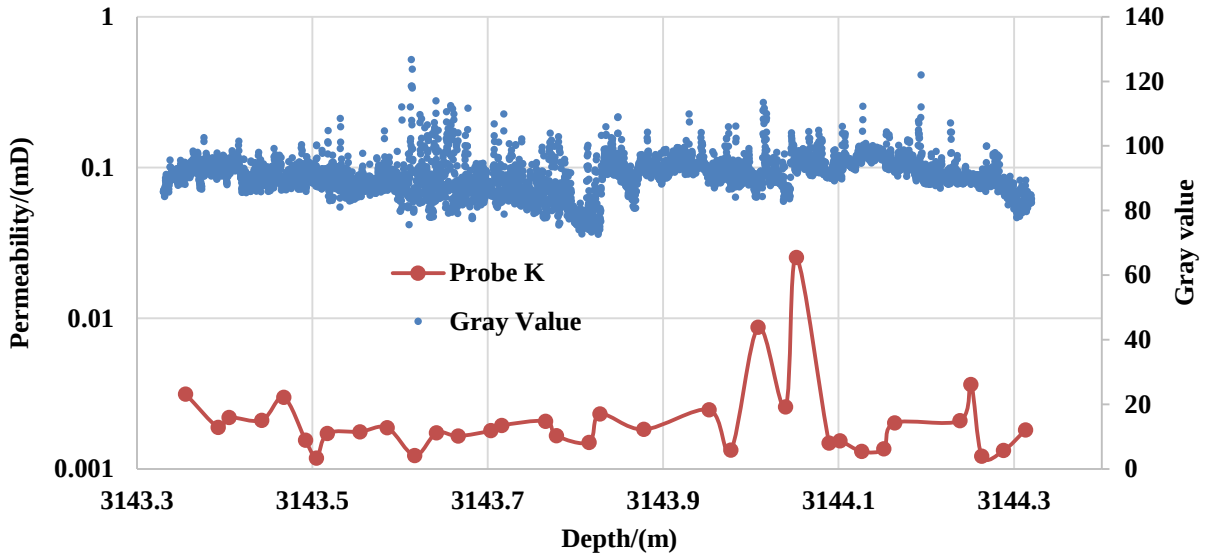

Figure 3.45 Corrected gray value plot of 3143.3 to 3145.9 m interval.

Figure 3.46 Core slab image from 3143.3 to 3145.9 m.

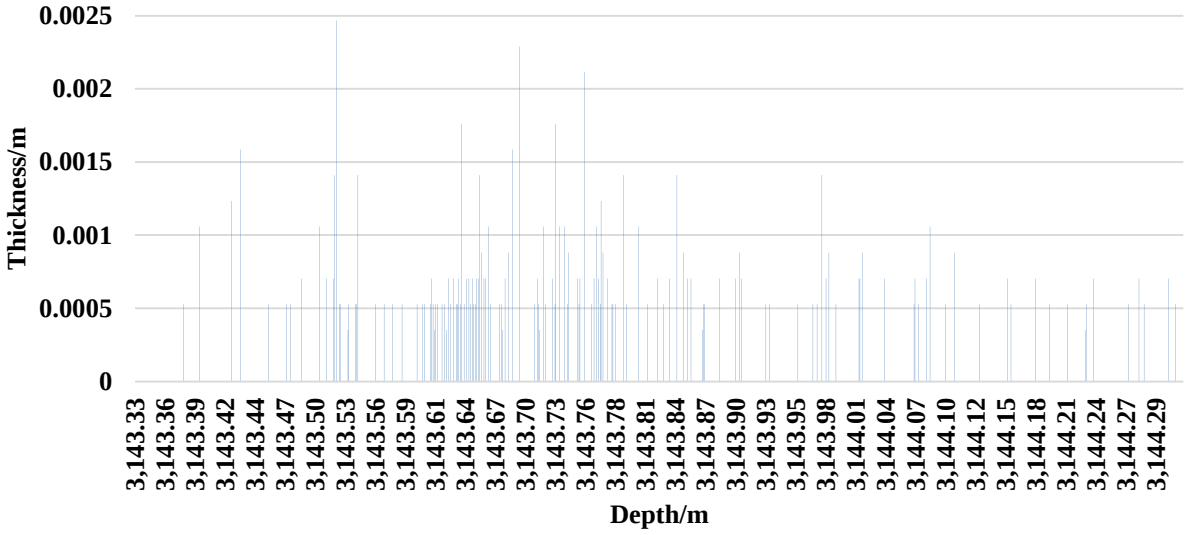

Figure 3.47 Lamination thickness profile of 3143.3 to 3145.9 m interval.

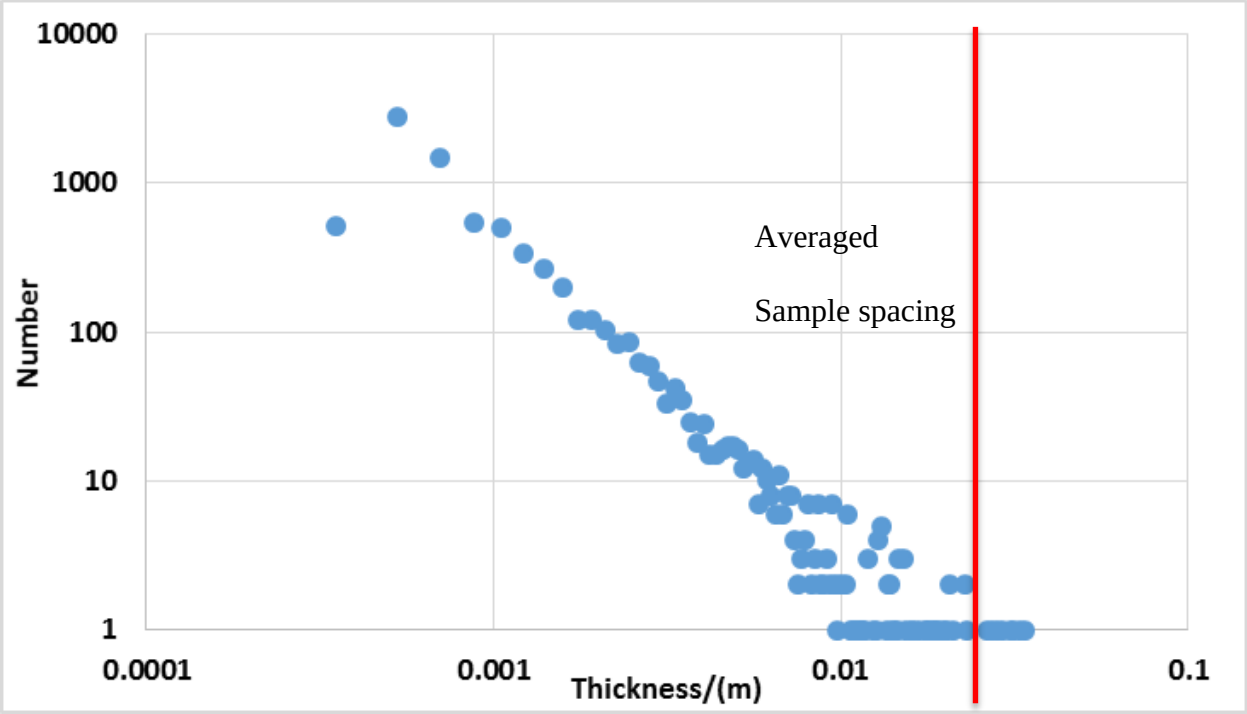

Figure 3.48 Relationship of lamination number and lamination thickness.

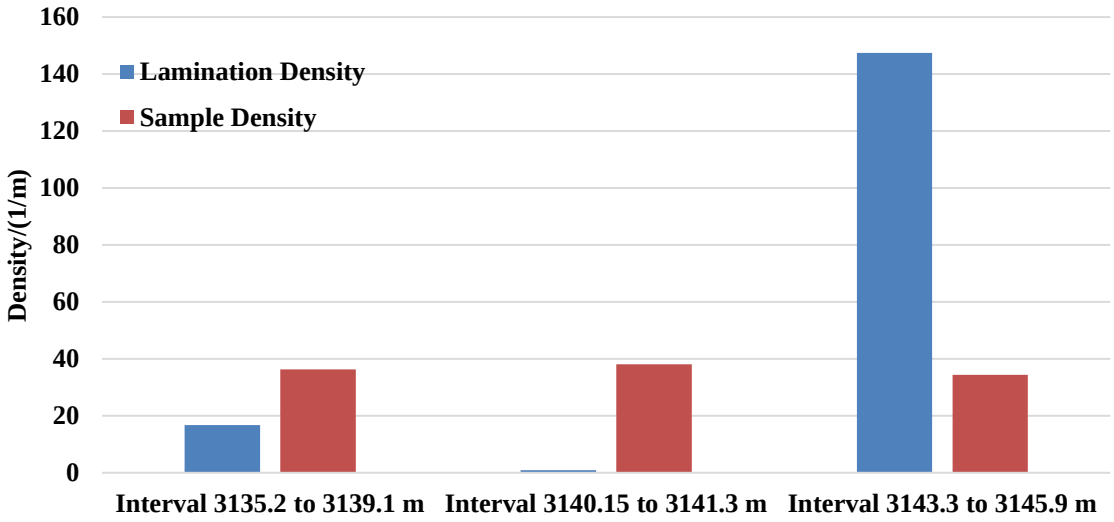

Figure 3.49 Lamination density and sample density comparison.

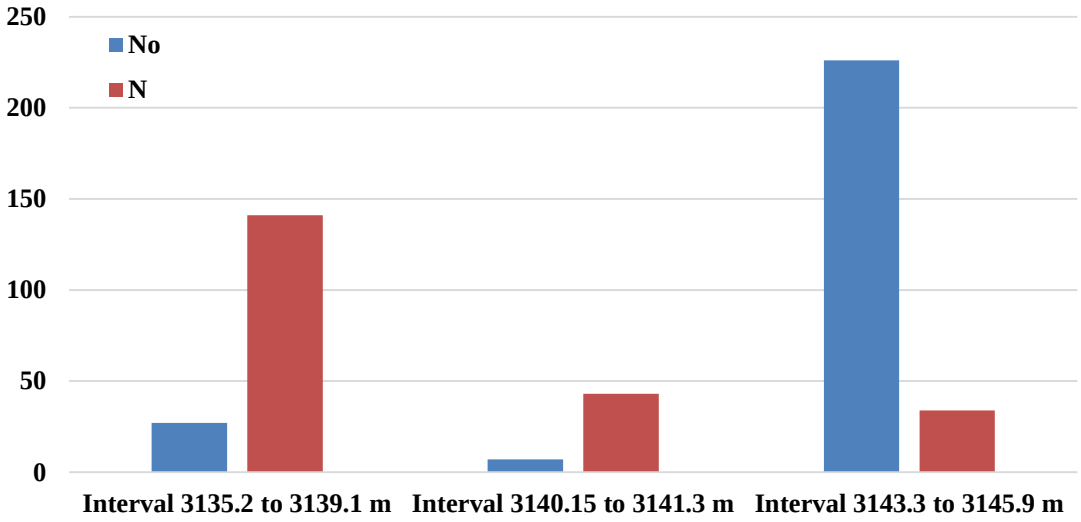

Figure 3.50 No and N comparison.

3.3.2 Comparison of Gray Value and Mechanical Hardness

From the previous study, the core slab images from the Montney have been digitized into gray value data with necessary adjustments and corrections. Mechanical hardness tests have been performed at the same core slabs using a portable Equotip Piccolo 2 hardness tool (Proceq SA, Switzerland) (Ghanizadeh et al, 2015). The mechanical hardness data have been collected at approximately 2.5-cm (1 inch) intervals with the diameter of the impact area of about 1 mm. This test can differentiate and quantify the mechanical properties of cm-scale heterogeneities in rocks, which are attributed to the variations in the mineralogy and pore network morphology/structure of the analyzed core slab (Ghanizadeh et al, 2015). Meanwhile, the gray value data can also quantitatively characterize the heterogeneities in the mineralogy because the color variations in minerals are captured during the digitization process. So the correlation between mechanical hardness and gray value data are investigated here.

The comparisons of gray value, hardness and core plug in different intervals are shown in Figs. 3.51 and 3.52. In these figures, the argillaceous laminations appear dark gray or black in color in core slab images and low gray value in the plots.

The cm-scale laminations in the black circles can be clearly captured by the low mechanical hardness (Fig. 3.51). In the heterogeneous part, the laminations are in mm-scale which are smaller than the sampling resolution so the mechanical hardness cannot fully characterize them (Fig. 3.51). In the homogeneous part, the hardness values are relatively constant.

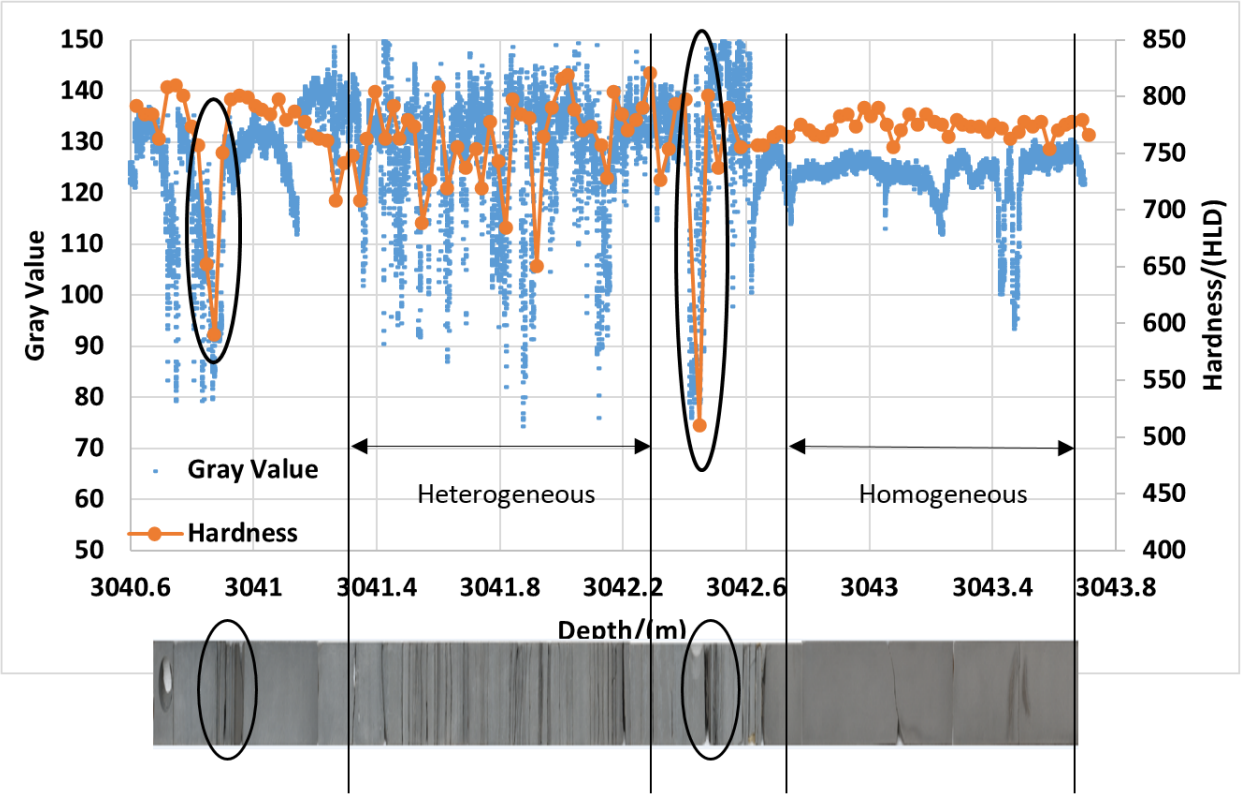

Figure 3.51 Comparison of gray value, hardness and core plug of 3040.6 to 3043.8m interval.

Fig. 3.52 shows a relatively homogeneous interval with cm-scale laminations. The mechanical hardness values correlate well with the gray value data (Fig. 3.52).

A point-by-point correlation was conducted after necessary depth shifts. Because the diameter of the impact area is about 1 mm, the gray value data within 1 mm distance to the measurement locations are averaged for the correlation (Fig. 3.53).

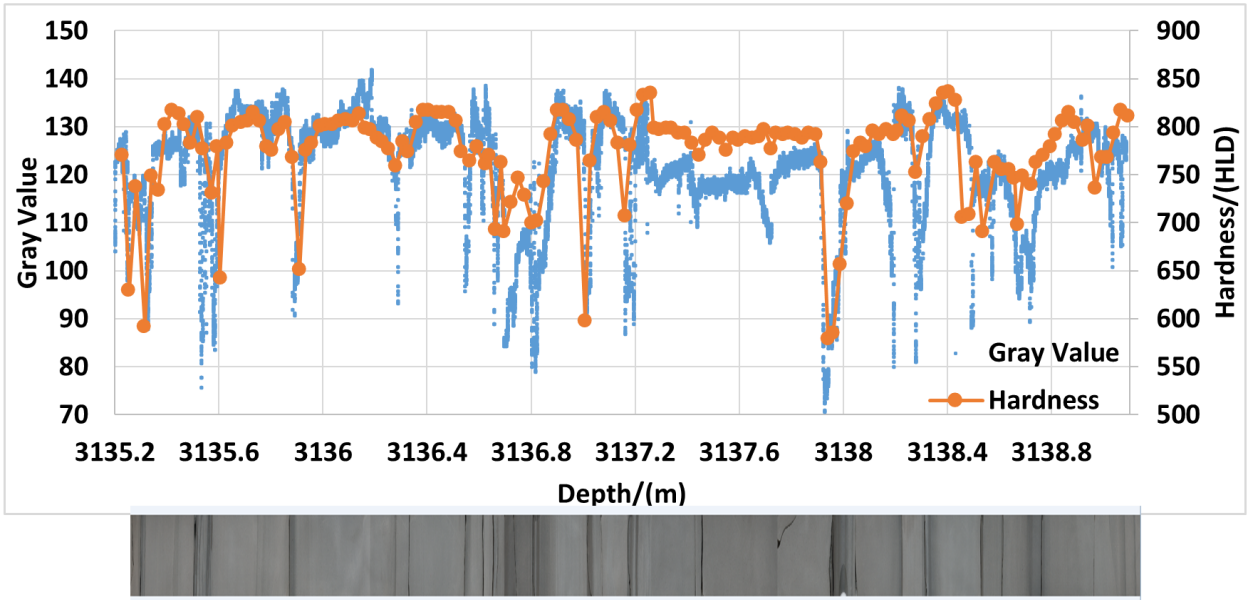

Figure 3.52 Gray value and hardness values compare favorably for this lower Montney interval, differing by about 2.5%.

Fig. 3.53 shows a linear trend between hardness and gray value. But when the gray value is below 100, most of the hardness values are overestimated using this relationship. So a 100 gray value is used as a threshold to divide the core into two subunits. For each subunit, the linear relationship is redeveloped and they are combined together to predict the hardness. The final results have an average error percentage of 2.4% (Fig. 3.54). So the final relationship is used on the gray value data from 3135.2 to 3139.1 m (Fig. 3.52). By this method, the hardness value can be predicted relatively accurately at a much smaller scale, which captures the missing characterization of heterogeneities due to the sampling resolution issues.

After necessary adjustments and comparison, a linear relationship is developed between hardness and gray value data. This relationship predicts relatively accurate hardness values on the interval with cm-scale heterogeneities.

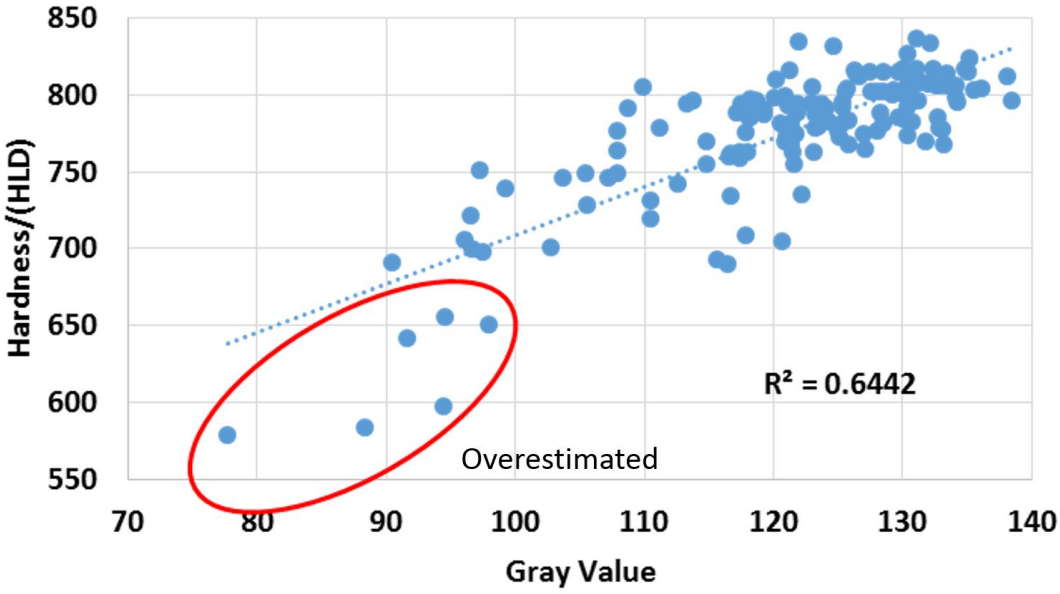

Figure 3.53 Correlation between hardness and gray value data.

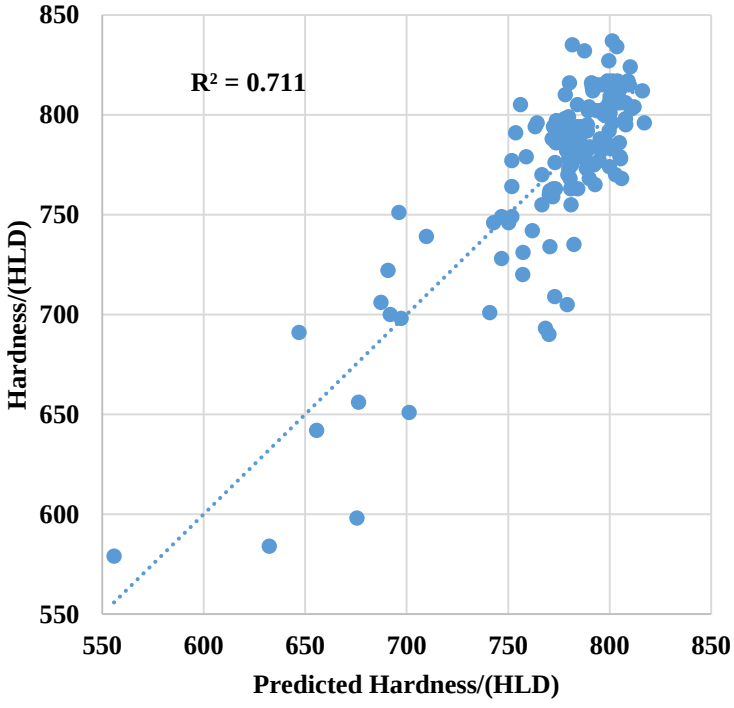

Figure 3.54 Hardness prediction results.

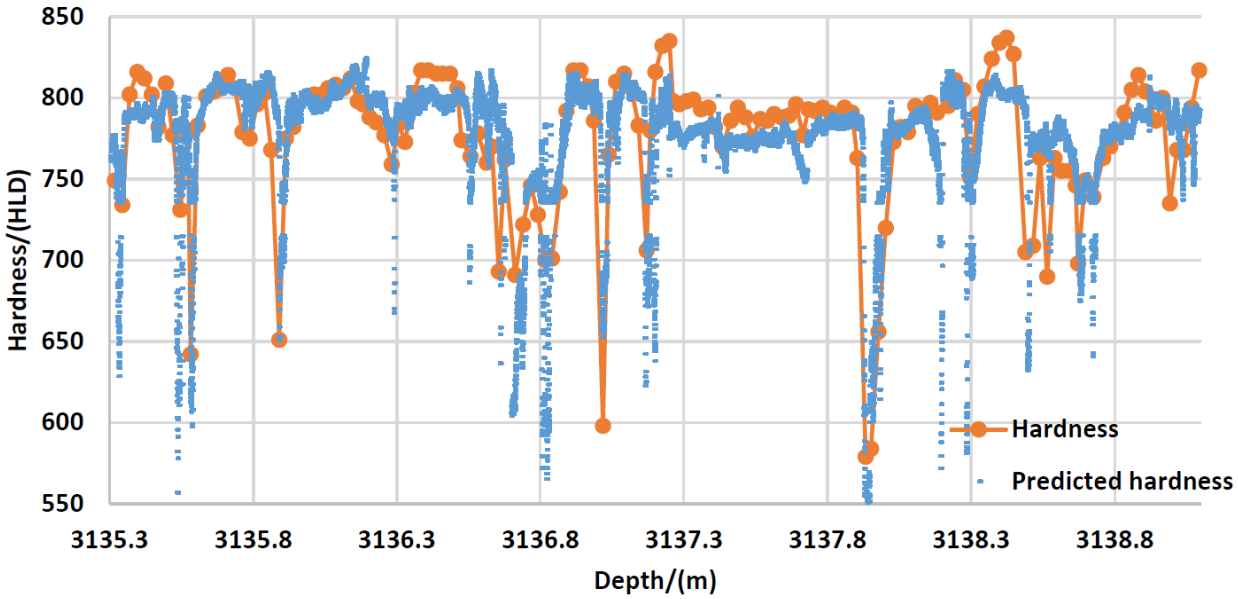

Figure 3.55 Hardness prediction results.

3.4 Discussion and Conclusions

Two assumptions need to be made to justify the digitization process and results. Firstly, the gray value of core slab images is assumed to be able to represent the mineralogy and lithology variations. Although lithology variations do not consistently affect the color and vice-versa, the digitization results can characterize the shaliness variation that highly affects the petrophysical properties like permeability, which is supported by the comparisons of gray value, probe permeability and core slab images.

Secondly, the average gray value of each pixel column is assumed to be able to represent the entire pixel column, so the two dimensional image is simplified into a one dimensional line. Most of the laminations present in cores are horizontal, where this simplification is effective. But there are still non-horizontal layers and features such as bioturbation and concretions, which cannot be so simply characterized.

These assumptions do neglect some issues which have relatively small influences on the results. So they are considered necessary to conduct this investigation.

This research concluded that probe permeability measurement is capable of differentiating and quantifying the permeability of laminaset scale heterogeneities in rock. But, for the sample spacing and resolution for this dataset, this technique cannot fully characterize the highly heterogeneous reservoirs such as the Montney, where lithological heterogeneities are abundant vertically along the cores on a lamination scale basis. Compared to this technique, the digitization method can quantitatively characterize the heterogeneity at the lamination scale correctly and efficiently, which can be a useful guide when planning core measurements and improving sampling efficiency, so that sample density is consistent with the heterogeneity. Also, the gray value data correlate well with the mechanical hardness data and can be used to predict the mechanical hardness because they partly reflects the variations of mineralogy.

CHAPTER 4: Permeability Prediction Using Artificial Neural Networks

4.1 Introduction

As described in Chapter 1, three different approaches have been proven practically useful for permeability determination from logs in conventional reservoirs: empirical models, multiple variable regression, and ANNs.

For the empirical model method, a significant amount of investigation was done by many researchers (Table 1.2). They tried to characterize the complex permeability relationship to other petrophysical properties using a model with general applicability. These studies have helped gain a better understanding of the factors controlling permeability. But it is still difficult to establish a "universal" model that can predict permeability (Balan et al., 1995), especially in the highly heterogeneous cases such as the Montney Formation.

The regression approach, by definition, predicts the expected value of permeability conditioned on a given set of parameters. The previous empirical models may provide the guidelines for choosing the explanatory variables which are used in the development of the predictor.

ANNs have been applied in the petroleum industry from exploration to drilling operations. For log interpretation, a neural network is trained with all the available data, including lab-measured measured permeability and well log data. During this process, the pattern of permeability distribution is recognized and then predicted by the network.

This chapter will compare these three methods of permeability prediction using well logs and build a workflow for ANN development. Meanwhile, the main controls on Montney permeability will also be investigated.

4.2 Multiple Variable Regression Method

Multiple variable regression is an extension of the regression analysis that incorporates additional explanatory variables in the predictive equation. In this study of the Montney, the model to be fitted is:

$$Y = B_0 + B_1X_1 + B_2X_2 + \cdots + B_pX_p + e \quad (4.1)$$

Where Y is the response variable. X_1, X_2,…, X_p are the explanatory variables and e is the residual which is the amount of variation in Y not accounted for by the linear relationship. The parameters B_1, B_2,…, B_p are the regression coefficients which are to be estimated.

In this study, seven types of well logs (Neutron porosity log, Sonic log, PE log, Density log, Resistivity log, GR log and Caliper log) are used as the explanatory variables. The logarithm of permeability is empirically used as the response variable since permeability seems to be log-normally distributed (Amaefule et al., 1993; Balan et al., 1995). In this case, Eq. 4.1 becomes:

$$\log_{10}(K) = B_0 + B_1(NPHS) + B_2(MDT) + B_3(PE) + B_4(RHOB) + B_5(R) + B_6(GR) + B_7(CALI) + e \quad (4.2)$$

where K is the Klinkenberg-corrected core plug permeability (millidarcy); NPHS is the neutron porosity in sandstone matrix (decimal porosity) from the Neutron log; MDT is the Monopole Delta Time (us/m) from the Sonic log; PE is the photoelectric factor from the PE log; RHOB is the formation density (kg/m^3) from the Density log; R is the resistivity (ohm-m) from the deep Resistivity log; GR is the gamma ray api (api) from the GR log; CALI is the caliper measurement (mm) from the Caliper log.

The regression results do not show a strong correlation between the explanatory variables and the response variable (Figs. 4.1, 4.2 and Table 4.2). The prediction will cause relatively significant errors with the mean square errors (MSE) = 0.6039 compared to the target (Fig. 4.1). But the

predictor works much better in the upper Montney which is less shaly than the lower Montney. So the regression is conducted on the upper Montney and the lower Montney respectively. The results show a better correlation than the previous one with smaller MSE = 0.4825 ((Figs. 4.3 and 4.4). The residuals tend to be symmetrically distributed along depth (Fig. 4.5). The lower Montney has much larger residuals compared to the upper Montney (Fig. 4.5). The residuals tend to increase by the increase of the Log(K), which means the low permeability tends to be overestimated and the high permeability tend to be underestimated (Fig. 4.6). The high permeability values of fractured core plugs in the lower Montney are difficult to characterize using linear regression, which might cause these issues.

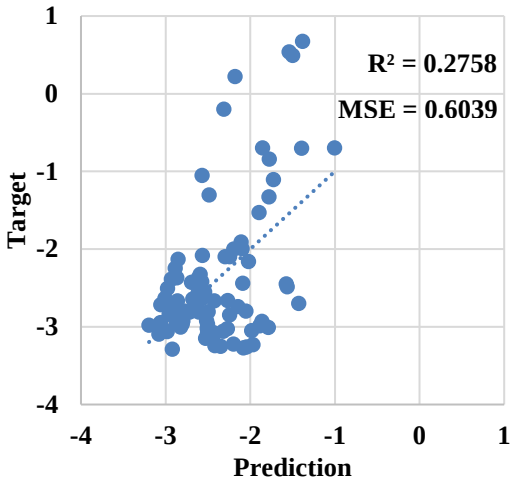

Figure 4.1 Multiple variables regression results.

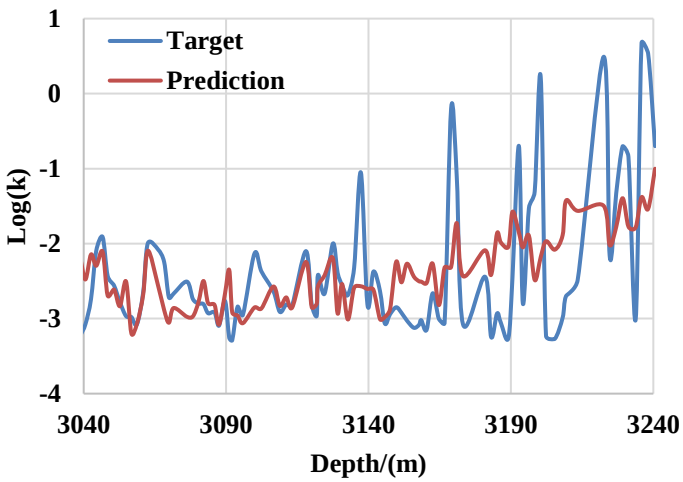

Figure 4.2 Multiple variables regression results.

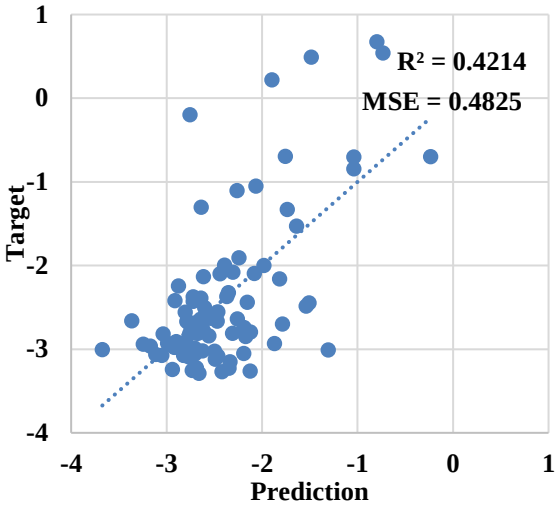

Figure 4.3 2-interval model results.

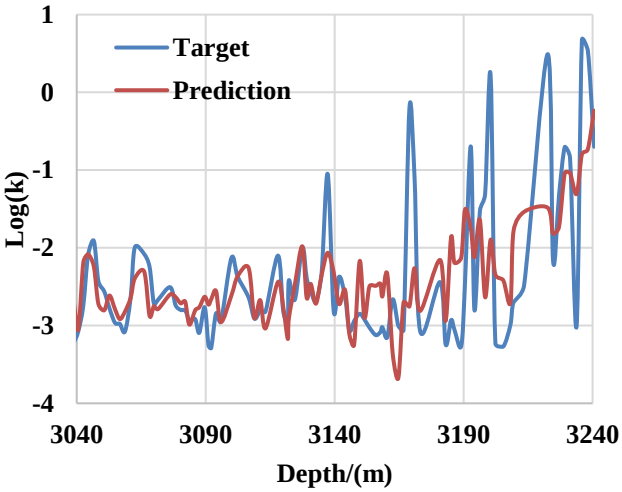

Figure 4.4 2-interval model results.

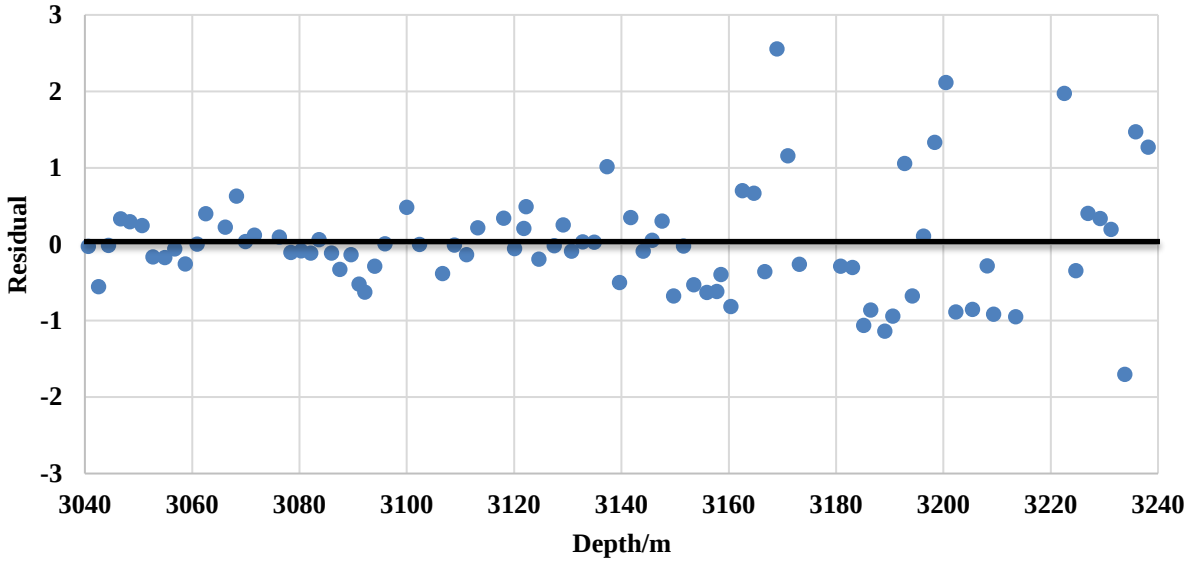

Figure 4.5 Residuals compared to depth show an increase in variability with depth.

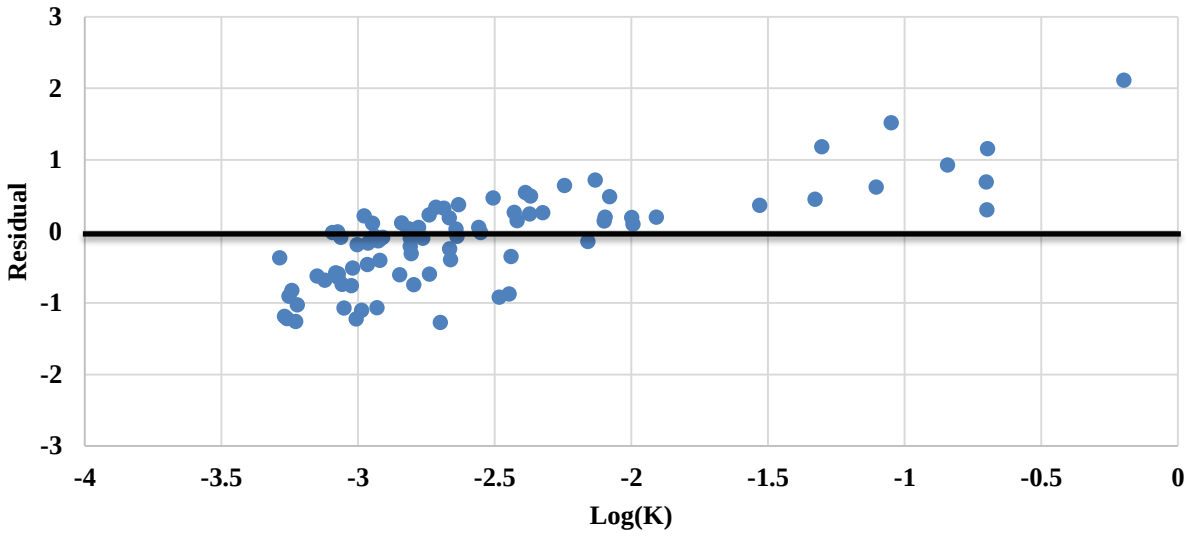

Figure 4.6 Residuals compared to Log(K) show a trend which the model does not include.

4.3 Empirical Model Method

Among the work done by many researchers (Table 4.1), Yao and Holditch (1993) presented a model, including porosity, deep and shallow resistivities, and GR (Eq. 4.2). Compared with other models (Table. 4.1), this model works better for lower permeability sandstone, which suggests it might be suitable for the tight Montney Formation.

$$K = \frac{U\emptyset^{e1}(1 - I_{GR})^{e2} R_{ild}^{e3}}{\left(R_{ild}/R_{sfl}\right)^{e4}} \tag{4.2}$$

where K is the permeability; U is a constant; $\emptyset$ is the porosity (decimal); I_{GR} is the GR log index; R_{ild} is the deep resistivity (ohmm); R_{sfl} is the shallow resistivity; $e1, e2, e3$ and $e4$ are model parameters. In this study, the core plug permeability and porosity are used as k and $\emptyset$, respectively. 90 inch radial resistivity and invaded zone resistivity will replace R_{ild} and R_{sfl}.

The overall correlation between the target and the prediction from the empirical model is weak with the R^2 = 0.1601 (Fig. 4.7). The prediction will cause relatively significant errors compared to the target for the section below 3140 m (Fig. 4.8). So the model is used on the upper Montney and the lower Montney respectively. The results show a better correlation than the previous one with smaller MSE = 0.6400 ((Figs. 4.9 and 4.10). The residuals tend to be symmetrically distributed along depth (Fig. 4.11). The lower Montney has much larger residuals compared to the upper Montney (Fig. 4.11). The residuals tend to increase with the increase of Log(K), which means the low permeability values tend to be overestimated and the high permeability values tend to be underestimated (Fig. 4.12). The high permeability values of fractured core plugs in the lower Montney are difficult to characterize using linear regression, which might cause these issues.

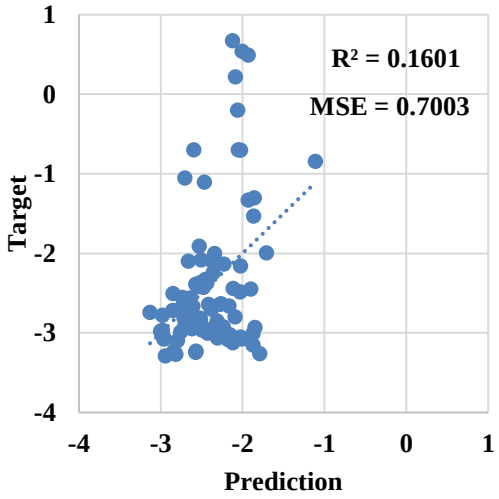

Figure 4.7 Empirical model results.

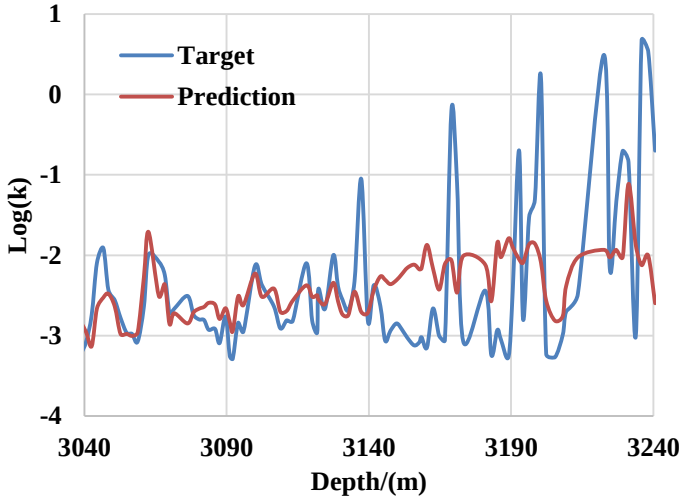

Figure 4.8 Empirical model results.

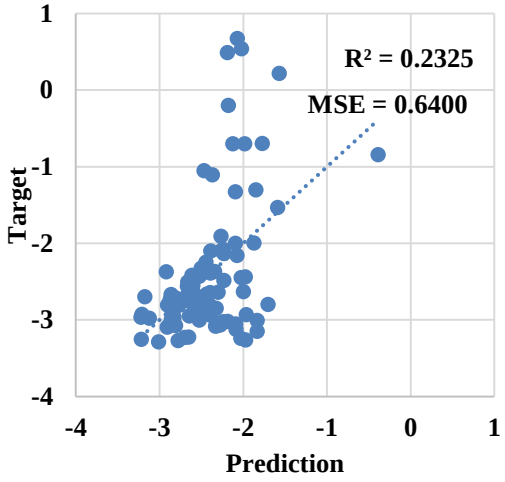

Figure 4.9 2-interval model results.

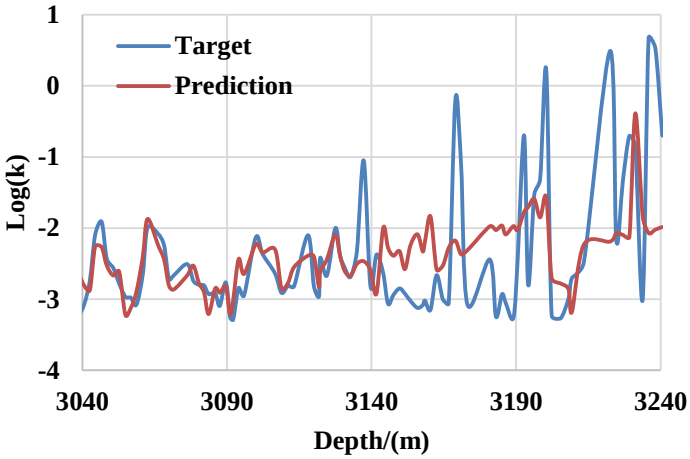

Figure 4.10 2-interval model results.

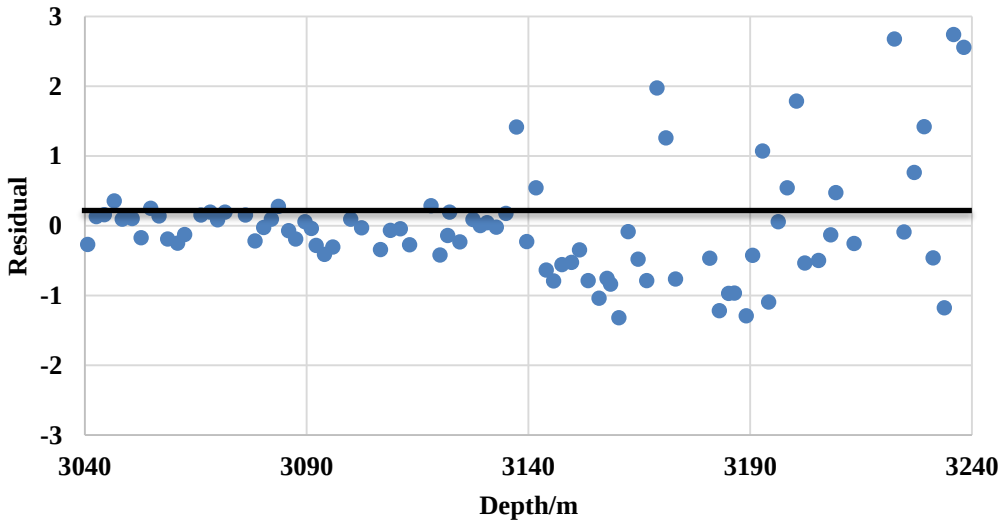

Figure 4.11 Residuals compared to depth show an increase in variability with depth.

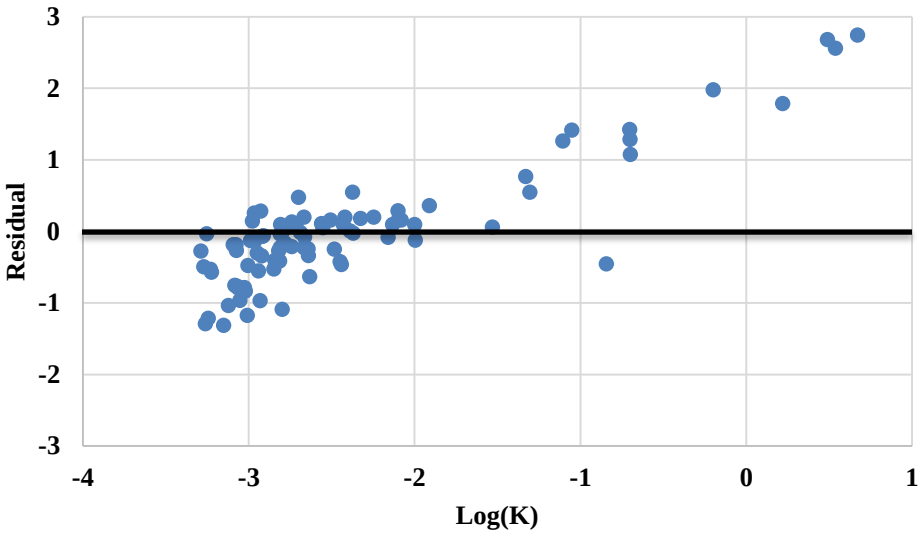

Figure 4.12 Residuals compared to Log(K) show a trend which the model does not include.

4.4 Artificial Neural Network Method

4.4.1 Introduction and Workflow

Neural networks are composed of simple elements (“neurons”) operating in parallel. These elements attempt to emulate biological nervous systems. As in nature, the connections between elements largely determine the network function. You can train a neural network to perform a particular function by adjusting the values of the connections (weights) between elements.

Typically, neural networks are adjusted, or trained, so that a particular input leads to a specific output (the ‘prediction’). The network is adjusted, based on a comparison of the output and the “target” i.e., the measured response values, until the network output matches the target (Fig. 4.13). Typically, many such input/target pairs are needed to train a network.

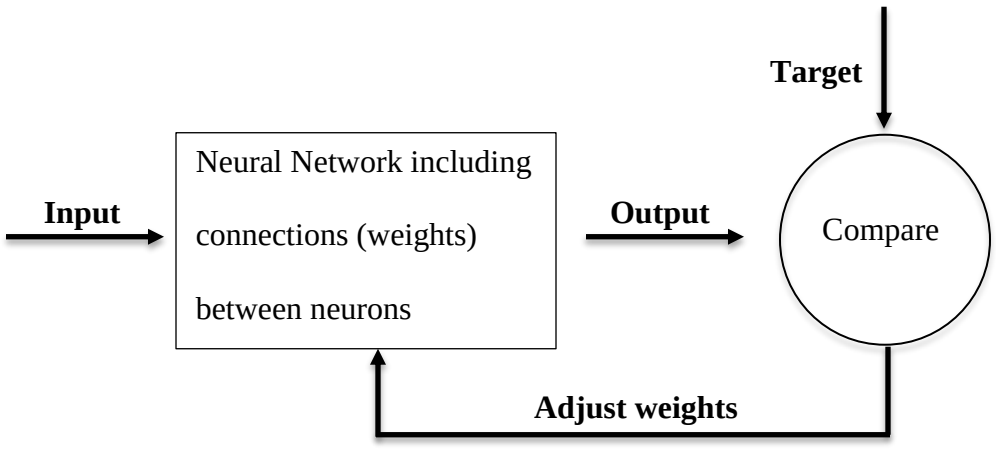

Figure 4.13 Neural network system.

In this study, the Neural Network Toolbox software from Matlab is used through basic command-line operations to develop the model. The workflow for the neural network design process has seven primary steps (Beale et al, 2016):

1) Collect data
2) Create the network
3) Configure the network
4) Initialize the weights and biases
5) Train the network
6) Validate the network
7) Use the network

After confirming the data quality, the data then need to be divided into three subsets (Step one), the "training" dataset, the "validation" dataset and the "test" dataset, before they are used. The roles of the three subsets will be discussed later. In Matlab, there are four functions provided for dividing data: 'dividerand' (the default), 'divideblock', 'divideint', and 'divideind' (Table 4.1). The dataset division is a significant factor for the ANN performance so it will be a part of the optimization workflow.

Table 4.1 Divide functions..

Function	Algorithm
dividerand	Divide the data randomly (default)
divideblock	Divide the data into contiguous blocks
divideint	Divide the data using an interleaved selection
divideind	Divide the data by index

After the data are collected and divided into subsets, a multiple-layer network needs to be created and configured (Steps two and three, Fig. 4.14). In Fig. 4.14, the colored dots represents neurons, and neurons are connected by weights and biases. The blue dots represent neurons in the hidden layer. According to the number of inputs and targets, the network will be configured and the weights and biases will be initialized randomly. The number of hidden neurons is a critical factor for the neural network performance, so it will be a part of the optimization workflow.

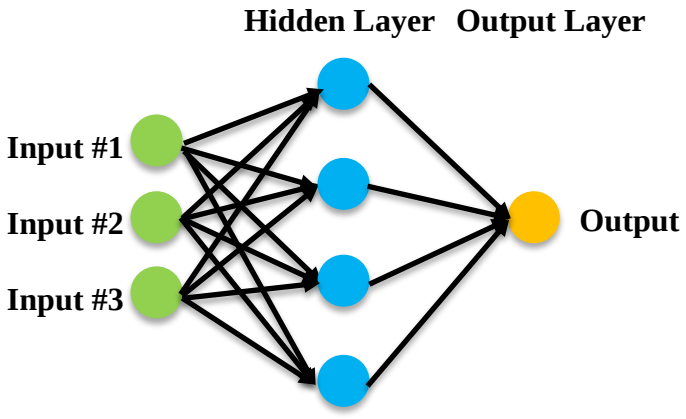

Figure 4.14 Example of ANN architecture with one hidden layer.

When the network weights and biases are initialized (Step four), the network is ready for training. The process of training a neural network involves tuning the values of the weights and

biases of the network to optimize network performance, which is measured by the mean square error (MSE) between the outputs and targets. It is defined as follows:

$$mse = \frac{1}{N}\sum_{i=1}^{N}(t_i - a_i)^2 \tag{4.3}$$

where t are the target values and a are the output values.

The "training" dataset is used for computing the gradient and updating the network weights and biases (Step five). Each time the network is trained, the validation dataset will be used to generate the validation performance which is also the MSE (Step six). When the network begins to overfit the data, the error on the validation dataset typically begins to rise. The number of validation checks represents the number of successive iterations that the validation performance fails to decrease. If this number reaches 6 (the default value), the training will stop, so the network weights and biases are saved at the minimum of the validation dataset error. The number of validation checks will also be a part of the optimization workflow. The "test" dataset, where error is not used during training, is used to compare different models.

After the network is trained and validated, the network object can be used to calculate the network response to any input. But different neural networks trained on the same problem can give different outputs for the same input due to different initial weight and bias values and different divisions of data. Thus, given the same input and target data, the neural network may need to be retrained many times to find the one with best accuracy (smallest MSE). Similarly, the ANN architecture (e.g., number of hidden nodes and hidden layers) may also affect the results. Thus, there is an inherent uncertainty associated with ANN models and this uncertainty is often overlooked when studies using ANNs report "successful" applications.

Based on the neural network design process, the model optimization workflow is summarized as follows:

1) Determine the optimum number of hidden neurons

2) Determine the optimum number of validation checks
3) Determine the optimum dataset division
4) Determine the optimum weights and biases

4.4.2 Determining the Optimum Number of Hidden Neurons

Larger numbers of neurons in the hidden layer give the network more flexibility because the network has more parameters it can optimize. But increasing the number of hidden neurons also requires more computation and is more likely to produce overfitting. Thus, an optimum number of hidden neurons should be determined to optimize the performance and efficiency. However, because the initial weight and bias values and division of dataset will be reset randomly when the network is retrained, the performance of the network will also be different each time it is retrained. Thus, the parameter setting of this optimization is summarized in Table 4.2, and the workflow is summarized as follows:

1) Set the number of hidden neurons from 1 to 25 for 25 networks respectively;
2) Retrain each network for N times and generate N groups of responses;
3) For each network, check whether N groups of responses are sufficient to represent their probability distribution or not. If not, increase N and repeat step 2 and 3;
4) Use the averaged R^2 values from N groups of responses as the standard to determine the optimum number of hidden neurons.

Table 4.2 Parameter setting.

Parameters	**Setting**
Input data	Neutron porosity log, Sonic log, PE log, Density log, Resistivity log, GR log and Caliper log
Target data	logarithm of core plug permeability

Divide function	dividerand (default, ratio: training 70%, validation 15%, test 15%)
Number of hidden neurons	To be determined
Number of validation checks	6 (default)
Other parameters	Default setting

In this study, N is set to 500, so each network is retrained for 500 times. The recommended number I of retraining times is calculated from the C_V of the responses' R^2 values to decide whether N is large enough or not, where $I = (10C_V)^2$.

The results show that N groups of responses is sufficient to represent their probability distribution (Table 4.3). The averaged R^2 value increases with the increase of the number of hidden neurons from 1 to 8 but from 8 to 25 there are no significant improvements for the averaged R^2 value (Fig. 4.15). Thus, the optimum number of hidden neurons is set to 8.

Table 4.3 The recommended number of training times

Number of hidden neurons	N	I	Number of hidden neurons	N	I
1	500	17.64	14	500	8.18
2	500	18.65	15	500	7.53
3	500	16.96	16	500	11.38
4	500	14.61	17	500	8.03
5	500	12.16	18	500	8.85
6	500	10.04	19	500	11.23
7	500	10.78	20	500	8.13
8	500	8.74	21	500	12.45

9	500	8.75	22	500	10.20
10	500	8.84	23	500	10.43
11	500	8.53	24	500	11.39
12	500	8.29	25	500	11.93
13	500	9.19			

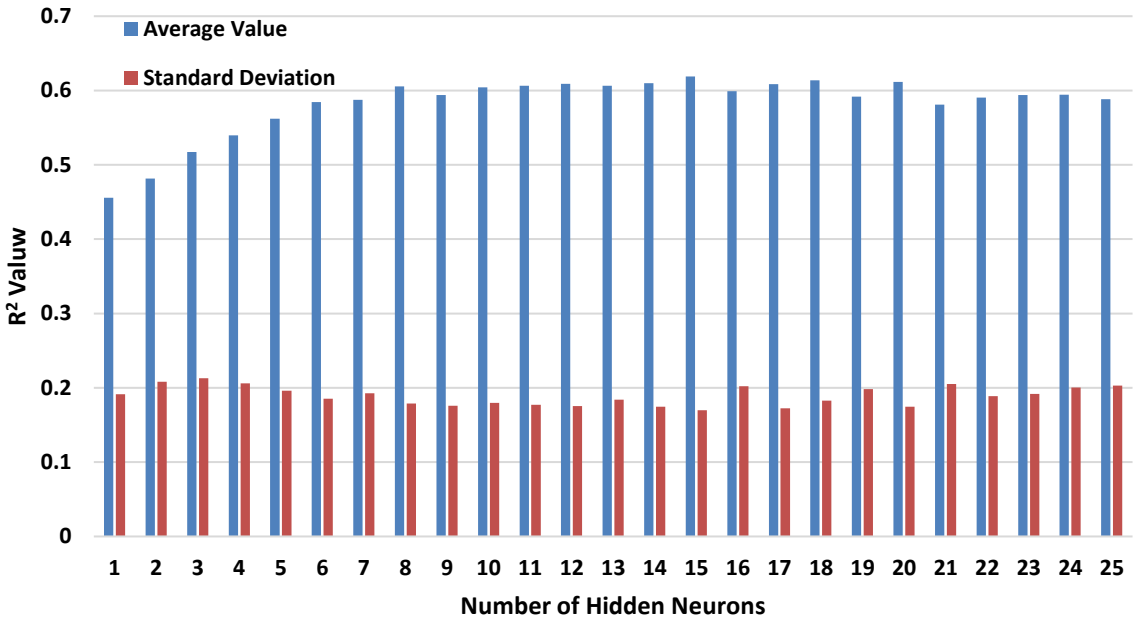

Figure 4.15 Comparison of the averaged R^2 values from N groups of responses.

4.4.3 Determining the Optimum Number of Validation Checks

Validation is a process to avoid overfitting, but the number of validation checks could be a significant factor on the network performance. Thus the influence of the number of validation checks is investigated and the optimum value is determined. The parameter setting of this optimization is summarized in Table 4.4, and the workflow is summarized as follows:

1) Set the number of validation checks from 1 to 25 for 25 networks respectively;
2) Retrain each network for N times and generate N groups of responses;
3) For each network, check whether N groups of responses are sufficient to represent their probability distribution or not. If not, increase N and repeat step 2 and 3;
4) Use the averaged R^2 values from N groups of responses as the standard to determine the optimum number of validation checks.

Table 4.4 Parameter setting.

Parameters	**Setting**
Input data	Neutron porosity log, Sonic log, PE log, Density log, Resistivity log, GR log and Caliper log
Target data	logarithm of core plug permeability
Divide function	dividerand (default, ratio: training 70%, validation 15%, test 15%)
Number of hidden neurons	8
Number of validation checks	To be determined
Other parameters	Default setting

The results show that N groups of responses is sufficient to represent their probability distribution (Table 4.5). The averaged R^2 value does not change significantly with the number of validation checks on a specific trend, so the number is not a significant factor on the network performance (Fig. 4.16). Thus, the optimum number of validation checks is set to the default value 6.

Table 4.5 The recommended number of training times

Number of hidden neurons	N	I	Number of hidden neurons	N	I
1	500	7.22	14	500	5.76
2	500	5.83	15	500	5.16
3	500	5.85	16	500	4.72
4	500	6.25	17	500	4.81
5	500	6.05	18	500	5.05
6	500	5.31	19	500	5.16
7	500	5.30	20	500	5.37
8	500	4.39	21	500	5.23
9	500	4.62	22	500	5.40
10	500	6.43	23	500	5.51
11	500	5.44	24	500	4.99
12	500	5.00	25	500	4.72
13	500	6.74			

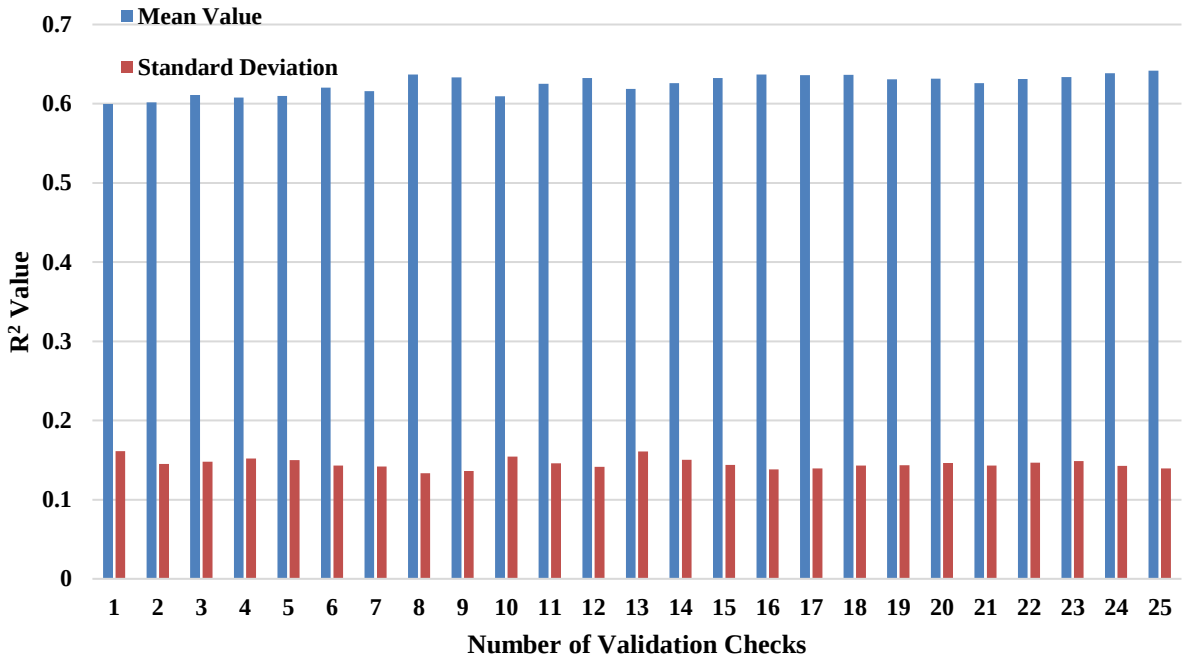

Figure 4.16 Comparison of the averaged R^2 values from N groups of responses.

4.4.4 Determining the Optimum Dataset Division

The default setting for the dataset divide function is 'dividerand' with a ratio of 70% to training, 15% to validation and 15% to test, which means the data are divided into three subsets randomly. Thus, an optimum dataset division which produces the best network performance can be determined from all random combinations of data. The parameter setting of this optimization is summarized in Table 4.6, and the workflow is summarized as follows:

1) Retrain each network for 10000 times and generate 10000 groups of responses;
2) Select the networks whose R^2 values are among the highest 500 of 10000 results;
3) For the training, validation and test subsets, calculate the frequency of each data point respectively from the best 500 results.
4) Develop the new dataset division based on the frequency of each data point. Priority is given to the data point with higher frequency in the subset.

Table 4.6 Parameter setting.

Parameters	**Setting**
Input data	Neutron porosity log, Sonic log, PE log, Density log, Resistivity log, GR log and Caliper log
Target data	logarithm of core plug permeability
Divide function	dividerand (default, ratio: training 70%, validation 15%, test 15%)
Number of hidden neurons	8
Number of validation checks	6
Other parameters	Default setting

After the optimum dataset division is determined, the divide function is changed from 'dividerand' to 'divideind'. With 'divideind', the dataset division will be fixed for the network, so it will not change each time the network is retrained.

4.4.5 Determining the Optimum Weights and Biases

Each network will be trained starting from randomly assigned weights and biases, so the optimum weights and biases can be determined from the comparison of network performance. Meanwhile, another way to improve performance is to train multiple neural networks and average their outputs. The parameter setting of this optimization is summarized in Table 4.7, and the workflow is summarized as follows:

1) Retrain each network for 10000 times and generate 10000 groups of responses;
2) Select the networks whose R^2 values are among the highest 10 of 10000 results;
3) Save these 10 network structures and parameters and average their outputs.

Table 4.7 Parameter setting.

Parameters	**Setting**
Input data	Neutron porosity log, Sonic log, PE log, Density log, Resistivity log, GR log and Caliper log
Target data	logarithm of core plug permeability
Divide function	divideind
Number of hidden neurons	8
Number of validation checks	6
Other parameters	Default setting

The correlation between the target and the prediction from the averaged neural network outputs is strong with the $R^2 = 0.94$ and the MSE = 0.0485 (Fig. 4.17). But test results show weak correlation with the $R^2 = 0.26$ and the MSE = 0.1609 (Fig. 4.18), which means the generalization power of this ANN is poor and also explain the large errors on some predictions (Fig. 4.19). Thus this ANN requires further work and will be improved by incorporating the digitized image data from Chapter 3. Nonetheless, this ANN gives considerably better performance than the empirical and regression approaches shown earlier (Figs. 4.20 and 4.21).

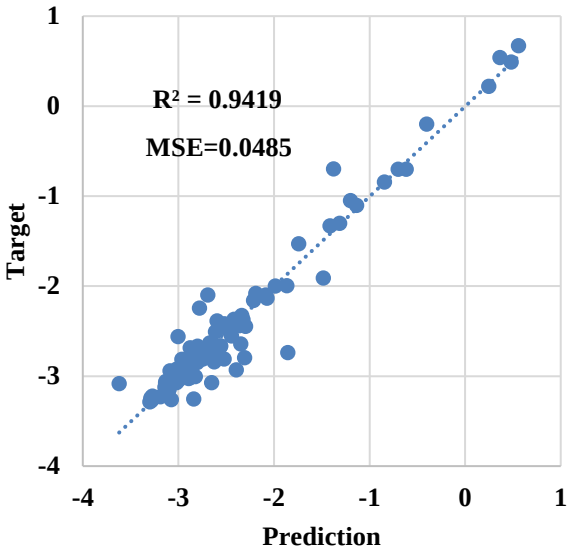

Figure 4.17 ANN results.

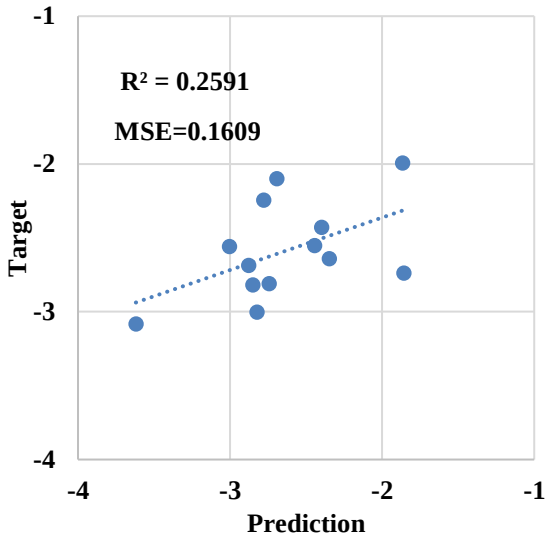

Figure 4.18 Test dataset results.

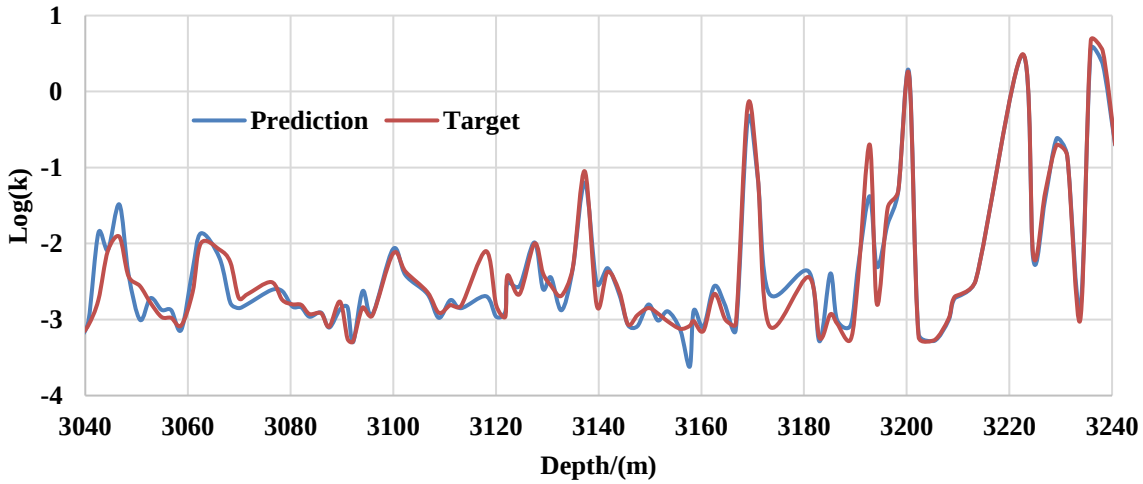

Figure 4.19 ANN results.

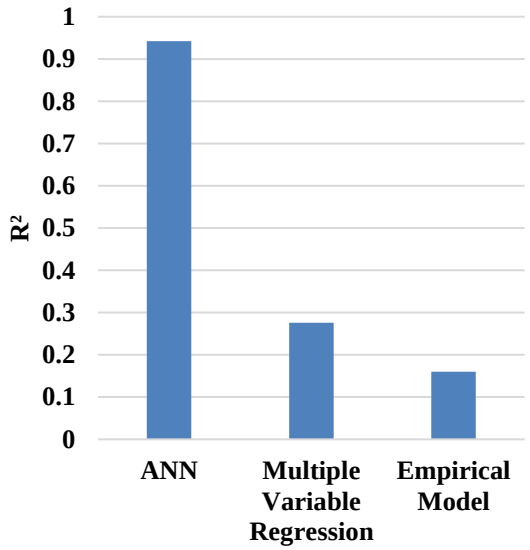

Figure 4.20 R^2 value of ANN, multiple variable regression and empirical model.

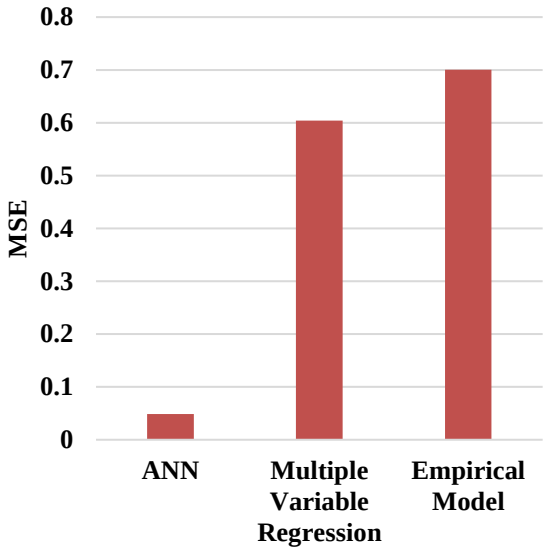

Figure 4.21 MSE of ANN, multiple variable regression and empirical model.

4.4.6 Incorporating the Digitized Image Data into the ANN

In Chapter 3, we described the digitizing and analysis of the core photographs. Now we seek to combine those data with the well log data using the ANN.

Firstly, the image data are incorporated as one input variable. Thus the new ANN using the inputs of image data and the same well log data is developed by following the optimization workflow. The results do not show improvements from the old one (Figs. 4.22 and 4.23), which might result from the scale issue. The image data represent the lamination scale variation on the core surface, while the core plug data represent the average quality of a 2.54 cm diameter core plug volume. This difference of investigation scale may cause that the image data cannot be directly used as an explanatory variable for core plug permeability.

Thus, image data are incorporated indirectly. Based on the GR log data, the core can be divided into 3 intervals (Fig. 4.24). Interval 1 is mainly comprised of clean siltstone and laminated shaly siltstone. Intervals 2 and 3 are composed of dispersed shaly siltstone. The image data can also identify the difference between these intervals. The average gray values, lamination density and average lamination thickness of Intervals 1 and 2 are very different from each other (Fig. 4.25). But for Intervals 2 and 3, the difference of the average gray values, lamination density and average lamination thickness are not significant (Fig. 4.26). Thus two cases are investigated. In the first case, three ANNs are developed respectively based on the three intervals following the model optimization workflow described above. In the second case, the data from Intervals 2 and 3 are combined into one dataset, and the data from Interval 1 is the other dataset, so two ANNs are developed respectively based on these two datasets following the model optimization workflow.

The results for first case show improvement compared to the former model with the $R^2 = 0.9657$ and the MSE = 0.0309 (Fig. 4.27). But the test results still show a poor generalization power of this ANN with MSE = 0.1460 (Figs. 4.28 and 4.29).

The results for the second case show better behavior of this model with the $R^2 = 0.9918$ and MSE = 0.0069 (Fig. 4.30), giving a 20-fold reduction in MSE. Meanwhile, the test results also

show improvement compared to the first case with the R^2 = 0.9526 and the MSE = 0.0217 (Fig. 4.31), which means the generalization power of this ANN is much greater than the former models (Fig. 4.32).

The accurate prediction and high generalization power of the second case model suggest that the gray value is helpful in evaluating the division of core intervals and the correlation of petrophysical data. Also, compared to the multiple variables regression and the empirical model methods, the ANN method is more accurate in capturing the relationship between well log data and permeability data (Table 4.8) so this model is used to investigate the main controls on permeability.

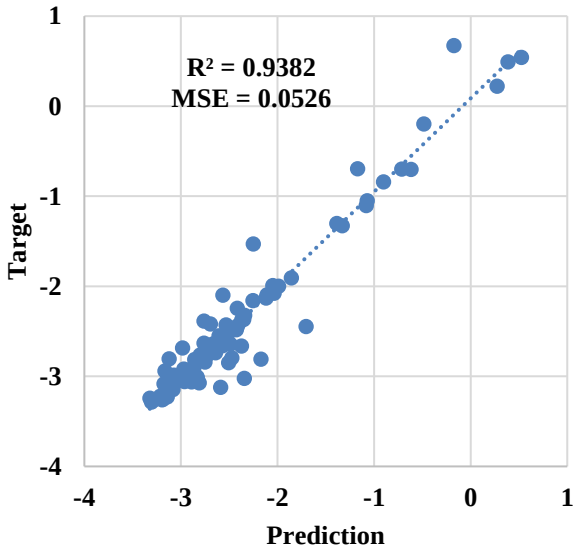

Figure 4.22 New ANN results.

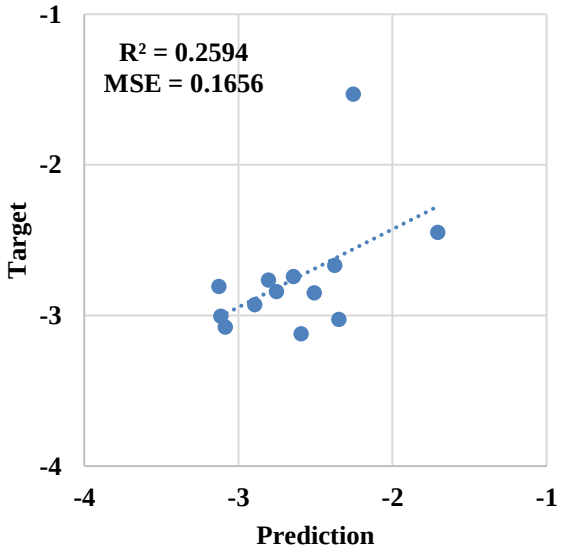

Figure 4.23 New ANN test results.

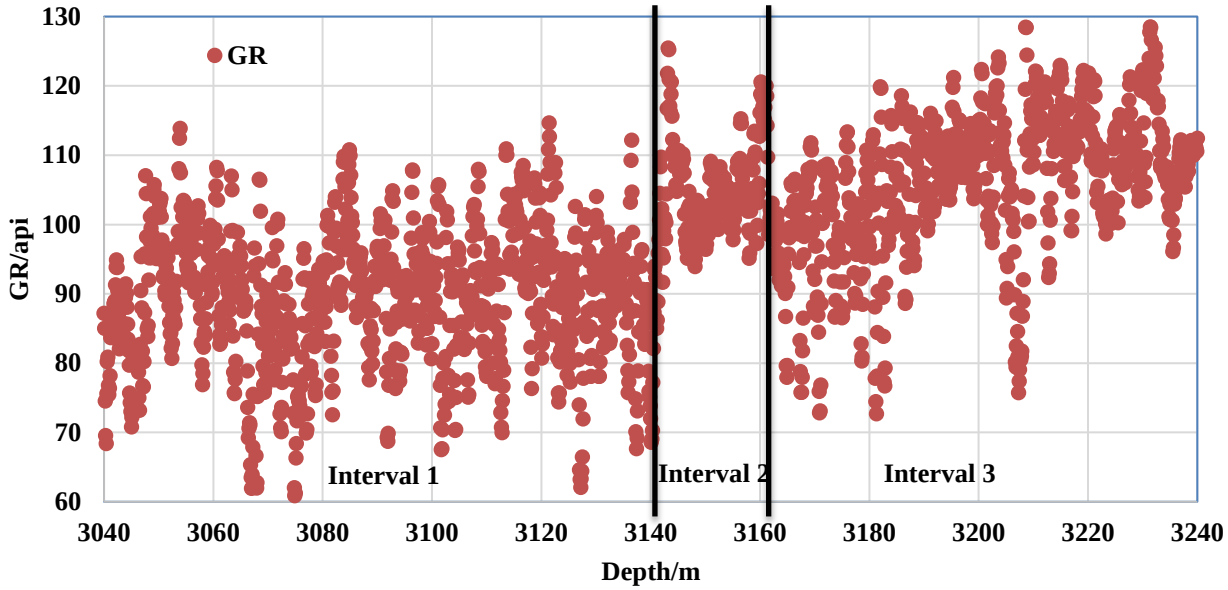

Figure 4.24 GR log data.

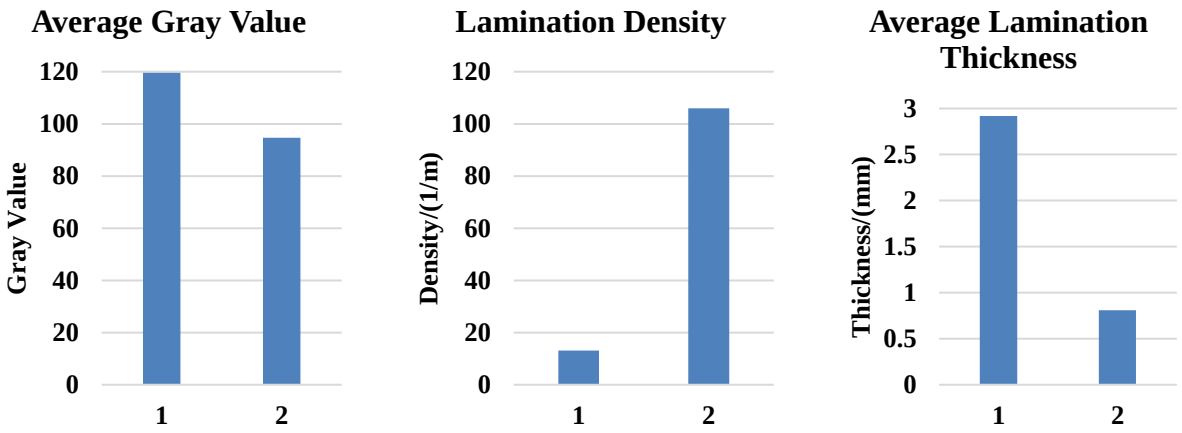

Figure 4.25 Comparison of Intervals 1 and 2.

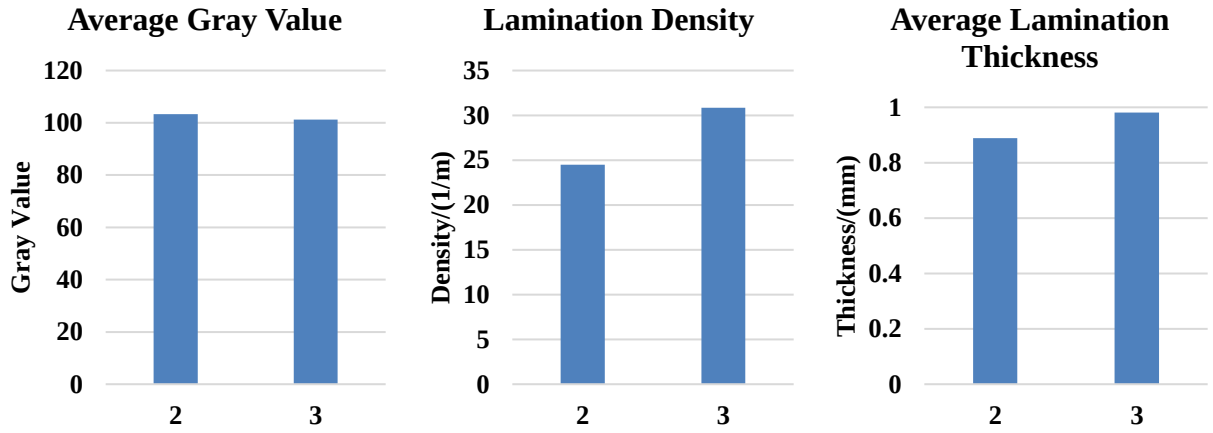

Figure 4.26 Comparison of Intervals 2 and 3.

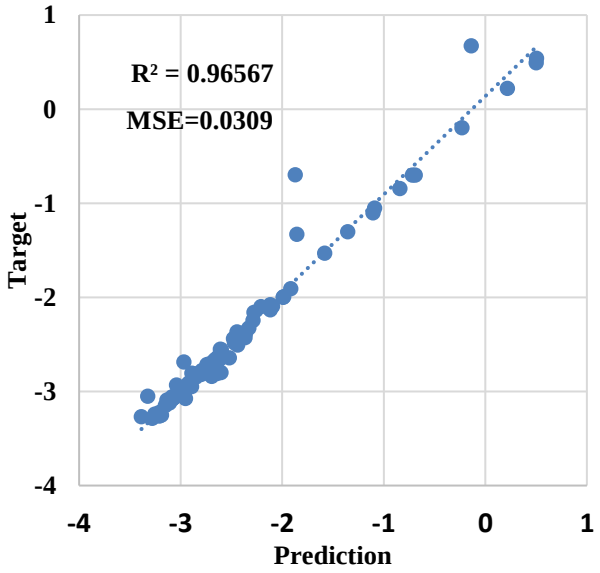

Figure 4.27 3-interval ANN results.

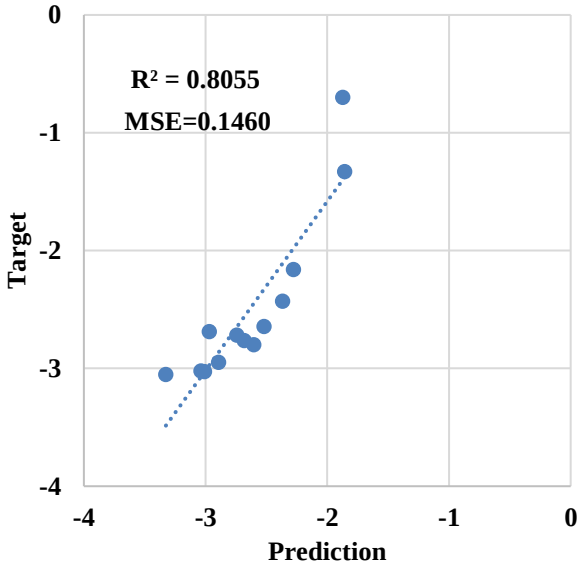

Figure 4.28 3-interval ANN test dataset results.

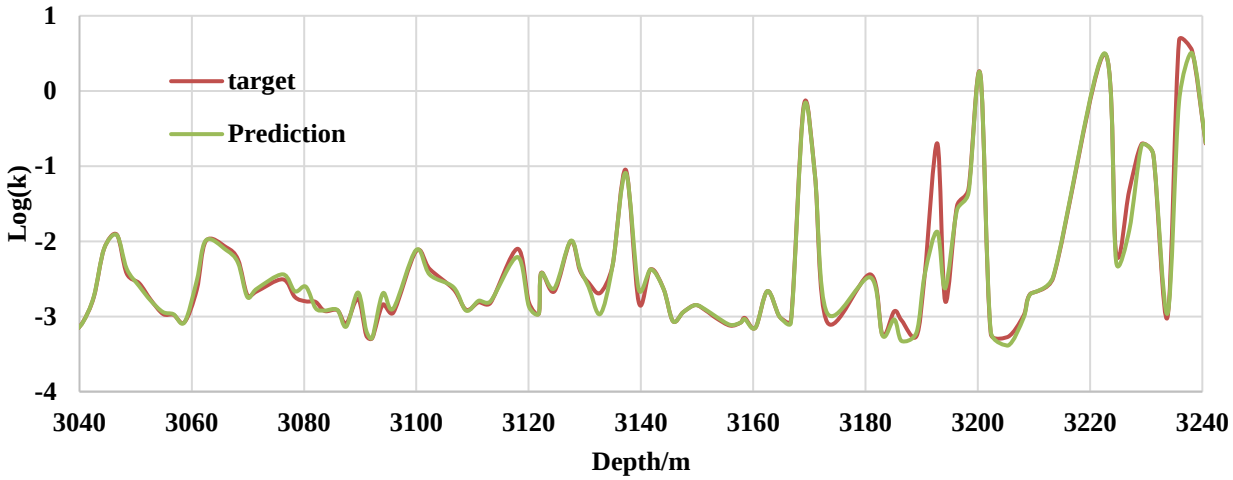

Figure 4.29 3-interval ANN results.

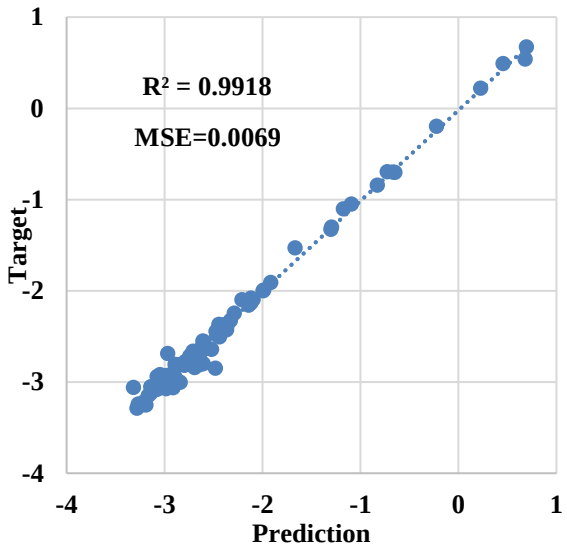

Figure 4.30 2-interval ANN results.

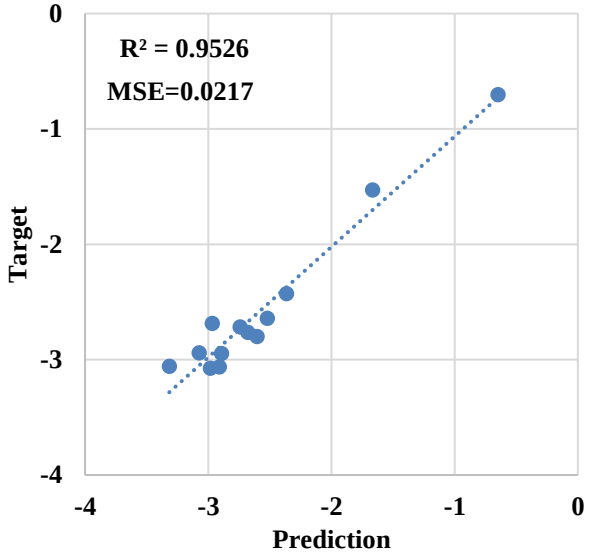

Figure 4.31 2-interval ANN test dataset results.

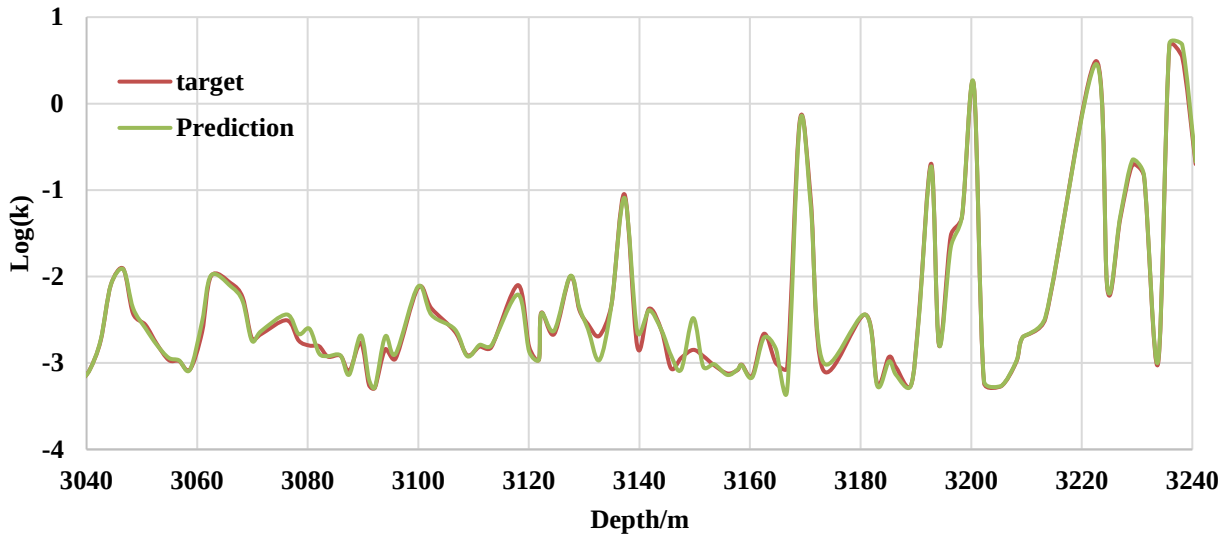

Figure 4.32 2-interval ANN results.

Table 4.8 Results summary.

			R^2	MSE
Multiple Variable Regression		1-interval model	0.28	0.60
		2-interval model	0.42	0.48
Empirical Model		1-interval model	0.16	0.70
		2-interval model	0.23	0.64
ANN	Without image data		0.94	0.049
	Incorporating image data as an explanatory variable		0.94	0.053
	Incorporating image data indirectly	3-interval model	0.97	0.031
		2-interval model	0.99	0.0069

4.5 Investigation of Main Controls on Permeability

The ANN model developed from the optimization workflow is effective in capturing the relationships between well log data and permeability data. So based on this model, a seven-factor design for ANOVA analysis is made to investigate the importance of the seven well log inputs on the permeability target. There are two levels, low level and high level, which are tested for each explanatory factor, so $2^7 = 128$ treatment combinations are obtained and each treatment combination has only one replication (Table 4.9). In this case, the low level value is defined as "the mean value minus one standard deviation" for each factor dataset and the high level value is defined as "the mean value plus one standard deviation" for each factor dataset.

Table 4.9 Parameter setting.

Factor name		NPHS	MDT (us/m)	PE	RHOB (kg/m^3)	R (ohmm)	GR (api)	CALI (mm)
Mean		0.103	196.056	3.322	2656.447	81.497	95.699	0.157
Standard Deviation		0.025	7.784	0.240	32.765	41.510	11.482	0.636
Low Level	Value	0.078	188.272	3.082	2623.682	39.988	84.217	-0.479
	Notation	1	1	1	1	1	1	1
High Level	Value	0.128	203.840	3.562	2689.211	123.007	107.181	0.794
	Notation	a	b	c	d	e	f	g
Treatment Combinations		1, a, b, c, d, e, f, g, ab, ac, ad, ae, af, ag, bc, bd, be, bf, bg, cd, ce, cf, cg, de, df, dg, ef, eg, fg, abc, abd, abe, abf, abg, acd, ace, acf, acg, ade, adf, adg, aef, aeg, afg, bcd, bce, bcf, bcg, bde, bdf, bdg, bef, beg, bfg, cde, cdf, cdg, cef, ceg, cfg, def, deg, dfg, efg, abcd, abce, abcf, abcg, abde, abdf, abdg, abef, abeg, abfg, acde, acdf, acdg, acef, aceg, acfg, adef, adeg, adfg, aefg,						

	bcde, bcdf, bcdg, bcef, bceg, bcfg, bdef, bdeg, bdfg, befg, cdef, cdeg, cdfg, cefg, defg, abcde, abcdf, abcdg, abcef, abceg, abcfg, abdef, abdeg, abdfg, abefg, acdef, acdeg, acdfg, acefg, adefg, bcdef, bcdeg, bcdfg, bcefg, bdefg, cdefg, abcdef, abcdeg, abcdfg, abcefg, abdefg, acdefg, bcdefg, abcdefg

MINITAB is used to obtain the sum of squares for each effect, then Eq. 4.4 (Devore et al, 2013) is used to determine the estimated effects.

$$Effect = \sqrt{\frac{sum\ of\ squares}{n'2^{k-2}}} \tag{4.4}$$

where k is the number of factors and n' is the number of replicates for each treatment combination.

Then the estimated effects for each treatment combination are plotted in a normal probability plot to determine whether or not the entire factors and interaction effects are due to chance or random error. Any factors or interactions whose effects are due to chance are expected to be randomly distributed around zero. These effects will tend to fall along a straight line on a normal probability plot. The effects that might be significant have average values different from zero and deviate from the line that represents no effect. So the interactions that do not appear to have any effects will be screened out and combined into a substitute measure of random error.

Based on the probability plot (Fig. 4.33), all third-order to seventh-order interactions and all second-order interactions (except ae, ef and de) are reasonably considered as consisting only of random error, so they are eliminated from the ANOVA model in MINITAB and combined to estimate an error term (Table 4.10). The factor RHOB has a p-value of 0.126 that is larger than the level of significance, 0.05, so this factor can be considered as an insignificant effect on the results, which also can be found on the main effects plot (Fig. 4.34). Thus, the factor RHOB is removed from the inputs, and a new ANN is developed following the model optimization

workflow proposed above. The results change only a small amount after removing RHOB with the new R^2 = 0.9817 and the new MSE = 0.0154 (Figs. 4.35 and 4.36). So the new ANN model without the factor RHOB is analyzed using a six-factor design.

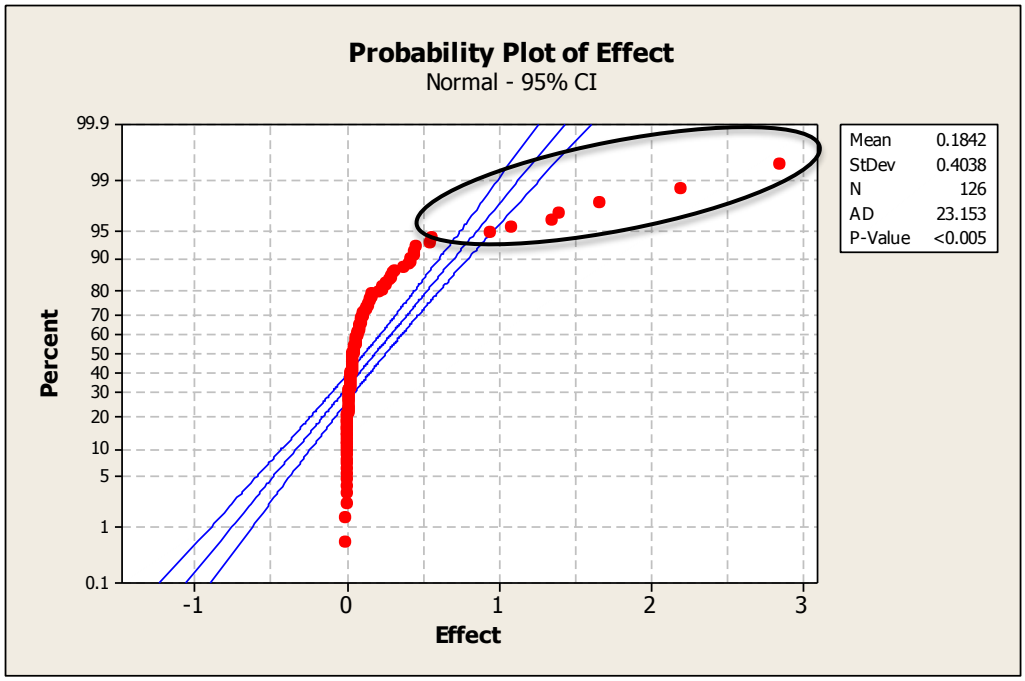

Figure 4.33 Normal probability plot obtained from MINITAB of the estimated effects.

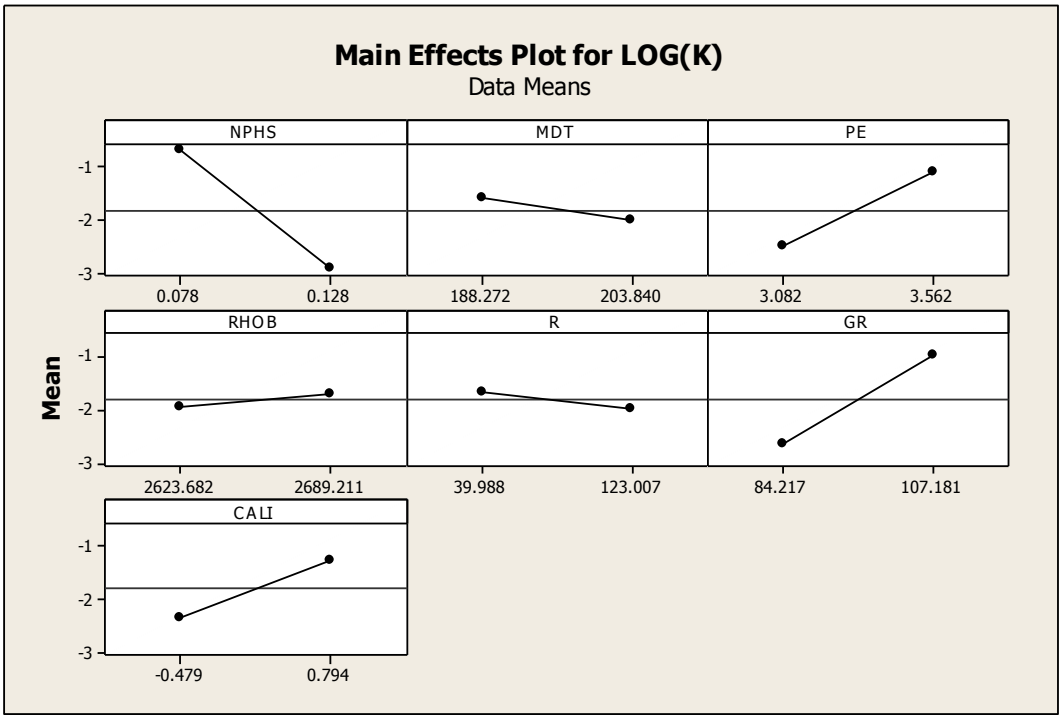

Figure 4.34 Main effects plot obtained from MINITAB for the 7-factor case.

Table 4.10 Seven-factor ANOVA model obtained from MINITAB.

Source	DF	SS	MS	F	P
NPHS	1	154.558	154.558	204.73	0
MDT	1	5.655	5.655	7.49	0.007
PE	1	62.436	62.436	82.7	0
RHOB	1	1.796	1.796	2.38	0.126
R	1	3.247	3.247	4.3	0.04
GR	1	89.311	89.311	118.3	0
CALI	1	37.397	37.397	49.54	0
NPHS*R	1	259.426	259.426	343.63	0
R*CALI	1	58.311	58.311	77.24	0
RHOB*R	1	28.521	28.521	37.78	0
Error	117	88.329	0.755		
Total	127	788.987			

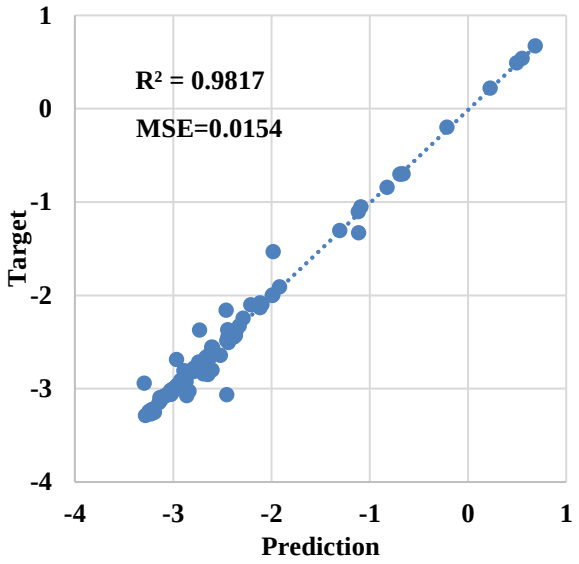

Figure 4.35 ANN results without RHOB.

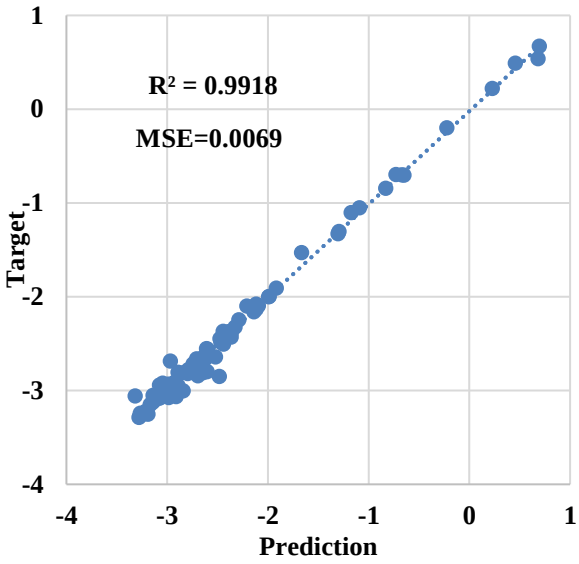

Figure 4.36 Original ANN results.

Based on the new model, a six-factor design for ANOVA analysis is made to investigate the importance of the remaining six well log inputs on the permeability target (Table 4.11). Based on the probability plot (Fig. 4.37), all third-order to sixth-order interactions and all second-order interactions (except ad, df, cd, af and ac) are reasonably considered as consisting only of random error, so they are eliminated from the ANOVA model in MINITAB and combined to estimate an error term (Table 4.12). The factor R (resistivity) has a p-value of 0.219 that is larger than the level of significance, 0.05. So this factor is considered as an insignificant effect on the results, which also can be found on the main effects plot (Fig. 4.38). Thus, the factor R is removed from the inputs, and a new ANN is developed following the model optimization workflow proposed above. The results again show little change after removing R with the new $R^2 = 0.9886$ and the new MSE = 0.0097 (Figs. 4.39 and 4.40). So the new ANN model without the factors RHOB and R is also analyzed using a five-factor design.

Table 4.11 Parameter setting.

Factor name		NPHS	MDT (us/m)	PE	R (ohmm)	GR (api)	CALI (mm)
Mean		0.103	196.056	3.322	81.497	95.699	0.157
Standard Deviation		0.025	7.784	0.240	41.510	11.482	0.636
Low Level	Value	0.078	188.272	3.082	39.988	84.217	-0.479
	Notation	1	1	1	1	1	1
High Level	Value	0.128	203.840	3.562	123.007	107.181	0.794
	Notation	a	b	c	d	e	f
Treatment Combinations		1, a, b, c, d, e, f, ab, ac, ad, ae, af, bc, bd, be, bf, cd, ce, cf, de, df, ef, abc, abd, abe, abf, acd, ace, acf, ade, adf, aef, bcd, bce, bcf, bde, bdf, bef, cde, cdf, cef, def, abcd, abce, abcf, abde, abdf, abef, acde, acdf, acef, adef, bcde, bcdf, bcef, bdef, cdef, abcde, abcdf, abcef, abdef, acdef, bcdef, abcdef					

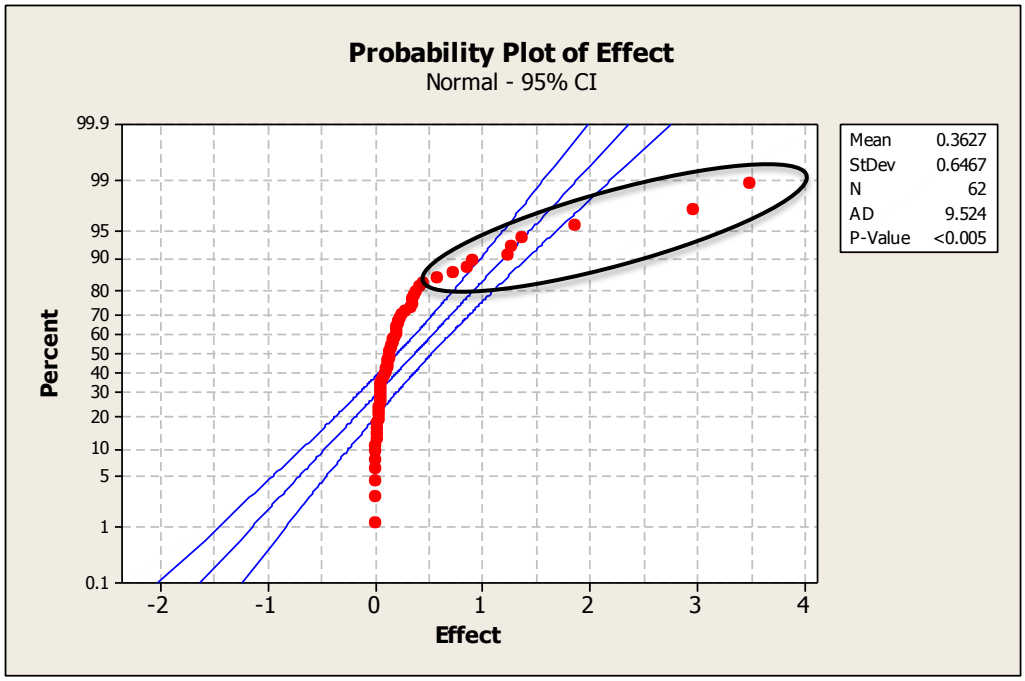

Figure 4.37 Normal probability plot obtained from MINITAB of the estimated effects.

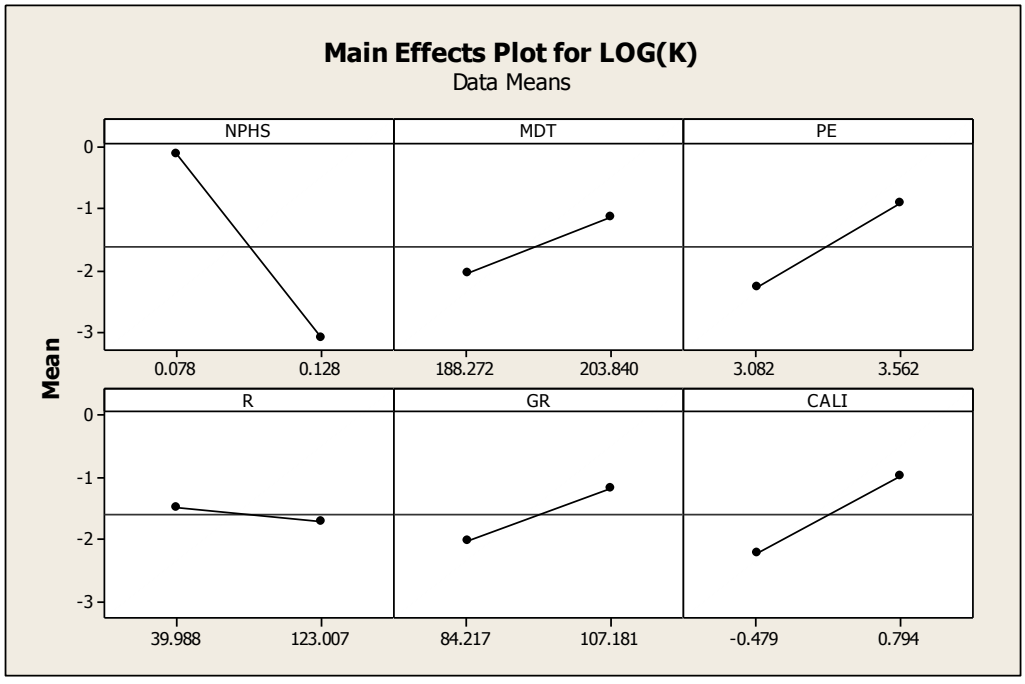

Figure 4.38 Main effects plot obtained from MINITAB for 6 factors.

Table 4.12 Six-factor ANOVA model obtained from MINITAB.

Source	DF	SS	MS	F	P
NPHS	1	140.85	140.85	266.52	0
MDT	1	13.275	13.275	25.12	0
PE	1	29.763	29.763	56.32	0
R	1	0.818	0.818	1.55	0.219
GR	1	11.713	11.713	22.16	0
CALI	1	24.65	24.65	46.64	0
NPHS*R	1	195.184	195.184	369.33	0
R*CALI	1	54.991	54.991	104.06	0
PE*R	1	26.116	26.116	49.42	0
NPHS*CALI	1	8.469	8.469	16.03	0
NPHS*PE	1	5.362	5.362	10.15	0.002
Error	32	27.481	0.528		
Total	63	538.673			

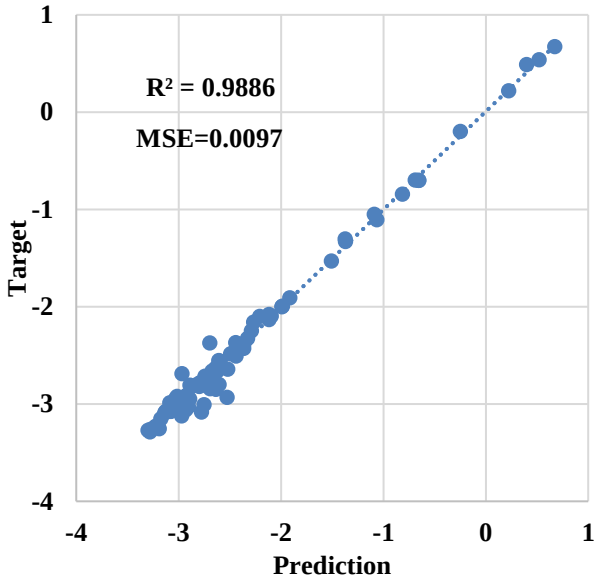

Figure 4.39 ANN results without RHOB and R.

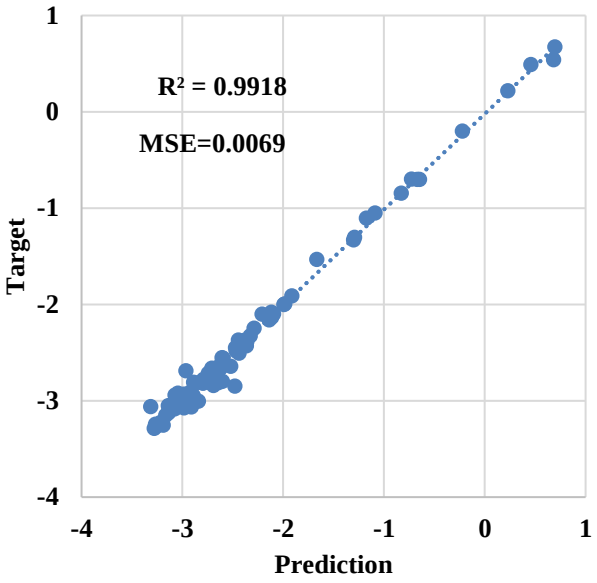

Figure 4.40 Original ANN results.

Based on the new model, a five-factor design for ANOVA analysis is made to investigate the importance of the five well log inputs on the permeability target (Table 4.13). Based on the probability plot (Fig. 4.41), all third-order to fifth-order interactions are reasonably considered as consisting only of random error, so they are eliminated from the ANOVA model in MINITAB and combined to estimate an error term (Table 4.14).

All factors are significant to results, and NPHS has the strongest influence on the prediction of permeability, which can also be explained in terms of the petrophysics.

1. The Neutron log measures a formation's hydrogen index, an indication of its richness in hydrogen. The relationship between the Neutron log porosity and the true porosity in clean formations is valid when either water or oil fills the pores because the two fluids have essentially the same hydrogen index. But shale has a relatively high hydrogen index due to the presence of both free water and bound water, which will make the Neutron log overestimate the porosity. Given the organic-rich character of some tight formations, organic matter might also affect Neutron log readings because of its high hydrogen index. But Ghanizadeh et al. (2015a) suggest that the Montney samples from this study

are organic-lean (<0.7 wt.% TOC content) which agrees with other Montney sample results from west-central Alberta (Beaton et al., 2010).

2. In the Montney Formation, the GR log has a downwards increasing trend, with the upper Montney being less shaly than the lower Montney (Fig. 4.24). This affects the relationship between porosity and permeability. Based on the plug porosity and permeability data, the correlation between porosity and permeability is much stronger in the upper Montney than in the lower Montney (Figs. 4.42 and 4.43).
3. In the upper Montney, NPHS can capture the trend of plug porosity (Fig. 4.44). In the lower Montney, NPHS increasingly deviates from plug porosity because of the increasing shale effects on neutron porosity (Figs. 4.44 and 4.45). So in the upper Montney, the significance of NPHS in predicting permeability may result from its capture of porosity. In the lower Montney, the shale effect on NPHS and permeability might be the factor that correlates them.

Table 4.13 Parameter settings.

Factor name		NPHS	MDT (us/m)	PE	GR (api)	CALI (mm)
Mean		0.103	196.056	3.322	95.699	0.157
Standard Deviation		0.025	7.784	0.240	11.482	0.636
Low Level	Value	0.078	188.272	3.082	84.217	-0.479
	Notation	1	1	1	1	1
High Level	Value	0.128	203.840	3.562	107.181	0.794
	Notation	a	b	c	d	e
Treatment Combinations		1, a, b, c, d, e, ab, ac, ad, ae, bc, bd, be, cd, ce, de, abc, abd, abe, acd, ace, ade, bcd, bce, bde, cde, abcd, abce, abde, acde, bcde, abcde				

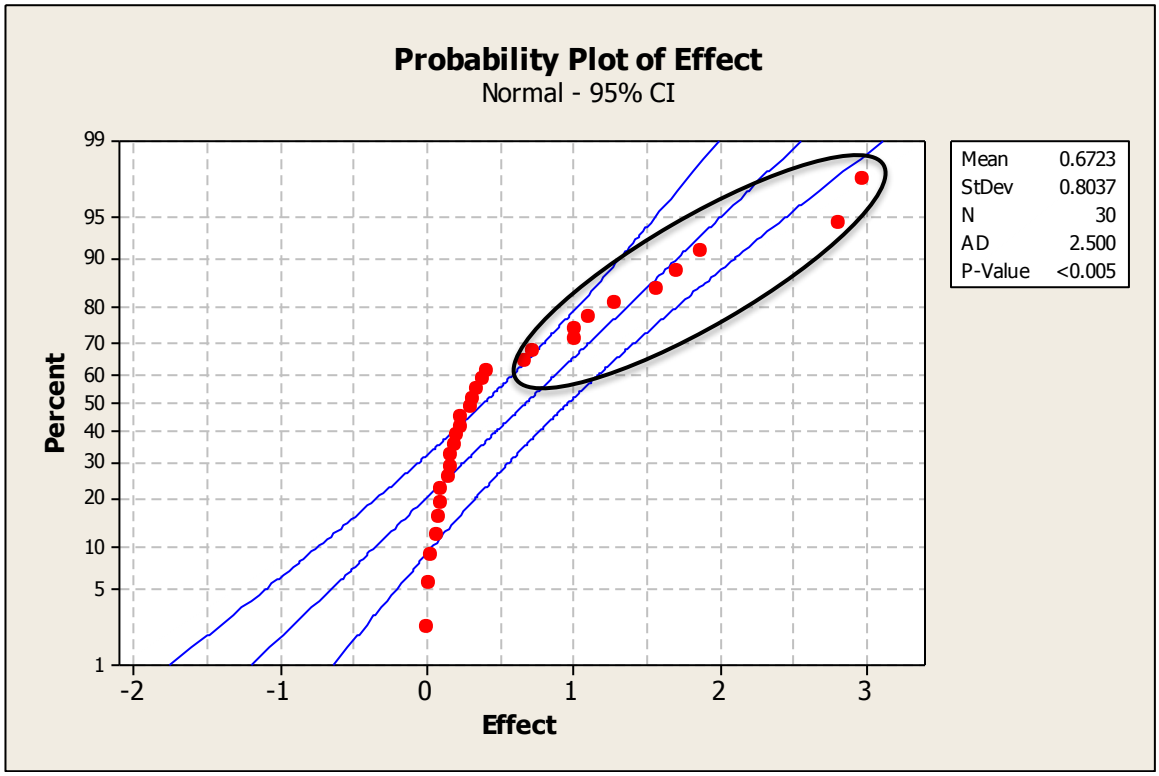

Figure 4.41 Normal probability plot obtained from MINITAB of the estimated effects.

Table 4.14 Five-factor ANOVA model obtained from MINITAB.

Source	DF	SS	MS	F	P
NPHS	1	70.402	70.402	66.11	0
MDT	1	9.788	9.788	9.18	0.008
PE	1	62.732	62.732	58.91	0
GR	1	13.185	13.185	12.38	0.003
CALI	1	19.615	19.615	18.42	0.001
NPHS*MDT	1	0.422	0.422	0.4	0.538
NPHS*PE	1	28.004	28.004	26.3	0
NPHS*GR	1	8.076	8.076	7.58	0.014
NPHS*CALI	1	3.63	3.63	3.41	0.083
MDT*PE	1	0.288	0.288	0.27	0.61
MDT*GR	1	0.712	0.712	0.67	0.426
MDT*CALI	1	23.303	23.303	21.88	0
PE*GR	1	0.062	0.062	0.06	0.812
PE*CALI	1	0.756	0.756	0.71	0.412
GR*CALI	1	0.951	0.951	0.89	0.359
Error	16	17.039	1.065		
Total	31	258.955			

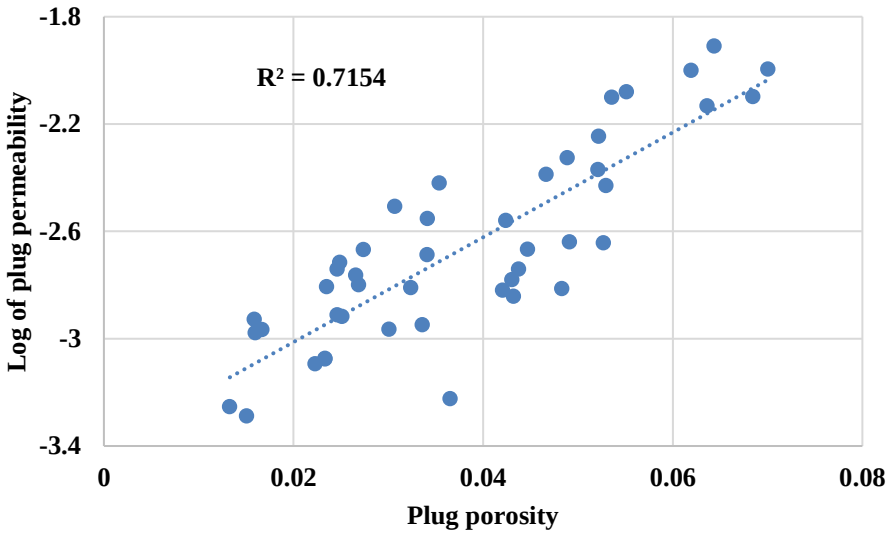

Figure 4.42 Correlation between plug porosity and plug permeability in the upper Montney.

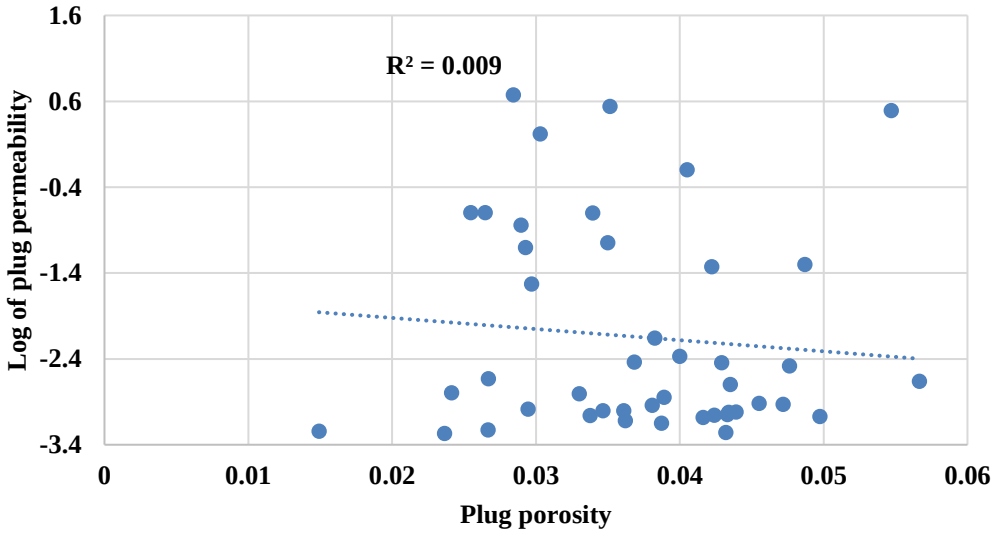

Figure 4.43 Correlation between plug porosity and plug permeability in the lower Montney.

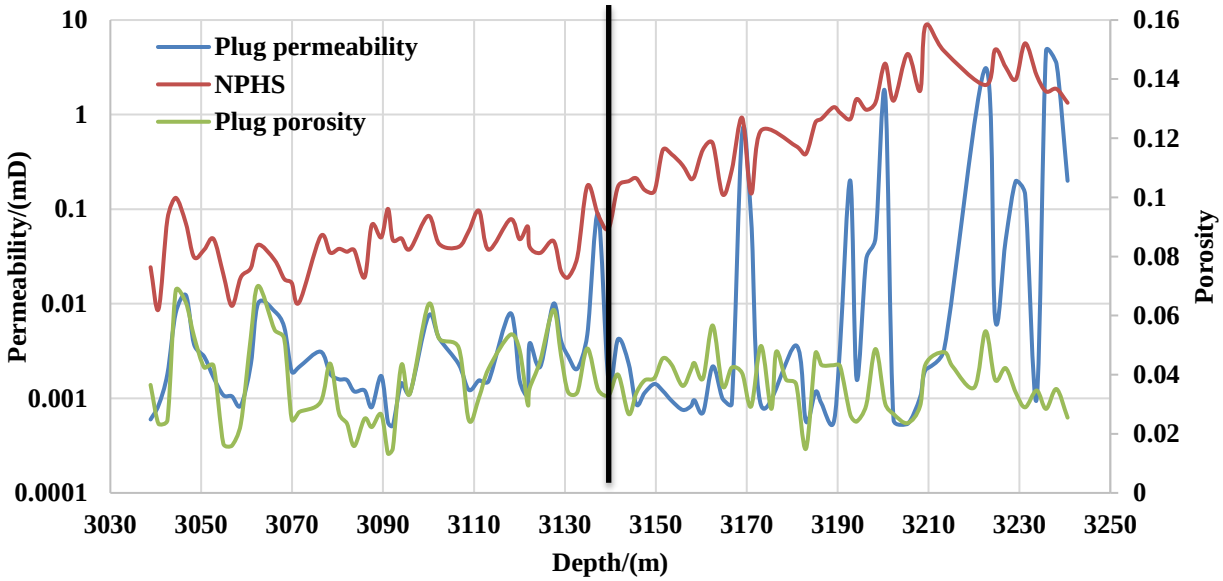

Figure 4.44 Comparison between plug permeability, plug porosity and NPHS.

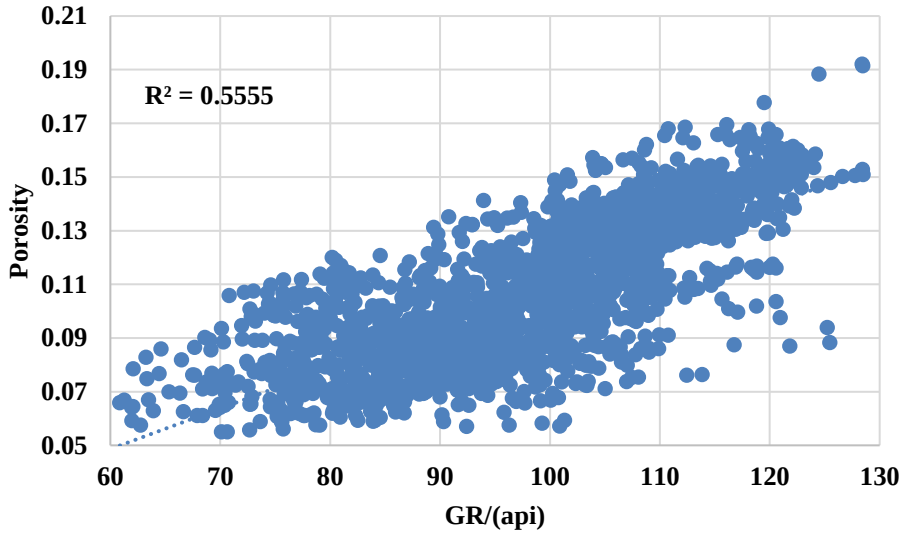

Figure 4.45 Correlation between NPHS and GR.

4.6 Discussion and Conclusions

In the ANN optimization workflow previously described, the optimized parameters are the number of hidden neurons, the number of validation checks, the dataset division, the weights and the biases with the other parameters set to the default setting. But the other parameters can also be optimized in different situations.

1) The default setting for the number of layers is 2, 1 hidden layer and 1 output layer. While two-layer networks can potentially learn virtually any input-output relationship, networks with more layers might learn complex relationships more quickly. For most problems, it is best to start with two layers, and then increase to three layers, if the performance with two layers is not satisfactory.

2) The default setting for the training function is trainlm using the Levenberg-Marquardt algorithm. A list of the training algorithms that are available in the Neural Network Toolbox is shown in the Table 4.15. For a detailed description of these techniques, see Demuth et al. (2014). It is very difficult to know which training algorithm will be the fastest for a given problem. It depends on many factors, including the complexity of the problem, the number of data points in the training set, the number of weights and biases in the network, the error goal, and whether the network is being used for pattern recognition (discriminant analysis) or function approximation (regression). In general, on function approximation problems, for networks that contain up to a few hundred weights, the Levenberg-Marquardt algorithm will have the fastest convergence. In many cases, trainlm is able to obtain lower mean square errors than the other algorithms. However, as the number of weights in the network increases, the advantage of trainlm decreases. In addition, trainlm performance is relatively poor on pattern recognition problems. The trainrp function is the fastest algorithm on pattern recognition problems. However, it does not perform well on function approximation problems.

Table 4.15 The training algorthims

Function	Algorithm
trainlm	Levenberg-Marquardt
trainbr	Bayesian Regularization
trainbfg	BFGS Quasi-Newton
trainrp	Resilient Backpropagation
trainscg	Scaled Conjugate Gradient
traincgb	Conjugate Gradient with Powell/Beale Restarts
traincgf	Fletcher-Powell Conjugate Gradient
traincgp	Polak-Ribiére Conjugate Gradient
trainoss	One Step Secant
traingdx	Variable Learning Rate Gradient Descent
traingdm	Gradient Descent with Momentum
traingd	Gradient Descent

This chapter presents a model optimization workflow for the ANN development, which produces accurate predictions for permeability with the $R^2 = 0.9918$ and the MSE = 0.0069. The image data are also indirectly incorporated into the model. It also compares prediction results of the multiple variables regression method, the empirical model method, and the ANN method to find that the ANN method is more accurate than the others in the permeability prediction using well logs because the ANN method can effectively capture the strong nonlinear relationship between permeability and well logs. The investigation of the main controls on permeability shows that NPHS, MDT, PE, GR and CALI are all significant to results, and NPHS has the highest influence on the prediction of permeability.

CHAPTER 5: Conclusions and Future Work

5.1 Conclusions

This study has discussed heterogeneity characterization and permeability prediction methods based on a dataset from one well in the Montney Formation. The main results are listed below:

1. The middle Montney is missing in the studied well. The core descriptions and images show that the Montney in the studied well is dominantly composed of variable amounts of interbedded fine-grained shale, sandstone and siltstone.
2. The whole core permeability measurements cannot characterize the dominant small permeability variations (10^{-3} to 10^{-4} mD) because of their resolution limit and censorship. The vertical permeability measurements are also censored, so the anisotropy (k_v/k_h ratio) also cannot be characterized.
3. The core plug permeability measurements are the only available data that are consistently measured on the entire Montney Formation and can characterize the dominant 10^{-3} to 10^{-4} mD level permeability variations at the same time, which will be the target data for permeability prediction.
4. Both core plug and whole core permeability are largely under-sampled at the 20% tolerance but are adequately sampled for a 91% tolerance and a 51% tolerance respectively. The five missing core plug permeability measurements might make the small permeability underrepresented.
5. The porosity-permeability relationship is weak in the Montney, suggesting that porosity (i.e., sorting) is not the unique permeability controlling factor. But the relationship is much stronger in the upper Montney than in the lower Montney, which may result from the increasing shale effects in the lower Montney.
6. The probe permeability measurement is capable of differentiating and quantifying the permeability of laminaset scale heterogeneities in rock. But, for the sample spacing and resolution for this dataset, this technique cannot fully characterize the highly heterogeneous reservoirs such as the Montney, where lithological heterogeneities are abundant vertically along the cores on a lamination scale basis.

7. A digital image analysis program (written in Matlab) can correct the effects of fractures, convert the image into gray value data and identify the locations and thicknesses of laminations.
8. The digital image analysis can quantitatively characterize lamination scale heterogeneity correctly and efficiently, which can be a useful guide when planning core measurements and improving sampling efficiency, so that sample density is consistent with the heterogeneity.
9. The semivariogram of the gray value data shows 20 – 30 cm cycles that can be identified on the core images. But the cycles are absent in the semivariogram of the probe permeability data because of the large sample sapcing. The cycles likely reflect the depositional bed-set thickness.
10. The gray value data correlate well with the mechanical hardness data with $R^2 = 0.71$ and can be used to predict the mechanical hardness with an error percentage of 2.4% because the images partly reflect the mineralogical variations.
11. The ANN developed from the workflow produces predictions for permeability with $R^2 = 0.99$ and the MSE = 0.0069, which is more accurate than the predictions from the multiple variables regression, $R^2 = 0.42$ and the MSE = 0.4828, and Yao and Holditch (1993) permeability models, $R^2 = 0.23$ and the MSE = 0.6400.
12. The incorporation of digital image data into the ANN does not change the overall training performance very much, but the test dataset performance is dramatically improved indicating that the generalization power of the ANN model is greatly improved.
13. The ANOVA analysis based on the ANN model indicates that NPHS, MDT, PE, GR and CALI are all significant to results, and NPHS has the strongest influence on the prediction of permeability.

5.2 Future Work

This work uses petrophysical and data analysis method to analyze the degree of heterogeneity of the Montney Formation and discusses the permeability prediction using multiple variable regression, empirical model and ANN, which needs lots of work to improve:

1. The depth match of core and well log data is the basis for much of this study and a reassessment might yield different results. The core depths are "stretched" compared to the well log depths. Take Interval 1 for example (Fig. 5.1). If we match the depth over the shallower part of this interval, there will be a consistent mismatch of 30-40 cm in the deeper part. In order to avoid an overlap with the next interval, we cannot shift it separately. Also because of this issue, Intervals 6, 7 and 8 are reconnected after the shift, but there should be missing cores between each interval. This issue may result from the mishandling of cores (Morton-Thompson and Woods, 1993) and needs to be improved.

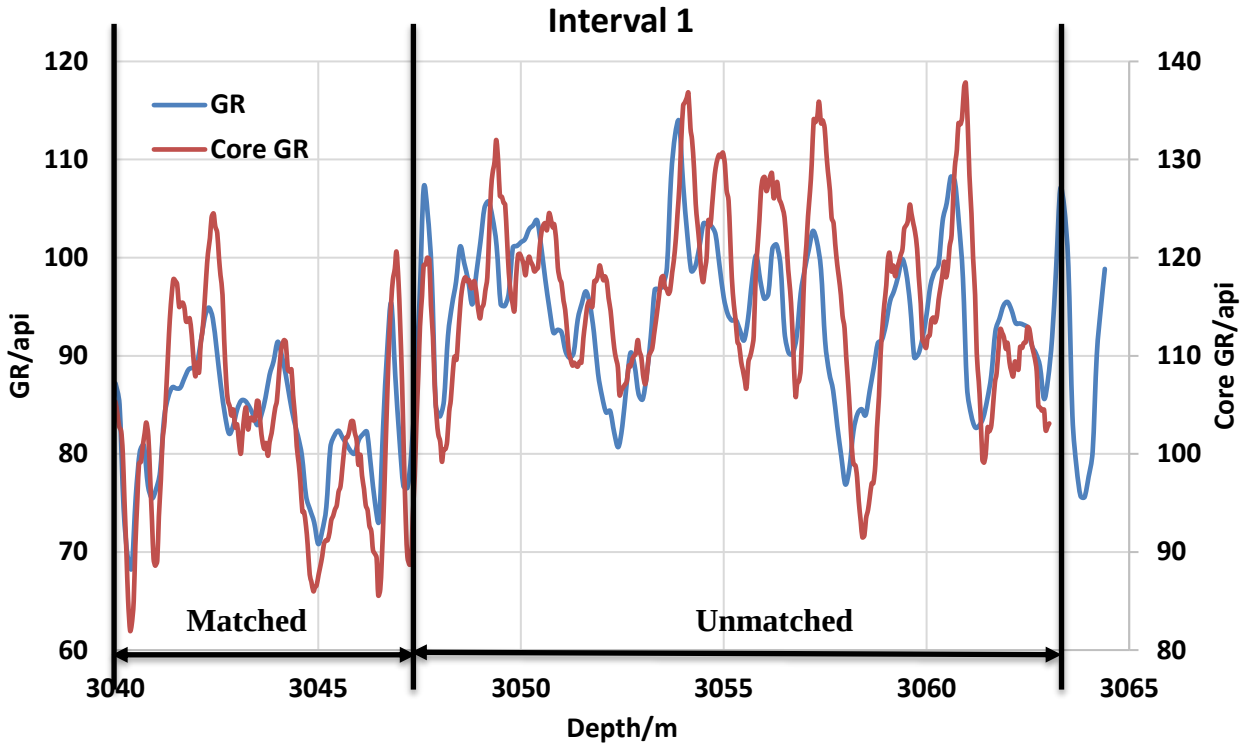

Figure 5.1 Depth shift example of Interval 1

2. The digital image results may also be used to improve the depth matching quality. The gray value data correlate well with other petrophysical measurements in Chapter 3, so they can be compared with well log data, especially GR log, for depth matching. But the scale issue has to be considered because the digital image results reflect the variations at the pixel scale (<1 mm) which is much smaller than the investigation scale (>10 cm) of well logs. The gray value data would likely need to be upscaled properly before they are compared with well log data.
3. In the digital analysis program, the two dimensional image is simplified into a one dimensional line by averaging the gray value of each pixel column in the central 1/3rd of the core. While most of the laminations present in cores are horizontal, where this simplification is effective, there are non-horizontal layers and features such as bioturbation and concretions, which cannot be so simply characterized. So it might be possible to use the 2 D image data to extract more features or, at least, assign

uncertainty values to the 1D line values to indicate the level of variability inherent in each average value.

4. In the lower Montney, there are 7 core plugs permeability measurements which may be affected by fractures. These 7 abnormally high permeability measurements may cause the ANN to overfit in the lower Montney. This influence needs to be further evaluated.
5. The investigation depth of well log measurements is from 10 to 20 cm (Rider and Kennedy, 2011) which is much greater than that of core plug measurements (2.54 cm). So the well log data represent the average quality of a much larger formation volume, which I correlated to that of smaller volume in the ANN. The difference of volume scales may cause the mismatch of data and overfitting of the model, which needs to be further investigated.
6. The incorporation of digital image data into ANN can be improved if images of ends of core plugs are also available. The digital image results of core plug ends correspond precisely to the core plug measurements without scale issues.
7. The applicability of this ANN model on other wells from the same geological context needs to be investigated, which needs the corresponding well log data and core data from other wells. By this method, the degree of overfitting could also be assessed.
8. Well logs are measured at in-situ conditions but core plug permeability is not. Clarkson et al. (2012a) suggest that permeability measurements have considerable stress-dependence using samples from the Montney, so "adjustment" of the core plug permeability to in-situ conditions may enhance the correlation between well logs and permeability measurements.
9. NPHS is determined to be the strongest influence on the prediction of permeability. But its significance needs to be further investigated and explained.

REFFERENCES

Ali Ahmadi, M., Zendehboudi, S., Lohi, A., Elkamel, A., and Chatzis, I. 2013. Reservoir permeability prediction by neural networks combined with hybrid genetic algorithm and particle swarm optimization. Geophysical Prospecting, 61(3), 582-598.

Amaefule, J. O., Altunbay, M., Tiab, D., Kersey, D. G., and Keelan, D. K. 1993, January. Enhanced reservoir description: using core and log data to identify hydraulic (flow) units and predict permeability in uncored intervals/wells. In SPE annual technical conference and exhibition. Society of Petroleum Engineers, SPE-26436-MS.

Aminian, K., Bilgesu, H. I., Ameri, S., & Gil, E. 2000, January. Improving the simulation of waterflood performance with the use of neural networks. In SPE Eastern Regional Meeting. Society of Petroleum Engineers, SPE-65630-MS.

Anselmetti, F. S., & Eberli, G. P. 1993. Controls on sonic velocity in carbonates. Pure and Applied geophysics, 141(2-4), 287-323.

Baraka-Lokmane, S., Main, I. G., Ngwenya, B. T., & Elphick, S. C. 2009. Application of complementary methods for more robust characterization of sandstone cores. Marine and Petroleum Geology, 26(1), 39-56.

Beale, M. H., Hagan, M. T., and Demuth, H. B. 2016. Neural Network Toolbox™ User's Guide. The Mathworks Inc.

Beaton, A. P., Pawlowicz, J. G., Anderson, S. D. A., Berhane, H., & Rokosh, C. D. 2010. Rock Eval^{TM}, total organic carbon and adsorption isotherms of the Montney formation in Alberta: shale gas data release. Energy Resources Conservation Board, ERCB/AGS Open File Report 2010-05

Bowden, G. J., Maier, H. R., & Dandy, G. C. 2002. Optimal division of data for neural network models in water resources applications. Water Resources Research, 38(2), 1010, doi:10.1029/2001WR000266.

Braddock, R. D., Kremmer, M. L., & Sanzogni, L. 1998. Feed-forward artificial neural network model for forecasting rainfall run-off. Environmetrics, 9(4), 419-432.

Chalmers, G. R., and Bustin, R. M. 2012. Geological evaluation of Halfway–Doig–Montney hybrid gas shale–tight gas reservoir, northeastern British Columbia. Marine and Petroleum Geology, 38(1), 53-72.

Clarkson, C. R., Jensen, J. L., and Chipperfield, S. 2012a. Unconventional gas reservoir evaluation: what do we have to consider? Journal of Natural Gas Science and Engineering, 8, 9-33.

Clarkson, C. R., Jensen, J. L., Pedersen, P. K., and Freeman, M. 2012b. Innovative methods for flow-unit and pore-structure analyses in a tight siltstone and shale gas reservoir. AAPG bulletin, 96(2), 355-374.

Crabtree Jr, S. J., Ehrlich, R., & Prince, C. 1984. Evaluation of strategies for segmentation of blue-dyed pores in thin sections of reservoir rocks. Computer vision, graphics, and image processing, 28(1), 1-18.

Coates, G. R., & Dumanoir, J. L. 1973, January. A new approach to improved log-derived permeability. In SPWLA 14th Annual Logging Symposium. Society of Petrophysicists and Well-Log Analysts, SPWLA-1973-R.

Coles, M. E., Muegge, E. L., and Sprunt, E. S. 1991. Applications of CAT scanning for oil and gas production. IEEE Transactions on nuclear science, 38(2), 510-515.

Coles, M. E., Hazlett, R. D., Spanne, P., Soll, W. E., Muegge, E. L., and Jones, K. W. 1998. Pore level imaging of fluid transport using synchrotron X-ray microtomography. Journal of Petroleum Science and Engineering, 19(1-2), 55-63.

Davies, G. R., Moslow, T. F., and Sherwin, M. D. 1997. The lower Triassic Montney formation, west-central Alberta. Bulletin of Canadian Petroleum Geology, 45(4), 474-505.

Davies, G. R. 1997. The Triassic of the Western Canada Sedimentary Basin: tectonic and stratigraphic framework, paleogeography, paleoclimate and biota. Bulletin of Canadian Petroleum Geology, 45(4), 434-460.

Davies, G. R., and Sherwin, M. D. 1997. Productive dolomitized coquinal facies of the Lower Triassic Montney Formation, Western Canada Sedimentary Basin. Canadian Society of Petroleum Geologists/Society for Sedimentary Geology (SEPM), Joint Convention, Calgary, Core Workshop Notes, p. 257-276.

Demuth, H. B., Beale, M. H., De Jess, O., and Hagan, M. T. 2014. Neural network design. Martin Hagan.

Derder, O. M. 2012. Characterizing reservoir properties for the Lower Triassic Montney Formation (units C and D) based on petrophysical methods (Masters thesis, University of Calgary).

Devore, J., Farnum, N., and Doi, J. 2013. Applied statistics for engineers and scientists. Nelson Education.

Di, J. 2015. Permeability characterization and prediction in a tight oil reservoir, Edson Field, Alberta (Masters thesis, University of Calgary).

Dixon, J. 2000. Regional lithostratigraphic units in the Triassic Montney Formation of western Canada. Bulletin of Canadian Petroleum Geology, 48(1), 80-83.

Egbobawaye, E. I. 2013. Tight Gas Reservoir Characterization in Montney Formation, Northeastern British Columbia, Western Canada. (Doctoral thesis, University of Calgary).

Ehrlich, R., Kennedy, S. K., Crabtree, S. J., and Cannon, R. L. 1984. Petrographic image analysis, I. Analysis of reservoir pore complexes. Journal of Sedimentary Research, 54(4), 1365-1378.

Ehrlich, R., Crabtree, S. J., Horkowitz, K. O., and Horkowitz, J. P. 1991a. Petrography and reservoir physics I: Objective classification of reservoir porosity (1). AAPG Bulletin, 75(10), 1547-1562.

Ehrlich, R., Etris, E. L., Brumfield, D., Yuan, L. P., and CRABTREE, S. J. 1991b. Petrography and Reservoir Physics III: Physical Models for Permeability and Formation Factor (1). AAPG Bulletin, 75(10), 1579-1592.

Flood, I., & Kartam, N. (1994). Neural networks in civil engineering. I: Principles and understanding. Journal of computing in civil engineering, 8(2), 131-148.

Gerard, R. E., Philipson, C. A., Manni, F. M., and Marschall, D. M. 1992. Petrographic image analysis: an alternate method for determining petrophysical properties. In Automated pattern analysis in petroleum exploration (pp. 249-263). Springer, New York, NY.

Gibson, D. W., and Edwards, D. E. 1990. An overview of Triassic stratigraphy and depositional environments in the Rocky Mountain foothills and western interior plains, Peace River arch area, northeastern British Columbia. Bulletin of Canadian Petroleum Geology, 38(1), 146-158.

Ghanizadeh, A., Clarkson, C. R., Aquino, S., Ardakani, O. H., and Sanei, H. 2015a. Petrophysical and geomechanical characteristics of Canadian tight oil and liquid-rich gas reservoirs: I. Pore network and permeability characterization. Fuel, 153, 664-681.

Ghanizadeh, A., Clarkson, C. R., Aquino, S., Ardakani, O. H., and Sanei, H. 2015b. Petrophysical and geomechanical characteristics of Canadian tight oil and liquid-rich gas reservoirs: II. Geomechanical property estimation. Fuel, 153, 682-691.

Hornik, K., Stinchcombe, M., & White, H. 1989. Multilayer feedforward networks are universal approximators. Neural networks, 2(5), 359-366.

Irani, R., and Nasimi, R. 2011. Evolving neural network using real coded genetic algorithm for permeability estimation of the reservoir. Expert Systems with Applications, 38(8), 9862-9866.

Jensen, J. L., Lake, L. W., Corbett, P., and Goggin, D. 2000. Statistics for petroleum engineers and geoscientists (Vol. 2). Gulf Professional Publishing.

Jones, S. C. 1994. A new, fast, accurate pressure-decay probe permeameter. SPE Formation Evaluation, 9(03), 193-199.

Keller, L. M., Holzer, L., Wepf, R., & Gasser, P. 2011. 3D geometry and topology of pore pathways in Opalinus clay: Implications for mass transport. Applied Clay Science, 52(1-2), 85-95.

Kendall, D. R. 1999. Sedimentology and stratigraphy of the lower triassic Montney Formation, Peace River basin, subsurface of northwestern Alberta. (Masters thesis, University of Calgary).

Ketcham, R. A., & Carlson, W. D. 2001. Acquisition, optimization and interpretation of X-ray computed tomographic imagery: applications to the geosciences. Computers & Geosciences, 27(4), 381-400.

Lesher, C. E., Wang, Y., Gaudio, S., Clark, A., Nishiyama, N., and Rivers, M. 2009. Volumetric properties of magnesium silicate glasses and supercooled liquid at high pressure by X-ray microtomography. Physics of the Earth and Planetary Interiors, 174(1-4), 292-301.

McCreesh, C. A., Ehrlich, R., and Crabtree, S. J. 1991. Petrography and Reservoir Physics II: Relating Thin Section Porosity to Capillary Pressure, the Association Between Pore Types and Throat Size (1). AAPG Bulletin, 75(10), 1563-1578.

Minns, A. W., & Hall, M. J. 1996. Artificial neural networks as rainfall-runoff models. Hydrological sciences journal, 41(3), 399-417.

Mohaghegh, S., Arefi, R., Ameri, S., & Rose, D. 1995. Design and development of an artificial neural network for estimation of formation permeability. SPE Computer Applications, 7(06), 151-154.

Mohaghegh, S., Balan, B., Ameri, S., 1997. Permeability determination from well log data: SPE Formation Evaluation, 12 (3):170-174.

Morton-Thompson, D., and Woods, A. M. (Eds.). 1993. Development geology reference manual: AAPG methods in exploration series, no. 10. AAPG.

Nederbragt, A. J., Dunbar, R. B., Osborn, A. T., Palmer, A., Thurow, J. W., & Wagner, T. 2006. Sediment colour analysis from digital images and correlation with sediment composition. Geological Society, London, Special Publications, 267(1), 113-128.

Nielson, D. L., Nash, G. D., & White, W. S. 1995. Reservoir characterization using image analysis of core examples from The Geysers geothermal field: E. Barbier, G. Frye, E. Iglesias, G. Pàlmason. In Proceedings of the World Geothermal Congress (pp. 3017-3021).

Orsi, T. H., Edwards, C. M., & Anderson, A. L. 1994. X-ray computed tomography: a nondestructive method for quantitative analysis of sediment cores. Journal of Sedimentary Research, 64(3), 690–693.

Prince, C. M., Dixon, M. W., & Haynes, L. L. 2005. The use of high-resolution core imagery in reservoir characterization: an example from unlithified Miocene turbidites. In paper SCA2005-02, Society of Core Analysts Annual International Symposium.

Rivard, C., Lavoie, D., Lefebvre, R., Séjourné, S., Lamontagne, C., and Duchesne, M. 2014. An overview of Canadian shale gas production and environmental concerns. International Journal of Coal Geology, 126, 64-76.

Rivers, M. L., Wang, Y., & Uchida, T. 2004, October. Microtomography at GeoSoilEnviroCARS. In Developments in X-Ray Tomography IV (Vol. 5535, pp. 783-792). International Society for Optics and Photonics.

Ruzyla, K., & Jezek, D. I. 1987. Staining method for recognition of pore space in thin and polished sections. Journal of Sedimentary Research, 57(4), 777-778.

Saemi, M., Ahmadi, M., and Varjani, A. Y. 2007. Design of neural networks using genetic algorithm for the permeability estimation of the reservoir. Journal of Petroleum Science and Engineering, 59(1-2), 97-105.

Sanei, H., Wood, J. M., Ardakani, O. H., Clarkson, C. R., & Jiang, C. 2015. Characterization of organic matter fractions in an unconventional tight gas siltstone reservoir. International Journal of Coal Geology, 150, 296-305.

Saner, S., Kissami, M., Nufaili, S.A., 1997. Estimation of permeability from well logs using resistivity and saturation data: SPE Formation Evaluation, 12 (3): 27-31.

Solano, N. A., Clarkson, C. R., and Krause, F. 2012, January. Quantification of cm-Scale Heterogeneities in Tight-Oil Intervals of the Cardium Formation at Pembina, WCSB, Alberta, Canada. In SPE Canadian Unconventional Resources Conference. Society of Petroleum Engineers. SPE-162837-MS.

Sondergeld, C. H., Newsham, K. E., Comisky, J. T., Rice, M. C., & Rai, C. S. 2010, January. Petrophysical considerations in evaluating and producing shale gas resources. In SPE Unconventional Gas Conference. Society of Petroleum Engineers. SPE-131768-MS.

Stone, M. 1974. Cross-validatory choice and assessment of statistical predictions. Journal of the royal statistical society. Series B (Methodological), 111-147.

Tixier, M. P. 1949. Evaluation of permeability from electric-log resistivity gradients. Oil and Gas Journal, 48(6), 113-123.

Tokar, A. S., & Johnson, P. A. 1999. Rainfall-runoff modeling using artificial neural networks. Journal of Hydrologic Engineering, 4(3), 232-239.

Van Vliet, L. J., & Verbeek, P. W. 1995, May. Estimators for orientation and anisotropy in digitized images. In ASCI (Vol. 95, pp. 16-18).

Wong, P. M., Jang, M., Cho, S., & Gedeon, T. D. 2000. Multiple permeability predictions using an observational learning algorithm. Computers & Geosciences, 26(8), 907-913.

Xue, G.P., Gupta, A.D., Valko, P., Blasingame, T., 1997. Optimal transformations for multiple regression: application to permeability estimation from well logs: SPE Formation Evaluation, 12 (2): 85-93.

Yanguas, J. E., & Dravis, J. J. 1985. Blue fluorescent dye technique for recognition of microporosity in sedimentary rocks. Journal of Sedimentary Research, 55(4), 600-603.

Yao, C.Y. and Holditch, S.A., 1993. Estimating Permeability Profiles Using Core and Log Data: SPE-26921-MS.

Zonneveld, J. P., Moslow, T. F., and Henderson, C. M. 1997. Lithofacies associations and depositional environments in a mixed siliciclastic-carbonate coastal depositional system, upper Liard Formation, Triassic, northeastern British Columbia. Bulletin of Canadian Petroleum Geology, 45(4), 553-575.

APPENDIX

The fracture correction program workflow is described as follows:

1. Use 50 as the cutoff value to identify the fracture zone.
2. The areas within the distance of 5 pixels from the boundary of the fracture zone are identified to be the transition zones.
3. The averaged gray value in the transition zones are used as the cutoff value for correction.
4. The fracture zone and the transition zones are combined into the correction zone.
5. From left to right, check the gray value of each pixel in the correction zone. If the gray value is smaller than the cutoff value, this pixel will be replaced by the one at its left.