

การใช้โครงข่ายประสาทเทียมทำนายการผลิตก๊าซในกรณีที่มีการเจาะหลุมผลิตเพิ่ม



นาย รัชฎาเทพ เต็มเกียรติวิเศษ

ศูนย์วิทยพัทยากร
จุฬาลงกรณ์มหาวิทยาลัย

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต

สาขาวิชาวิศวกรรมปิโตรเลียม ภาควิชาวิศวกรรมเหมืองแร่และปิโตรเลียม

คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

ปีการศึกษา 2553

ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

USING ARTIFICIAL NEURAL NETWORK TO PREDICT GAS PRODUCTION
FOR INFILL WELLS

Mr. Nattaphon Temkiatvises



ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Engineering Program in Petroleum Engineering

Department of Mining and Petroleum Engineering

Faculty of Engineering


Chulalongkorn University

Academic Year 2010

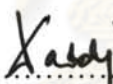
Copyright of Chulalongkorn University


Thesis Title USING ARTIFICIAL NEURAL NETWORK TO
 PREDICT GAS PRODUCTION FOR INFILL
 WELLS
By Mr. Nattaphon Temkiatvises
Field of Study Petroleum Engineering
Thesis Advisor Assistant Professor Suwat Athichanagorn, Ph.D.


Accepted by the Faculty of Engineering, Chulalongkorn University in
Partial Fulfillment of the Requirements for the Master's Degree


.....Dean of the Faculty of Engineering
(Associate Professor Boonsom Lerdhirunwong, Dr.Ing.)

THESIS COMMITTEE


.....Chairman
(Associate Professor Sarithdej Pathanasethpong)


.....Thesis Advisor
(Assistant Professor Suwat Athichanagorn, Ph.D.)


.....External Examiner
(Siree Nasakul, Ph.D.)

ณัฐพน เต็มเกียรติวิเศษ : การใช้โครงข่ายประสาทเทียมทำนายการผลิตก๊าซในกรณีที่มีการเจาะหลุมผลิตเพิ่ม (USING ARTIFICIAL NEURAL NETWORK TO PREDICT GAS PRODUCTION FOR INFILL WELLS) อ. ที่ปรึกษาวิทยานิพนธ์หลัก: ผศ.ดร. สุวัฒน์ อธิชนากร, 152 หน้า.

ในปัจจุบันนี้ แหล่งผลิตก๊าซทั่วโลกพยายามที่จะเพิ่มปริมาณการผลิตก๊าซขึ้นมาจากแหล่งกักเก็บด้านล่างให้ได้มากที่สุด การเจาะหลุมผลิตเพิ่มนั้นถือว่าเป็นวิธีที่มีประสิทธิภาพวิธีหนึ่ง ซึ่งตัวแปรที่จะเป็นตัวช่วยในการตัดสินใจว่าจะทำการเจาะหลุมผลิตเพิ่มหรือไม่นั้นก็คืออัตราการไหลเริ่มต้นและปริมาณก๊าซสะสมที่ผลิตได้นั้นเอง การเจาะหลุมผลิตเพิ่มจะประสบความสำเร็จมากน้อยเพียงใดนั้นก็ขึ้นอยู่กับความแม่นยำของการทำนายค่าเหล่านี้ ยิ่งทำนายได้แม่นยำเท่าไรก็ยิ่งทำให้การตัดสินใจในการเจาะนั้นทำได้มากขึ้นเท่านั้น ซึ่งมีวิธีหนึ่งที่น่าจะช่วยในการทำนายตำแหน่งที่เหมาะสมในการเจาะได้ก็คือโครงข่ายประสาทเทียมนั่นเอง โดยโครงข่ายประสาทเทียมนี้นั้นสามารถที่จะเรียนรู้ข้อมูลในอดีตเพื่อที่จะหาความสัมพันธ์ที่ซับซ้อนระหว่างตัวแปรที่ป้อนเข้าไปให้กับโครงข่ายกับตัวแปรที่ออกมาจากโครงข่ายได้อย่างมีประสิทธิภาพ

ในการศึกษานี้จะทำการร่วมกับการสร้างแบบจำลองของแหล่งกักเก็บก๊าซธรรมชาติ โดยจะนำข้อมูลทางการผลิตจากแบบจำลองมาใช้ในการสอนโครงข่ายเพื่อที่จะทำนายการผลิตก๊าซที่คาดว่าจะได้ในบริเวณที่ยังไม่ได้เจาะ มีตัวแปรที่จะป้อนเข้าไปในโครงข่ายมากมายที่ถูกพิจารณา คัดกรองและเลือกเพื่อก่อนนำมาใช้งาน ทำการทดลองโดยแบ่งเป็นหลายๆกรณีเพื่อศึกษาถึงความสำคัญของตัวแปรแต่ละตัวที่มีต่อผลการทดลอง และท้ายสุดทำการประเมินความแม่นยำของค่าที่ทำนายได้จากโครงข่ายประสาทเทียม จากผลที่ได้จะเห็นว่าโครงข่ายประสาทเทียมสามารถทำนายการผลิตก๊าซได้อย่างแม่นยำ อย่างไรก็ตามพบว่ามีคลาดของการทำนายอยู่มากเมื่อนำมาใช้ทำนายผลการทดลองที่ค่อนข้างมีการประมาณค่าในช่วงของข้อมูลที่มี

ภาควิชา วิศวกรรมเหมืองแร่และปิโตรเลียมลายมือชื่อนิติศ ณัฐพน เต็มเกียรติวิเศษ
สาขาวิชา วิศวกรรมปิโตรเลียมลายมือชื่อ อ.ที่ปรึกษาวิทยานิพนธ์หลัก *สม อธิ*
ปีการศึกษา 2553

5171606521: MAJOR PETROLEUM ENGINEERING

KEYWORDS: GAS PRODUCTION / ARTIFICIAL NEURAL NETWORK /
INFILL WELLS

NATTAPHON TEMKIATVISES : USING ARTIFICIAL NEURAL
NETWORK TO PREDICT GAS PRODUCTION FOR INFILL WELLS.

ADVISOR: ASST. PROF. SUWAT ATHICHANAGORN, Ph.D., 152 pp.

Nowadays, more and more fields all over the world are trying to improve gas recovery factor. Infill drilling is one effective method for this purpose. The expected gas production rate and cumulative gas production is a key component in determining whether or not to drill a well. The success of adding a new well to the field depends on the accuracy of prediction of the gas production, as the more accurate the prediction is, the better the decision on drilling location will be. One interesting technique to predict an appropriate location for infill well is Artificial Neural Network (ANN) which can learn from the historical data to create a representation of complex relationship between input and output samples.

In this study, the ANN is applied in conjunction with numerical reservoir simulation. Production data generated from a numerical simulator were used to train the network to forecast gas production at undrilled location. Many input parameters were considered, screened, and chosen in order to study their effect on the result. A few case studies were performed to highlight the importance of these input parameters. Finally, the prediction performance of ANN was evaluated. The results show that the ANN can be effectively used to predict gas production with accurate prediction performance. However, substantial errors still occurred at some well locations due to the inaccuracies when using ANN to predict the output based on an extrapolation basis.

Department: Mining and Petroleum Engineering..... Student's Signature: *ณัฐพนธ์ เตมกีเกียรติยศ*

Field of Study: Petroleum Engineering..... Advisor's Signature: *Suwat Athichanagorn*

Academic Year: 2010.....

ACKNOWLEDGEMENTS

First of all, I would like to express my gratitude to Assistant Professor Dr. Suwat Athichanagorn, my advisor, for giving knowledge of petroleum engineering and invaluable guidance, advice, and encouragement throughout the development of this thesis.

I also would like to extend my thanks to PTT Exploration and Production Public Company Limited for providing the valuable information for this study to develop and study the reservoir model.

I would like to thank the thesis committee members for their comments and recommendations.

I would like to thank all faculty members in the Department of Mining and Petroleum Engineering who have offered their petroleum knowledge, technical advice, and invaluable consultation.

I would like to thank Schlumberger for providing ECLIPSE license to the Department of Mining and Petroleum Engineering. Without the software, this study would not have been completed.

Lastly, I would like to express my gratitude to my family and friends for their understanding, encouragement and continuing support throughout the course of my studies.

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

CONTENTS

	Page
Abstract (Thai).....	iv
Abstract (English).....	v
Acknowledgements.....	vi
Contents.....	vii
List of Tables.....	x
List of Figures.....	xi
List of Abbreviations.....	xviii
Nomenclature.....	xix
Chapter	
I INTRODUCTION.....	1
1.1 Outline of Methodology.....	3
1.2 Thesis Outline.....	4
II LITERATURE REVIEW.....	5
2.1 Prediction of Infill Drilling Performance.....	5
2.2 Artificial Neural Network (ANN).....	7
III THEORIES AND CONCEPTS.....	9
3.1 Artificial Neural Network (ANN).....	9
3.1.1 Biological Neuron.....	10
3.1.2 Artificial Neuron.....	11
3.1.3 Network Structure.....	13
3.1.4 Learning Algorithm.....	16
IV RESERVOIR SIMULATION MODEL.....	20
V ARTIFICIAL NEURAL NETWORK MODEL DEVELOPMENT.....	27
5.1 Output of ANN.....	27
5.2 Input of ANN.....	28
5.3 Partitioning Data Sets.....	32
5.4 Artificial Neural Network Case Study.....	37

	Page
5.4.1 Initial Gas Rate Prediction.....	37
5.4.1.1 Case 1-1	38
5.4.1.1.1 Data Preprocessing.....	38
5.4.1.1.2 Model Training	43
5.4.1.1.3 Model Testing Results and Discussion	51
5.4.1.2 Case 1-2-1	53
5.4.1.2.1 Data Preprocessing.....	54
5.4.1.2.2 Model Training	57
5.4.1.2.3 Model Testing Results and Discussion	62
5.4.1.3 Case 1-2-2.....	64
5.4.1.3.1 Data Preprocessing.....	65
5.4.1.3.2 Model Training	68
5.4.1.3.3 Model Testing Results and Discussion	73
5.4.1.4 Case 1-3	75
5.4.1.4.1 Data Preprocessing.....	75
5.4.1.4.2 Model Training	84
5.4.1.4.3 Model Testing Results and Discussion	89
5.4.1.5 Performance of ANN Prediction	91
5.4.2 Cumulative Gas Production Prediction	100
5.4.2.1 Case 2-1	101
5.4.2.1.1 Data Preprocessing.....	101
5.4.2.1.2 Model Training	102
5.4.2.1.3 Model Testing Results and Discussion	107
5.4.2.2 Case 2-2-1	109
5.4.2.2.1 Data Preprocessing.....	109

	Page
5.4.2.2.2 Model Training	110
5.4.2.2.3 Model Testing Results and Discussion	115
5.4.2.3 Case 2-2-2	117
5.4.2.3.1 Data Preprocessing	117
5.4.2.3.2 Model Training	118
5.4.2.3.3 Model Testing Results and Discussion	123
5.4.2.4 Case 2-3	125
5.4.2.4.1 Data Preprocessing	125
5.4.2.4.2 Model Training	126
5.4.2.4.3 Model Testing Results and Discussion	131
5.4.2.5 Performance of ANN Prediction	133
5.4.3 Location for Infill Well Number 101	140
VI CONCLUSIONS AND RECOMMENDATIONS	141
6.1 Conclusions	141
6.2 Recommendations	142
References	143
Appendices	145
Appendix A	146
Appendix B	147
Appendix C	151
Vitae	152

LIST OF TABLES

	Page
Table 4.1: Range of porosity and calculated permeability	21
Table 4.2: PVT properties at surface conditions (temperature = 60°F, pressure = 14.7psia)	24
Table 4.3: PVT properties at reservoir conditions (temperature = 250°F)	24
Table 4.4: Downhole equipment.....	26
Table 4.5: Vertical flow performance by “PROSPER” program	26
Table 5.1: Summary of parameters that should be the input for ANN	32
Table 5.2: Summary of statistics of all data sets (Case 1-1).....	42
Table 5.3: Model configuration for Case 1-1.....	46
Table 5.4: Summary of statistics of all data sets (Case 1-2-1).....	56
Table 5.5: Model configuration for Case 1-2-1	57
Table 5.6: Summary of statistics of all data sets (Case 1-2-2).....	67
Table 5.7: Model configuration for Case 1-2-2	68
Table 5.8: Summary of statistics of all data sets (Case 1-3).....	83
Table 5.9: Model configuration for Case 1-3.....	84
Table 5.10: Summary of best performance model for each case.....	91
Table 5.11: Order of candidate location from a higher to lower initial gas rate.....	98
Table 5.12: Summary of error in initial gas rate predicted for location 1.	99
Table 5.13: Model configuration for Case 2-1.....	102
Table 5.14: Model configuration for Case 2-2-1	110
Table 5.15: Model configuration for Case 2-2-2	118
Table 5.16: Model configuration for Case 2-3.....	126
Table 5.17: Summarized of best performance model for each case.	133
Table 5.18: Order of candidate location from a higher to lower 1 year cumulative gas production	138
Table 5.19: Summary of error in 1 year cumulative gas production predicted for location 9.....	139
Table 5.20: Summary of best candidate well location for each case.....	140

LIST OF FIGURES

	Page
Figure 3.1: Schematic of biological neuron.....	10
Figure 3.2: Schematic of multiple neurons in one layer	11
Figure 3.3: Linear transfer function	12
Figure 3.4: Log-sigmoid transfer function.....	13
Figure 3.5: Tan-sigmoid transfer function	13
Figure 3.6: Schematic of multilayer multiple neurons.....	14
Figure 3.7: Error surface (MSE in function of w_1 and w_2)	17
Figure 3.8: Schematic of learning rate effect on error surface	19
Figure 3.9: Schematic of local and global minimum.....	19
Figure 4.1: Side view of reservoir model.....	21
Figure 4.2: Top view of reservoir model	22
Figure 4.3: 3D view of reservoir model.....	22
Figure 4.4: Porosity distribution map	23
Figure 4.5: Drilled location on porosity distribution map	25
Figure 5.1: Schematic of porosity data mask.....	31
Figure 5.2: An example of target function fitted to the data.....	33
Figure 5.3: An example of overfitting behavior	33
Figure 5.4: Learning process of ANN.....	34
Figure 5.5: Locations of well number 1 and surrounding wells on porosity map	36
Figure 5.6: Locations of well number 22 and surrounding wells on porosity map	36
Figure 5.7: Schematic diagram of ANN for Case 1-1	38
Figure 5.8: Histogram of permeability for training sets	39
Figure 5.9: Histogram of permeability for validating sets.....	40
Figure 5.10: Histogram of permeability for of testing sets	40
Figure 5.11: Histogram of pressure for training sets	41
Figure 5.12: Histogram of pressure for validating sets.....	41
Figure 5.13: Histogram of pressure for testing sets	42
Figure 5.14: An example of MATLAB source code of ANN model	43

	Page
Figure 5.15: Training result window from ANN	45
Figure 5.16: Performance curve of model 13 (Case 1-1).....	47
Figure 5.17: Performance curve of model 16 (Case 1-1).....	47
Figure 5.18: Cross plot of predicted vs actual initial gas rates of training sets of model 13 (Case 1-1)	49
Figure 5.19: Cross plot of predicted vs actual initial gas rates of validating sets of model 13 (Case 1-1)	49
Figure 5.20: Cross plot of predicted vs actual initial gas rates of training sets of model 16 (Case 1-1)	50
Figure 5.21: Cross plot of predicted vs actual initial gas rates of validating sets of model 16 (Case 1-1)	50
Figure 5.22: Cross plot of predicted vs actual initial gas rates of testing sets of model 13 (Case 1-1)	51
Figure 5.23: Cross plot of predicted vs actual initial gas rates of testing sets of model 16 (Case 1-1)	52
Figure 5.24: Schematic diagram of ANN for Case 1-2-1	53
Figure 5.25: Histogram of arithmetic average pressure for training sets.....	54
Figure 5.26: Histogram of arithmetic average pressure for validating sets	55
Figure 5.27: Histogram of arithmetic average pressure for testing sets	55
Figure 5.28: Performance curve of model 1 (Case 1-2-1)	58
Figure 5.29: Performance curve of model 4 (Case 1-2-1)	59
Figure 5.30: Cross plot of predicted vs actual initial gas rate of training sets of model 1 (Case 1-2-1)	60
Figure 5.31: Cross plot of predicted vs actual initial gas rate of validating sets of model 1 (Case 1-2-1)	60
Figure 5.32: Cross plot of predicted vs actual initial gas rate of training sets of model 4 (Case 1-2-1)	61
Figure 5.33: Cross plot of predicted vs actual initial gas rate of validating sets of model 4 (Case 1-2-1)	61
Figure 5.34: Cross plot of predicted vs actual initial gas rate of testing sets of model 1 (Case 1-2-1)	62

Figure 5.35: Cross plot of predicted vs actual initial gas rate of testing sets of model 4 (Case 1-2-1)	63
Figure 5.36: Schematic diagram of ANN for Case 1-2-2	64
Figure 5.37: Histogram of inverse-distance average pressure for training sets	65
Figure 5.38: Histogram of inverse-distance average pressure for validating sets	66
Figure 5.39: Histogram of inverse-distance average pressure for testing sets	66
Figure 5.40: Performance curve of model 13 (Case 1-2-2)	69
Figure 5.41: Performance curve of model 14 (Case 1-2-2)	70
Figure 5.42: Cross plot of predicted vs actual initial gas rates of training sets of model 13 (Case 1-2-2)	71
Figure 5.43: Cross plot of predicted vs actual initial gas rates of validating sets of model 13 (Case 1-2-2)	71
Figure 5.44: Cross plot of predicted vs actual initial gas rates of training sets of model 14 (Case 1-2-2)	72
Figure 5.45: Cross plot of predicted vs actual initial gas rates of validating sets of model 14 (Case 1-2-2)	72
Figure 5.46: Cross plot of predicted vs actual initial gas rates of testing sets of model 13 (Case 1-2-2)	73
Figure 5.47: Cross plot of predicted vs actual initial gas rates of testing sets of model 14 (Case 1-2-2)	74
Figure 5.48: Schematic diagram of ANN for Case 1-3	75
Figure 5.49: Histogram of 1 st ring porosity for training sets	76
Figure 5.50: Histogram of 1 st ring porosity for validating sets	76
Figure 5.51: Histogram of 1 st ring porosity for testing sets	77
Figure 5.52: Histogram of 2 nd ring porosity for training sets	77
Figure 5.53: Histogram of 2 nd ring porosity for validating sets	78
Figure 5.54: Histogram of 2 nd ring porosity for testing sets	78
Figure 5.55: Histogram of 3 rd ring porosity for training sets	79
Figure 5.56: Histogram of 3 rd ring porosity for validating sets	79
Figure 5.57: Histogram of 3 rd ring porosity for testing sets	80
Figure 5.58: Histogram of drill date for training sets	80

	Page
Figure 5.59: Histogram of drill date for validating sets.....	81
Figure 5.60: Histogram of drill date for testing sets	81
Figure 5.61: Histogram of number of surrounding wells for training sets	82
Figure 5.62: Histogram of number of surrounding wells of validating sets	82
Figure 5.63: Histogram of number of surrounding wells for testing sets	83
Figure 5.64: Performance curve of model 3 (Case 1-3).....	85
Figure 5.65: Performance curve of model 4 (Case 1-3).....	86
Figure 5.66: Cross plot of predicted vs actual initial gas rates of training sets of model 3 (Case 1-3)	87
Figure 5.67: Cross plot of predicted vs actual initial gas rates of validating sets of model 3 (Case 1-3)	87
Figure 5.68: Cross plot of predicted vs actual initial gas rates of training sets of model 4 (Case 1-3)	88
Figure 5.69: Cross plot of predicted vs actual initial gas rates of validating sets of model 4 (Case 1-3)	88
Figure 5.70: Cross plot of predicted vs actual initial gas rates of testing sets of model 3 (Case 1-3)	89
Figure 5.71: Cross plot of predicted vs actual initial gas rates of testing sets of model 4 (Case 1-3)	90
Figure 5.72: Cross plot of predicted vs actual initial gas rate for candidate well locations (Case 1-1)	92
Figure 5.73: Cross plot of predicted vs actual initial gas rate for candidate well locations (Case 1-2-1).....	92
Figure 5.74: Cross plot of predicted vs actual initial gas rate for candidate well locations (Case 1-2-2).....	93
Figure 5.75: Cross plot of predicted vs actual initial gas rate for candidate well locations (Case 1-3)	93
Figure 5.76: Cross plot of predicted vs actual initial gas rate for candidate well locations (positive value of Case 1-2-1)	95
Figure 5.77: Cross plot of predicted vs actual initial gas rate for candidate well locations (positive value of Case 1-2-2)	95

	Page
Figure 5.78: Cross plot of predicted vs actual initial gas rate for candidate well locations (Case 1-2-1) for different drill date.....	96
Figure 5.79: Candidate well locations on porosity distribution map.....	97
Figure 5.80: Schematic diagram of ANN for Case 2-1	101
Figure 5.81: Performance curve of model 1 (Case 2-1).....	103
Figure 5.82: Performance curve of model 2 (Case 2-1).....	104
Figure 5.83: Cross plot of predicted vs actual 1 year cumulative gas of training sets of model 1 (Case 2-1)	105
Figure 5.84: Cross plot of predicted vs actual 1 year cumulative gas of validating sets of model 1 (Case 2-1)	105
Figure 5.85: Cross plot of predicted vs actual 1 year cumulative gas of training sets of model 2 (Case 2-1)	106
Figure 5.86: Cross plot of predicted vs actual 1 year cumulative gas of validating sets of model 2 (Case 2-1)	106
Figure 5.87: Cross plot of predicted vs actual 1 year cumulative gas of testing sets of model 1 (Case 2-1)	107
Figure 5.88: Cross plot of predicted vs actual 1 year cumulative gas of testing sets of model 2 (Case 2-1)	108
Figure 5.89: Schematic diagram of ANN for Case 2-2-1	109
Figure 5.90: Performance curve of model 1 (Case 2-2-1)	111
Figure 5.91: Performance curve of model 14 (Case 2-2-1)	112
Figure 5.92: Cross plot of predicted vs actual 1 year cumulative gas of training sets of model 1 (Case 2-2-1).....	113
Figure 5.93: Cross plot of predicted vs actual 1 year cumulative gas of validating sets of model 1 (Case 2-2-1).....	113
Figure 5.94: Cross plot of predicted vs actual 1 year cumulative gas of training sets of model 14 (Case 2-2-1).....	114
Figure 5.95: Cross plot of predicted vs actual 1 year cumulative gas of validating sets of model 14 (Case 2-2-1).....	114
Figure 5.96: Cross plot of predicted vs actual 1 year cumulative gas of testing sets of model 1 (Case 2-2-1).....	115

	Page
Figure 5.97: Cross plot of predicted vs actual 1 year cumulative gas of testing sets of model 14 (Case 2-2-1).....	116
Figure 5.98: Schematic diagram of ANN Case 2-2-2.....	117
Figure 5.99: Performance curve of model 4 (Case 2-2-2).....	119
Figure 5.100: Performance curve of model 13 (Case 2-2-2).....	120
Figure 5.101: Cross plot of predicted vs actual 1 year cumulative gas of training sets of model 4 (Case 2-2-2).....	121
Figure 5.102: Cross plot of predicted vs actual 1 year cumulative gas of validating sets of model 4 (Case 2-2-2).....	121
Figure 5.103: Cross plot of predicted vs actual 1 year cumulative gas of training sets of model 13 (Case 2-2-2).....	122
Figure 5.104: Cross plot of predicted vs actual 1 year cumulative gas of validating sets of model 13 (Case 2-2-2).....	122
Figure 5.105: Cross plot of predicted vs actual 1 year cumulative gas of testing sets of model 4 (Case 2-2-2).....	123
Figure 5.106: Cross plot of predicted vs actual 1 year cumulative gas of testing sets of model 13 (Case 2-2-2).....	124
Figure 5.107: Schematic diagram of ANN for Case 2-3.....	125
Figure 5.108: Performance curve of model 2 (Case 2-3).....	127
Figure 5.109: Performance curve of model 15 (Case 2-3).....	128
Figure 5.110: Cross plot of predicted vs actual 1 year cumulative gas of training sets of model 2 (Case 2-3).....	129
Figure 5.111: Cross plot of predicted vs actual 1 year cumulative gas of validating sets of model 2 (Case 2-3).....	129
Figure 5.112: Cross plot of predicted vs actual 1 year cumulative gas of training sets of model 15 (Case 2-3).....	130
Figure 5.113: Cross plot of predicted vs actual 1 year cumulative gas of validating sets of model 15 (Case 2-3).....	130
Figure 5.114: Cross plot of predicted vs actual 1 year cumulative gas of testing sets of model 2 (Case 2-3).....	131

Figure 5.115: Cross plot of predicted vs actual 1 year cumulative gas of testing sets of model 15 (Case 2-3)	132
Figure 5.116: Cross plot of predicted vs actual cumulative gas production for candidate well location (Case 2-1)	134
Figure 5.117: Cross plot of predicted vs actual cumulative gas production for candidate well location (Case 2-2-1)	134
Figure 5.118: Cross plot of predicted vs actual cumulative gas production for candidate well location (Case 2-2-2)	135
Figure 5.119: Cross plot of predicted vs actual cumulative gas production for candidate well location (Case 2-3)	135
Figure 5.120: Cross plot of predicted vs actual initial gas rate for candidate well locations (positive value of Case 2-2-1)	137
Figure 5.121: Cross plot of predicted vs actual initial gas rate for candidate well locations (positive value of Case 2-2-2)	137

LIST OF ABBREVIATIONS

ANN	artificial neural network
BPN	feed forward back propagation network
MSE	mean square error
scf	standard cubic feet
STB	stock tank barrel



ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

NOMENCLATURE

A	area of reservoir
a	predicted output of ANN
B_g	gas formation volume factor
b	bias of ANN
f	transfer function
G	initial gas in place
G_p	cumulative gas production
h	reservoir thickness
k	permeability
lr	learning rate
mc	momentum
n	output of summing function (input of transfer function)
p	input of ANN
P_R	average reservoir pressure
P_w	well flowing pressure
P_i	initial reservoir pressure
P	reservoir pressure at certain time
Q	gas production rate
Q	number of layers
R	number of inputs
r_e	radius to external boundary
r_w	wellbore radius
S	number of neurons in layer
S_w	water saturation
t	target output
w	weight of ANN
Z_i	initial gas compressibility factor
Z	gas compressibility at certain time

GREEK LETTERS

Φ	porosity
μ	viscosity
Σ	summing function



ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

CHAPTER I

INTRODUCTION

With an increasing demand for energy and rising gas prices around the world, petroleum producing companies are trying to improve the recovery factor of gas production. There are many methods that have been used for this purpose, with varying levels of difficulty. The easiest method that has been used in recent years is the drilling of infill wells. It is one effective method that is widely used to increase gas production in an existing large oil field. The advantage of this method is that drilling more wells results in increase in connecting paths from the reservoir to the surface. This means gas can be transported out to the surface easier, and we are able to speed up the production of gas. Moreover, it helps reduce a pressure drop from friction loss in the pore space of rock when gas travels from the location that is far away from a well. And in the case that there are faults obstructing the flow path of gas to the existing wells, drilling more wells will help increase the amount of gas produced. However, even though the drilling of infill wells is easy to perform, but the cost of drilling a well is very expensive. Thus, before making a decision to drill, we have to ensure that the location will give enough gas production to justify the cost.

Consequently, a key component in making the final decision of whether or not to drill the well is the expected production profiles including the gas production rate and cumulative gas production. The success of adding a new well to the field depends on the prediction accuracy of these parameters. The more accurate the prediction is, the better the decision on drilling location will be. However, a large variability in rock properties, well spacing, and the large number of wells involved make a prediction to be difficult. From the past until now, many methods have been performed in order to yield the best prediction. In the past, analysis based on conventional decline curve and pressure transient analysis which is only valid for production data from homogeneous reservoirs often produce inaccurate results which lead to wrong conclusions. Presently, with the growing availability of new technologies, conducting the reservoir simulation through computer software is the most accurate way to determine infill-drilling potential. However, complete reservoir evaluation involves geological, geophysical, reservoir analyses, and a large set of input parameters needed in the software. This

includes the developing of a geological model of the studied area, estimating distributions of reservoir properties such as porosity and permeability, constructing and calibrating a reservoir simulation model, and then using the reservoir model to predict future production at potential infill well locations. While it may be accurate, this technique is time consuming and expensive, especially in a large basin when a large number of wells are involved. Therefore, a new technique that is inexpensive and less time consuming needs to be developed in order to produce accurate predictions.

In recent years, there has been a growing interest in applying Artificial Neural Network (ANN) to various fields such as science, engineering, and finance. The major advantage of the technique is that the ANN has the ability to learn a complex relationship between input and output. The ANN is a mathematical model that has a working algorithm applied from a working process of the biological nervous system in the human brain to form a learning ability. With this ability, the ANN has been widely used to solving various kinds of problems in many applications such as pattern classification, clustering, function approximation, forecasting, optimization, and control.

The objective of this study is to predict gas production for infill wells development project in closed boundary depletion drive gas reservoirs by using the Artificial Neural Network (ANN) technique.

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

1.1 Outline of Methodology

1. Gather and prepare data by referring most properties such as fluid and geological properties from an actual gas field in the Gulf of Thailand to create reasonable reservoir simulation model.
2. Specify the output parameters of ANN based on the purpose of study.
3. Select appropriate input parameters for the ANN using knowledge from past literature and theories.
4. Set up multiple case studies based on combination of input and output parameters.
5. Run reservoir simulation to create production data used to train the network.
6. Generate pair datasets between input of ANN and target output.
7. Partition all datasets into 3 main sets that are 1) training sets, 2) validating sets, and 3) testing sets.
8. Set up various kinds of network configuration by varying the number of hidden node, number of hidden layer, learning rate, and momentum.
9. Train all network configurations for each case study using Artificial Neural Network Toolbox of “MATLAB” software.
10. Choose 2 best performance models which represent the lowest MSE of validating sets.
11. Test the accuracy of trained ANN by using testing sets, compare the accuracy, and then choose only one best performance model to predict the performance of the infill well.
12. Analyze the accuracy for predicted result of the infill well.

1.2 Thesis Outline

This thesis paper consists of six chapters, and the outline of each chapter is listed below:

Chapter II reviews previous work related to infill drilling project and Artificial Neural Network.

Chapter III introduces the theory and concept related to this study.

Chapter IV shows simulation grid model used in this study.

Chapter V describes the development of the Artificial Neural Network model to predict gas production in terms of initial gas rate and cumulative gas production.

Chapter VI provides conclusions of the study and recommendations for further study.



ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

CHAPTER II

LITERATURE REVIEW

This chapter describes past studies that are related to various methods to predict performance of infill wells and Artificial Neural Network (ANN) studies.

2.1 Prediction of Infill Drilling Performance

Coats^[1] proposed a two dimensional numerical calculation of semi-steady-state flow to determine an optimum drilling schedule for remaining field development, starting from an initial time corresponding to an arbitrary degree of depletion and arbitrary number and locations of existing wells. This optimum schedule consists of specified well locations to drill and the time at which each is to be drilled. The additional well requirements for maintenance of field productivity are determined from back-pressure curves that relate well deliverability to the difference between average field pressure and flowing wellbore pressures.

McCain^[2] used a statistical technique, “Moving Window” method to determine infill potential in a complex, low-permeability gas reservoir (Cotton Valley). “Moving Window” is the technique that evaluates the performance of each well with surrounding wells within the same window, compares new wells to old wells for signs of depletion, calculates effective well density, and once linked to a scattering of conventional estimates of drainage area, provides estimates of undrained acreage and infill reserves. In his study, three productivity indicators selected were 1) maximum monthly production rate, 2) average monthly production rate for the most productive twelve month period, and 3) monthly production rate at the time when a cumulative production of 250 MMscf was reached. The best year of production is simply the best 12 consecutive months of production divided by 12. The “Best Year” indicator can be used to provide a rough estimate of the gas recovery per well by making a plot of the best year production indicator versus the 10 year cumulative. This is a very good correlation between best year indicator and long term performance.

Voneiff and Cipolla^[3] further developed the “Moving Window” technique and applied it for rapid assessment of infill and recompletion potential in the Ozona field.

They chose “Best Year” and “Decline Ratio” (maximum month of production)/(Best Year) as indicators. The production indicator plot can be used to rapidly identify areas of depletion before drainage areas had been calculated through conventional reservoir engineering techniques.

Guan^[4] assessed the accuracy of the moving window technique for selecting infill candidate wells in low permeability gas reservoirs by analyzing synthetically generated production data. He extended the method described by Voneiff and Cipolla by using the model based on a combination of the material balance equation and pseudo steady state flow equation, simplified by assuming that many properties are constant within an individual window. He showed that the moving window technique can predict average infill performance of a group of candidate wells reasonably well often to within 10%. Thus, it can serve as a useful screening tool. However, predictions for individual wells can be off by more than 50%. He summarized the parameters that affect accuracy as permeability, heterogeneity, well spacing, and number of wells.

Gao^[5] proposed “Rapid Inversion Method”, a new simulation-based inversion approach for rapid assessment of infill well potential. Instead of focusing on small-scale, high-resolution problem, he focused on large-scale, coarse-resolution studies consisting of hundreds or, potentially, thousand of wells. This method is able to identify potential areas or groups of wells for infill development quickly and inexpensively. The result showed that this method is more accurate than moving window statistic methods in synthetic cases.

Guan et.al.^[6] discussed the two recently developed techniques which are moving window technique and rapid inversion method to determine the infill drilling potential in large tight gas reservoirs. The paper summarizes what petroleum engineers have learnt about the application of those two techniques.

Soto et.al.^[11] used multiple mathematical techniques to develop primary ultimate oil recovery, initial waterflooding ultimate oil recovery, and infill drilling ultimate oil recovery models for carbonate reservoirs in West Texas. These techniques are 1) non-linear regression, 2) non-parametric regression, and 3) neural network model.

2.2 Artificial Neural Network (ANN)

Boomer^[7] used the Artificial Neural Network technique as a new method to predict the oil production rate profile which are more accurate than human prediction in infill drilling field development of the Permian Basin of West Texas and Southeast New Mexico. He used the concept of “Data Mask” to collect the production data of every well within each concentric ring to be the input of the neural network on a per-ring basis instead of per-well basis. The output is the production rate of well that Data Mask is placed over.

AI-Fattah and Startzman^[8] developed a neural network model to forecast U.S. natural gas supply to the year 2020. The Network was developed with an initial large pool of input parameters. After applying reduction techniques, the number of input parameters was decreased. A three-layer neural network was successfully trained with yearly data starting from 1950 to 1989 using the quick-propagation learning algorithm. The target output is the production rate of natural gas. A test set, not used to train the network and containing data from 1990 to 1998, was used to verify and validate the network performance for prediction.

Sampaio et.al.^[9] proposed an alternative to speed up the history matching process by using the application of feed-forward neural networks as nonlinear proxies of reservoir simulation. The focus of their study is to show the steps of choosing the best number of hidden layers, the neurons and the training method.

Doraisamy^[10] proposed a methodology for optimizing field development. The objective of his study is to structure the field development schemes using Artificial Neural Network (ANN) in conjunction with numerical reservoir simulation, a process he call neuro-simulation. In neuro-simulation, a few field development scenarios were examined using a numerical simulator. The results of these studies were then used to train the ANN. Using neuro-simulation, the number of numerical simulations is significantly reduced.

Jalali^[12] proposed uncertainty quantification of a complex coalbed methane production enhancement reservoir model. He proposed a new technique by

developing a Surrogate Reservoir Model that can accurately mimic the behavior of commercial reservoir model by using an Artificial Neural Network.



ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

CHAPTER III

THEORIES AND CONCEPTS

3.1 Artificial Neural Network (ANN)

Artificial Neural Network (ANN) is a mathematical model that has an ability to learn complex relationship between a pair of input and output samples namely, "Training Dataset". The idea of Artificial Neural Network is motivated from the biological nerve cell called "Neuron". Billions of neurons are connected to each other to form a biological nervous system like in the human brain. The human brain has an ability to memorize a mistake in the past and learning to improve itself to avoid making the same mistake in the future. ANN is completely different when compared with a conventional computer that operates by following the order which has been programmed in by the human user. Therefore, the same mistake can still occur as long as the error of program is not fixed yet. In contrast, a working algorithm of ANN was applied from a working process of biological nervous system to form a learning system. With this ability, the ANN has been widely used to solve various kinds of problems in many applications such as pattern classification, clustering, function approximation, forecasting, optimization, and control.

The following section describes the fundamentals of ANN. The first part describes the main units and a working process of biological neuron. The second part describes the main units and working process of artificial neuron that is applied from the biological neuron. The third part describes the structure of the network, showing the algorithm of data transfer from input through hidden layer and output, and the fourth part describes the learning process to determine appropriate weights.

3.1.1 Biological Neuron

The biological neuron is a unit of the nervous system in the human brain which has a simple or uncomplicated working process. A task of these neurons is to receive signal from a previous neuron, change amplitude, and send it to the next neuron that connects to each other. Figure 3.1 shows a schematic of a biological neuron. Each neuron consists of four main units that are the dendrite, cell body, axon, and synapse. The dendrite is a unit that receives signal from other neuron to cell body. The cell body is a unit that sums a total incoming signal from dendrite, changing the amplitude of signal, and fire an electrochemical to the axon. The axon is a unit that receives signal from cell body and passes it to the synapse. The synapse is a gap between each neuron. A signal will pass from the axon to a dendrite of the neighboring neuron through this gap. Many millions of biological neurons are connected to each other to compose a biological neural network like in the human brain. A learning process can be performed by updating the synapse's strength which connects the neurons.

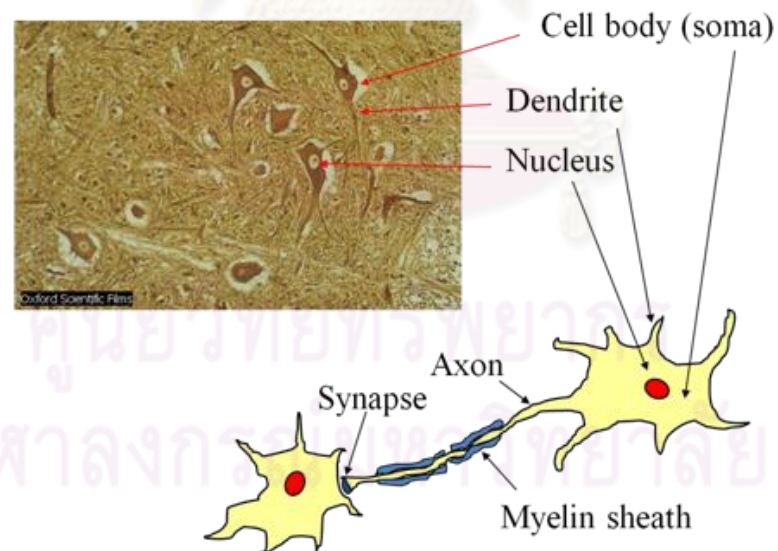


Figure 3.1: Schematic of biological neuron^[14]

3.1.2 Artificial Neuron

The artificial neuron is a simple processing unit that applies a working algorithm from the biological neuron. The schematic of multiple neurons in one layer is shown in Figure 3.2.

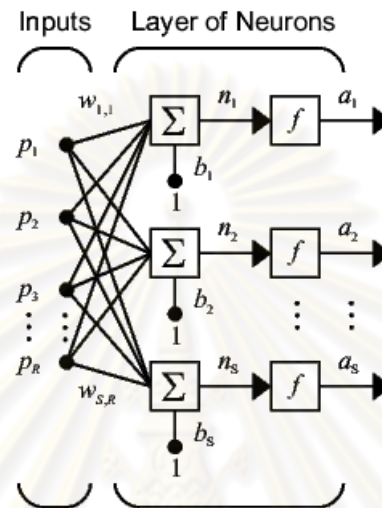


Figure 3.2: Schematic of multiple neurons in one layer^[13]

The equation used to express the output of each neuron is

$$a_i = f(n_i) = f[\sum(w_{i,j}p_j)] + b_i \quad (3.1)$$

where	p_j	=	input number j
	f	=	transfer function
	\sum	=	summing function
	n_i	=	output of summing function (input of transfer function) from neuron number i
	$w_{i,j}$	=	weight connected between input number j and neuron number i
	a_i	=	predicted output of neuron number i
	b_i	=	bias of neuron number i
	R	=	number of inputs
	S	=	number of neurons in layer
	i	=	[1, 2, 3, ..., S]
	j	=	[1, 2, 3, ..., R]

Each neuron consists of six main units as shown in Figure 3.2 from the left to the right, which are input, weight, bias, summing function, transfer function, and the output. Input value (p_j) can be transferred to a neuron from 2 sources that are 1) inputted by a user in case that the neuron is located in the first layer and 2) from previous neuron in case that the neuron is located at layer greater than or equal to 2. The number of input can be any values based on the number of available data. After that, input values are multiplied by weights ($w_{i,j}$) which are values that connect between neurons in different layers. The weights are used for adjusting the amplitude of input values. Using the same concept as the biological neurons, the network can learn by adjusting these weights that behave like synapse's strengths of biological neurons. Then, the product of input multiplied with weight ($w_{i,j} p_j$) from all nodes in the previous layer is moved to a summing function. At the same time, a bias (b_i) which behaves like a weight except that the input that is multiplied with bias will be only "1" all the time is sent to summing function to sum together with product $w_{i,j} p_j$. Each neuron has only one bias, the task of a bias is that it helps to speed up a training process and enables a network to get more accurate output. Next, a summed value (n_i) is passed through a transfer function to change the amplitude. There are many transfer functions studied by many researchers:

1) Linear Transfer Function - The value of output from the function is the same value as the input parameters as shown in Figure 3.3.

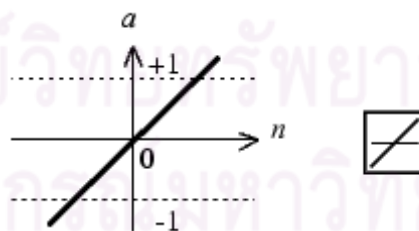


Figure 3.3: Linear transfer function ^[13]

The equation for linear transfer function is written as

$$a = n \quad (3.2)$$

where a = output of transfer function
 n = input of transfer function

2) **Log-sigmoid Transfer Function** - The value of output will be in the range of 0 to 1. Figure 3.4 depicts log-sigmoid transfer function.

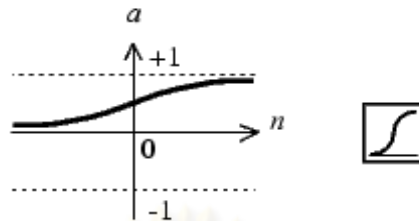


Figure 3.4: Log-sigmoid transfer function ^[13]

The log-sigmoid transfer function is expressed as:

$$a = \frac{1}{1+e^{-n}} \quad (3.3)$$

3) **Tan-sigmoid transfer function** - The value of output will be in the range of -1 to 1. Figure 3.5 depicts tan-sigmoid transfer function.

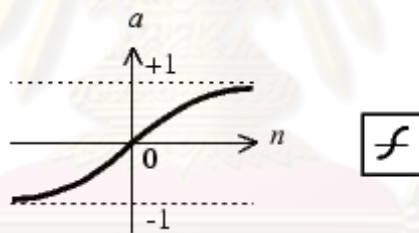


Figure 3.5: Tan-sigmoid transfer function ^[13]

The tan-sigmoid transfer function is expressed as:

$$a = \frac{e^n - e^{-n}}{e^n + e^{-n}} \quad (3.4)$$

The transfer function will generate the output (a_i) and this output will be passed to the neighboring neuron in the next layer as input.

3.1.3 Network Structure

Using the same concept as biological neural network, many artificial neurons connected to each other compose an Artificial Neural Network. The network with multilayer multiple neurons can be created by assembling several networks of one

layer multiple neurons together. The schematics of the network with multilayer multiple neurons and associated mathematical equations are shown in Figure 3.6 and Equation 3.5.

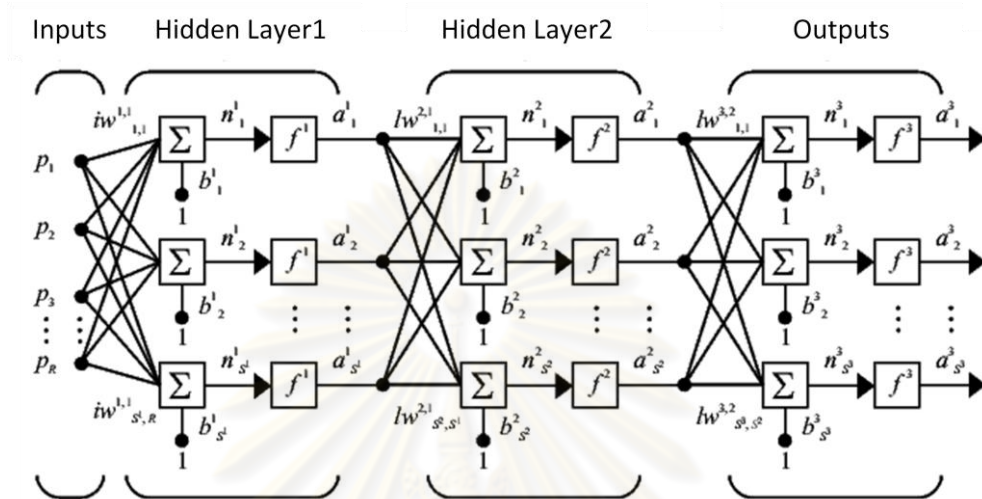


Figure 3.6: Schematic of multilayer multiple neurons ^[13]

$$a_i^1 = f^1(n_i^1) = f^1[\sum(iw_{i,j}^{1,1} p_j) + b_i^1]$$

$$a_i^2 = f^2(n_i^2) = f^2[\sum(lw_{i,j}^{2,1} a_j^1) + b_i^2]$$

$$a_i^3 = f^3(n_i^3) = f^3[\sum(lw_{i,j}^{3,2} a_j^2) + b_i^3]$$

⋮

$$a_i^k = f^k(n_i^k) = f^k[\sum(lw_{i,j}^{k,k-1} a_j^{k-1}) + b_i^k] \quad (3.5)$$

where p_j = input number j
 f_k = transfer function at layer number k
 \sum = summing function
 n_i^k = output of summing function (input of transfer function) from neuron number i of layer number k
 $iw^{1,1}$ = weight connect between input and 1st hidden layer (iw stand for input weight), superscript means that connection is within hidden layer number 1

$lw_{ij}^{k,k-1}$	=	weight connected between output of neuron number j of layer $k-1$ and neuron number i of layer k
a_i^k	=	predicted output of neuron number i of layer k
b_i^k	=	bias of neuron number i of layer number k
R	=	number of inputs
S	=	number of neurons in layer (can be different for each layer)
Q	=	number of layers (all hidden layers and output)
i	=	[1, 2, 3, ..., S]
j	=	[1, 2, 3, ..., R]
k	=	[1, 2, 3, ..., Q]

The main structure of the network consists of at least three main layers as shown in the above schematic from the left to the right that are 1) one input layer, 2) a certain number of hidden layers, and 3) one output layer. The training data set will be input through the input layer. The number of neurons in this layer is not limited, as it can be any value depending on the number of input parameters which have a strong relationship with the output. Too small a number of input neurons can cause the accuracy of prediction to be low, whereas too large a number of input neurons will cause the network to require more number of training data sets. After that, a data set is passed to the hidden layer. Because this layer is located within a network, we cannot see any output value from this layer. So it is called the "hidden layer". The number of neurons in each layer and the number of layers in a network is not limited. A large number of neurons and layers would be useful to solve a difficult problem with a complex relationship between input and output. However, the larger the number of neurons and layers, the larger the number of weights and biases there will be, resulting in a time consuming calculation process. The last layer is the output layer. The prediction result will be shown at this stage. A number of output neurons can be any value depending on the number of the answers we need.

There are many types of networks studied by many researchers, "Feed forward back propagation network (BPN)", known as the most famous and effective network

will be used in this study. The name "feed forward network" is used to explain the behavior of network in which the data will be transferred in only forward direction to the next layer, no data transfer across a node in a same layer and transfer backward to a node in a previous layer. The name "back propagation" is used to explain the behavior of network to adjust weights. After the network compares predicted output with desired output, the error will be calculated, and will be back propagated to adjust the weights step by step from the last layer until the first layer. The adjustment is made until the error meets a target value set up by the user. For BPN, the transfer function of all hidden layers are log-sigmoid transfer function which limits the output in the range of 0 and 1, and the transfer function of the output layer is set to be linear transfer function which has no limit of output value.

3.1.4 Learning Algorithm

After the input is passed onto the network, then the predicted output will be calculated through a network. The target output must be prepared and shown to the network. Then, the error between predicted and target outputs will be calculated. The BPN uses "Mean Square Error (MSE)" to be the error criteria as shown in equation 3.6 as follow:

$$MSE = \frac{1}{N} \sum_{i=1}^N (t_i - a_i)^2 \quad (3.6)$$

where

MSE	=	Mean Square Error
a_i	=	predicted output
t_i	=	target output
N	=	number of training set

The error (MSE) will be used for updating of all connection weights in order to minimize the error. The methodology to update the weights is the "Gradient Descent Method". This method changes the weights in the direction that descent the error surface. An example of error surface is shown in Figure 3.7. In this case, there are only 2 weights. The weights must be updated until the error has reached the lowest point in the error surface which has MSE equal to 0. The network will decide to increase or decrease a weight by calculating the surface slope at that point. If the slope

has a minus sign when compared with the direction of increasing weight, then the network will increase that weight. On the otherhand, if the sign of the slope is positive, then the network will decrease that weight. With this method, the error is moving towards the minimum point step by step everytime that weight is updated. This behavior can be described in easier way by the following example. Let the error surface represent a mountain and there is a lake located at the bottom of mountain which represents the location where MSE is 0. A blind man stays on the mountain and needs to find a way to the lake. Since he cannot see anything, he knows just only which direction the place that he stays inclines to. Therefore, what he can do to find the lake is climb down following the direction to which that mountain inclines.

So with this method, each weight is updated by using the error between target and predicted outputs to calculate a slope. The equation to calculate the change in weight and bias is written in Equations 3.7 to 3.10.

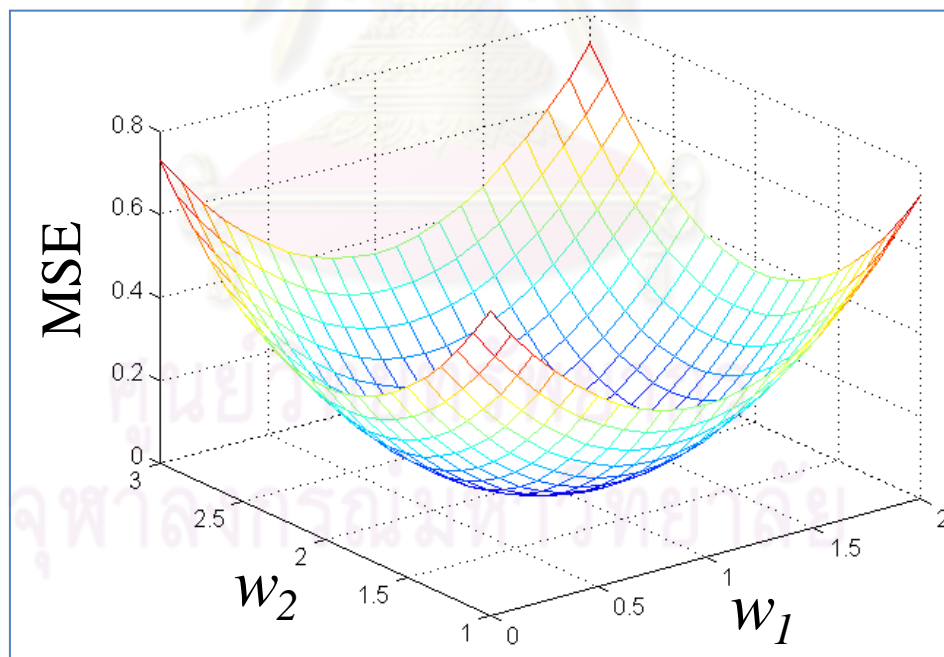


Figure 3.7: Error surface (MSE in function of w_1 and w_2)^[14]

$$\Delta w_i = lr \left(t_i - f(\sum_i^N w_i p_i + b) \right) \left(f'(\sum_i^N w_i p_i + b) \right) p_i + mc(w_{old}) \quad (3.7)$$

$$\Delta b = lr \left(t_i - f(\sum_i^N w_i p_i + b) \right) \left(f'(\sum_i^N w_i p_i + b) \right) + mc(w_{old}) \quad (3.8)$$

$$w_{new} = w_{old} + \Delta w_i \quad (3.9)$$

$$b_{new} = b_{old} + \Delta b \quad (3.10)$$

where lr = learning rate
 mc = momentum

The learning rate determines the acceleration of the weight and bias updating. Normally, this value is between 0 and 1. Figure 3.8 illustrates the effect of learning rate on the error surface (top view) which has the target point where MSE is 0 at the middle. The rightmost figure shows the effect of setting the learning rate too low. The weight will be updated with a small rate, resulting in slow training while too large of it could result in an unstable training as shown in the leftmost of the figure. Sometimes, instability may result in nonconvergence. That means the network will never reach the target.

Figure 3.9 shows the schematic of local and global minimum. The target error where MSE is 0 is located at global minimum. But sometimes the network never knows which one is global minimum and may get stuck in the local minimum area. Momentum is commonly used in weight updating to help prevent the training being stuck in the local minima area. The momentum takes a value between 0 and 1. A high momentum will reduce the risk of getting stuck. However, it may increase the risk of overshooting the solution.

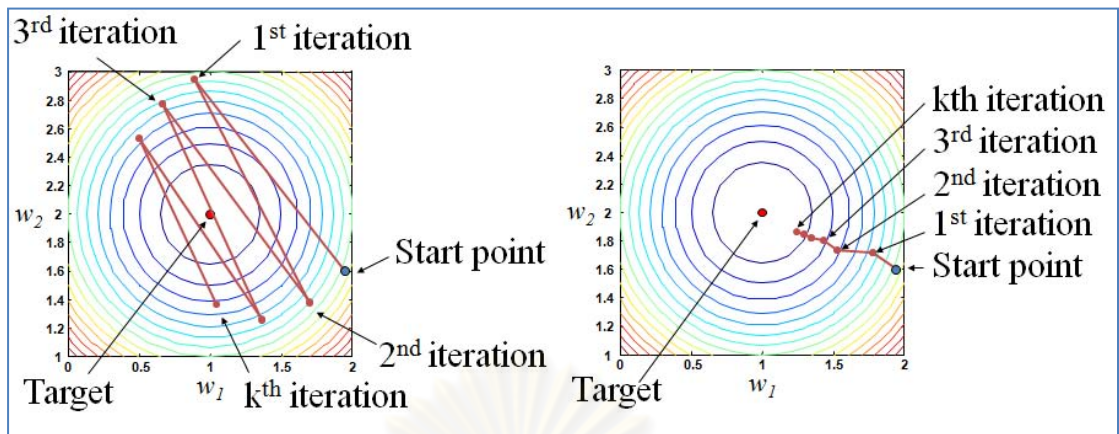


Figure 3.8: Schematic of learning rate effect on error surface ^[14]

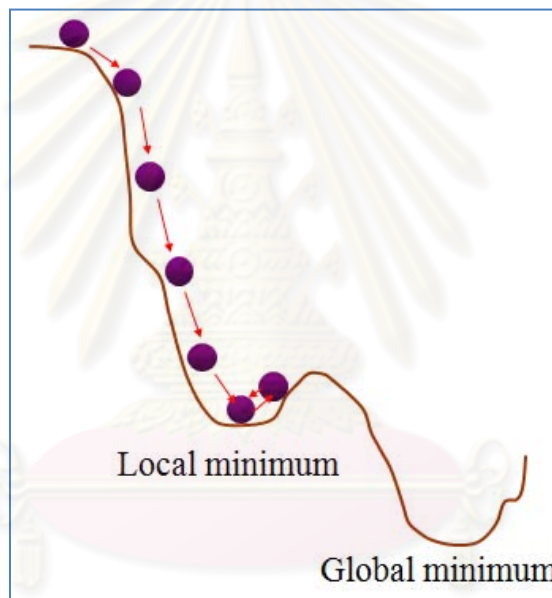


Figure 3.9: Schematic of local and global minimum ^[14]

จุฬาลงกรณ์มหาวิทยาลัย

CHAPTER IV

RESERVOIR SIMULATION MODEL

In order to perform the ANN model development, data sets (pairs of input and target output) to train a network needs to be prepared first. As the purpose of this study is to infill a well, so the data sets must be prepared from a large gas field which already has some number of wells drilled within. These data sets can be taken from the real gas field or taken from the reservoir simulation model. Using data sets taken from the real field is more realistic. However, in many cases there are not enough data sets to train the network. Reservoir simulation is a good alternative to create pairs of input and output for a given oil field given that reservoir and fluid properties are known. This study will use a production profile created by reservoir simulation.

The reservoir simulation is performed by using the reservoir simulator called “ECLIPSE E100” from Schlumberger to simulate a synthetic case of field production. Most of the reservoir properties and fluid properties are obtained from a real gas field in the Gulf of Thailand.

The reservoir model is set up to be a closed boundary depletion drive gas reservoir, consisting of 200 x 200 x 5 grid blocks with grid size 100 ft x 100 ft x 20 ft along the x, y, and z direction, respectively. These dimensions make up a reservoir of the size 20,000 ft in width, 20,000 ft in length, and 100 ft in thickness. The top face of reservoir is located at depth 5,000 ft. Figure 4.1, 4.2, and 4.3 depict the side view, top view, and a 3D view of gas reservoir, respectively.

The permeability for this paper is calculated by using an equation obtained from one formation of a gas field in Gulf of Thailand as shown in Equation 4.1. The range of porosity used in this study and the calculated permeability are shown in Table 4.1.

$$k = 10^{[-2.9971 + 0.2089 * (\% \Phi)]} \quad (4.1)$$

where k = permeability (mD)
 Φ = porosity (%)

Table 4.1: Range of porosity and calculated permeability

Parameter	Value	Unit
Porosity	10 - 30	%
X Perm = Y Perm	0.12 - 1862	mD
Z Perm = 0.1 x X Perm	0.012 - 186.2	mD

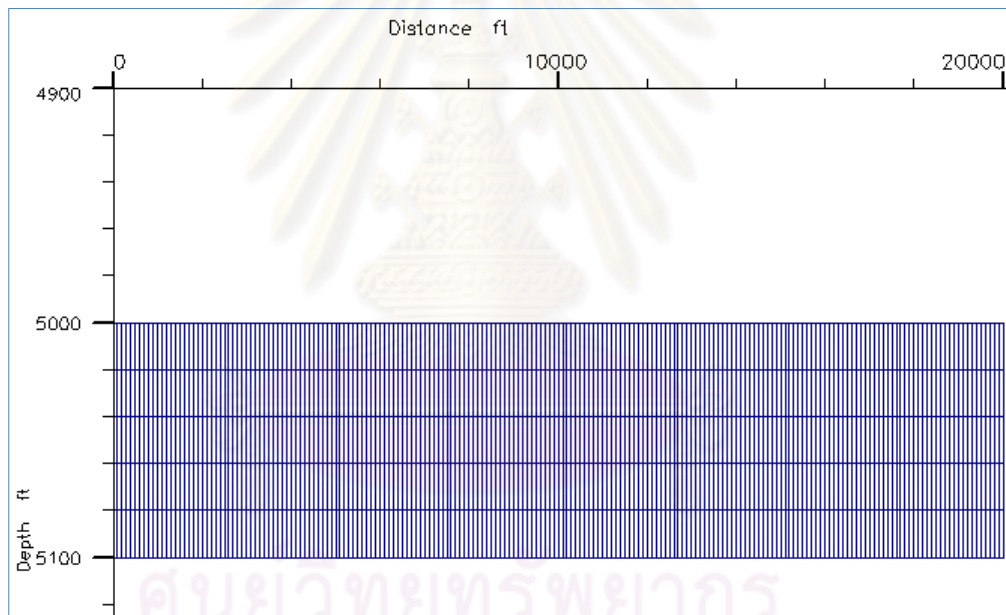


Figure 4.1: Side view of reservoir model

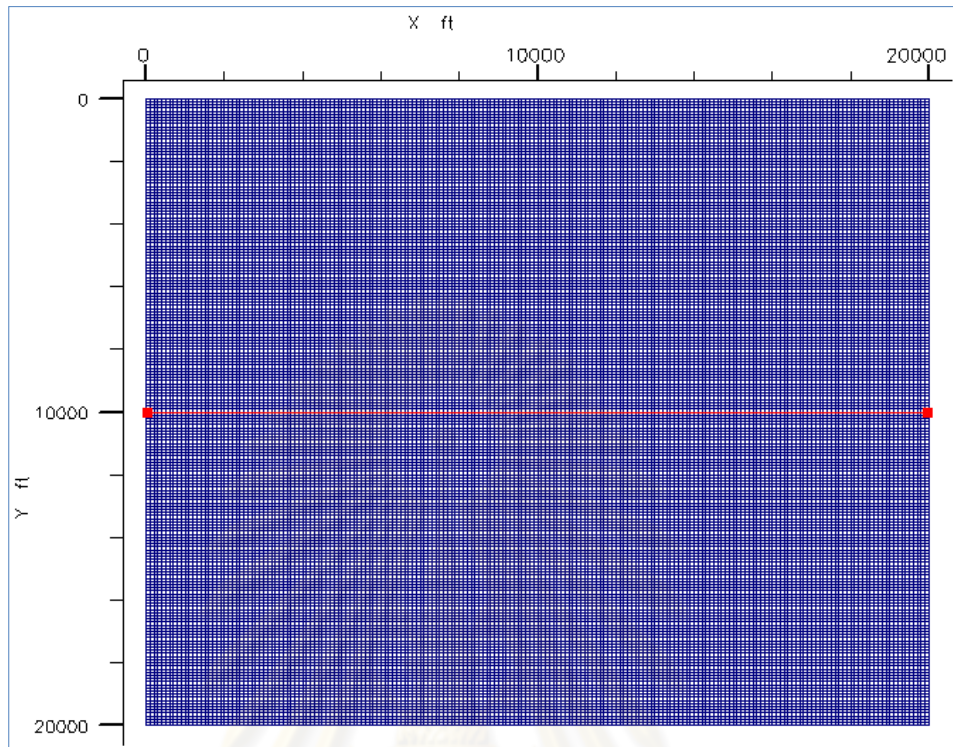


Figure 4.2: Top view of reservoir model

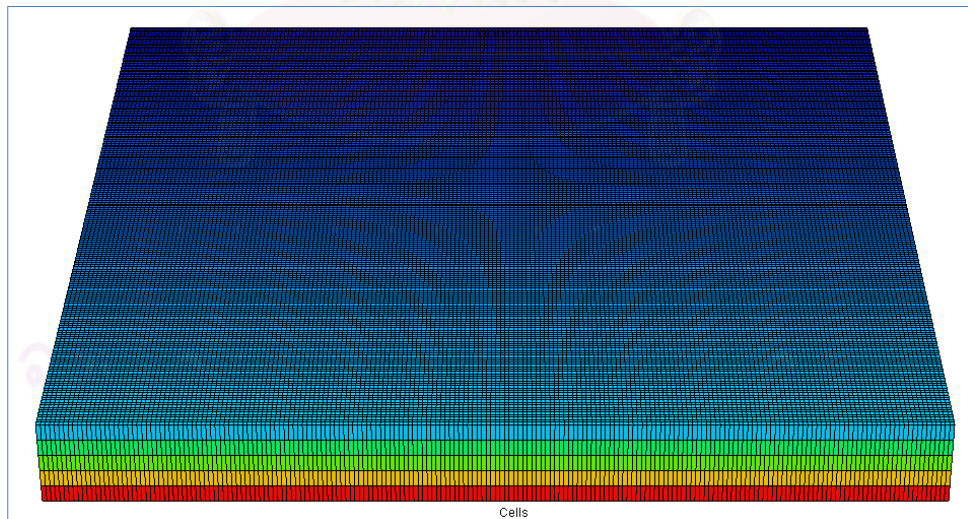


Figure 4.3: 3D view of reservoir model

In order to create reasonable porosity distribution for the entire area, Geostatistics needs to be considered. First, a certain number of grid blocks are selected. In this study, 100 grid blocks from total 40,000 grid blocks spreading around the entire area are selected. Second, the porosity value is randomly assigned to each selected grid block. Third, use a geostatistic simulator called “SGeMS” to make Gaussian simulation and then simulate a porosity of remaining grid blocks. The porosity distribution map created by Geostatistics is shown in Figure 4.4.

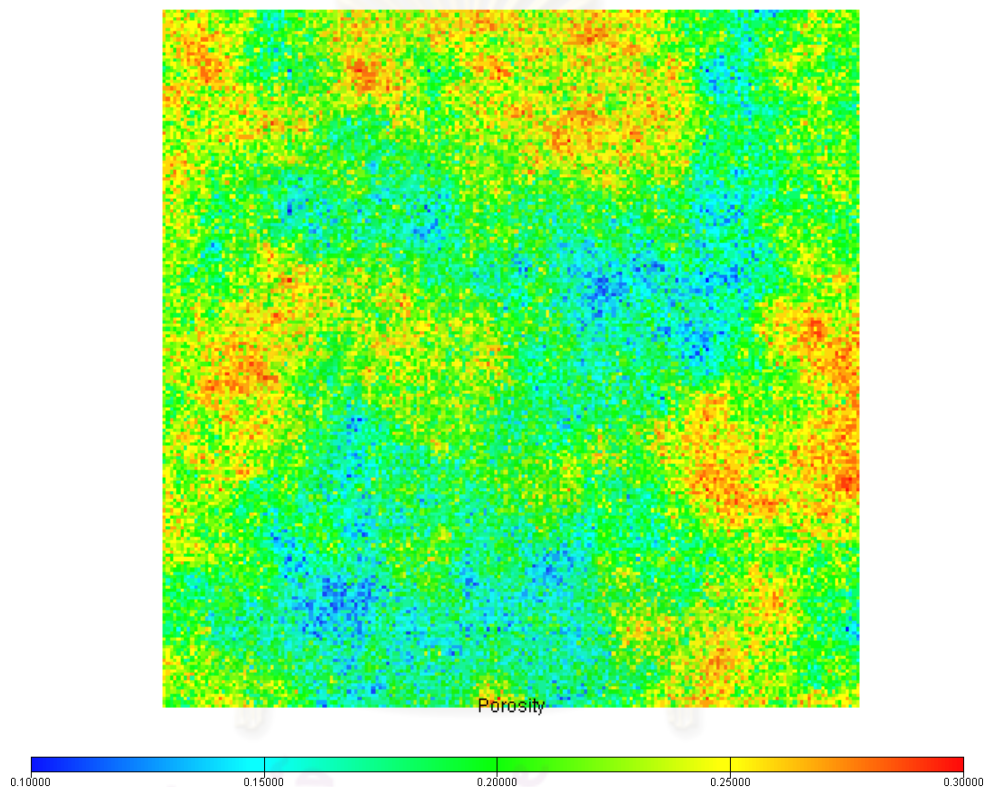


Figure 4.4: Porosity distribution map

The initial conditions of reservoir are calculated by using pressure and temperature gradient equal to 0.433 psi/ft and 3.8°F/100ft respectively which make the pressure and temperature at the top depth of reservoir (5,000ft) to be 2,180 psia and 250°F, respectively.

The PVT properties that are used in this study are generated by using a program called “PROSPER”, assuming that the gas gravity is 0.9 and condensate to gas ratio is 0 STB/MMscf. The PVT at surface and reservoir conditions are calculated and shown in Table 4.2 and 4.3, respectively.

Table 4.2: PVT properties at surface conditions (temperature = 60°F, pressure = 14.7psia)

Temp (°F)	Pressure (psig)	Gas density (lb/ft ³)	Gas viscosity (cp)	Gas FVF (RB/Mscf)	Z factor	Water density (lb/ft ³)	Water viscosity (cp)	Water FVF (RB/STB)	Water compressibility (1/psi)
60	0	0.069125	0.009612	177.159	0.99494	62.5	1.22509	0.99957	4.70E-06

Table 4.3: PVT properties at reservoir conditions (temperature = 250°F)

Temp (°F)	Pressure (psig)	Gas density (lb/ft ³)	Gas viscosity (cp)	Gas FVF (RB/Mscf)	Z factor	Water density (lb/ft ³)	Water viscosity (cp)	Water FVF (RB/STB)	Water compressibility (1/psi)
250	0	0.050457	0.013269	242.701	0.9981	58.9483	0.2349	1.05979	5.75E-06
250	421.053	1.57848	0.013716	7.75813	0.94512	58.9841	0.2349	1.05915	5.72E-06
250	842.105	3.27212	0.014524	3.74255	0.89648	59.02	0.2349	1.05851	5.69E-06
250	1263.16	5.11209	0.015678	2.39551	0.8558	59.056	0.2349	1.05786	5.65E-06
250	1684.21	7.03176	0.01718	1.74154	0.82717	59.092	0.2349	1.05722	5.62E-06
250	2105.26	8.92502	0.018985	1.3721	0.81322	59.128	0.2349	1.05657	5.59E-06
250	2526.32	10.6923	0.020996	1.14531	0.81362	59.1641	0.2349	1.05593	5.55E-06
250	2947.37	12.2793	0.023109	0.99729	0.82586	59.2002	0.2349	1.05528	5.52E-06
250	3368.42	13.6753	0.02524	0.89549	0.84697	59.2364	0.2349	1.05464	5.48E-06
250	3789.47	14.8938	0.027337	0.82223	0.87447	59.2726	0.2349	1.054	5.45E-06
250	4210.53	15.9575	0.029368	0.76742	0.90651	59.3088	0.2349	1.05335	5.42E-06
250	4631.58	16.89	0.031319	0.72505	0.94181	59.3451	0.2349	1.05271	5.38E-06
250	5052.63	17.7129	0.033184	0.69137	0.97944	59.3815	0.2349	1.05206	5.35E-06
250	5473.68	18.4445	0.034965	0.66394	1.01874	59.4179	0.2349	1.05142	5.32E-06
250	5894.74	19.1001	0.036666	0.64115	1.05925	59.4543	0.2349	1.05077	5.28E-06
250	6315.79	19.692	0.038293	0.62188	1.10061	59.4908	0.2349	1.05013	5.25E-06
250	6736.84	20.2303	0.039852	0.60533	1.14259	59.5273	0.2349	1.04948	5.22E-06
250	7157.89	20.7229	0.041348	0.59094	1.18499	59.5639	0.2349	1.04884	5.18E-06
250	7578.95	21.1765	0.042788	0.57829	1.22768	59.6005	0.2349	1.0482	5.15E-06
250	8000	21.5964	0.044176	0.56704	1.27055	59.6372	0.2349	1.04755	5.11E-06

For drilling schedule, this study simulated the situation that there are 100 existing wells drilled in the field. We plan to use the information from these wells to make a prediction of well number 101. There are 2 groups of drilled wells. The 1st group consisting of 25 wells with at least 3,000 ft of well spacing are located randomly throughout the entire field area. The 2nd group consisting of 75 wells with at least 1,500 ft of well spacing located randomly throughout the entire field. This group represents the latest infill development plan by using half of the well spacing of the 1st group. Moreover, in real situations it is unlikely that a well is drilled in location close

to a reservoir boundary because the boundary effect will cause the pressure to rapidly drop faster than normal. Therefore, both groups of wells are located at least 750 ft away from any reservoir boundary to avoid the boundary effect. These 100 wells were drilled and produced at different times, i.e., Well number 1 was drilled first, and next subsequent wells were drilled and completed 10 days afterward and so on until a total of 100 wells have been drilled. The schematic of drilling location of all 100 wells on porosity distribution map is shown in Figure 4.5.

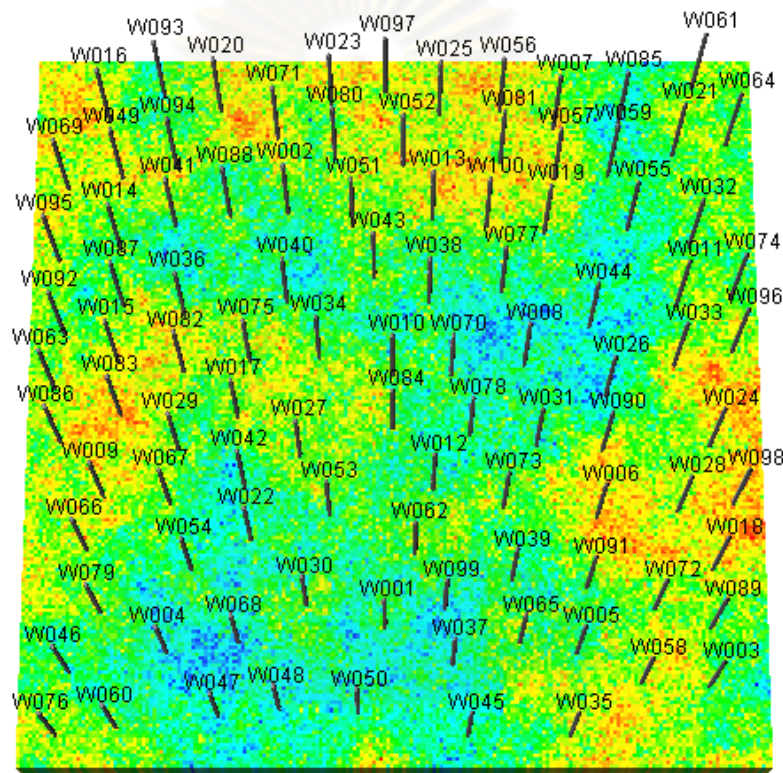


Figure 4.5: Drilled location on porosity distribution map

Downhole equipments of all wells are the same as shown in Table 4.4. A specification of downhole equipment, range of perforation depth, production well control, well economic limit, and vertical flow performance of the wells drilled in this field were set to be the same. All wells were perforated at depth all along a thickness of reservoir ($Z = 1$ to 5 in the simulation model). The minimum THP (Tubing Head Pressure) is set at 314.7 psia. Well economic limit is set to be 0.5 Mscf/d. The well will be automatically shut in after reaching this value.

Table 4.4: Downhole equipment

Parameters	Value	Unit
Wellbore ID	6.125	inch
Tubing ID	2.992	inch

The ECLIPSE program itself is able to calculate only the inflow of gas. Therefore, we need to provide the vertical lift performance table to the program, and it will calculate the flow from the bottom hole up to the surface. The vertical lift performance is calculated by using “PROSPER” program and is shown in Table 4.5.

Table 4.5: Vertical flow performance by “PROSPER” program

Gas rate (Mscf/day)	WellHead pressure (psig)	VLP pressure (psig)
1	300	351.74
10	300	351.73
20	300	351.73
30	300	351.72
40	300	351.72
50	300	351.72
60	300	351.72
70	300	351.72
80	300	351.72
90	300	351.72
100	300	351.73
500	300	353.04
1000	300	357.51
5000	300	485.88
10000	300	764.96
15000	300	1080.14
20000	300	1403.91
30000	300	2055.87
40000	300	2713.46
50000	300	3385.04

CHAPTER V

ARTIFICIAL NEURAL NETWORK MODEL DEVELOPMENT

This chapter describes a methodology to develop the Artificial Neural Network model to predict the gas production of the next infill well. The production data generated by the reservoir simulation in Chapter 4 are used as input for the network. In order to choose appropriate input and output of the network, related theories and methods from past literature need to be reviewed. Several cases were performed in order to investigate the effect of input parameters on the output. The ANN was trained with many network configuration models in each case. The accuracy of each case was examined by comparing the predicted output obtained from the ANN with the target output taken from reservoir simulations. Finally, the trained network was used to predict the best location for the next infill well, well number 101. The accuracy of prediction is examined again by comparing the predicted output with the simulation result.

5.1 Output of ANN

Since the purpose of this study is to predict a gas production at potential locations for drilling so that comparison among these locations can be made in order to find the best location to infill, the output of the ANN is expected gas production at each candidate well location. Gas production considered in this study includes 1) initial gas production rate and 2) cumulative gas production.

1) Initial gas production rate. This parameter represents the performance of drilling location at the moment when the well is put on production. It can be used as the output of ANN because it can roughly indicate a long term performance of the well located at a specific location. The location which has a higher initial gas rate will be more likely to yield a large amount of gas production in the long run than locations with lower initial gas rates.

2) Cumulative gas production. To determine locations that have better long-term performance, cumulative gas production is a more appropriate output. But the

problem is how long the cumulative production should be. The answer is that it depends on the drilling schedule of existing wells. For example, suppose that 10 wells have been drilled in the field and well number 8, 9, and 10 were drilled 2, 1.5, and 1 year ago. If we decide to predict 2 years of cumulative production, then we cannot use well number 9 and 10 to train the network. In this study, we use 1 year cumulative production as the output of ANN model such that data from more wells could be included in the training.

Consequently, this study uses 2 parameters as the output of ANN that are 1) initial gas production rate, and 2) 1 year cumulative gas production to indicate short term and long term performance, respectively.

5.2 Input of ANN

After appropriate output has been chosen, then appropriate inputs need to be identified. In order to train a network and use it to predict accurate output, we need to choose the right inputs, i.e., the inputs must have a good relationship with the output. With the right inputs, training a network becomes less difficult and good results can be produced. Therefore, many parameters that may have a relationship with the selected output need to be considered. This section describes how to choose these input parameters.

In order to choose appropriate inputs for the ANN for prediction of initial gas rate and 1 year cumulative gas production, we need to identify parameters that affect the gas production. In general, there are 3 sets of parameter that affect the gas production which are 1) reservoir parameters representing inflow, 2) well completion parameters representing outflow, and 3) control parameters representing surface conditions. In this study, we assume that all wells have the same completion and surface conditions. Thus, only reservoir parameters affect the initial gas production rate. The inflow of gas from the reservoir to the bottomhole can be analyzed via Darcy equation as written in Equation 5.1.

$$Q = (2\pi kh(P_R^2 - P_w^2))/(\mu \ln \left(\frac{r_e}{r_w}\right) - 0.75) \quad (5.1)$$

where	Q	=	gas production rate
	k	=	permeability
	h	=	reservoir thickness
	μ	=	viscosity
	P_R	=	average reservoir pressure
	P_w	=	well flowing pressure
	r_e	=	radius to external boundary
	r_w	=	wellbore radius

The equation shows that many parameters influence the gas production rate but most of these parameters are constant throughout the field. The parameter that should be used as input for the ANN is the one that is different from location to location. Therefore, viscosity and thickness are excluded in this study since the reservoir fluids and thickness are assumed to be the same throughout the field. Wellbore radius is also excluded because all wells have the same diameter. The drainage radius is represented by well spacing. Since all wells are subjected to the same controlled at the wellhead, i.e., the same wellhead pressure, the back pressure is the same for all wells. Thus, the bottomhole flowing pressure should not be considered as an input. As a result, only two parameters remain as input for the ANN, namely, permeability and average reservoir pressure.

The exact value of permeability at each location can be obtained after the well is drilled and the core sample is investigated. However, the permeability can be estimated using Geostatistics. In our case, we already know the permeability at at least 100 locations where existing wells are located. So we can use this information to estimate permeability in areas that are not drilled yet through a geostatistical method. Generally, the gas producing companies seem to have this information already. So it is not a big problem to find a permeability value in an undrilled area. But the average reservoir pressure is not easy to find because it cannot be known unless we drill and measure. Therefore, other parameters that affect the pressure should be considered instead.

The change in pressure after producing gas for a certain cumulative production can be analyzed via material balance equation. The material balance equation for depletion drive gas reservoir is written as Equation 5.2 as follows:

$$\frac{P}{Z} = \left(\frac{P_i}{Z_i}\right) \left(1 - \frac{G_p}{G}\right) \quad (5.2)$$

where	P_i	=	initial reservoir pressure
	Z_i	=	initial gas compressibility factor
	P	=	reservoir pressure at certain time
	Z	=	gas compressibility factor at certain time
	G	=	initial gas in place
	G_p	=	cumulative gas production

In this study, the initial pressure was set to be the same at 2,180 psia for all locations. Since gas is assumed to have the same gas properties at all locations, the initial gas compressibility factor must be the same for all locations. If 2 different areas with the same initial pressure and gas properties produce the same amount of cumulative gas but the pressure drops are different, then the original gas in place for the areas must be different. The area with higher initial gas in place has a smaller pressure drop than the area with lower initial gas in place. So, the amount of gas in place in the area strongly affects the change in pressure of the area. The amount of gas in place can be computed using Equation 5.3.

$$G = \frac{[Ah\Phi(1-S_w)]}{B_g} \quad (5.3)$$

where	G	=	initial gas in place
	A	=	area of reservoir
	h	=	reservoir thickness
	Φ	=	porosity
	S_w	=	water saturation
	B_g	=	gas formation volume factor

The equation shows that many parameters influence estimation of initial gas in place but most of these parameters are constant throughout the field. As described earlier, the parameter that should be used as input for the ANN is the one that is

different from location to location. Therefore, thickness of reservoir, and water saturation are excluded in this study since the thickness and water saturation are assumed to be the same throughout the field. Area is represented by well spacing. Gas formation volume factor is also excluded because the reservoir fluid is the same everywhere. As a result, only porosity is different for each location. So the amounts of initial gas in place are different depending on the porosity at each location. The area with high degree of porosity will result in a high amount of initial gas in place, and the rate of pressure drop is low when compared to the other areas with low porosity. So the porosity should be used as an input of ANN as it affects the change in pressure of the area the well is drilled in.

Not only the porosity of the grid where the well is located which affects the changing of reservoir pressure but also porosities of grid blocks surrounding the well. In order for the ANN to include porosities of blocks around the well, a “Data Mask” concept proposed by Boomer^[7] is used. With this concept, the area of interest is divided into 3 concentric rings as shown in Figure 5.1. The 1st ring consists of only grid block number 1. It represents porosity of well location. The 2nd ring consists of a group of grid blocks from grid block number 2 to number 9. This ring represents the porosities in an area that is a bit further away from the well. The 3rd ring consists of a group of grid blocks from grid block number 10 to number 25. It represents porosities in an area that is further away from the well. The average porosity for each concentric ring is calculated and used as representative porosity of each ring. These values will be used as input in one of the ANN trained in this study.

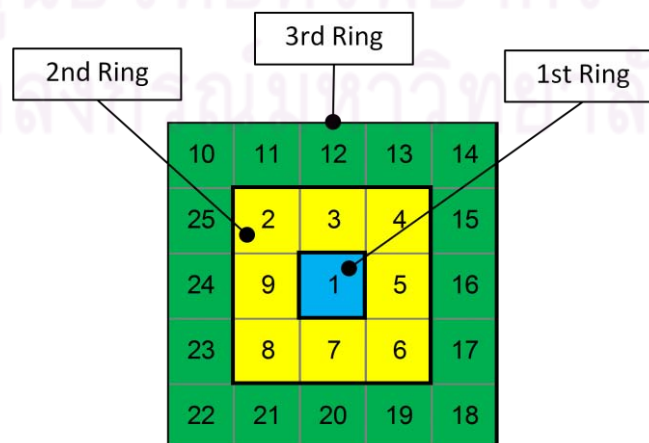


Figure 5.1: Schematic of porosity data mask

The next parameter that affects the changing of pressure is the drilling date of the well. Since the reservoir model in this study is closed boundary depletion drive gas reservoir, no drive mechanism such as water is available to support the reservoir pressure after the production starts. The reservoir pressure cannot be maintained and will continue to drop as long as the production continues. Therefore, the reservoir pressure at drilling date will be different from well to well

The other parameter that strongly affects the changing of reservoir pressure is the number of surrounding wells. The higher the amount of these wells, the lower the reservoir pressure at the predicting well location will be.

In summary, parameters that should be the input for ANN are tabulated in Table 5.1.

Table 5.1: Summary of parameters that should be the input for ANN

Item No	Parameters	Unit
1	Permeability	mD
2	Pressure	psia
3	Porosity	%
4	Starting date for drilling	day
5	Number of surrounding wells	well

5.3 Partitioning Data Sets

With an ability of ANN to learn complex relationship between input and output from a training data set fed to the network, in the first few iterations, the ANN may not give a good result because the weights and biases have not been updated to the right values yet. But after the training process continues to a certain number of iterations, the ANN will generally make a good prediction. However, it often presents inaccurate results when used to predict the output that the network has never seen before. This behavior is called “Overfitting”. Figure 5.2 and 5.3 show an example of this behavior. Shown in Figure 5.2, the ANN was trained until the approximation line is close to the target function. However, there are still some errors present at certain data points. But after continued training with a large number of iterations, then the error at all training points become zero as shown in Figure 5.3. But the approximation

line is completely different from the target function. So this network will surely not give good prediction because it is overfitted to the training data set.

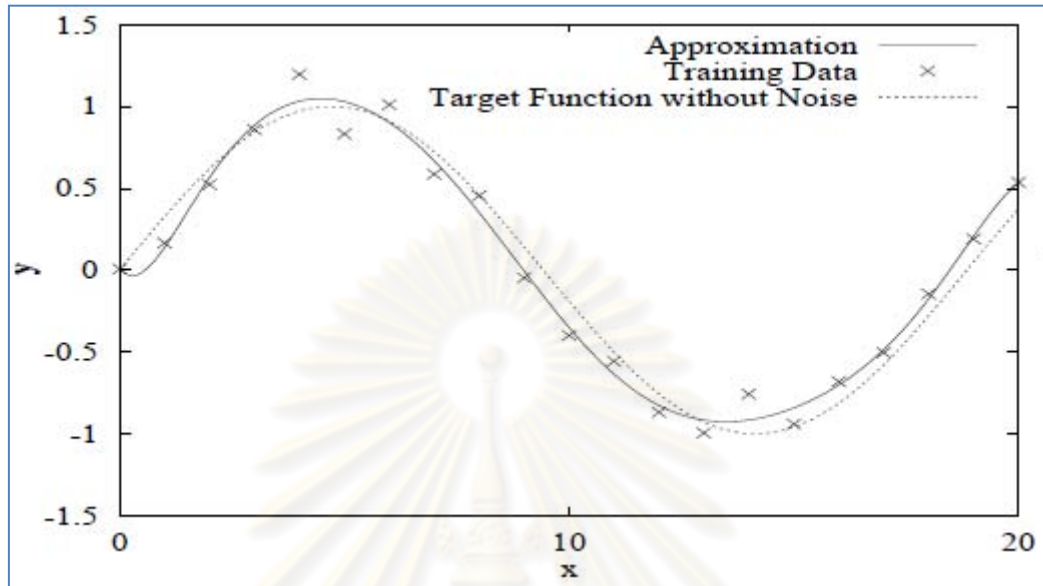


Figure 5.2: An example of target function fitted to the data ^[15]

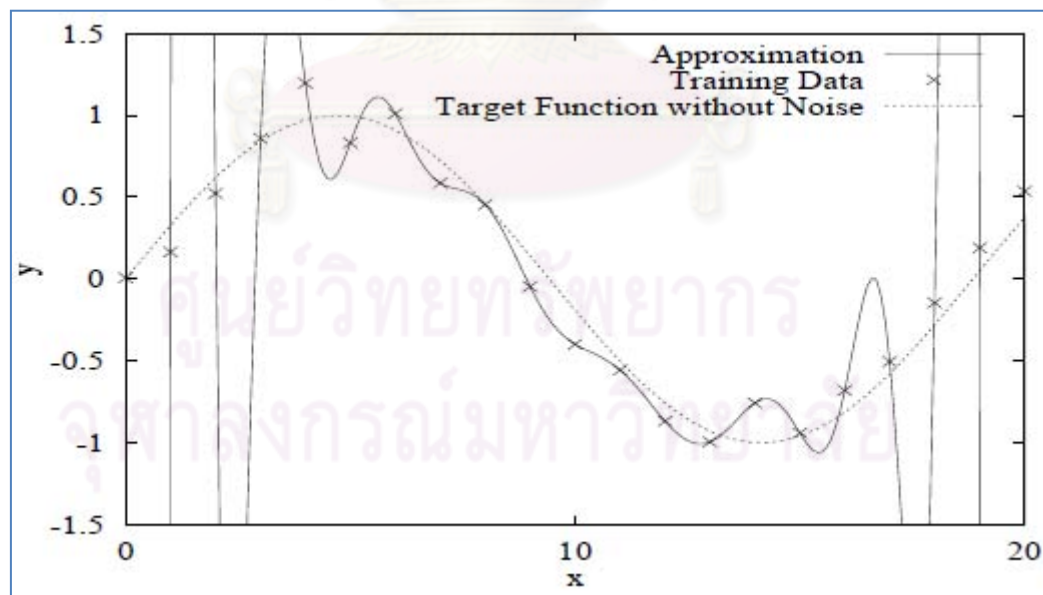


Figure 5.3: An example of overfitting behavior ^[15]

In order to avoid overfitting, the training process needs to stop before the network becomes overfitted by using the second group of data sets called “validating sets”. In the training process, both training and validating data sets are given to the

network at the same time but the weights and biases are updated only from the error of the training set. No weight updates are made from the error of validating set. With this method, the network will treat the validating set as a data set which it has never seen before. At each iteration (epoch), the errors of both training and validating sets are monitored, and the network will compare a new error of the validating set with the error from the previous epoch. If it is decreasing, then the training process will continue. The training process will keep going as long as the trend of the error of validating set is still decreasing. And at the moment that the errors start to increase, the network will know that further training will cause overfitting and the training process should be stopped. Therefore, we will have better opportunities to predict accurate output when using this network for unseen group of data set. Figure 5.4 is an example of error while training. From this figure, to avoid overfitting, we simply stop training at epoch 12, where performance on the validating set is optimal.

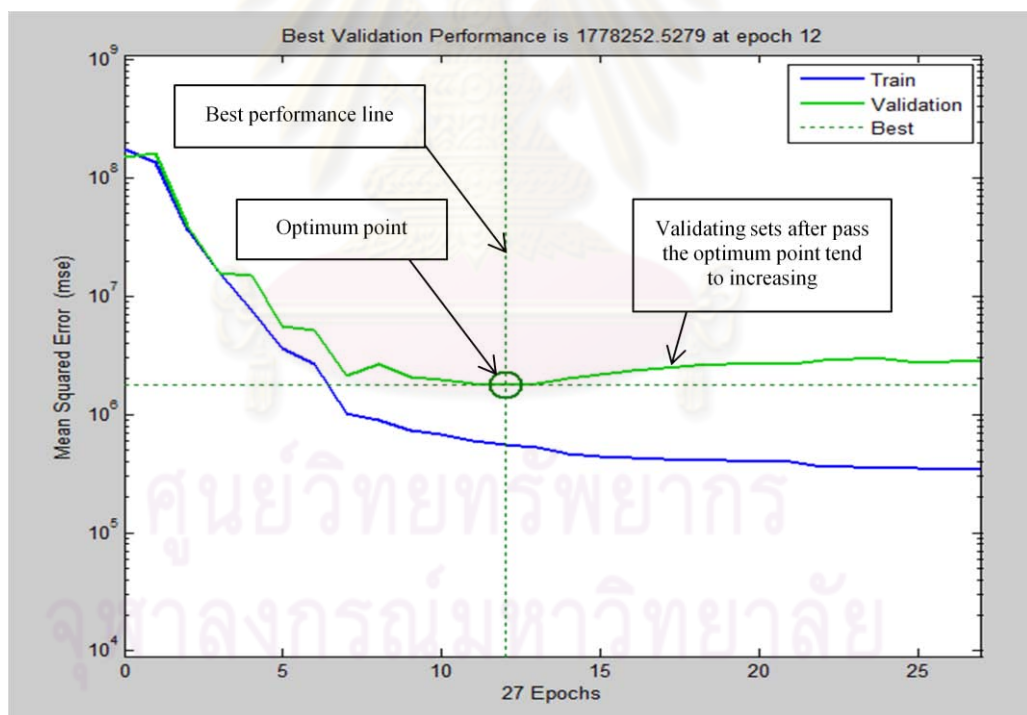


Figure 5.4: Learning process of ANN

Although the validating set is unseen by the network, the network can still be dependent on data chosen for validation. Therefore, if we wish to test the trained network on a set of independent data to measure its ability to generalize, we need

another set of independent data, a third group of data set called “Testing set”. With this type of data set, we can ensure the accuracy of the prediction of unseen data.

So before training a network, the pool of data needs to be divided into 3 main sets, namely, training set, validating set, and testing set. There is no restriction for the ratio of these 3 main sets. The famous and widely used ratio is 4:1:1 for training, validating, and testing sets, respectively. This study will use this ratio for partitioning the data.

Even if we have the production data of all 100 wells which seem to be plenty of information, we cannot use all of them to train a network. The reason is that in a real situation we want to use the information of surrounding wells to predict the performance of well in the middle. But some locations do not have enough number of surrounding wells to refer to or need large area for the moving window to include more wells. For example, if we look at the map after 100 wells have been drilled with the well name in chronological order (well 1 is drilled prior to well 2, well 2 is drilled prior to well 3, so on and so forth), we may find that there are well number 30, 37, 50, 62, and 99 surrounding well number 1 as shown in Figure 5.5. But in the real situation at the time when predicting the initial rate for well number 1, there is no well surrounding well number 1 because well number 1 is drilled prior to other wells. Therefore, well number 1 cannot be used to train the network. Another example is concerned with well number 22. There are well number 30, 42, 53, 54, and 67 surrounding well number 22 as shown in Figure 5.6. In fact, at the time of the drilling of well number 22, there is no surrounding well because well number 22 is drilled before the others. Therefore, well number 22 cannot be included in the training.

In this study, we used the data sets from the 2nd round of infill wells to train a network because we plan to infill the 101st well at the same well spacing that is 1500 ft. Therefore, we have a total of 75 data sets (from well number 25 to 100) to train the network. The data sets are divided into 3 main sets, namely, training, validating, and testing sets with ratio of 4:1:1. As a result, the numbers of training, validating, and testing data sets are 51, 12, and 12, respectively.

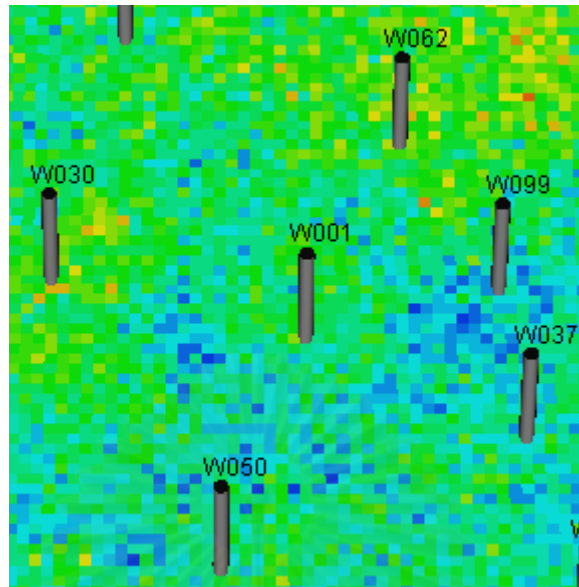


Figure 5.5: Locations of well number 1 and surrounding wells on porosity map

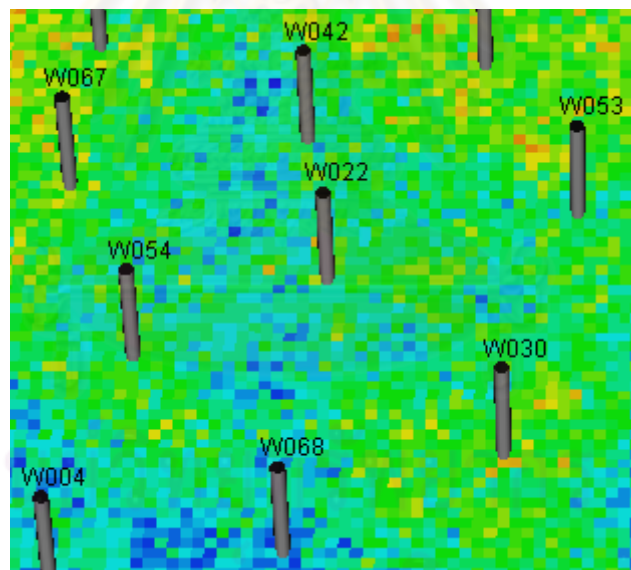


Figure 5.6: Locations of well number 22 and surrounding wells on porosity map

5.4 Artificial Neural Network Case Study

After the input and output parameters have been selected in Section 5.1 and 5.2, they are used to train the network and evaluate accuracy of prediction. Two kinds of prediction are used in this studies: an initial gas production rate and one-year cumulative gas production.

5.4.1 Initial Gas Rate Prediction

As described earlier via Darcy equation (Equation 5.1), the initial gas production rate has a relationship only with pressure and permeability at each well location since others parameters are constant. So in order to prove this conclusion and test how well a prediction performance of ANN is, Case 1-1 is set up using the pressure at the date of drilling and permeability at well location as ANN input parameters. However, in reality, the pressure at the location to be drilled is not known prior to drilling. Therefore, shut-in pressures from surrounding wells should be used to represent pressure at the location to be drilled. In this study, there are 2 averaging methods. 1) arithmetic average and 2) inverse distance average, closer wells will affect the average value more than further well. Case 1-2-1 and Case 1-2-2 are set up using average pressure. This study also includes a case that does not use pressure as an input but use other parameters as a proxy of pressure instead. These parameters are porosity, drilling date, and the number of surrounding wells. Case 1-3 is set up for this scenario.

In each case study, the ANN is trained with training and validating sets by varying network configurations that are number of hidden nodes, number of hidden layers, learning rate, and momentum on a trial and error basis. Only 2 models which give the lowest and next to lowest error of validating sets are used to test the accuracy with testing set. After that the only one model which gives the lowest error of the testing set will be chosen to be the best performance model for each case study. Finally, best performance model will be used to predict the initial gas production rate of well number 101 which is scheduled to be drilled one year after well number 100.

5.4.1.1 Case 1-1

In this case, the network consists of 2 input parameters which are permeability and pressure at the date of drilling at well location. The number of hidden layers and the number of neurons in each hidden layer are varied based on trial and error to get the best performance of predicted output. Figure 5.7 illustrates the schematic diagram of ANN in this case.

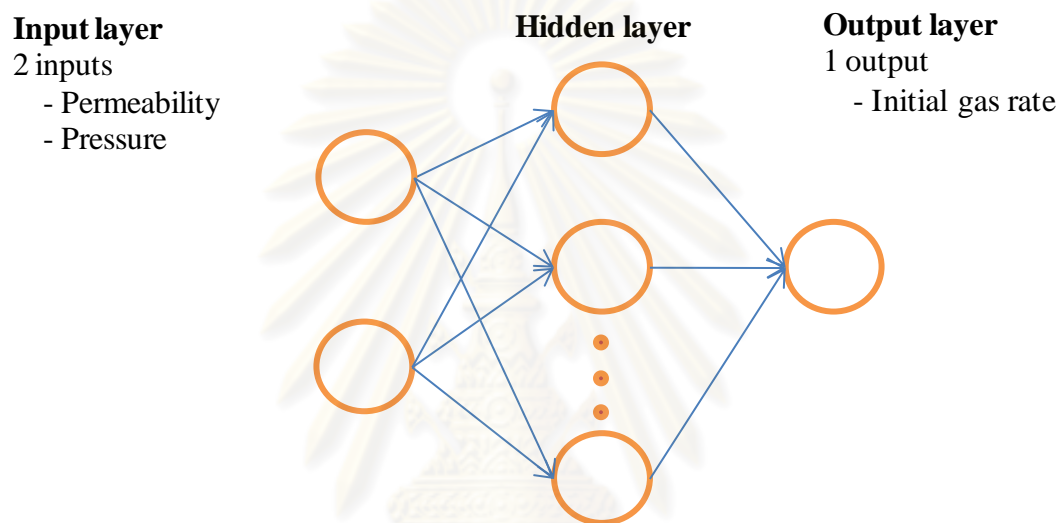


Figure 5.7: Schematic diagram of ANN for Case 1-1

In real situation, the pressure at the well location prior to drilling is unknown. As a result, it cannot be used as an input to predict performance of the well to be drilled. The purpose of this case study is just to prove a relationship between the input that are permeability and pressure with the output that is initial gas production rate.

5.4.1.1.1 Data Preprocessing

Before starting a training process, a total of 75 data sets taken from well number 26 to 100 need to be divided into three main sets as described earlier. With the ratio 4:1:1, the number of training, validating, and testing sets will be 51, 12, and 12, respectively. In order to train a network that can produce accurate output when used to predict output for unseen input, each parameter of all 3 data set should exhibit a similar distribution as much as possible. If the data sets are divided based on well number sequence (first 51 wells are training set, next 12 wells are validating set, and

last 12 wells are testing set), then the testing set will contain low values of reservoir pressure because the reservoir pressure of closed boundary depletion drive gas reservoir decreases as a function of time. So the distribution of reservoir pressure for the three types of data sets will not have a similar trend. Therefore, we randomly rearranged the data sets, divided into 3 sets, and then plotted their distribution to observe their behaviors. If the trends look similar, then we use the three divided sets of data to train the ANN. The histograms of the three data sets are shown in Figures 5.8 to 5.13 and their statistics are summarised in Table 5.2.

The histograms and statistics of the 3 data sets look similar. For example, the mean values of pressure of training, validating, and testing sets are 1251.43, 1195.37, 1248.57 psia, respectively. They are close to each other. This grouping of data sets are kept unchanged throughout the study because we want to compare the accuracy by varying only the network configuration. So, all parameters except the network configuration need to be controlled to be the same for all models in order to ensure that the only changes applied are network configuration.

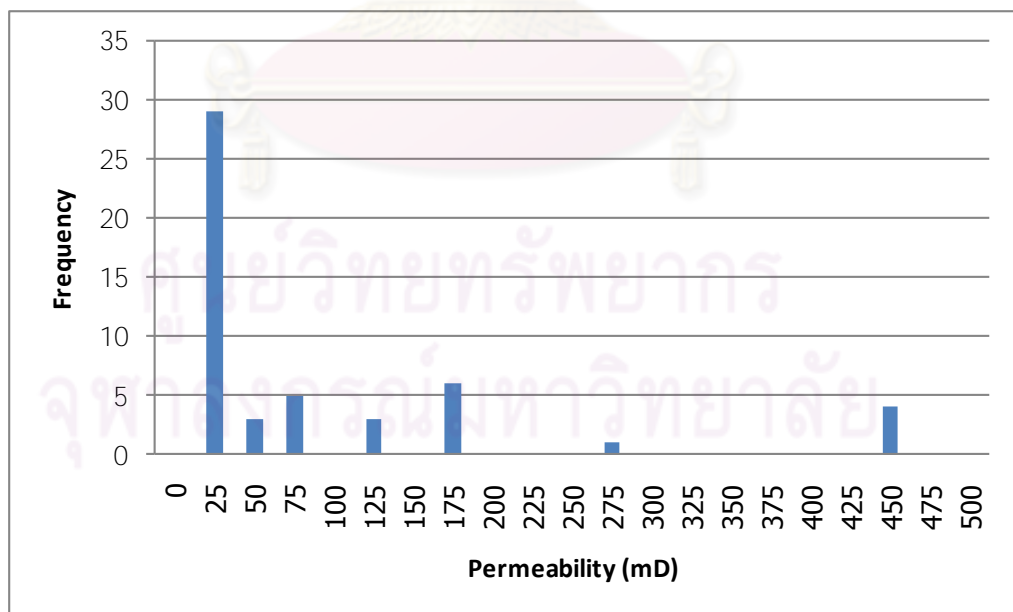


Figure 5.8: Histogram of permeability for training sets

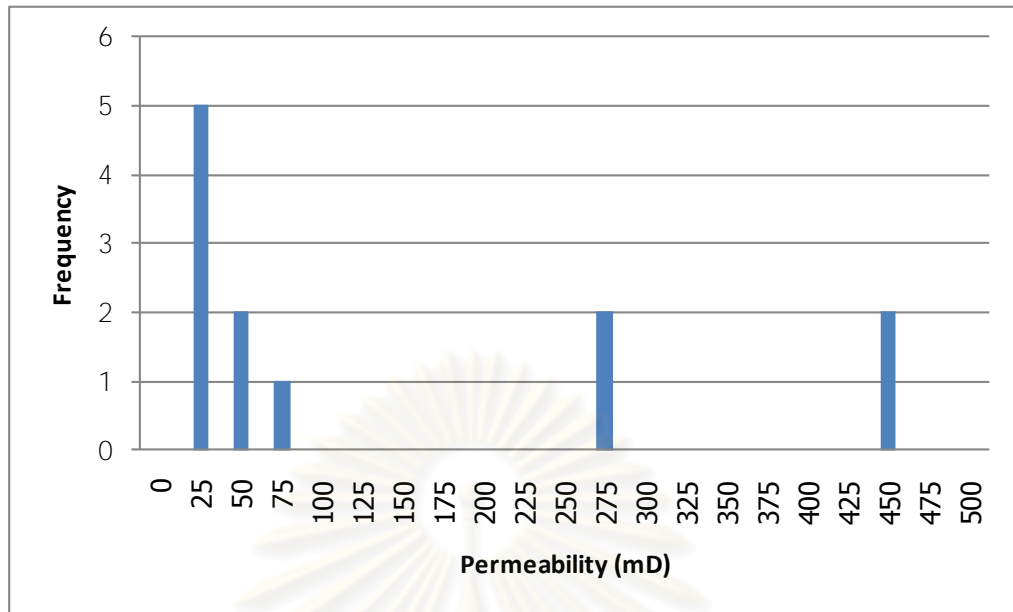


Figure 5.9: Histogram of permeability for validating sets

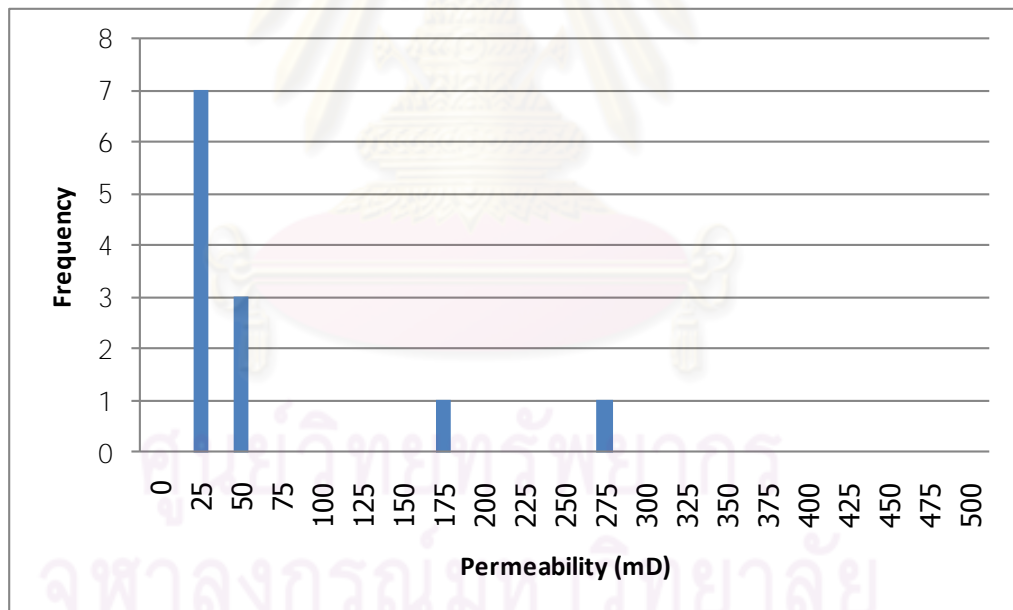


Figure 5.10: Histogram of permeability for of testing sets

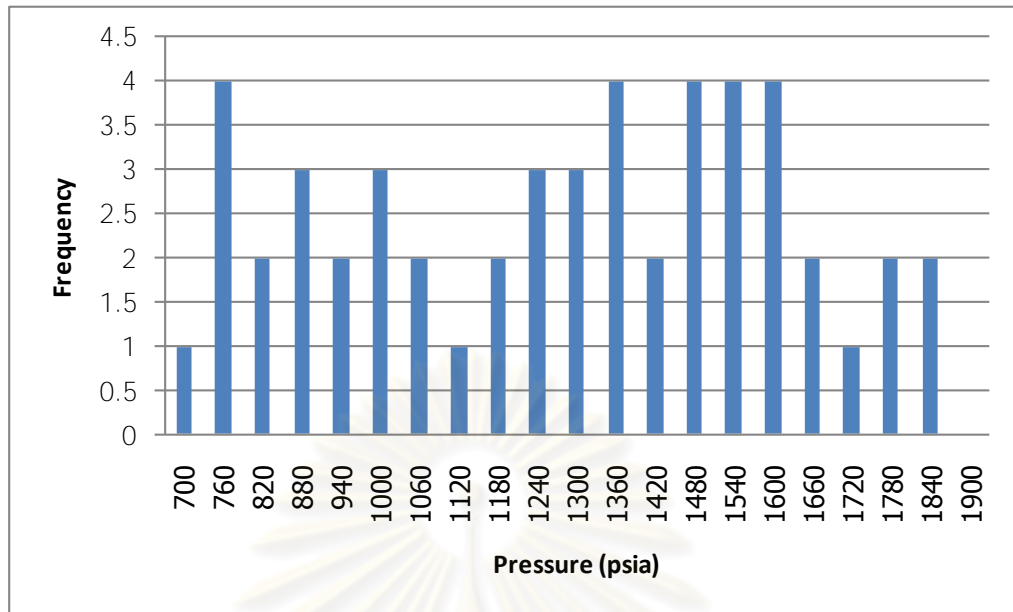


Figure 5.11: Histogram of pressure for training sets

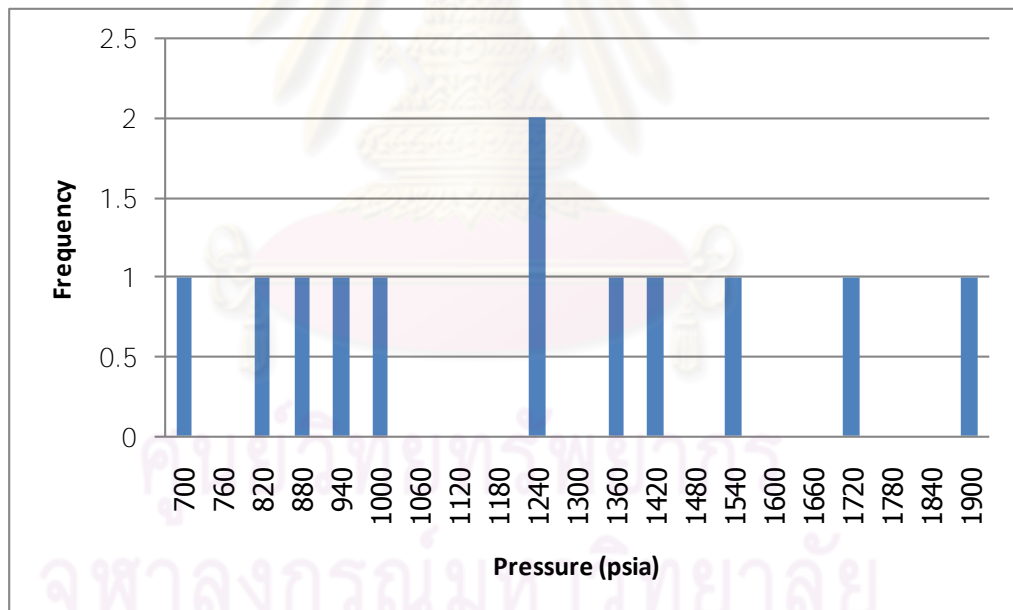


Figure 5.12: Histogram of pressure for validating sets

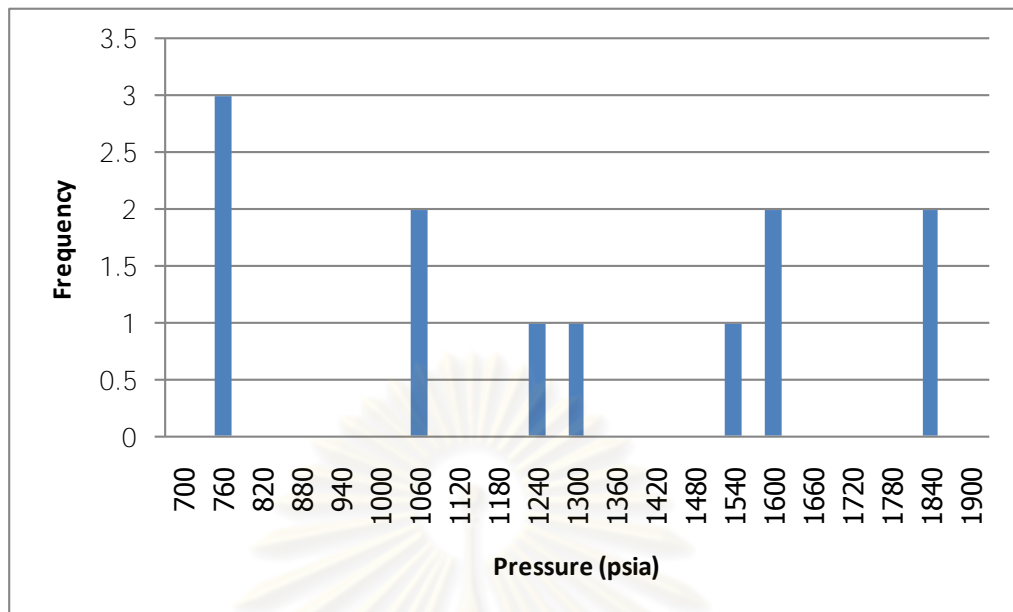


Figure 5.13: Histogram of pressure for testing sets

Table 5.2: Summary of statistics of all data sets (Case 1-1)

Statistical parameter	Dataset type	Permeability (mD)	Pressure (psia)
Maximum	Training	439.74	1,817.56
	Validating	439.74	1,858.05
	Testing	271.83	1,811.70
Minimum	Training	0.52	693.40
	Validating	0.85	676.13
	Testing	2.21	704.13
Mean	Training	78.91	1,251.43
	Validating	133.76	1,195.37
	Testing	51.93	1,248.57
1st quartile	Training	5.80	974.86
	Validating	5.24	908.57
	Testing	5.24	952.33
Median	Training	15.17	1,282.27
	Validating	39.69	1,183.70
	Testing	19.85	1,249.92
3rd quartile	Training	103.87	1,513.42
	Validating	271.83	1,403.99
	Testing	39.69	1,561.89
SD	Training	123.30	328.43
	Validating	172.62	371.58
	Testing	82.83	407.96

5.4.1.1.2 Model Training

In this study, the ANN model was developed by using the program “MATLAB”. This program is widely used to calculate and solve problems in various branches such as engineering and science. It consists of many useful tool boxes. One of them is the neural network toolbox. With the toolbox, we can design the network structure the way we want through the source code. An example of source code in this study is shown in Figure 5.14.

```

1      %Input & Target Output
2 -    input=[];
3 -    target=[];
4
5      %Construct a network
6 -    net = newff(input, target,[5],{'logsig'},'trainlm');
7
8      %Network configuration (Learning rate & Momentum)
9 -    net.trainParam.lr = 0.1;
10 -   net.trainParam.mc = 0.5;
11
12     %Criteria to stop training
13 -   net.trainParam.min_grad = 1e-20;
14 -   net.trainParam.epochs = 1000;
15 -   net.trainParam.mu_max = 1e2000;
16 -   net.trainParam.max_fail = 15;
17
18     %Partitioning ratio for training, validating, and testing sets
19 -   net.divideParam.trainRatio = 4;
20 -   net.divideParam.valratio = 1;
21 -   net.divideParam.testratio = 0;
22
23     %Start to train a network
24 -   [net,tr]=train(net,input,target);

```

Figure 5.14: An example of MATLAB source code of ANN model

After a training process is finished, the training result window appears as shown in Figure 5.15. The ANN training result window of MATLAB software can be divided into 4 main sections as follows:

Section 1 represents roughly schematic diagram of network configuration. In our case, the network consists of 2 layers (1 hidden layer and 1 output layer). Note that MATLAB software does not count the input as a layer. A logsig and linear transfer function is used for the hidden and output layers, respectively.

Section 2 represents the algorithms that were used to train a network. In our case, the training method is Levenberg-Marquardt with the Mean Squared Error (MSE) as a criterion to check the performance of the network and the data sets are divided by choosing from a sequence that we input to the network by a method called "divideblock".

Section 3 represents the progress of the training process. Once one of these criteria reaches its target, the training process is then stopped. There are 3 main criteria used in this study as follows:

1) Target MSE criterion (performance) – The training will continue until the MSE of training set reaches the target MSE. Setting the target MSE to be a high value will cause the network to have low accuracy. On the other hand, setting it too low will result in a long time to train a network or even non-convergence.

2) Number of epoch – This value is specified to limit the number of iterations in the training. The advantage is to stop the training in case that the performance cannot reach the target MSE. In this study, the maximum epoch was set quite high (1,000 epochs) because we would like to ensure that the MSE is as low as possible.

3) Validation Checks – As described earlier, the validating data set is used to avoid a network to overfit training data set by stopping the training process when the MSE of the validating set starts to increase. However, this criterion is too rigid since the trend of the MSE may change from increasing to decreasing again in subsequent epochs. Therefore, we should allow for some flexibility for the network to continue training even the MSE of the validating set is increasing. In this study, the number of epochs used for this purpose is 15.

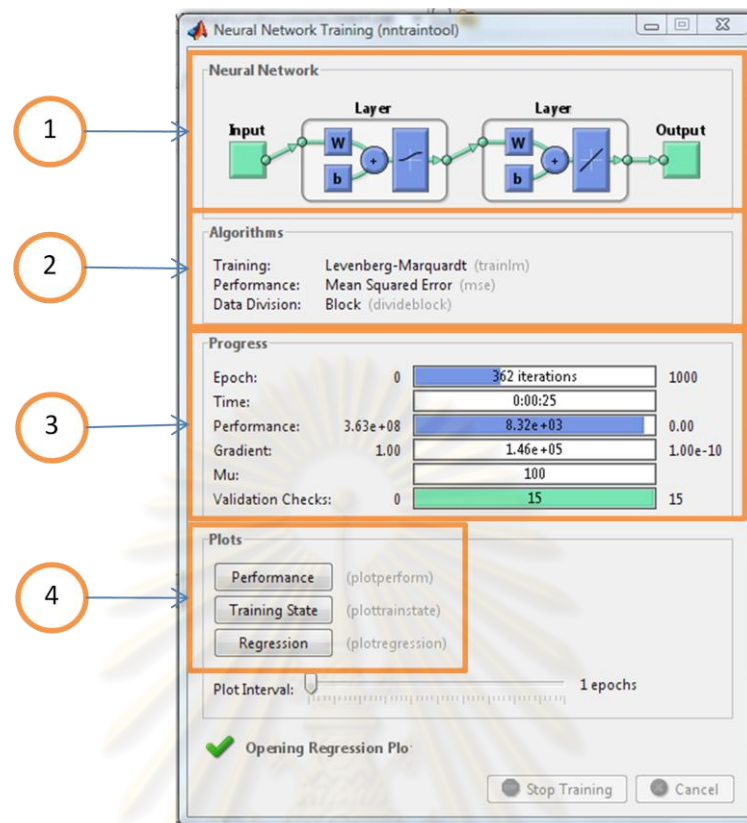


Figure 5.15: Training result window from ANN

Section 4 provides the buttons to plot the training results. There are 3 items which are 1) performance showing the MSE of all data sets at each epoch while the training process is still running, 2) training state showing the progress of each stop criterion, and 3) regression showing cross plot between predicted and actual result.

The ANN model was trained with various network configurations based on a trial and error basis. There are 20 models that were run in this case. Each model was trained many times to obtain the lowest MSE possible. The network configurations and their MSE of validating set are summarized in Table 5.3.

Table 5.3: Model configuration for Case 1-1

Model No	Number of neurons		Learning rate	Momentum	MSE
	Hidden Layer 1	Hidden Layer 2			
1	5	0	0.1	0.1	10,599
2	5	0	0.5	0.1	51,291
3	5	0	0.1	0.5	42,843
4	5	0	0.5	0.5	54,947
5	10	0	0.1	0.1	57,338
6	10	0	0.5	0.1	74,058
7	10	0	0.1	0.5	106,387
8	10	0	0.5	0.5	92,292
9	20	0	0.1	0.1	771,261
10	20	0	0.5	0.1	1,457,879
11	20	0	0.1	0.5	311,343
12	20	0	0.5	0.5	1,562,568
13	5	5	0.1	0.1	6,437
14	5	5	0.5	0.1	7,183
15	5	5	0.1	0.5	10,455
16	5	5	0.5	0.5	3,881
17	10	10	0.1	0.1	37,516
18	10	10	0.5	0.1	320,963
19	10	10	0.1	0.5	678,858
20	10	10	0.5	0.5	216,053

From a total 20 model configurations, two models with the lowest and next to lowest MSE (model 13 and 16) which has the MSE of 6,437 and 3,881, respectively, are chosen. Model 13 consists of 2 hidden layers with 5 neurons in each layer. The learning rate and momentum of 0.1 were used. Similar to model 13, model 16 consists of 2 hidden layers with 5 neurons in each layer. However, the learning rate and momentum were set to be 0.5. The performance curves of model 13 and 16 are shown in Figures 5.16 and 5.17, respectively. Model 13 was trained until epoch 107 but the weight and bias were updated until epoch 92 only because the MSE of validating set started to increase in this epoch. The training was continued for 15 more epochs for validation check. The lowest MSE of model 13 is 6,437. For model 16, the training was performed until epoch 123. However, the weight and bias were not updated after epoch 108 due to the same reason for model 13. The lowest MSE of model 16 is 3,881 in epoch 108.

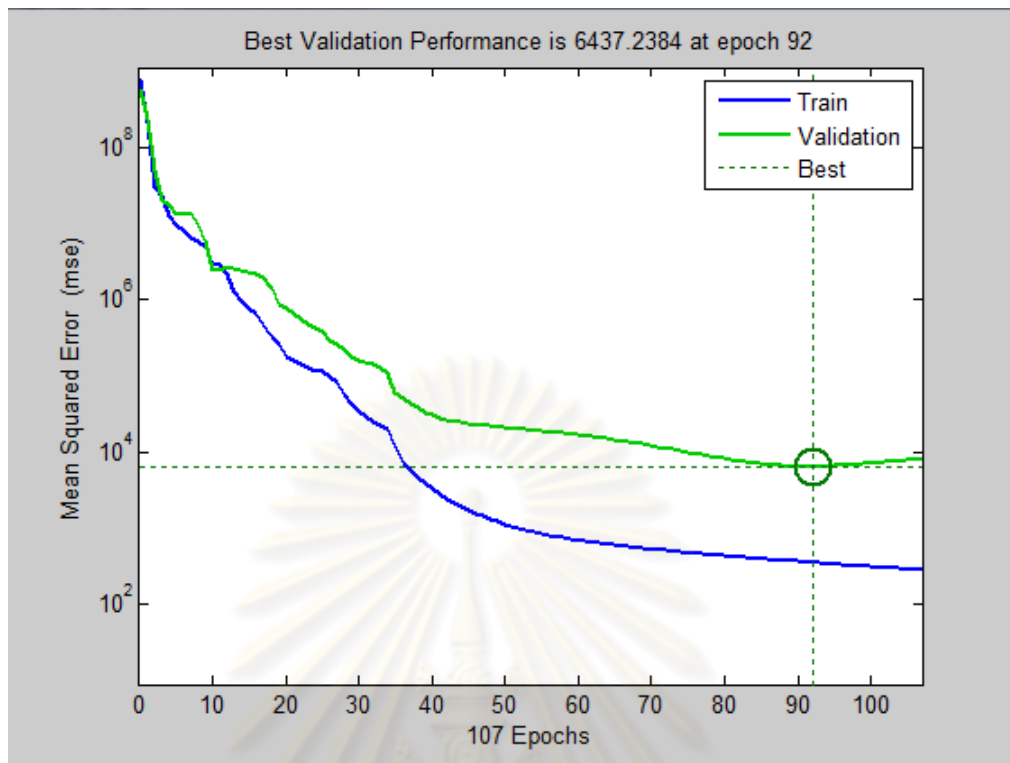


Figure 5.16: Performance curve of model 13 (Case 1-1)

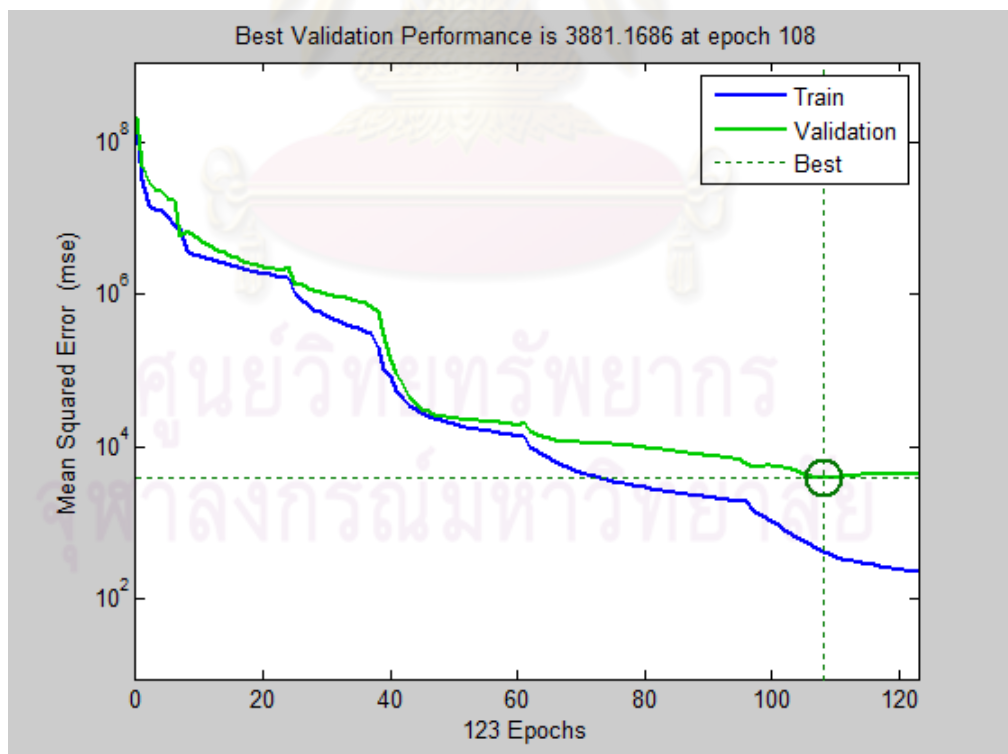


Figure 5.17: Performance curve of model 16 (Case 1-1)

Next, both models that produce the lowest MSE are further checked to ensure the accuracy of prediction. The outputs predicted by the ANN are compared with the target outputs of the training and validating sets by cross plotting them. Figures 5.18 and 5.19 represent the cross plots for the training and validating sets for model 13, respectively, while Figures 5.20 and 5.21 represent the cross plots for the training and validating sets for model 16, respectively. From the graph, the line $Y = X$ refers to correct prediction, i.e., each point on the 45-degree line is where predicted output is matched with the target output. So the closer the data points are located near the $Y = X$ line, the higher the accuracy of prediction. We can determine the accuracy of the prediction using regression coefficient of determination (R^2) as a criterion. R^2 equal to 1 represents a perfect fit to the $Y = X$ line. ANN will predict accurate output when R^2 is close to 1.

From Figures 5.18 to 5.21, we can see that the ANN can predict accurate output for both training and validating sets for both model 13 and 16. Model 13 gives R^2 of training and validating sets equal to 1 and 0.9998, respectively, and model 16 gives R^2 of the training and validating sets equal to 1 and 0.9999, respectively as well.

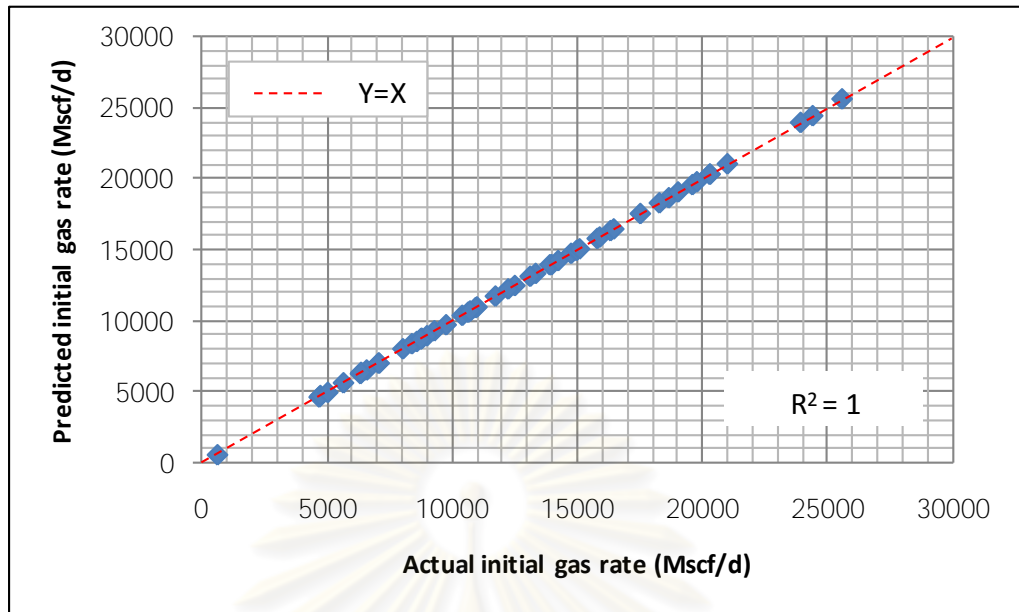


Figure 5.18: Cross plot of predicted vs actual initial gas rates of training sets of model 13 (Case 1-1)

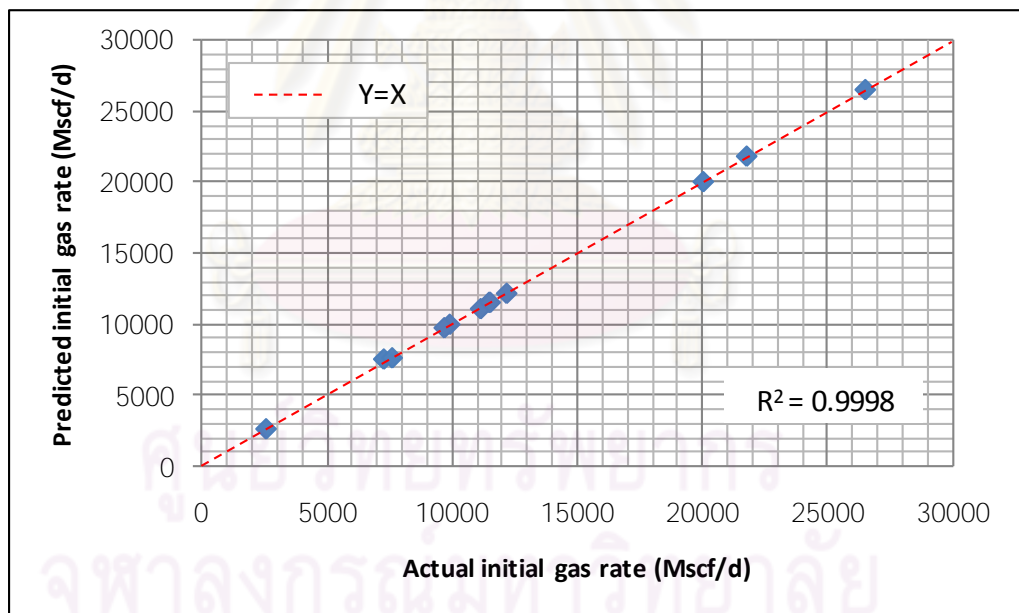


Figure 5.19: Cross plot of predicted vs actual initial gas rates of validating sets of model 13 (Case 1-1)

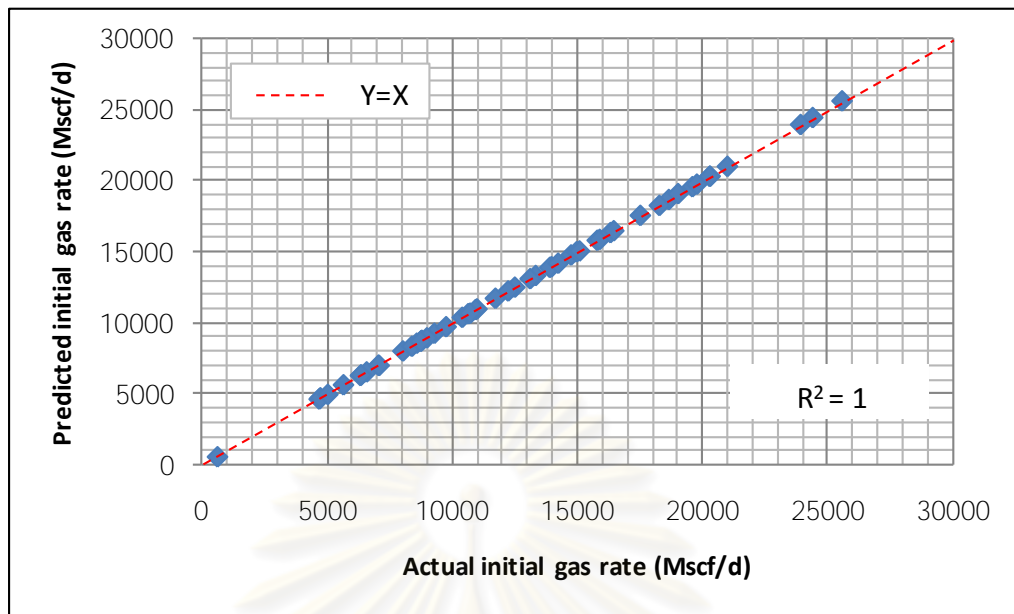


Figure 5.20: Cross plot of predicted vs actual initial gas rates of training sets of model 16 (Case 1-1)

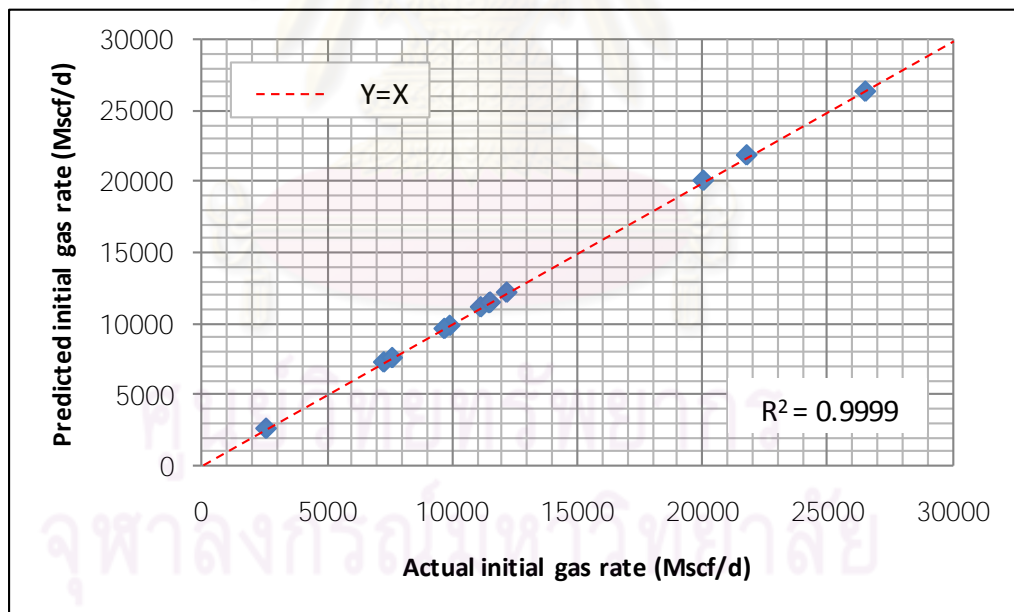


Figure 5.21: Cross plot of predicted vs actual initial gas rates of validating sets of model 16 (Case 1-1)

5.4.1.1.3 Model Testing Results and Discussion

In order to ensure the accuracy of ANN prediction when faced with unseen data sets, model 13 and 16 which yield the lowest MSE were then tested for accuracy by using testing data sets. After testing the ANN with testing sets, the outputs predicted by ANN are then compared with the target outputs by cross plotting them. Figures 5.22 and 5.23 represent the cross plots for model 13 and 16, respectively. From the graphs, R^2 of model 13 and 16 are equal to 0.9997 and 0.9998, respectively.

This means that both model 13 and 16 good performance of predicting the initial gas production rate. However, the coefficient of determination for model 16 is higher. Therefore, the best performance model which produces the most accurate predicted output is model 16.

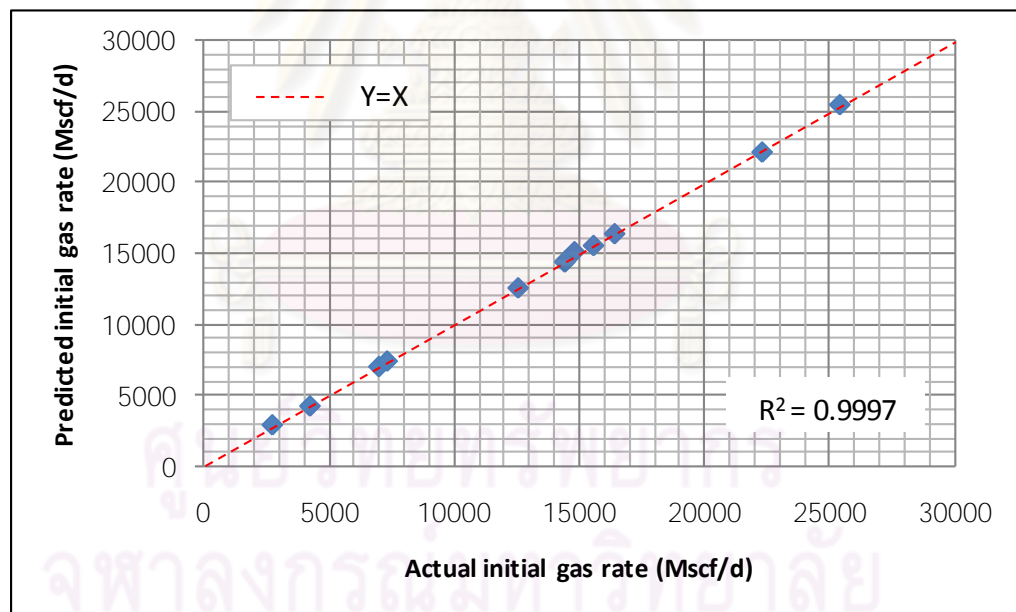


Figure 5.22: Cross plot of predicted vs actual initial gas rates of testing sets of model 13 (Case 1-1)

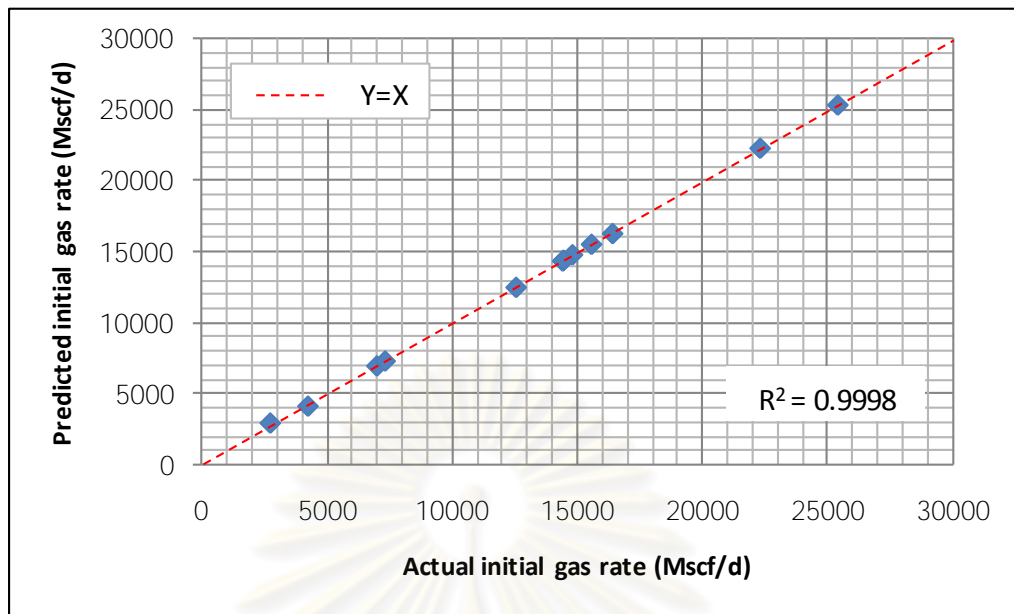


Figure 5.23: Cross plot of predicted vs actual initial gas rates of testing sets of model 16 (Case 1-1)

In summary, we have proved that the ANN can be used to predict initial gas rate when necessary input information is available. However, in practice, even the best model from this study cannot be used to predict the output of well number 101 because it is impossible to know the pressure at well number 101 prior to drilling. In any case, this case study helps to ensure that initial gas production rate is strongly related to permeability and pressure at predicting well location.

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

5.4.1.2 Case 1-2-1

Due to circumstances in real situation, we are unable to know the pressure at the predicting well location. However, the average pressure from the surrounding wells' location can be obtained. These pressures were taken from surrounding wells location at the time drilling started for the predicting well. For example, we are predicting the flow rate for well number 20 and the well is surrounding by well number 15 and 16. In this case, we have to use the pressures from well number 15 and 16 at the time when we start to drill well number 20 to find the average pressure and feed it to the ANN as an input parameter. Figure 5.24 illustrates the schematic diagram of ANN in this case.

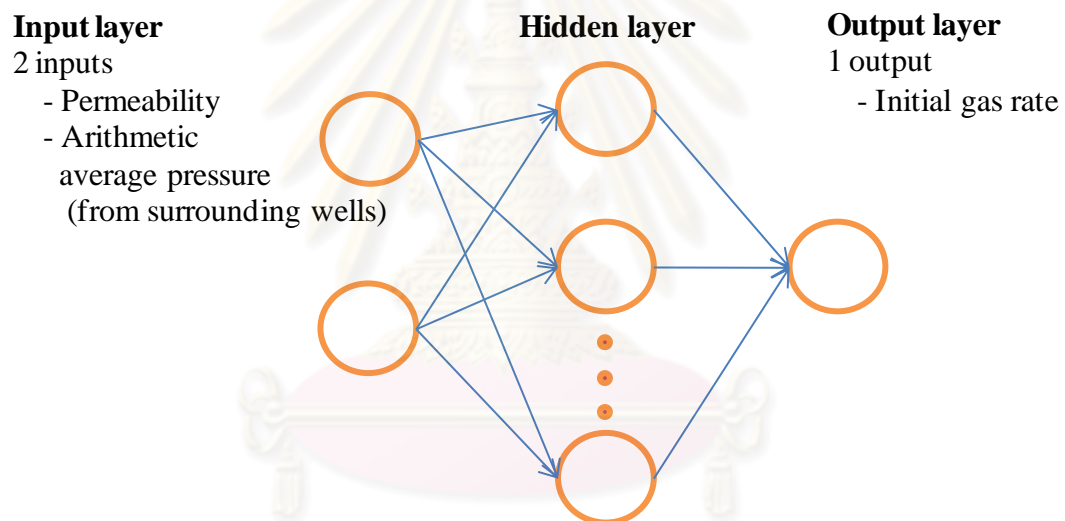


Figure 5.24: Schematic diagram of ANN for Case 1-2-1

In order to average the pressure, the number of surrounding wells needs to be counted first. Therefore, the area of influence needs to be specified. It is necessary to choose the right size of area of influence. Choosing the area to be too large can result in larger error in estimating the pressure because pressure in different areas in the reservoir are different due to heterogeneity. The pressure should be referenced from nearest surrounding wells. On the other hand, choosing the area to be too small can result in the lack of number of surrounding wells. From a trial and error basis, we found that the area of the size 5,300 ft x 5,300 ft around the predicting well location is the minimum size that has at least 1 well located around all predicting wells. Note that is quite large when compared with the total reservoir size (20,000 ft x 20,000 ft).

5.4.1.2.1 Data Preprocessing

Similar to the previous case, a total of 75 data sets taken from well number 26 to 100 were divided into three main sets namely, training, validating, and testing sets with ratio of 4:1:1 (51:12:12 data sets). The wells in each data set are still the same as the ones in the previous case study. Input parameters from all data sets are plotted to observe the distribution. Since the distribution of permeabilities are the same as in the previous case study, only the distributions of average pressures are then plotted. The histograms are shown in Figures 5.25 to 5.27 and their statistics are summarized in Table 5.4. From the result, we found that the histograms and statistics of the 3 data sets represent similar distributions.

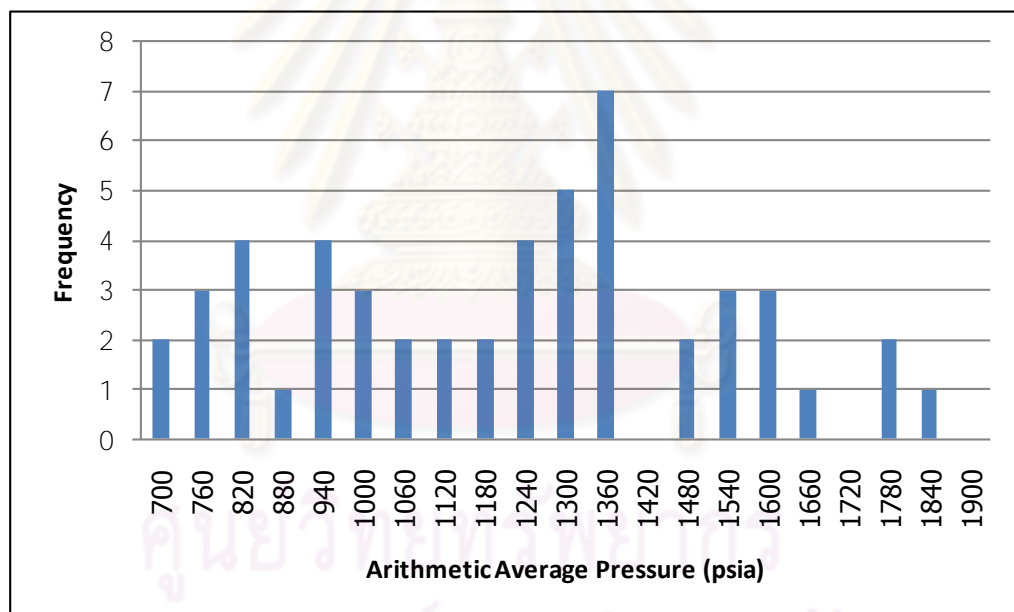


Figure 5.25: Histogram of arithmetic average pressure for training sets

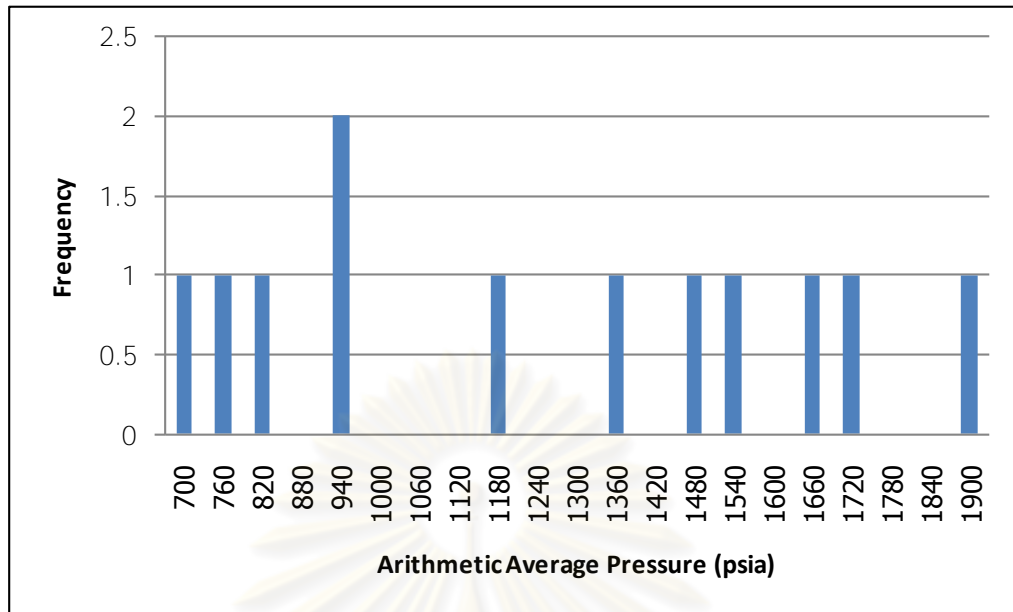


Figure 5.26: Histogram of arithmetic average pressure for validating sets

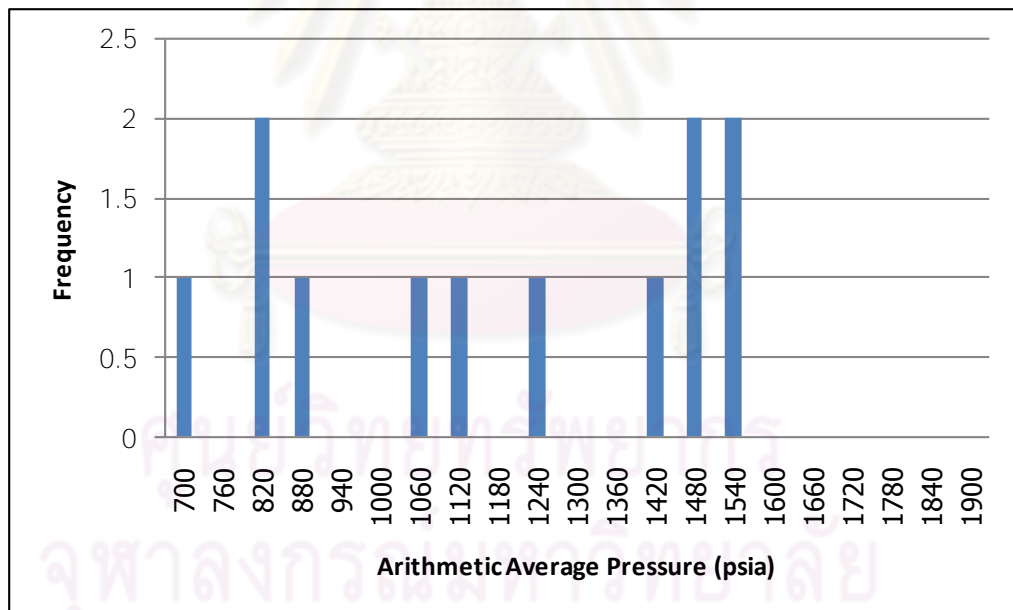


Figure 5.27: Histogram of arithmetic average pressure for testing sets

Table 5.4: Summary of statistics of all data sets (Case 1-2-1)

Statistical parameters	Dataset type	Arithmetic average pressure (psia)
Maximum	Training	1,798.89
	Validating	1,857.74
	Testing	1,533.35
Minimum	Training	690.99
	Validating	663.15
	Testing	669.20
Mean	Training	1,181.31
	Validating	1,219.51
	Testing	1,148.48
1st Quartile	Training	931.72
	Validating	879.81
	Testing	859.91
Median	Training	1,221.89
	Validating	1,247.85
	Testing	1,155.25
3rd Quartile	Training	1,338.24
	Validating	1,527.48
	Testing	1,433.24
SD	Training	301.00
	Validating	412.13
	Testing	319.22

5.4.1.2.2 Model Training

The ANN model was trained with various network configurations based on a trial and error basis. There are 20 models that were run in this case. Each model was trained many times to obtain the lowest MSE possible. The network configurations and their MSE of validating set are summarized in Table 5.5.

Table 5.5: Model configuration for Case 1-2-1

Model No	Number of neurons		Learning rate	Momentum	MSE
	Hidden Layer 1	Hidden Layer 2			
1	5	0	0.1	0.1	1,342,442
2	5	0	0.5	0.1	1,680,652
3	5	0	0.1	0.5	1,375,500
4	5	0	0.5	0.5	1,270,276
5	10	0	0.1	0.1	1,869,682
6	10	0	0.5	0.1	1,642,784
7	10	0	0.1	0.5	1,842,319
8	10	0	0.5	0.5	2,025,834
9	20	0	0.1	0.1	4,231,747
10	20	0	0.5	0.1	3,595,650
11	20	0	0.1	0.5	2,799,095
12	20	0	0.5	0.5	3,255,664
13	5	5	0.1	0.1	1,735,086
14	5	5	0.5	0.1	1,592,139
15	5	5	0.1	0.5	1,344,642
16	5	5	0.5	0.5	1,436,845
17	10	10	0.1	0.1	2,155,072
18	10	10	0.5	0.1	1,303,863
19	10	10	0.1	0.5	1,962,537
20	10	10	0.5	0.5	1,964,948

From a total 20 model configurations, two models with the lowest and next to lowest MSE (model 1 and 4) which has the MSE of 1,342,442 and 1,270,276, respectively, are chosen. Model 1 consists of only one hidden layers with 5 neurons. The learning rate and momentum of 0.1 were used. Similar to model 1, model 4 consists of only one hidden layers with 5 neurons. However, the learning rate and momentum were set to be 0.5. The performance curves of model 1 and 4 are shown in Figures 5.28 and 5.29, respectively. Model 1 was trained until epoch 29 but the

weight and bias were updated until epoch 14 only because the MSE of validating set started to increase in this epoch. The training was continued for 15 more epochs for validation check. The lowest MSE of model 1 is 1,342,442. For model 4, the training was performed until epoch 32. However, the weight and bias were not updated after epoch 17 due to the same reason for model 1. The lowest MSE of model 4 is 1,270,276 in epoch 17.

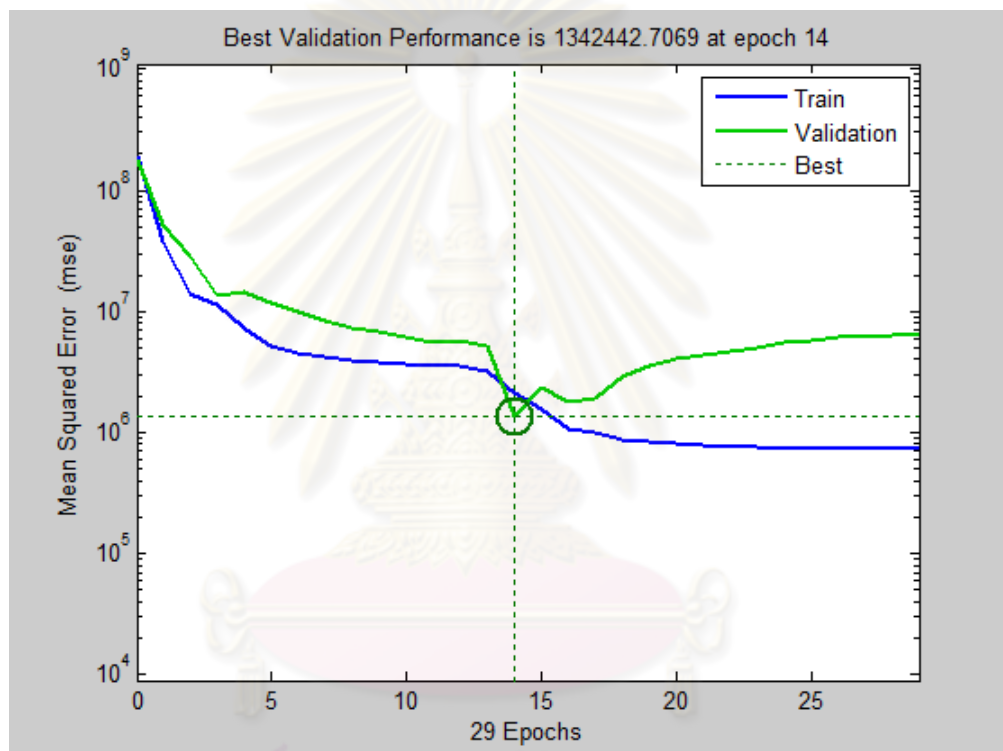


Figure 5.28: Performance curve of model 1 (Case 1-2-1)

จุฬาลงกรณ์มหาวิทยาลัย

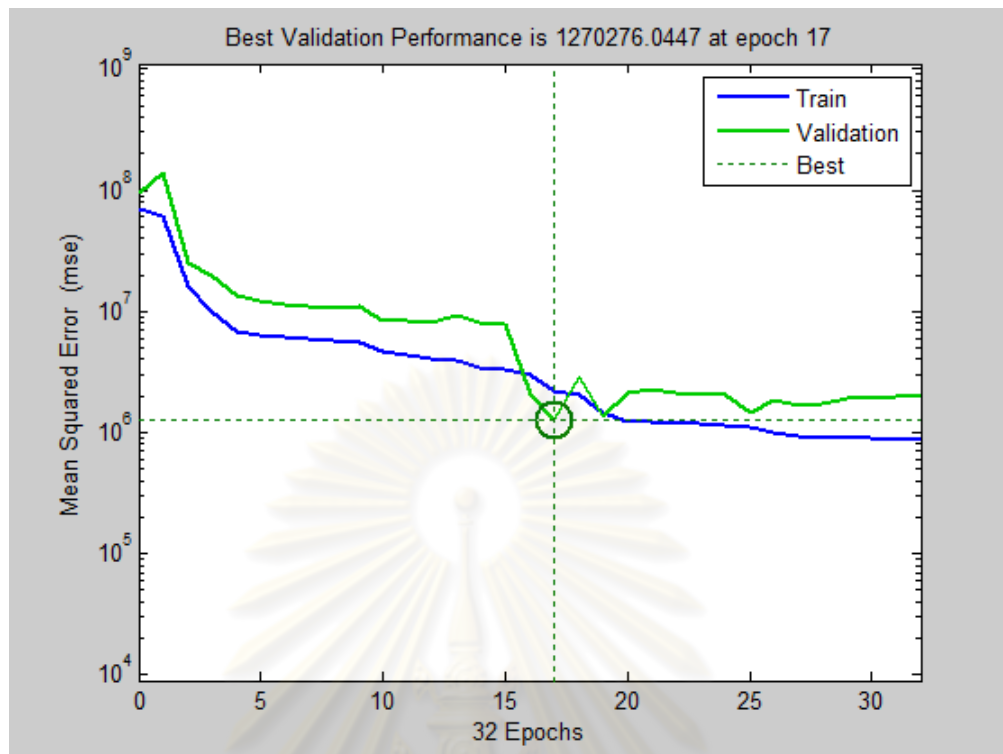


Figure 5.29: Performance curve of model 4 (Case 1-2-1)

Next, both models that produce the lowest MSE are further checked to ensure the accuracy of prediction. The outputs predicted by the ANN are compared with the target outputs of the training and validating sets by cross plotting them. Figures 5.30 and 5.31 represent the cross plots for the training and validating sets for model 1, respectively, while Figures 5.32 and 5.33 represent the cross plots for the training and validating sets for model 4, respectively. From the graph, the line $Y = X$ refers to correct prediction, i.e., each point on the 45-degree line is where predicted output is matched with the target output. So the closer the data points are located near the $Y = X$ line, the higher the accuracy of prediction. We can determine the accuracy of the prediction using regression coefficient of determination (R^2) as a criterion. R^2 equal to 1 represents a perfect fit to the $Y = X$ line. ANN will predict accurate output when R^2 is close to 1.

From Figures 5.30 to 5.33, we can see that the ANN can predict accurate output for both training and validating sets for both model 1 and 4. Model 1 gives R^2 of training and validating sets equal to 0.9154 and 0.9614, respectively, and model 4

gives R^2 of the training and validating sets equal to 0.9097 and 0.9628, respectively as well.

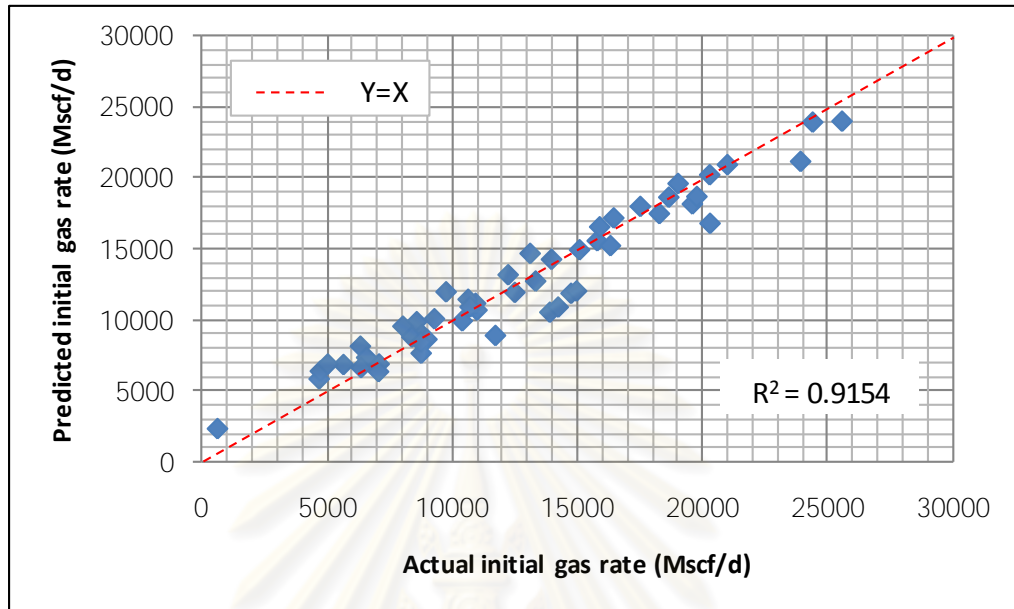


Figure 5.30: Cross plot of predicted vs actual initial gas rate of training sets of model 1 (Case 1-2-1)

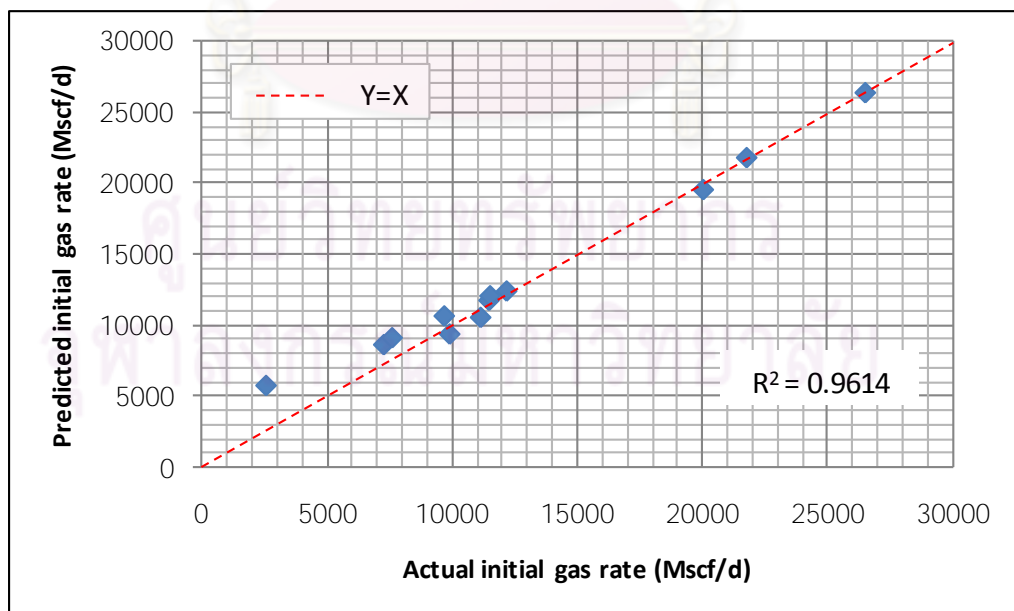


Figure 5.31: Cross plot of predicted vs actual initial gas rate of validating sets of model 1 (Case 1-2-1)

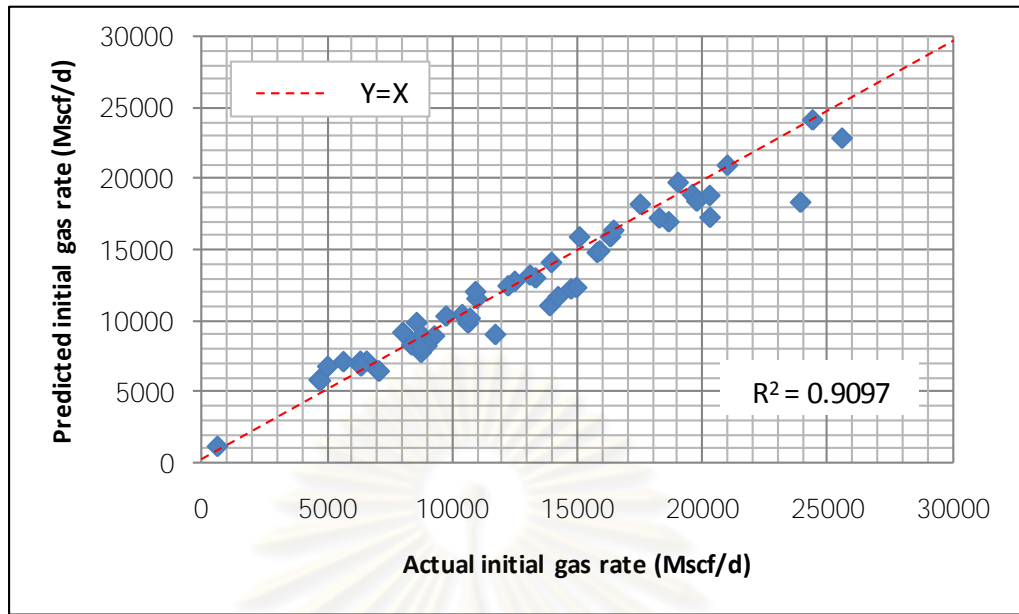


Figure 5.32: Cross plot of predicted vs actual initial gas rate of training sets of model 4 (Case 1-2-1)

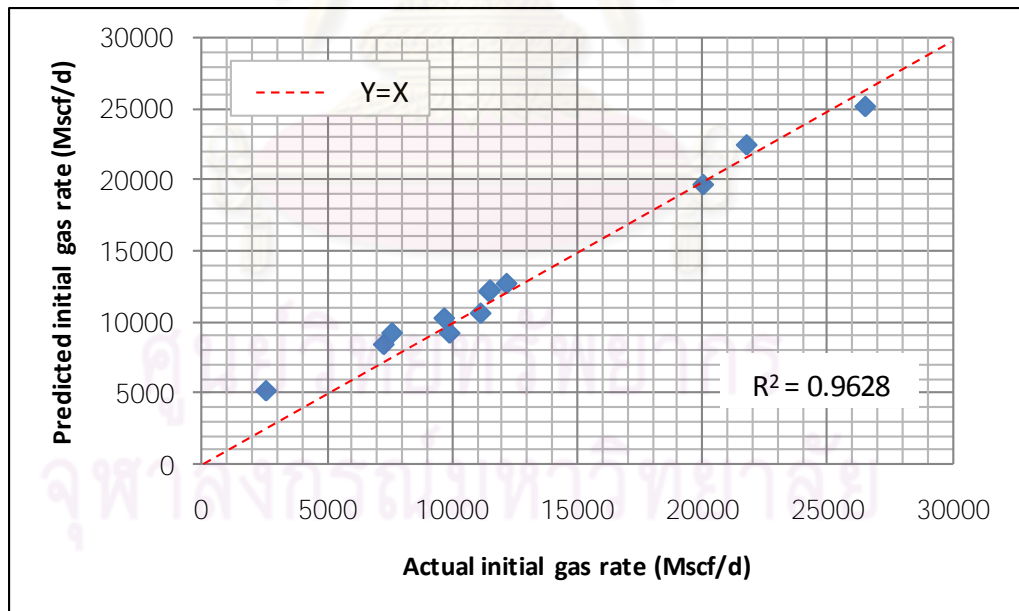


Figure 5.33: Cross plot of predicted vs actual initial gas rate of validating sets of model 4 (Case 1-2-1)

5.4.1.2.3 Model Testing Results and Discussion

In order to ensure the accuracy of ANN prediction when faced with unseen data sets, model 1 and 4 which yield the lowest MSE were then tested for accuracy by using testing data sets. After testing the ANN with testing sets, the outputs predicted by ANN are then compared with the target outputs by cross plotting them. Figures 5.34 and 5.35 represent the cross plots for model 1 and 4, respectively. From the graphs, R^2 of model 1 and 4 are equal to 0.8368 and 0.7374, respectively.

This means that both model 1 and 4 good performance of predicting the initial gas production rate. However, the coefficient of determination for model 1 is higher. Therefore, the best performance model which produces the most accurate predicted output is model 1. Consequently, this model will be used to predict the initial gas production rate for well number 101 which is drilled 1 year after the well number 100.

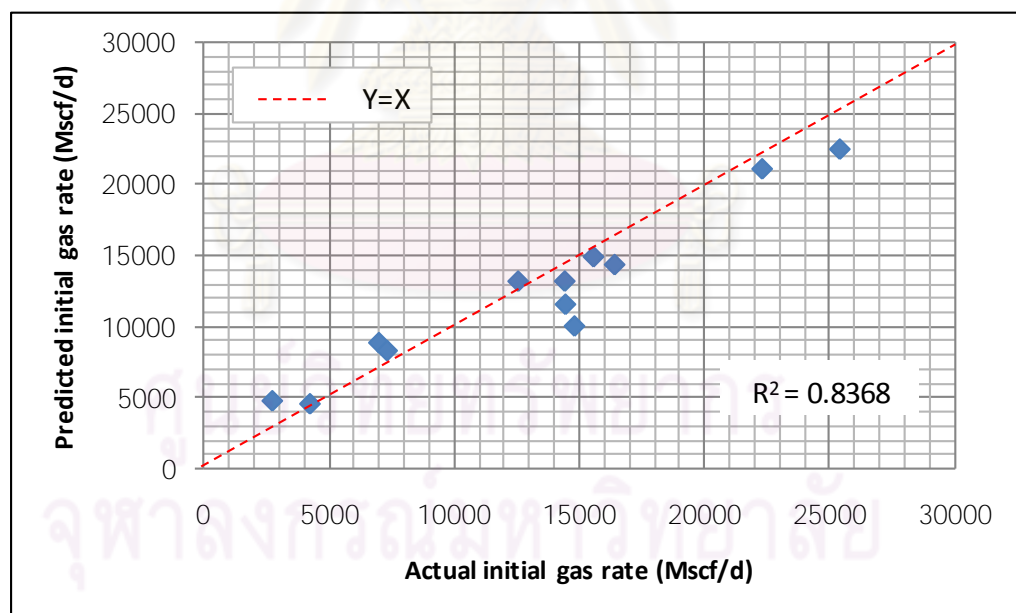


Figure 5.34: Cross plot of predicted vs actual initial gas rate of testing sets of model 1
(Case 1-2-1)

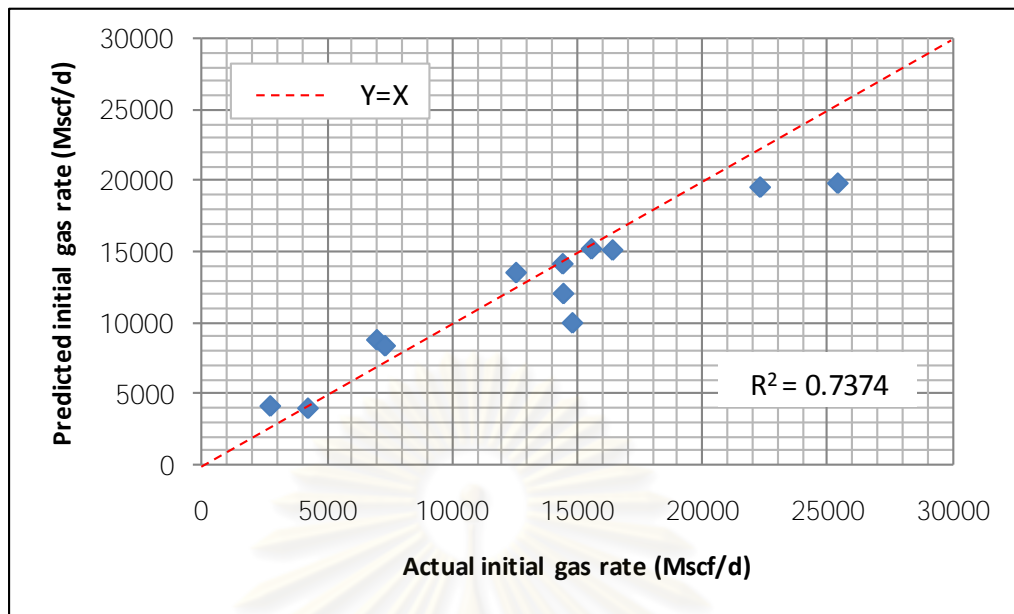


Figure 5.35: Cross plot of predicted vs actual initial gas rate of testing sets of model 4
(Case 1-2-1)

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

5.4.1.3 Case 1-2-2

This case study is the same as case 1-2-1 except that the average pressure used in this case is inverse-distance average pressure, i.e., pressure from the closer wells will affect the average value more than pressure from wells that are further away. Using this method to average the pressure should allow us to estimate the pressure around the predicting well location more accurately than the previous case study. The inverse-distance average pressure can be calculated via Equation 5.4. Figure 5.36 illustrates the schematic diagram of ANN in this case. The size of area of influence is the same as in the previous case which is 5,300 ft x 5,300 ft.

$$P_{average} = \sum_{i=1}^N \frac{P_i}{D_i} / \sum_{i=1}^N \frac{1}{D_i} \quad (5.4)$$

where	$P_{average}$	=	inverse-distance average pressure
	P_i	=	block pressure at well number i
	D_i	=	distance from well number i to predicting well
	N	=	number of surrounding wells

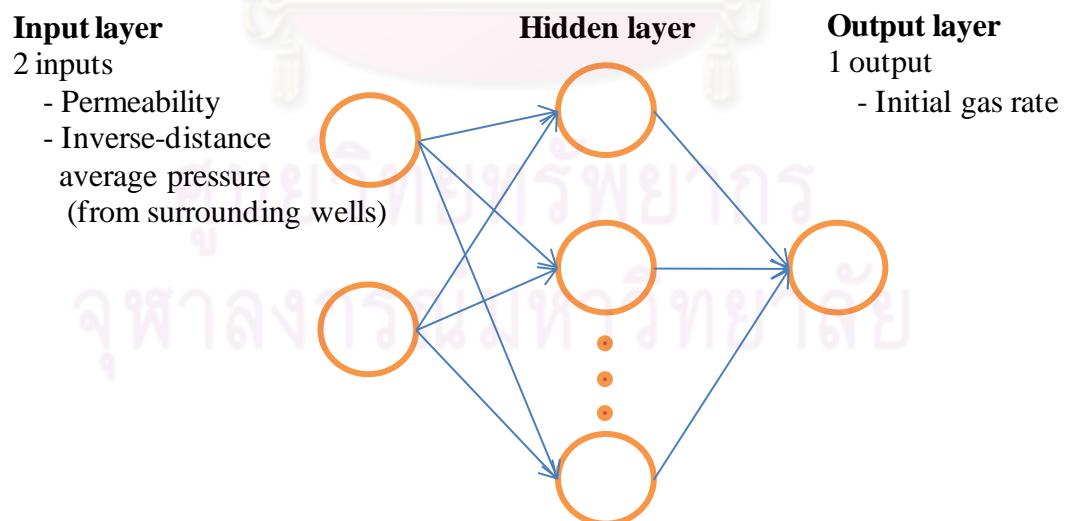


Figure 5.36: Schematic diagram of ANN for Case 1-2-2

5.4.1.3.1 Data Preprocessing

Similar to the previous case, a total of 75 data sets taken from well number 26 to 100 were divided into three main sets namely, training, validating, and testing sets with ratio of 4:1:1 (51:12:12 data sets). The wells in each data set are still the same as the ones in the previous case study. Input parameters from all data sets are plotted to observe the distribution. Since the distribution of permeabilities are the same as in the previous case study, only the distributions of inverse-distance average pressures are then plotted. The histograms are shown in Figures 5.37 to 5.39 and their statistics are summarized in Table 5.6. From the result, we found that the histograms and statistics of the 3 data sets represent similar distributions.

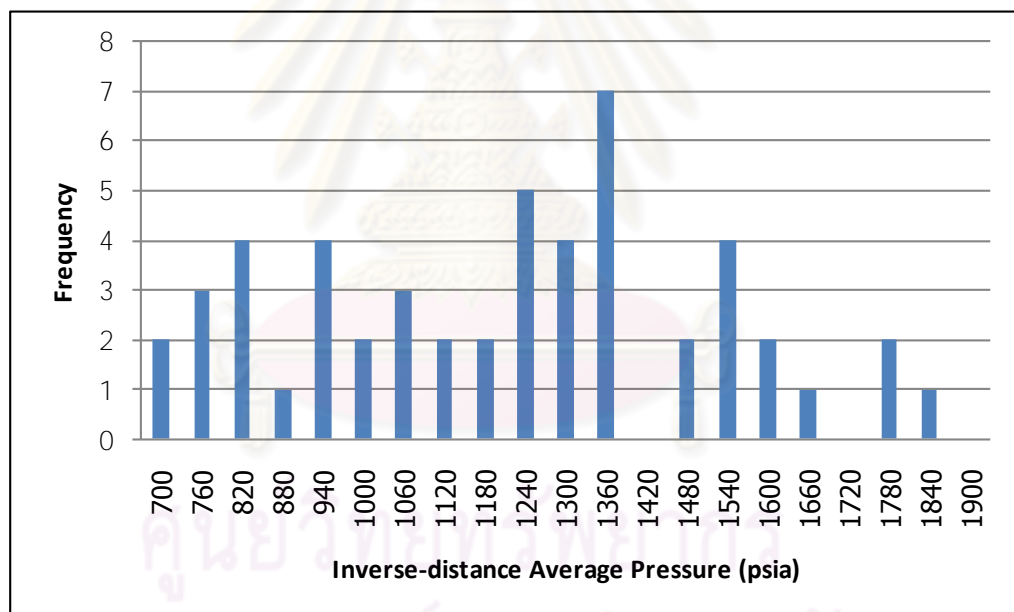


Figure 5.37: Histogram of inverse-distance average pressure for training sets

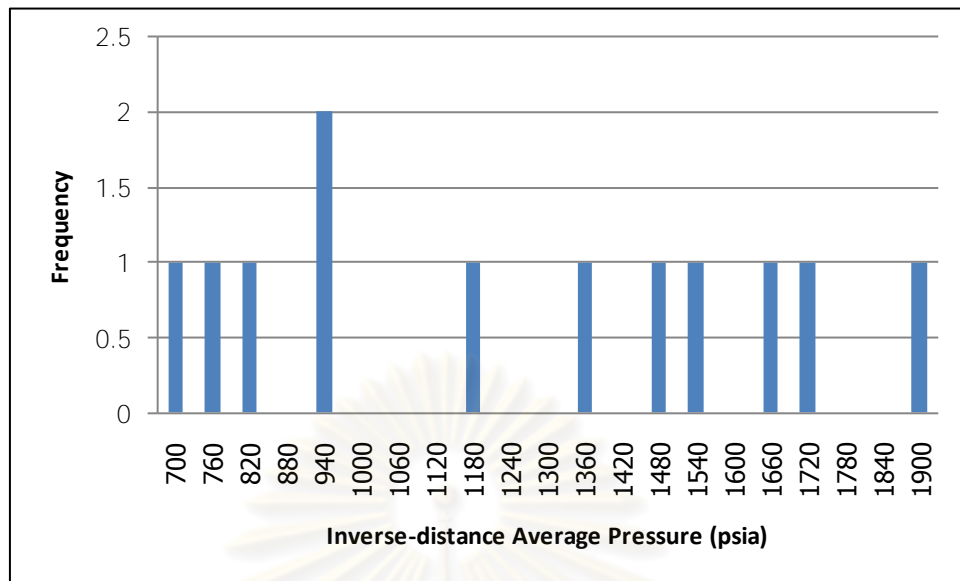


Figure 5.38: Histogram of inverse-distance average pressure for validating sets

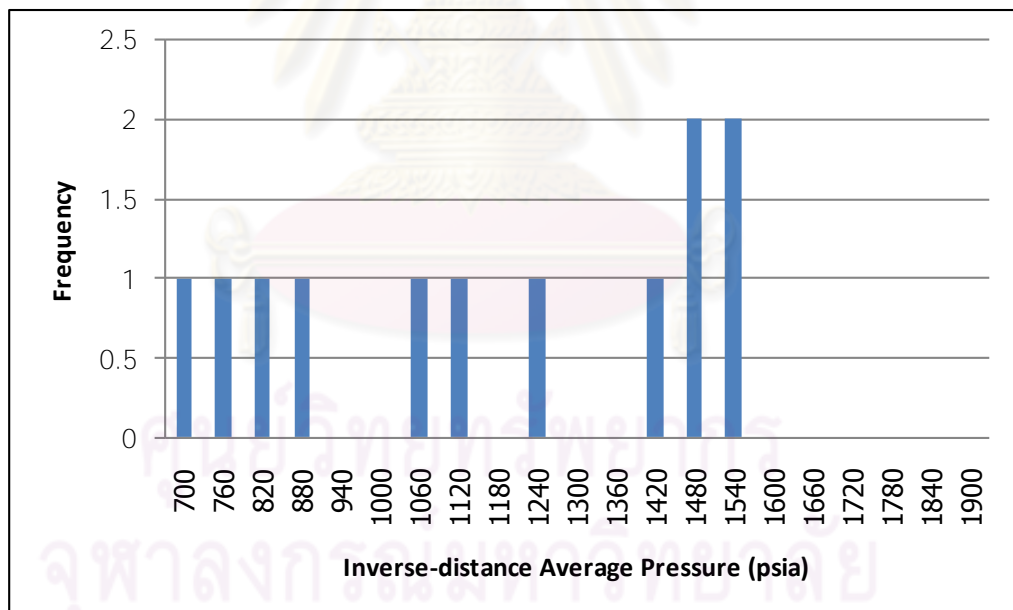


Figure 5.39: Histogram of inverse-distance average pressure for testing sets

Table 5.6: Summary of statistics of all data sets (Case 1-2-2)

Statistical parameters	Dataset type	Inverse-distance average pressure (psia)
Maximum	Training	1,798.89
	Validating	1,857.74
	Testing	1,533.35
Minimum	Training	690.99
	Validating	663.15
	Testing	669.20
Mean	Training	1,181.31
	Validating	1,219.51
	Testing	1,148.48
1st Quartile	Training	931.72
	Validating	879.81
	Testing	859.91
Median	Training	1,221.89
	Validating	1,247.85
	Testing	1,155.25
3rd Quartile	Training	1,338.24
	Validating	1,527.48
	Testing	1,433.24
SD	Training	301.00
	Validating	412.13
	Testing	319.22

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

5.4.1.3.2 Model Training

The ANN model was trained with various network configurations based on a trial and error basis. There are 20 models that were run in this case. Each model was trained many times to obtain the lowest MSE possible. The network configurations and their MSE of validating set are summarized in Table 5.7.

Table 5.7: Model configuration for Case 1-2-2

Model No	Number of neurons		Learning rate	Momentum	MSE
	Hidden Layer 1	Hidden Layer 2			
1	5	0	0.1	0.1	1,241,372
2	5	0	0.5	0.1	1,280,630
3	5	0	0.1	0.5	1,303,887
4	5	0	0.5	0.5	1,270,706
5	10	0	0.1	0.1	1,548,582
6	10	0	0.5	0.1	2,239,006
7	10	0	0.1	0.5	1,944,546
8	10	0	0.5	0.5	1,848,544
9	20	0	0.1	0.1	3,469,081
10	20	0	0.5	0.1	5,672,744
11	20	0	0.1	0.5	2,950,202
12	20	0	0.5	0.5	3,808,074
13	5	5	0.1	0.1	857,552
14	5	5	0.5	0.1	561,859
15	5	5	0.1	0.5	1,438,182
16	5	5	0.5	0.5	1,565,883
17	10	10	0.1	0.1	2,282,179
18	10	10	0.5	0.1	2,104,654
19	10	10	0.1	0.5	1,917,763
20	10	10	0.5	0.5	1,888,612

From a total 20 model configurations, two models with the lowest and next to lowest MSE (model 13 and 14) which has the MSE of 857,552 and 561,859, respectively, are chosen. Model 13 consists of 2 hidden layers with 5 neurons in each layer. The learning rate and momentum of 0.1 were used. Similar to model 13, model 14 consists of 2 hidden layers with 5 neurons in each layer. However, the learning rate and momentum were set to be 0.5 and 0.1, respectively. The performance curves of model 13 and 14 are shown in Figures 5.40 and 5.41, respectively. Model 13 was

trained until epoch 33 but the weight and bias were updated until epoch 18 only because the MSE of validating set started to increase in this epoch. The training was continued for 15 more epochs for validation check. The lowest MSE of model 13 is 857,552. For model 14, the training was performed until epoch 24. However, the weight and bias were not updated after epoch 9 due to the same reason for model 13. The lowest MSE of model 14 is 561,859 in epoch 9.

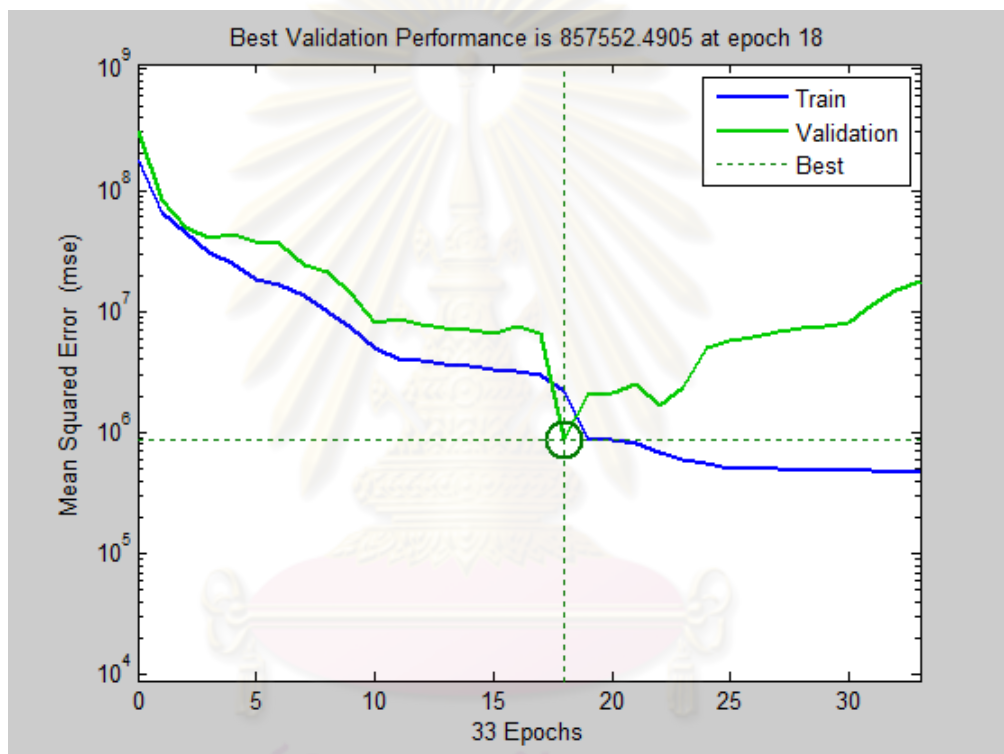


Figure 5.40: Performance curve of model 13 (Case 1-2-2)

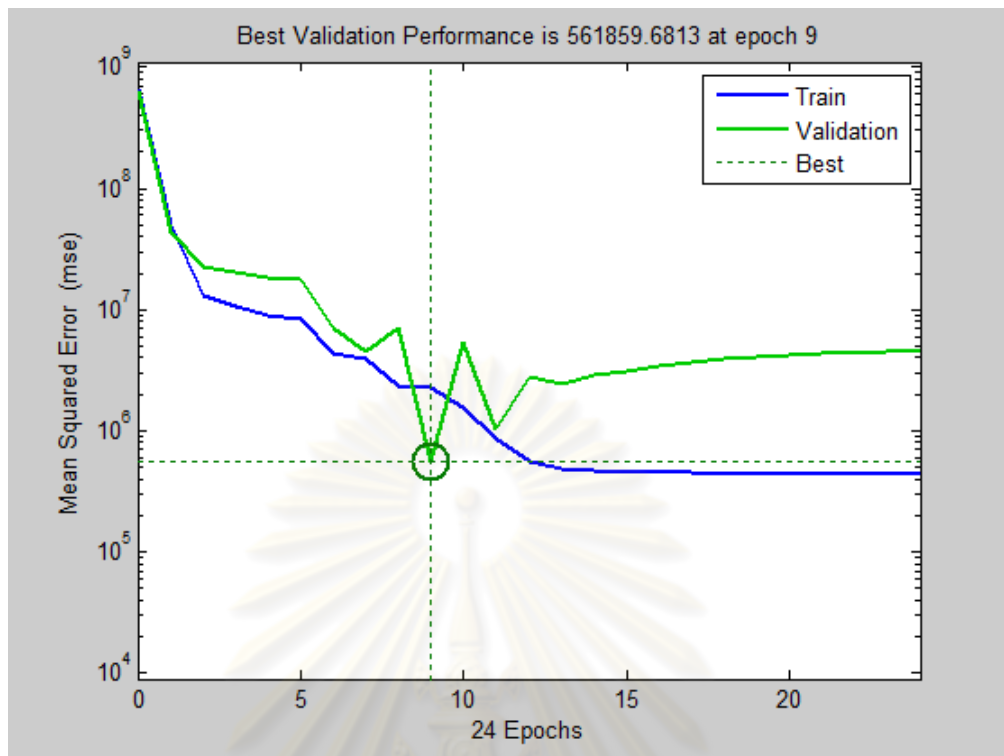


Figure 5.41: Performance curve of model 14 (Case 1-2-2)

Next, both models that produce the lowest MSE are further checked to ensure the accuracy of prediction. The outputs predicted by the ANN are compared with the target outputs of the training and validating sets by cross plotting them. Figures 5.42 and 5.43 represent the cross plots for the training and validating sets for model 13, respectively, while Figures 5.44 and 5.45 represent the cross plots for the training and validating sets for model 14, respectively. From the graph, the line $Y = X$ refers to correct prediction, i.e., each point on the 45-degree line is where predicted output is matched with the target output. So the closer the data points are located near the $Y = X$ line, the higher the accuracy of prediction. We can determine the accuracy of the prediction using regression coefficient of determination (R^2) as a criterion. R^2 equal to 1 represents a perfect fit to the $Y = X$ line. ANN will predict accurate output when R^2 is close to 1.

From Figures 5.42 to 5.45, we can see that the ANN can predict accurate output for both training and validating sets for both model 13 and 14. Model 13 gives R^2 of training and validating sets equal to 0.9379 and 0.9776, respectively, and model

14 gives R^2 of the training and validating sets equal to 0.9365 and 0.9852, respectively as well.

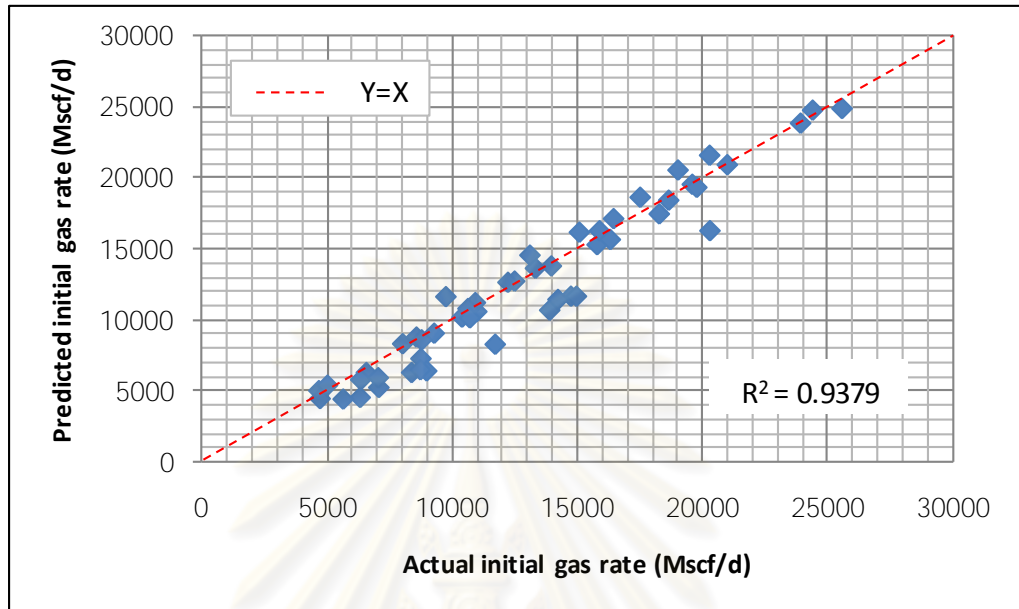


Figure 5.42: Cross plot of predicted vs actual initial gas rates of training sets of model 13 (Case 1-2-2)

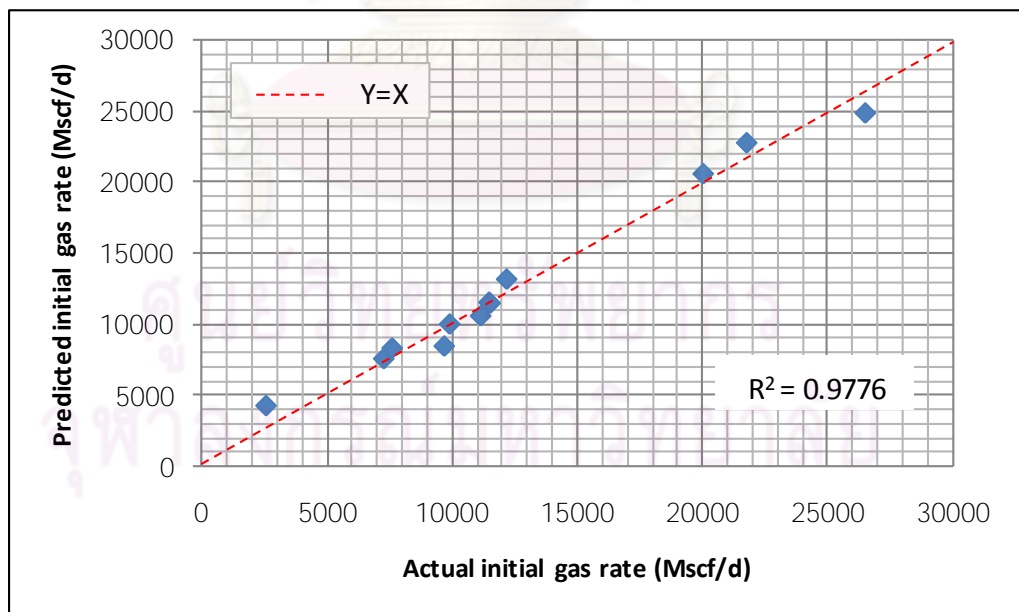


Figure 5.43: Cross plot of predicted vs actual initial gas rates of validating sets of model 13 (Case 1-2-2)

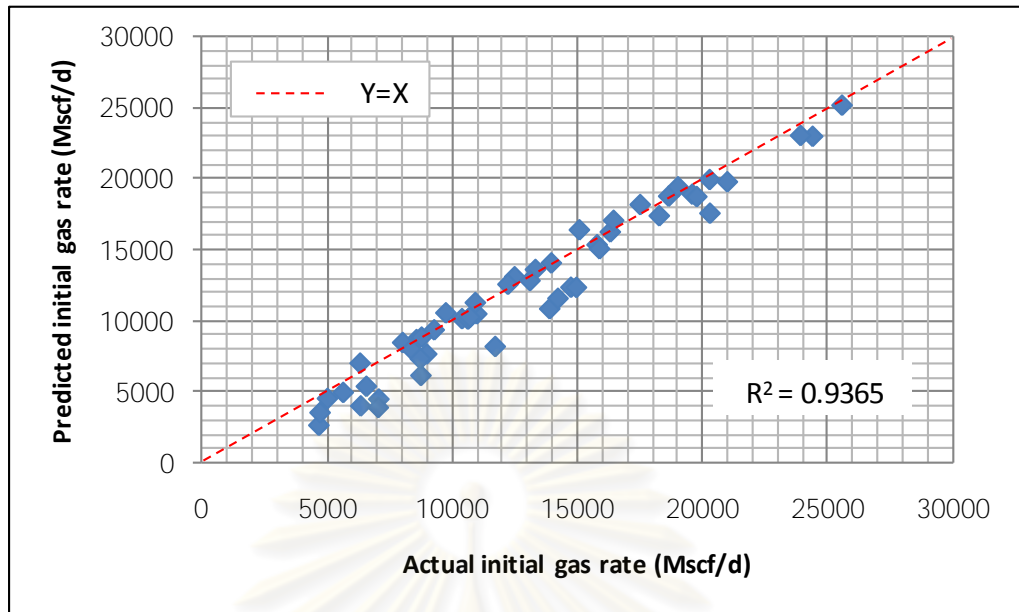


Figure 5.44: Cross plot of predicted vs actual initial gas rates of training sets of model 14 (Case 1-2-2)

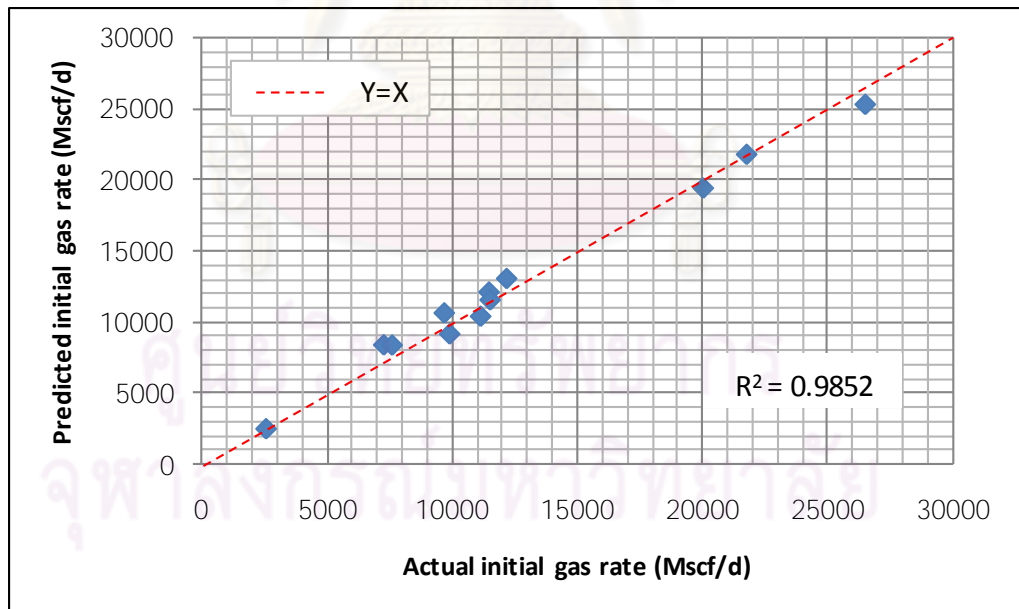


Figure 5.45: Cross plot of predicted vs actual initial gas rates of validating sets of model 14 (Case 1-2-2)

5.4.1.3.3 Model Testing Results and Discussion

In order to ensure the accuracy of ANN prediction when faced with unseen data sets, model 13 and 14 which yield the lowest MSE were then tested for accuracy by using testing data sets. After testing the ANN with testing sets, the outputs predicted by ANN are then compared with the target outputs by cross plotting them. Figures 5.46 and 5.47 represent the cross plots for model 13 and 14, respectively. From the graphs, R^2 of model 13 and 14 are equal to 0.8921 and 0.9078, respectively.

This means that both model 13 and 14 good performance of predicting the initial gas production rate. However, the coefficient of determination for model 14 is higher. Therefore, the best performance model which produces the most accurate predicted output is model 14. Consequently, this model will be used to predict the initial gas production rate for well number 101 which is drilled 1 year after well number 100.

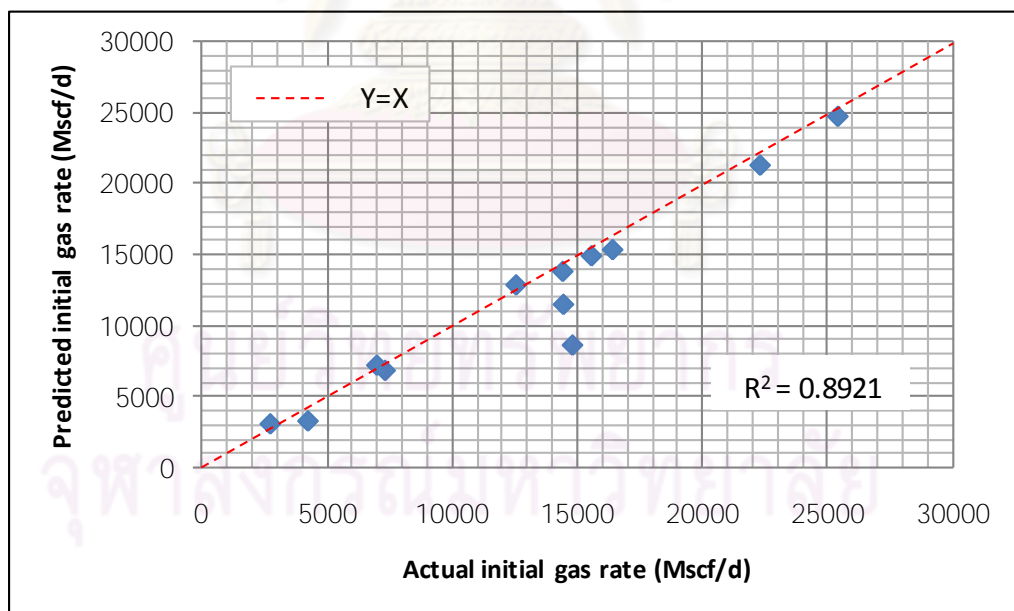


Figure 5.46: Cross plot of predicted vs actual initial gas rates of testing sets of model 13 (Case 1-2-2)

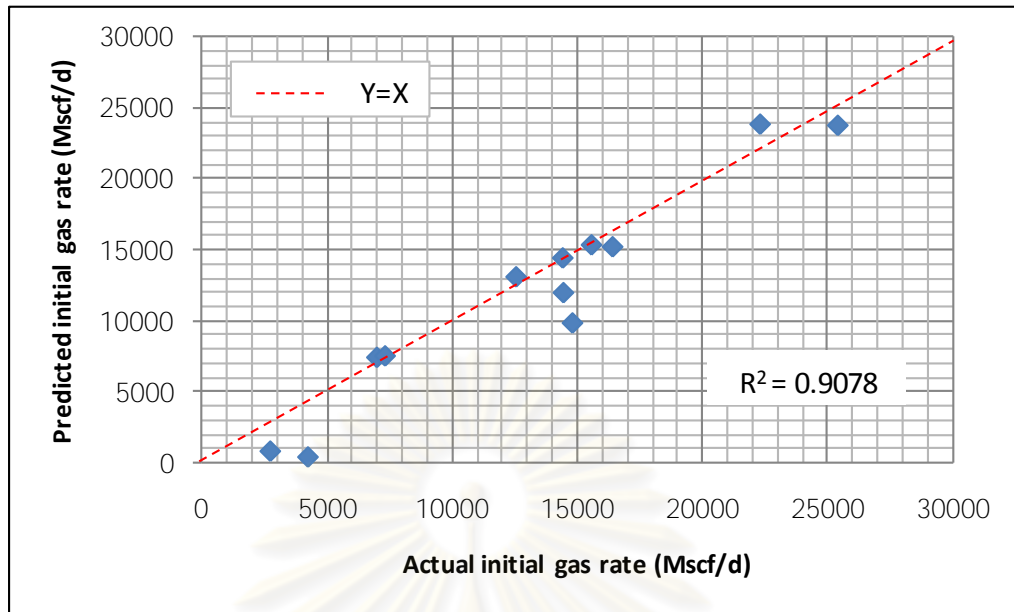


Figure 5.47: Cross plot of predicted vs actual initial gas rates of testing sets of model 14 (Case 1-2-2)

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

5.4.1.4 Case 1-3

In this case study, we use parameters which can refer to pressure at the predicting well location instead. These parameters are permeability, porosities around the well which are obtained by using Geostatistics, drill date, and the number of surrounding wells. The porosities around the well are divided into 3 rings. The first ring covers an area of 100 x 100 ft at the center. The boundary of the second ring is located at 500 ft from the center in the x-and y-directions while the boundary of the third ring is 700 ft away from the center. Figure 5.48 illustrates the schematic diagram of ANN in this case.

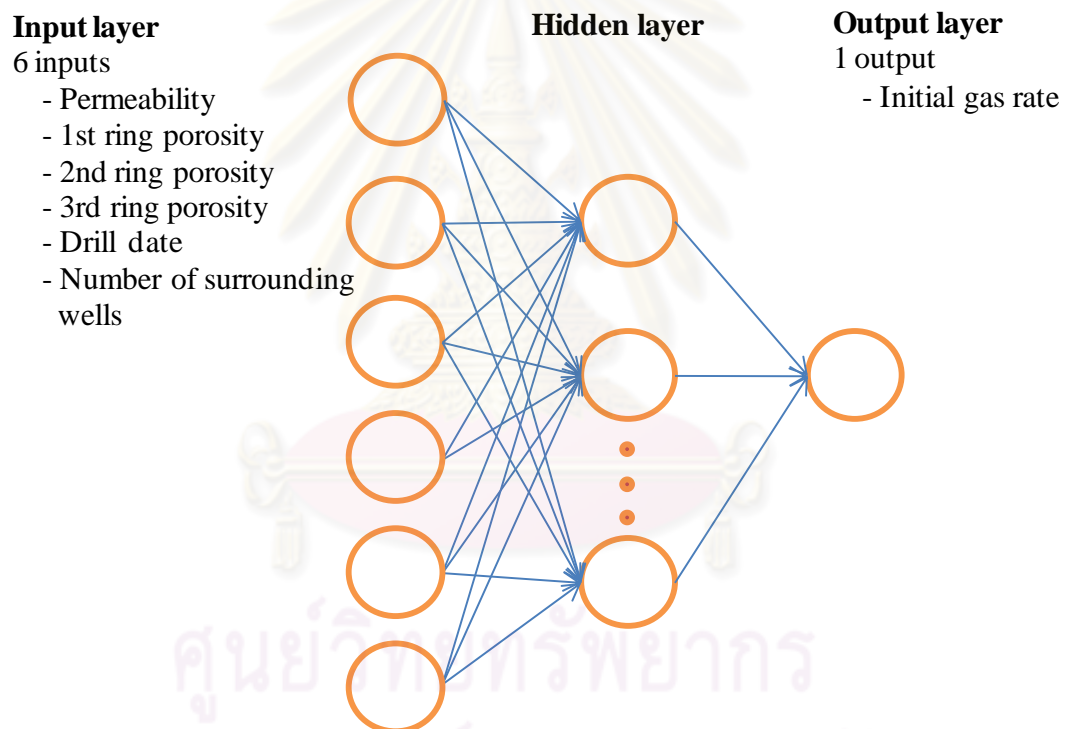


Figure 5.48: Schematic diagram of ANN for Case 1-3

5.4.1.4.1 Data Preprocessing

Similar to the previous case, a total of 75 data sets taken from well number 26 to 100 were divided into three main sets namely, training, validating, and testing sets with ratio of 4:1:1 (51:12:12 data sets). The wells in each data set are still the same as the ones in the previous case study. Input parameters from all data sets are plotted to observe the distribution. Since the distribution of permeabilities are the same as in the

previous case study, only the distributions of average porosity for each ring, drill date, and number of surrounding wells are then plotted. The histograms are shown in Figures 5.49 to 5.63 and their statistics are summarized in Table 5.8. From the result, we found that the histograms and statistics of the 3 data sets represent similar distributions.

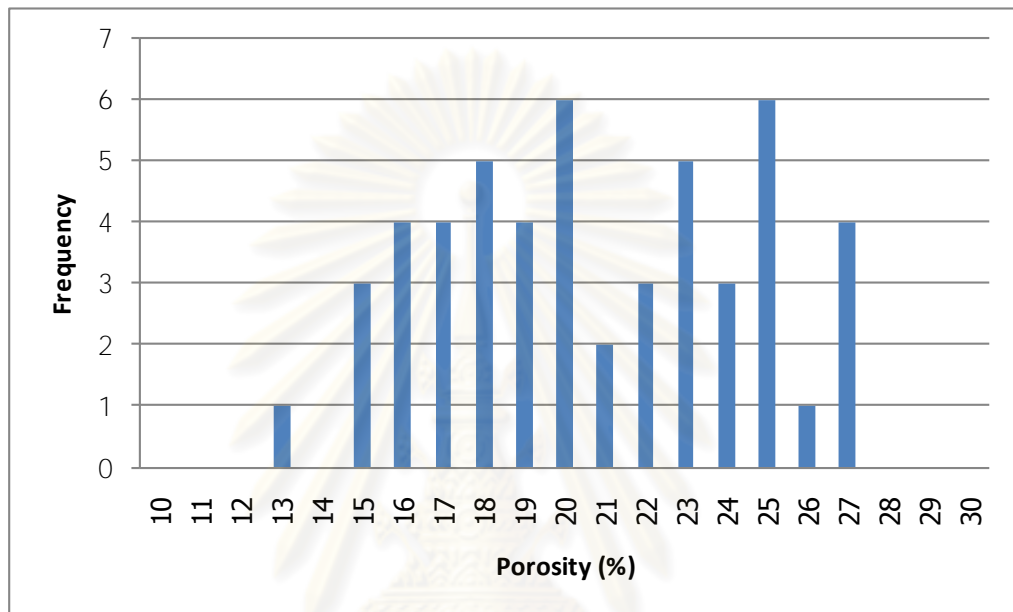


Figure 5.49: Histogram of 1st ring porosity for training sets

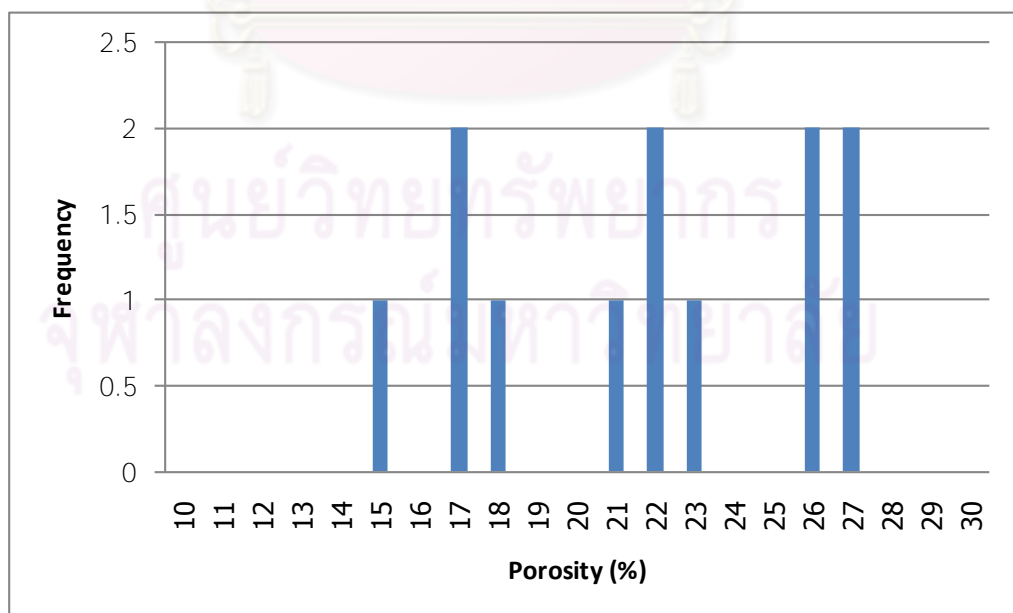


Figure 5.50: Histogram of 1st ring porosity for validating sets

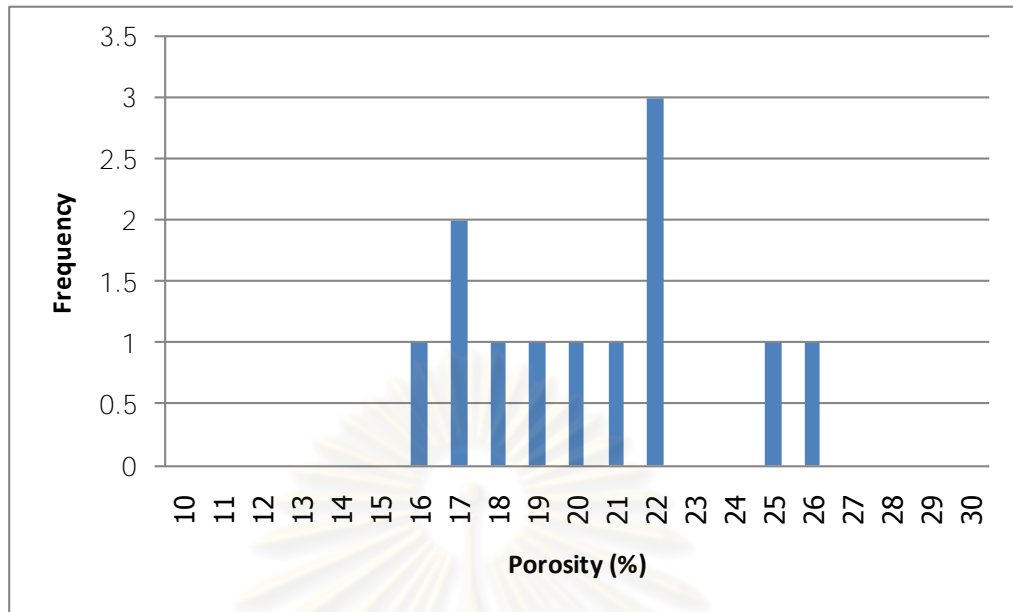


Figure 5.51: Histogram of 1st ring porosity for testing sets

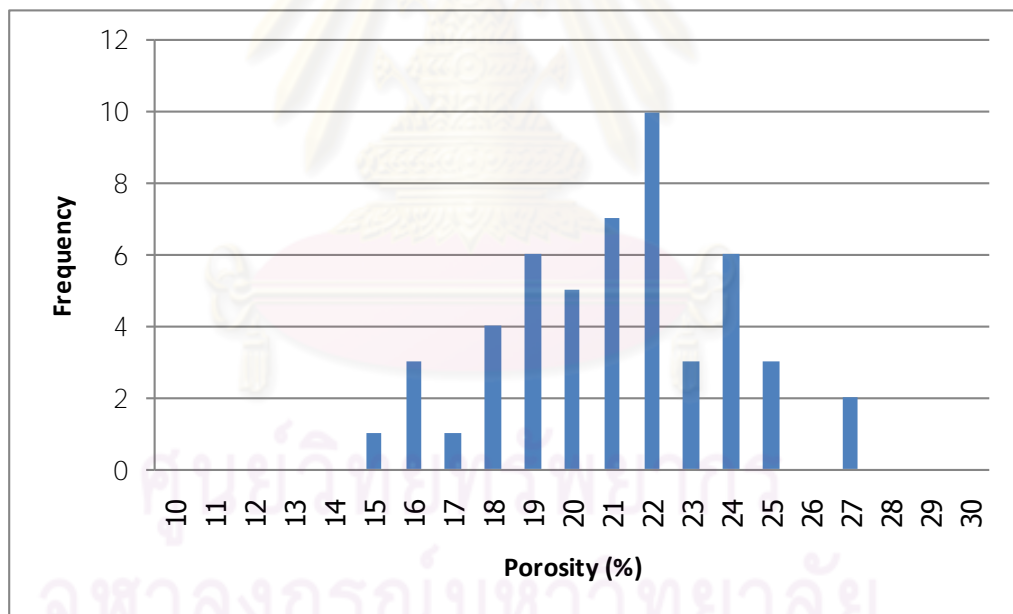


Figure 5.52: Histogram of 2nd ring porosity for training sets

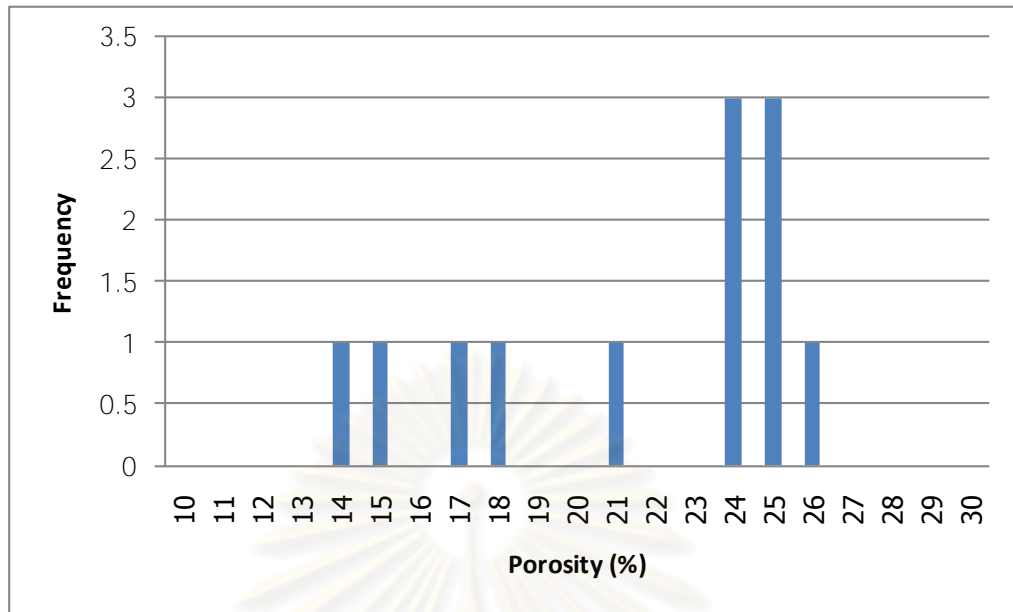


Figure 5.53: Histogram of 2nd ring porosity for validating sets

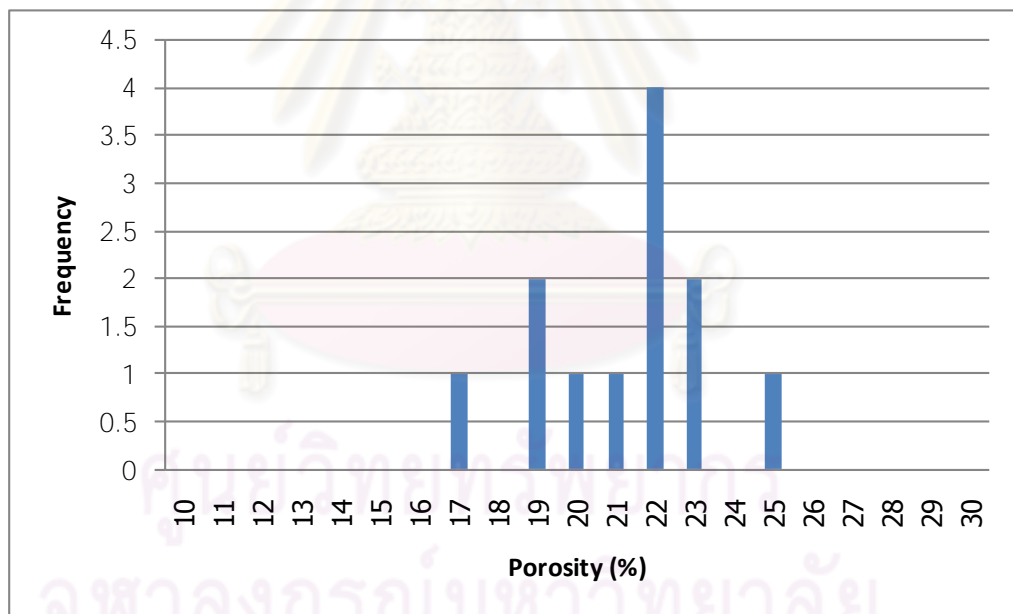


Figure 5.54: Histogram of 2nd ring porosity for testing sets

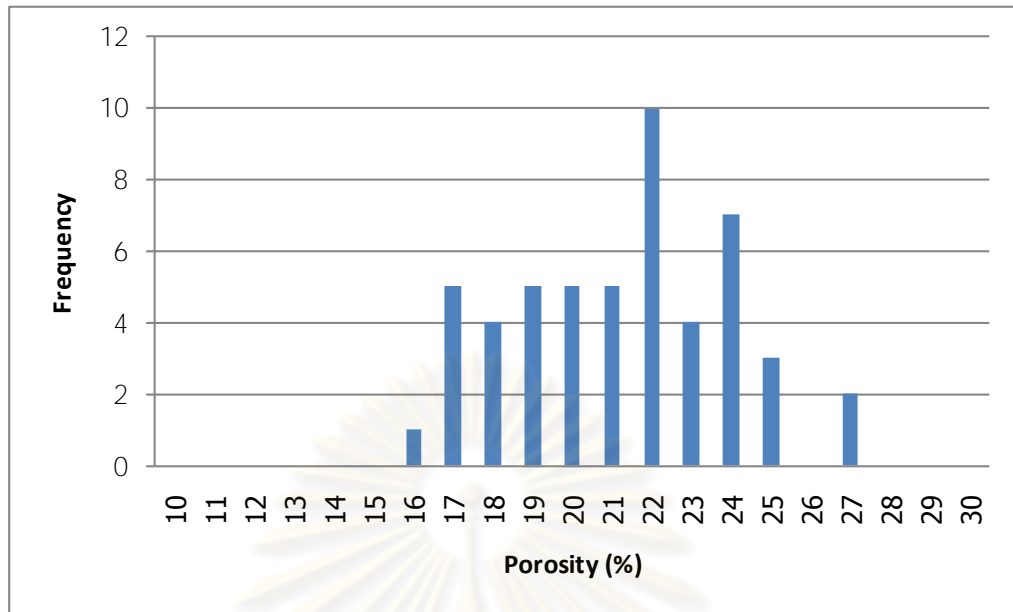


Figure 5.55: Histogram of 3rd ring porosity for training sets

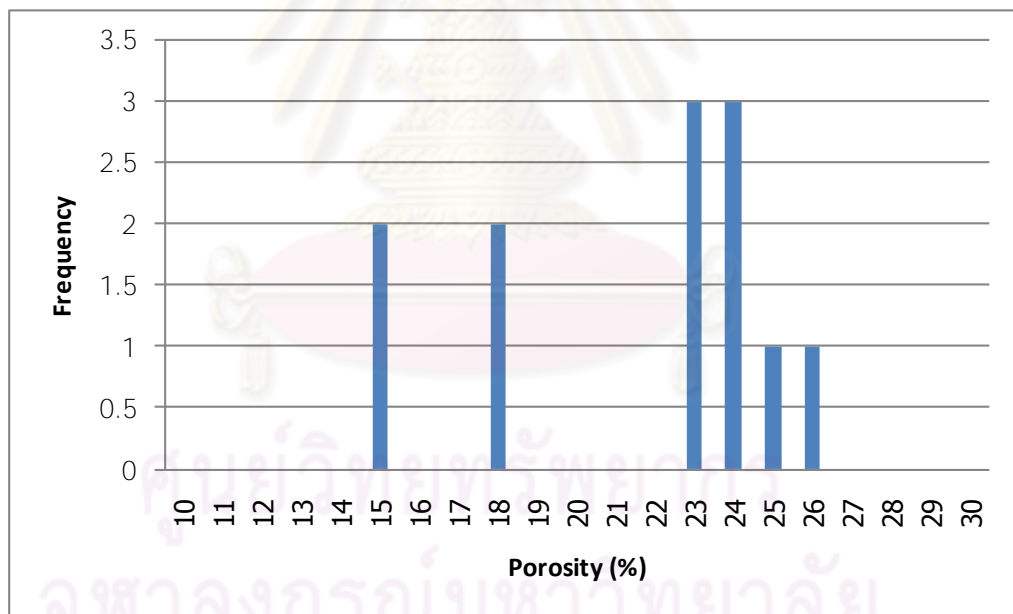


Figure 5.56: Histogram of 3rd ring porosity for validating sets

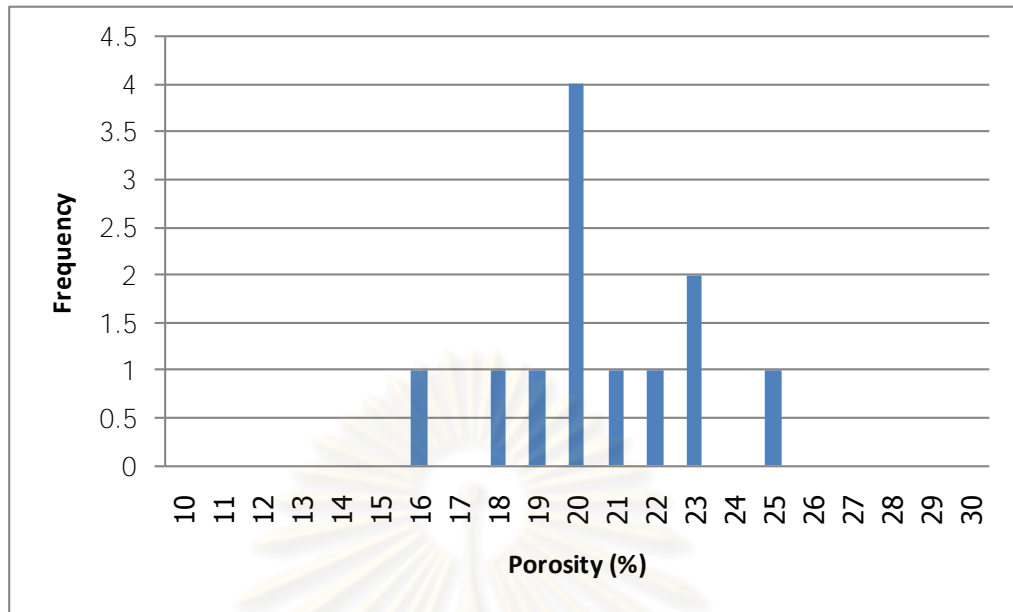


Figure 5.57: Histogram of 3rd ring porosity for testing sets

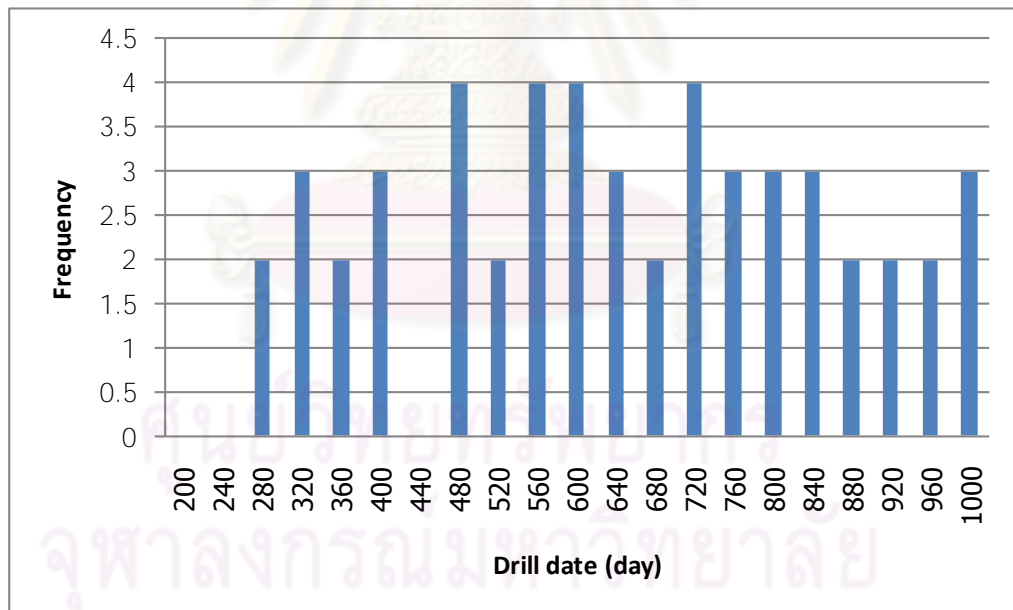


Figure 5.58: Histogram of drill date for training sets

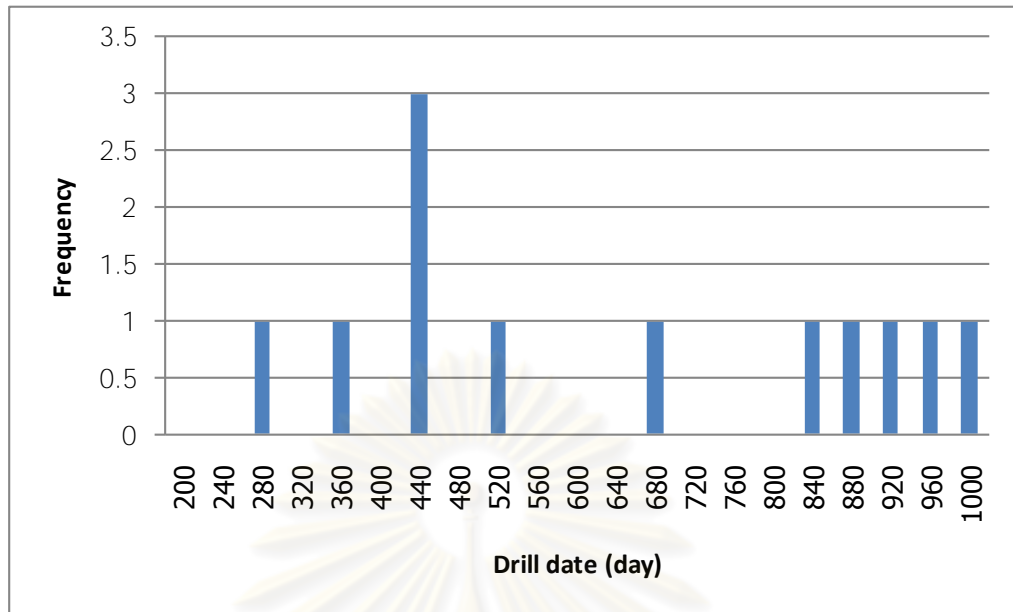


Figure 5.59: Histogram of drill date for validating sets

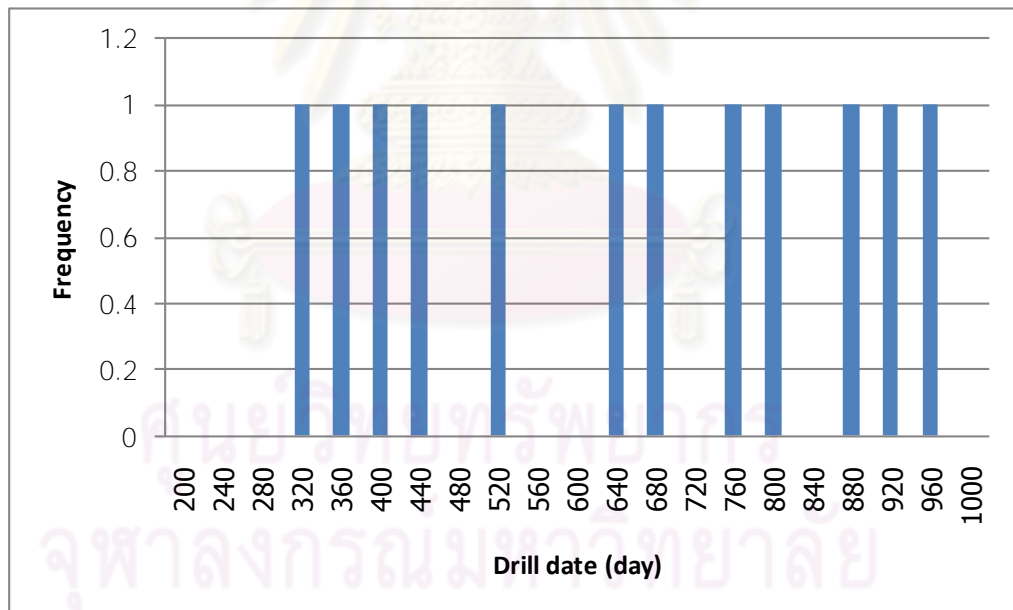


Figure 5.60: Histogram of drill date for testing sets

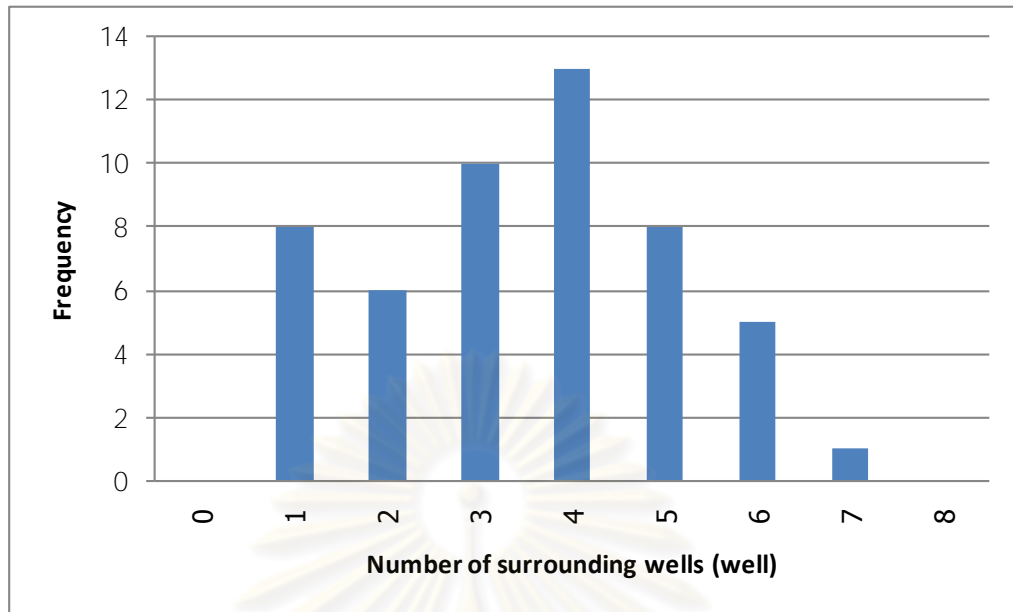


Figure 5.61: Histogram of number of surrounding wells for training sets

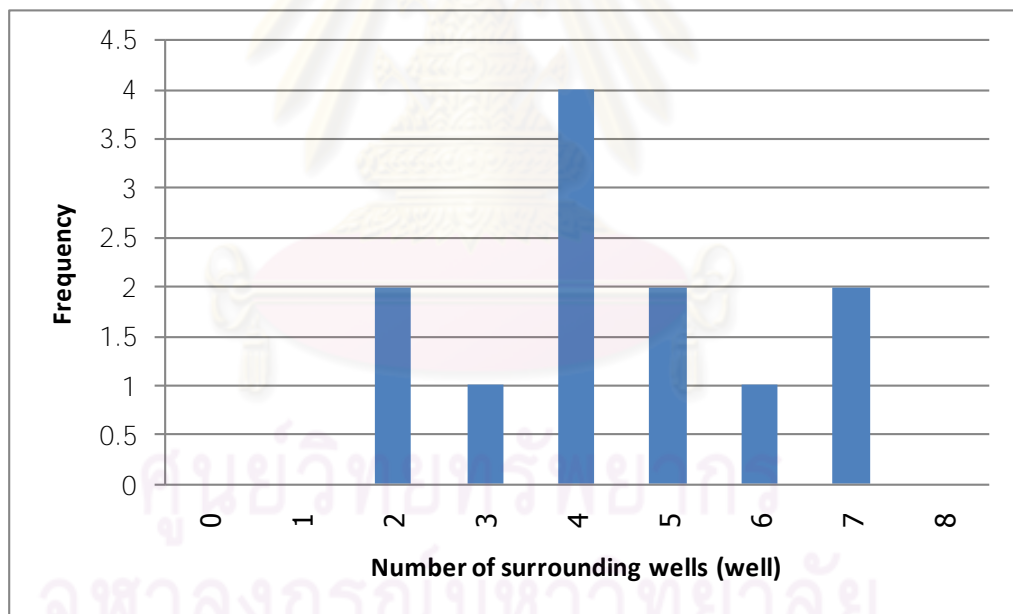


Figure 5.62: Histogram of number of surrounding wells of validating sets

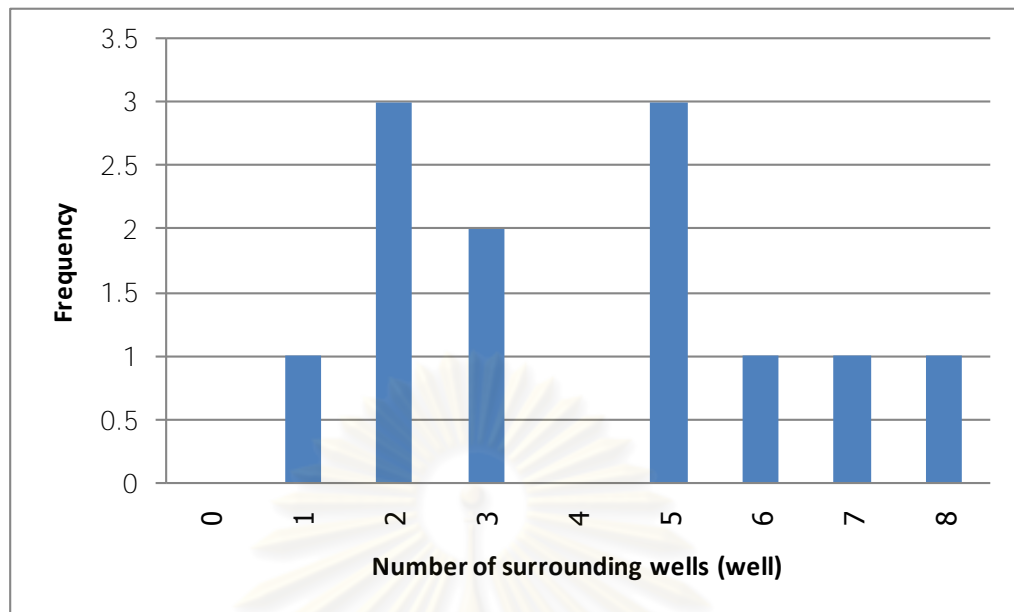


Figure 5.63: Histogram of number of surrounding wells for testing sets

Table 5.8: Summary of statistics of all data sets (Case 1-3)

Statistical parameters	Dataset type	1 st ring porosity (%)	2 nd ring porosity (%)	3 rd ring porosity (%)	Drill date (day)	Number of surrounding wells
Maximum	Training	27.00	27.00	26.42	990.00	7
	Validating	27.00	25.88	25.38	1,000.00	7
	Testing	26.00	24.75	24.25	960.00	8
Minimum	Training	13.00	14.54	15.75	260.00	1
	Validating	14.00	13.33	14.58	280.00	2
	Testing	16.00	16.21	15.75	300.00	1
Mean	Training	20.67	20.59	20.70	629.80	3.51
	Validating	21.67	21.06	20.98	634.17	4.42
	Testing	20.42	20.65	20.23	626.67	4.08
1st Quartile	Training	18.00	18.75	18.46	475.00	2
	Validating	17.75	17.19	17.40	417.50	3.75
	Testing	17.75	19.45	19.07	422.50	2
Median	Training	20.00	20.88	21.08	620.00	4
	Validating	22.00	23.38	22.69	600.00	4
	Testing	20.50	21.06	19.75	655.00	4
3rd Quartile	Training	24.00	22.48	22.40	795.00	5
	Validating	26.00	24.25	23.34	870.00	5.25
	Testing	22.00	21.58	22.00	790.00	5.25
SD	Training	3.81	2.83	2.70	209.72	1.62
	Validating	4.40	4.36	3.87	260.16	1.68
	Testing	3.18	2.23	2.41	227.85	2.23

5.4.1.4.2 Model Training

The ANN model was trained with various network configurations based on a trial and error basis. There are 20 models that were run in this case. Each model was trained many times to obtain the lowest MSE possible. The network configurations and their MSE of validating set are summarized in Table 5.9.

Table 5.9: Model configuration for Case 1-3

Model No	Number of neurons		Learning rate	Momentum	MSE
	Hidden Layer 1	Hidden Layer 2			
1	5	0	0.1	0.1	386,566
2	5	0	0.5	0.1	494,701
3	5	0	0.1	0.5	319,218
4	5	0	0.5	0.5	315,566
5	10	0	0.1	0.1	589,673
6	10	0	0.5	0.1	673,924
7	10	0	0.1	0.5	453,403
8	10	0	0.5	0.5	634,285
9	20	0	0.1	0.1	867,483
10	20	0	0.5	0.1	1,289,344
11	20	0	0.1	0.5	948,323
12	20	0	0.5	0.5	1,047,584
13	5	5	0.1	0.1	759244
14	5	5	0.5	0.1	820502
15	5	5	0.1	0.5	430542
16	5	5	0.5	0.5	600533
17	10	10	0.1	0.1	1,232,456
18	10	10	0.5	0.1	984,637
19	10	10	0.1	0.5	948,335
20	10	10	0.5	0.5	1,023,572

From a total 20 model configurations, two models with the lowest and next to lowest MSE (model 3 and 4) which has the MSE of 319,218 and 315,566, respectively, are chosen. Model 3 consists of only one hidden layers with 5 neurons. The learning rate and momentum of 0.1 and 0.5 were used. Similar to model 3, model 4 consists of only one hidden layers with 5 neurons. However, the learning rate and momentum were set to be 0.5. The performance curves of model 3 and 4 are shown in Figures 5.64 and 5.65, respectively. Model 3 was trained until epoch 21 but the

weight and bias were updated until epoch 6 only because the MSE of validating set started to increase in this epoch. The training was continued for 15 more epochs for validation check. The lowest MSE of model 3 is 319,218. For model 4, the training was performed until epoch 21. However, the weight and bias were not updated after epoch 6 due to the same reason for model 3. The lowest MSE of model 4 is 315,566 in epoch 6.

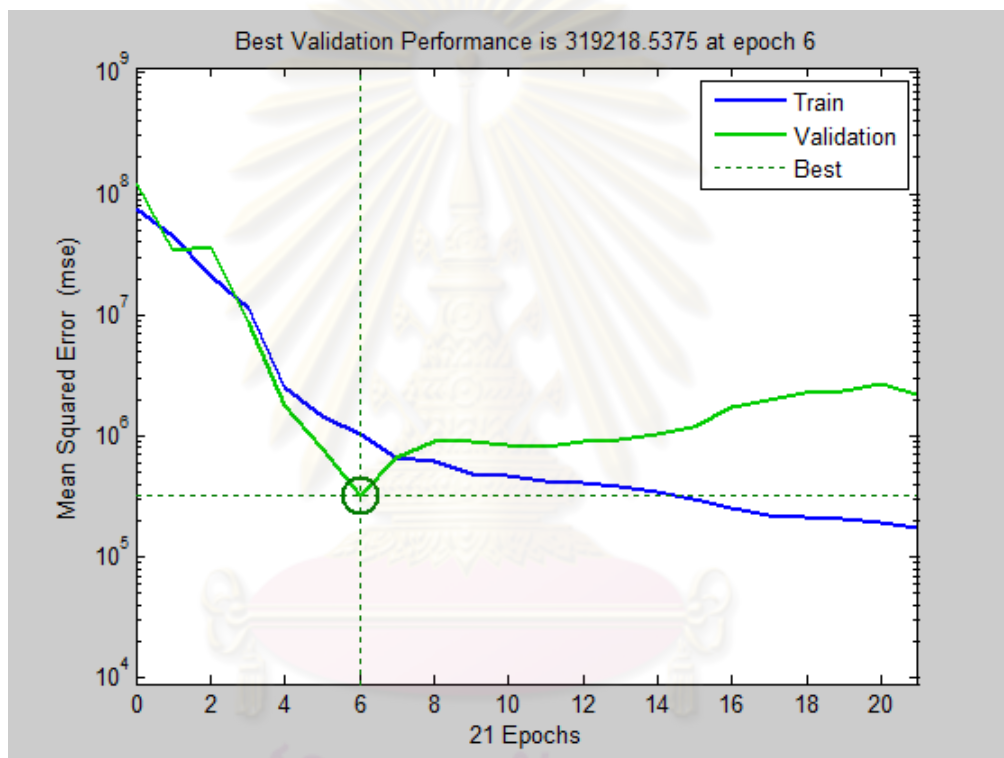


Figure 5.64: Performance curve of model 3 (Case 1-3)

จุฬาลงกรณ์มหาวิทยาลัย

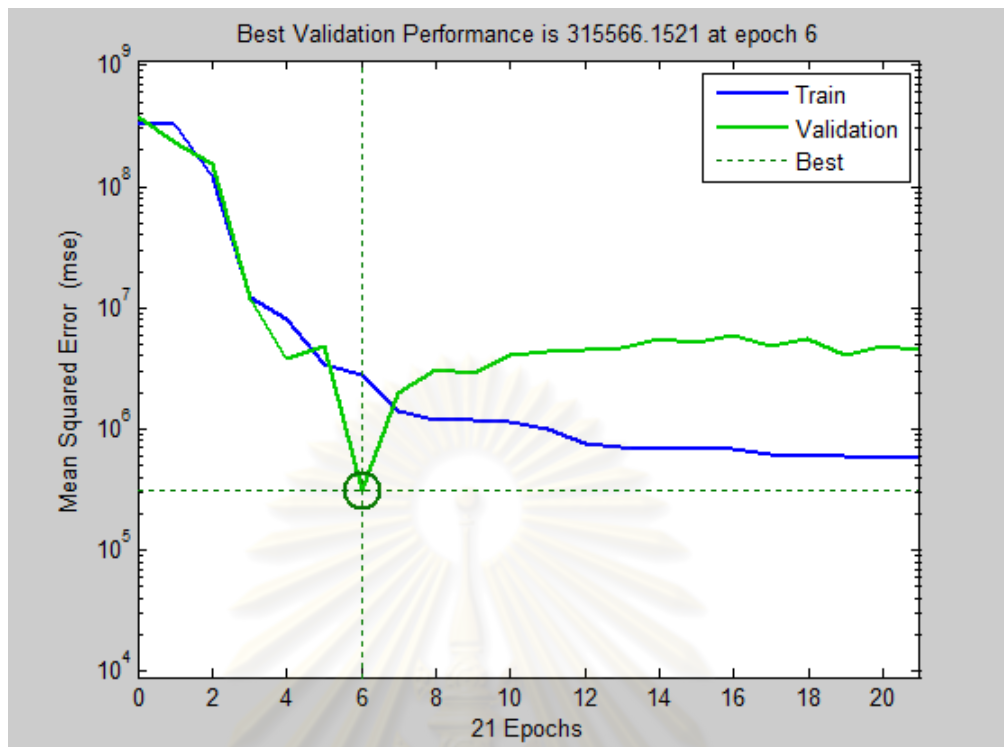


Figure 5.65: Performance curve of model 4 (Case 1-3)

Next, both models that produce the lowest MSE are further checked to ensure the accuracy of prediction. The outputs predicted by the ANN are compared with the target outputs of the training and validating sets by cross plotting them. Figures 5.66 and 5.67 represent the cross plots for the training and validating sets for model 3, respectively, while Figures 5.68 and 5.69 represent the cross plots for the training and validating sets for model 4, respectively. From the graph, the line $Y = X$ refers to correct prediction, i.e., each point on the 45-degree line is where predicted output is matched with the target output. So the closer the data points are located near the $Y = X$ line, the higher the accuracy of prediction. We can determine the accuracy of the prediction using regression coefficient of determination (R^2) as a criterion. R^2 equal to 1 represents a perfect fit to the $Y = X$ line. ANN will predict accurate output when R^2 is close to 1.

From Figures 5.66 to 5.69, we can see that the ANN can predict accurate output for both training and validating sets for both model 3 and 4. Model 3 gives R^2 of training and validating sets equal to 0.9590 and 0.9920, respectively, and model 4

gives R^2 of the training and validating sets equal to 0.8909 and 0.9922, respectively as well.

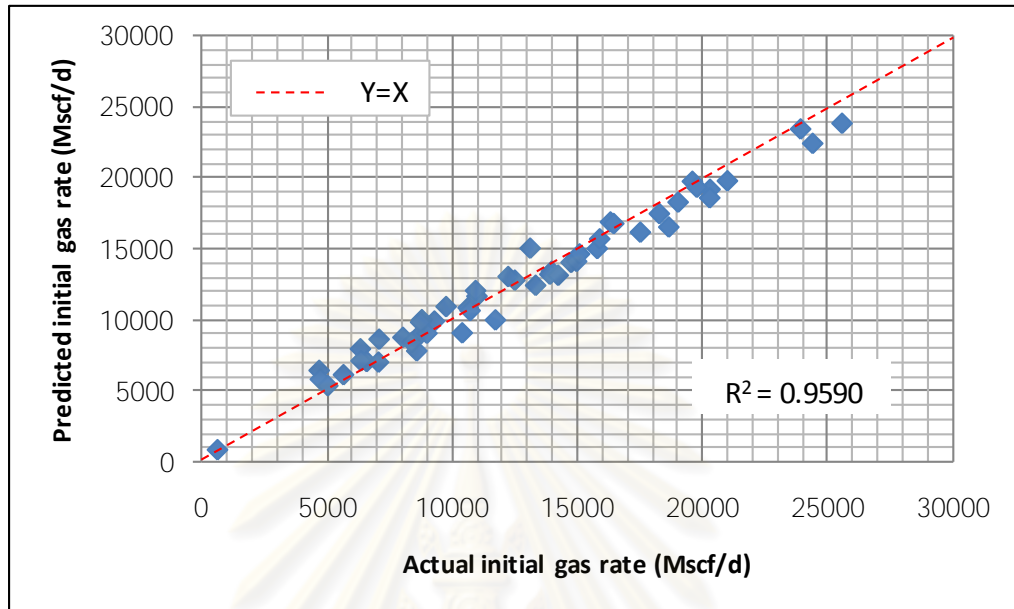


Figure 5.66: Cross plot of predicted vs actual initial gas rates of training sets of model 3 (Case 1-3)

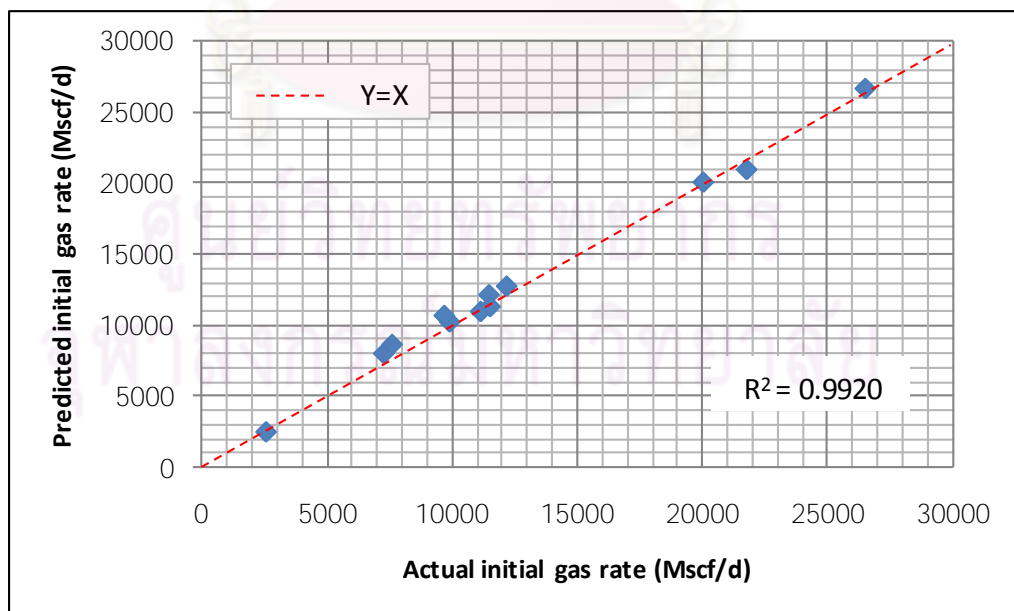


Figure 5.67: Cross plot of predicted vs actual initial gas rates of validating sets of model 3 (Case 1-3)

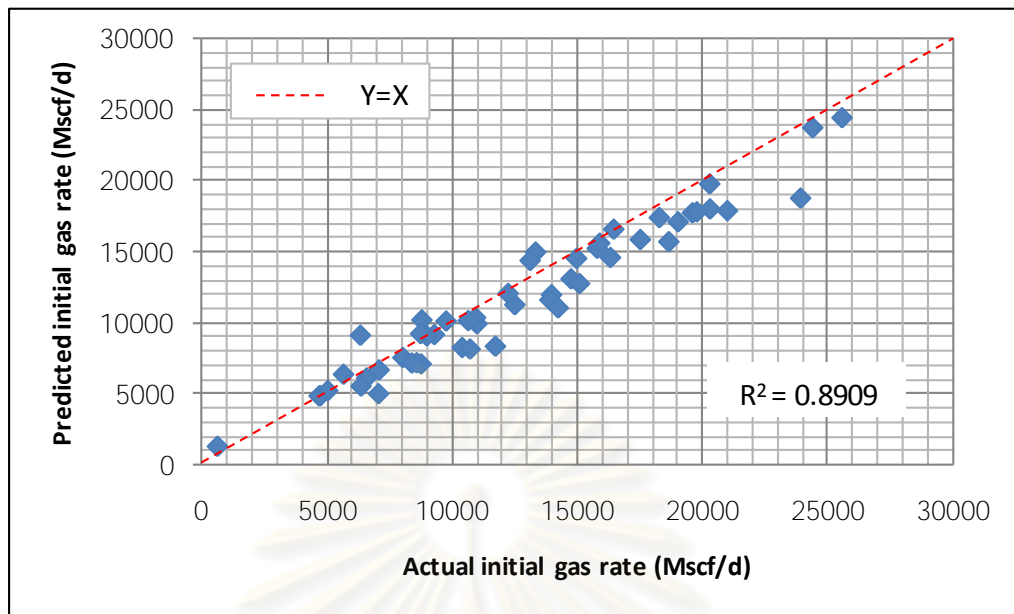


Figure 5.68: Cross plot of predicted vs actual initial gas rates of training sets of model 4 (Case 1-3)

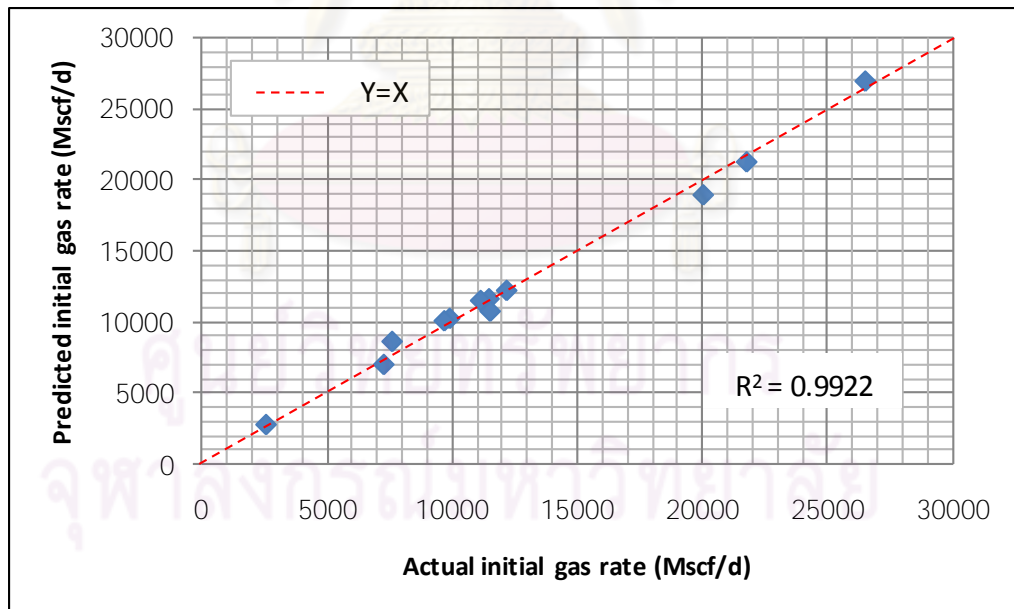


Figure 5.69: Cross plot of predicted vs actual initial gas rates of validating sets of model 4 (Case 1-3)

5.4.1.4.3 Model Testing Results and Discussion

In order to ensure the accuracy of ANN prediction when faced with unseen data sets, model 3 and 4 which yield the lowest MSE were then tested for accuracy by using testing data sets. After testing the ANN with testing sets, the outputs predicted by ANN are then compared with the target outputs by cross plotting them. Figures 5.70 and 5.71 represent the cross plots for model 3 and 4, respectively. From the graphs, R^2 of model 3 and 4 are equal to 0.9240 and 0.8909, respectively.

This means that both model 3 and 4 good performance of predicting the initial gas production rate. However, the coefficient of determination for model 3 is higher. Therefore, the best performance model which produces the most accurate predicted output is model 3. Consequently, this model will be used to predict the initial gas production rate for well number 101 which is drilled 1 year after well number 100.

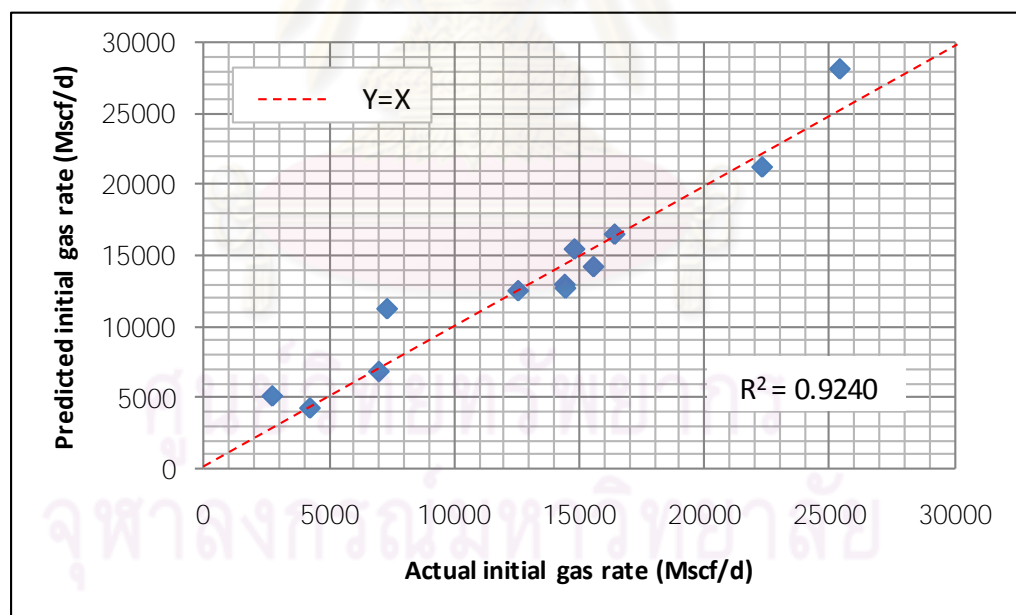


Figure 5.70: Cross plot of predicted vs actual initial gas rates of testing sets of model 3 (Case 1-3)

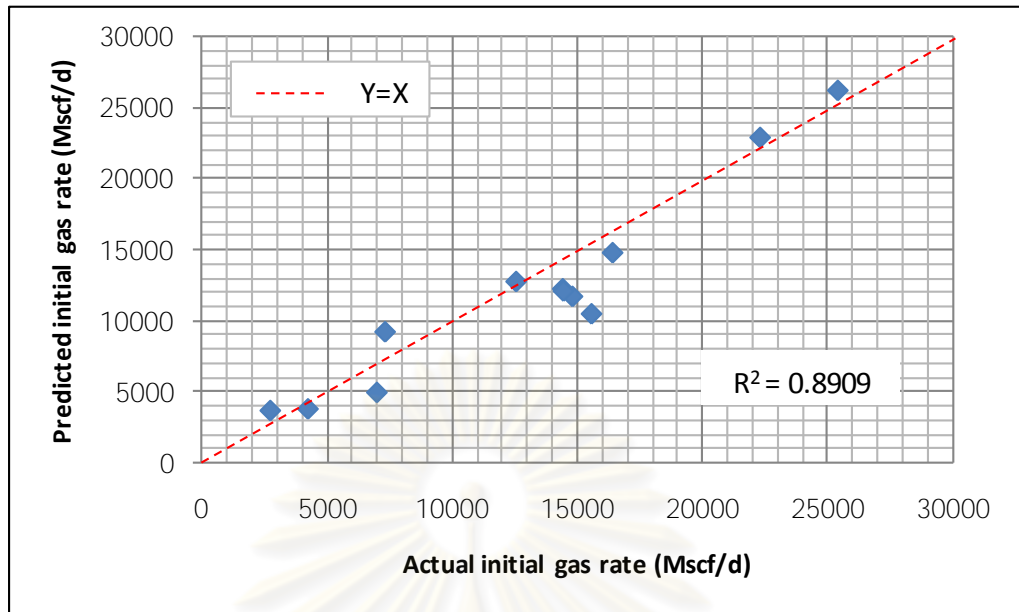


Figure 5.71: Cross plot of predicted vs actual initial gas rates of testing sets of model 4 (Case 1-3)

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

5.4.1.5 Performance of ANN Prediction

After several case studies with different input parameters were performed, we use each best performance model to predict the initial gas rate of the new well planned to be drill (well number 101) which is drilled 1 year after well number 100. Table 5.10 summarize the best performance model for each case study.

Table 5.10: Summary of best performance model for each case.

Case No.	Model No.	R ²		
		Training sets	Validating sets	Testing sets
1-1	16	1	0.9999	0.9998
1-2-1	1	0.9154	0.9614	0.8368
1-2-2	14	0.9365	0.9852	0.9078
1-3	3	0.9590	0.9920	0.9240

First of all, candidate locations for well number 101 must be specified. As described earlier, we plan to drill well number 101 in the 2nd round of drilling. And the well must be located at least 1,500 ft away from other wells and at least 750 ft away from any boundary to avoid boundary effect. After randomly placing a well in the remaining area of the field, only 19 locations can be used as candidate locations for well number 101. Then, the three ANN models were used to predict the initial flow rate for the candidate well locations.

In order to evaluate the accuracy of ANN prediction, we need to determine the initial flow rate of the well drilled at the 19 candidate locations. This was accomplished using ECLIPSE reservoir simulator. Well number 101 was added to the reservoir simulation and started to produce 1 year after well number 100 was drilled. Nineteen separate simulation runs were needed for 19 candidate well locations.

Finally, the outputs predicted by the ANN are compared with the target outputs taken from reservoir simulation by cross plotting them to each other. Figure 5.72, 5.73, 5.74, and 5.75, represent the graph for Case 1-1, Case 1-2-1, 1-2-2, and 1-3, respectively.

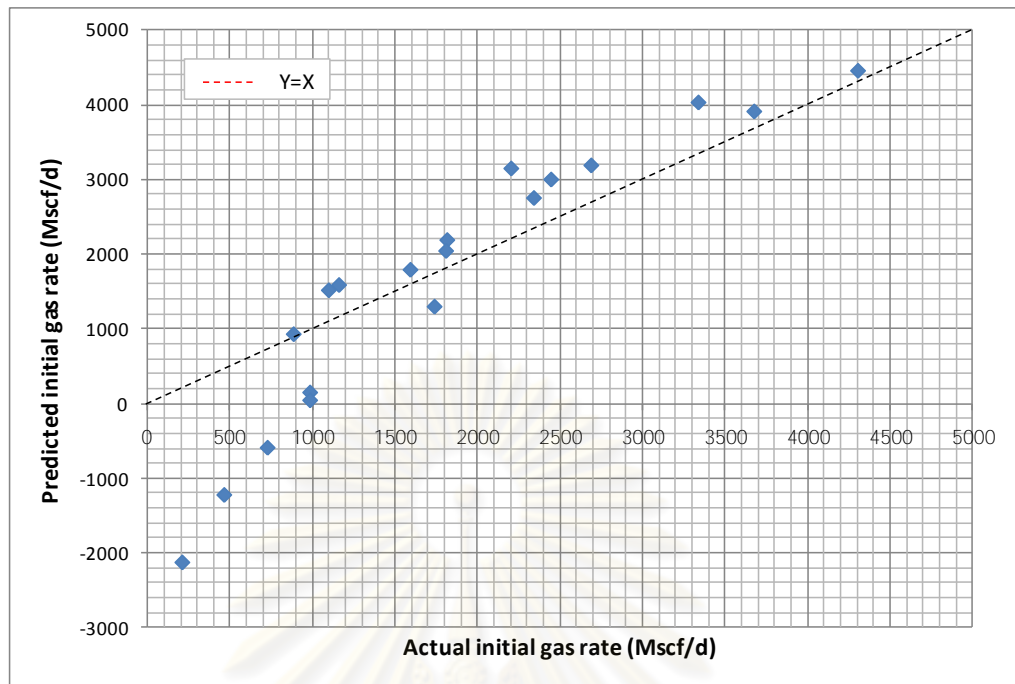


Figure 5.72: Cross plot of predicted vs actual initial gas rate for candidate well locations (Case 1-1)

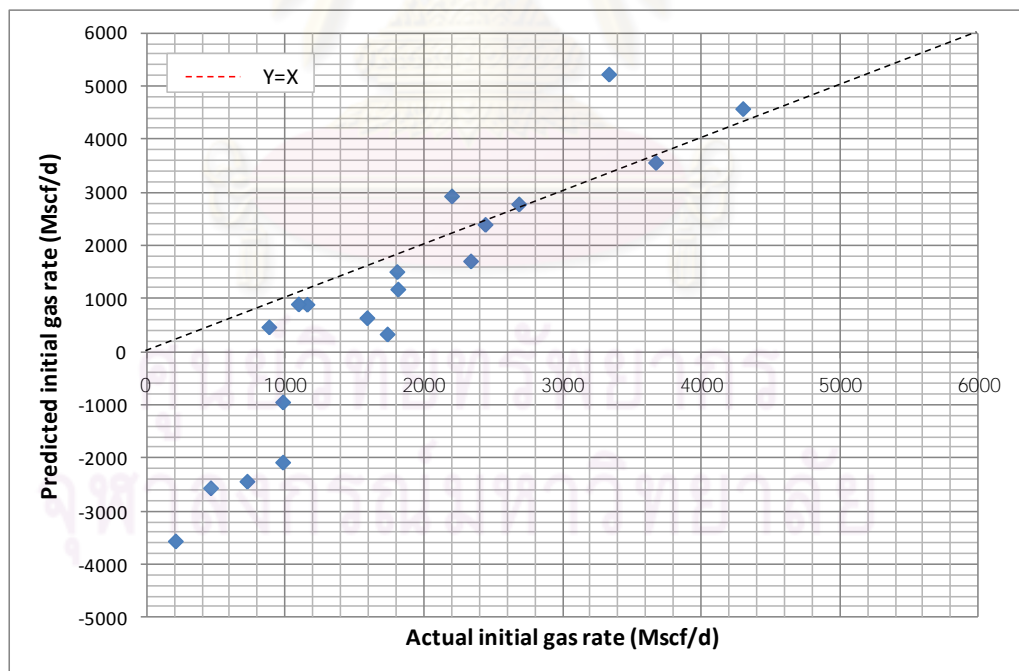


Figure 5.73: Cross plot of predicted vs actual initial gas rate for candidate well locations (Case 1-2-1)

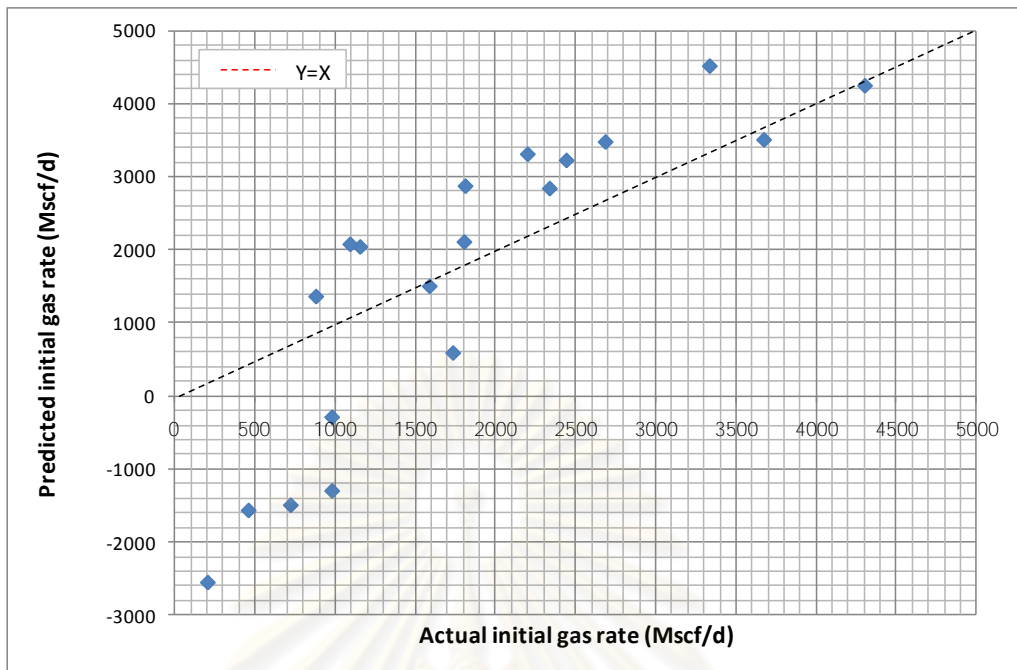


Figure 5.74: Cross plot of predicted vs actual initial gas rate for candidate well locations (Case 1-2-2)

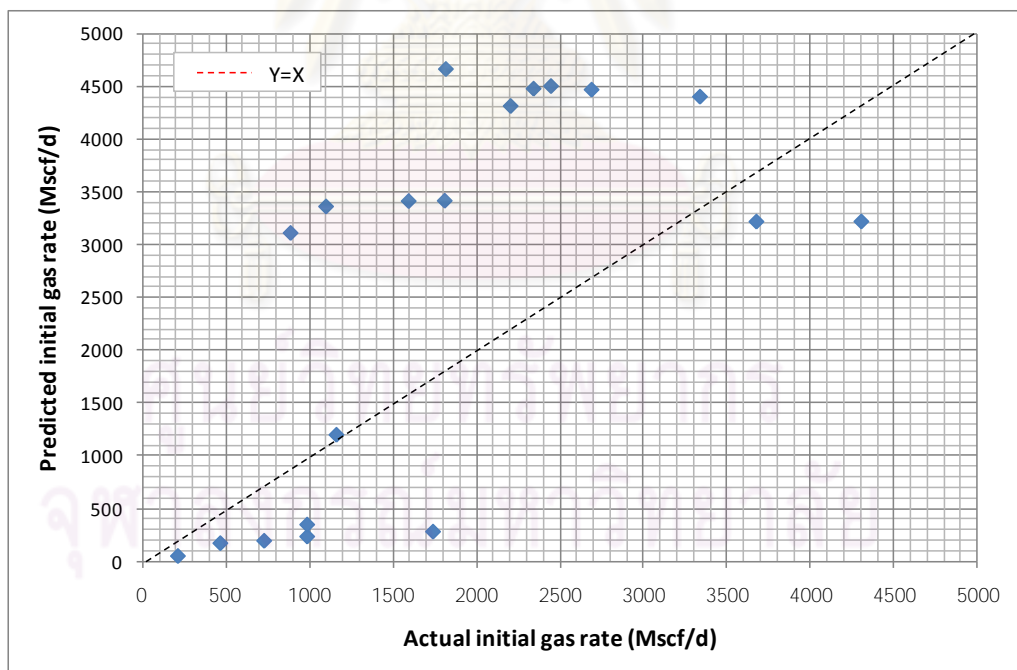


Figure 5.75: Cross plot of predicted vs actual initial gas rate for candidate well locations (Case 1-3)

From Figures 5.72-5.75, each ANN model of all case studies does not present accurate prediction as we expected. Several points are located near the $Y=X$ line, representing good prediction. However, many points are located far away from this line (some points even have negative values). Predictions of negative flow rates occur when the actual flow rates are small (less than 1,000 Mscf/d for Case 1-1, Case 1-2-1, and Case 1-2-2).

From the study of Hettiarachchi et.al^[17], they found a problem when using the ANN to predict the relationship between rainfall and streamflow. A problem arises in extrapolation, i.e., the prediction is not accurate when the training set does not contain the maximum or minimum possible input and output values. Therefore, if we use the trained ANN to predict the output that is out of range or using the input which is out of range, the prediction is inaccurate.

In this study, the type of reservoir is closed boundary depletion drive gas reservoir. Both the pressure and gas production rate continues to decrease as a function of time. The pressure at well number 101 which is drilled 1 year afterward may be lower than the minimum value of input in training sets. Therefore, the prediction is based on extrapolation, causing inaccurate prediction. With this reason, the ANN model that uses pressure as an input parameter (Case 1-1, 1-2-1, 1-2-2) will show a good prediction only when the pressure at that location is not much lower than the minimum pressure in the training data set.

From 4 cases, Case 1-1 is the best predictive model for this study. But as described earlier that the pressure at the location to be drilled is not known prior to drilling. Therefore, the trained ANN of Case 1-1 cannot be used as a tool to predict the initial gas production rate. After comparing all remaining cases, Case 1-3 presents the most inaccurate model. So, we will choose the best prediction model between Case 1-2-1 and Case 1-2-2. Figures 5.76 and 5.77 show the cross plot (only positive value) for Case 1-2-1 and Case 1-2-2, respectively.

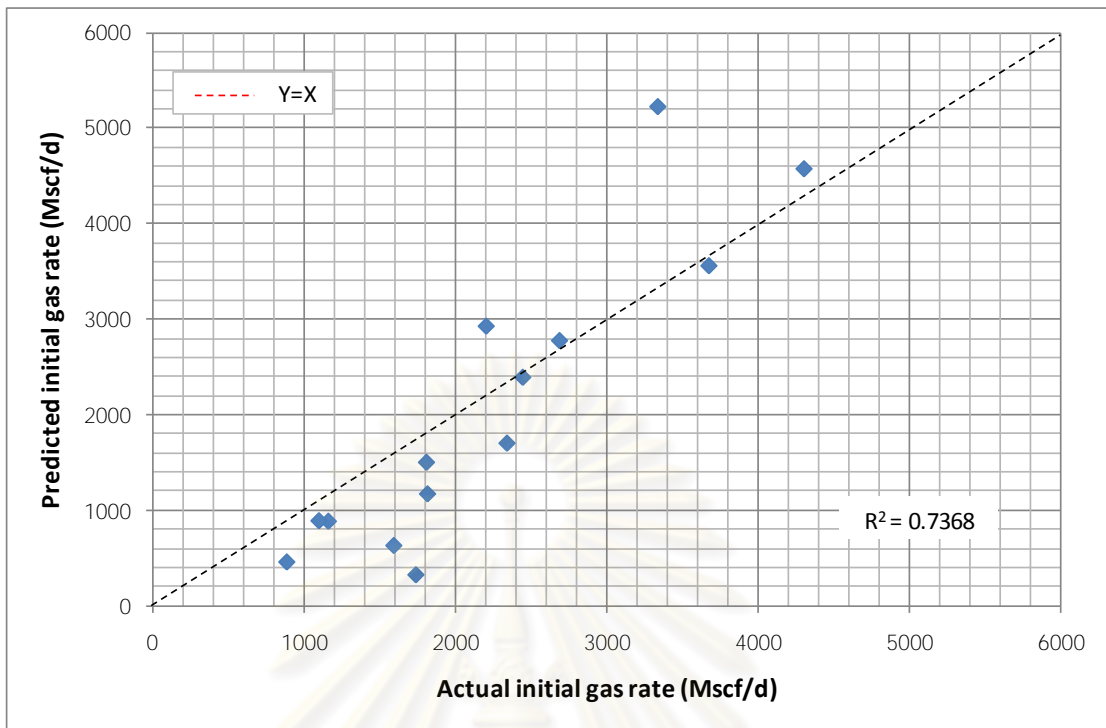


Figure 5.76: Cross plot of predicted vs actual initial gas rate for candidate well locations (positive value of Case 1-2-1)

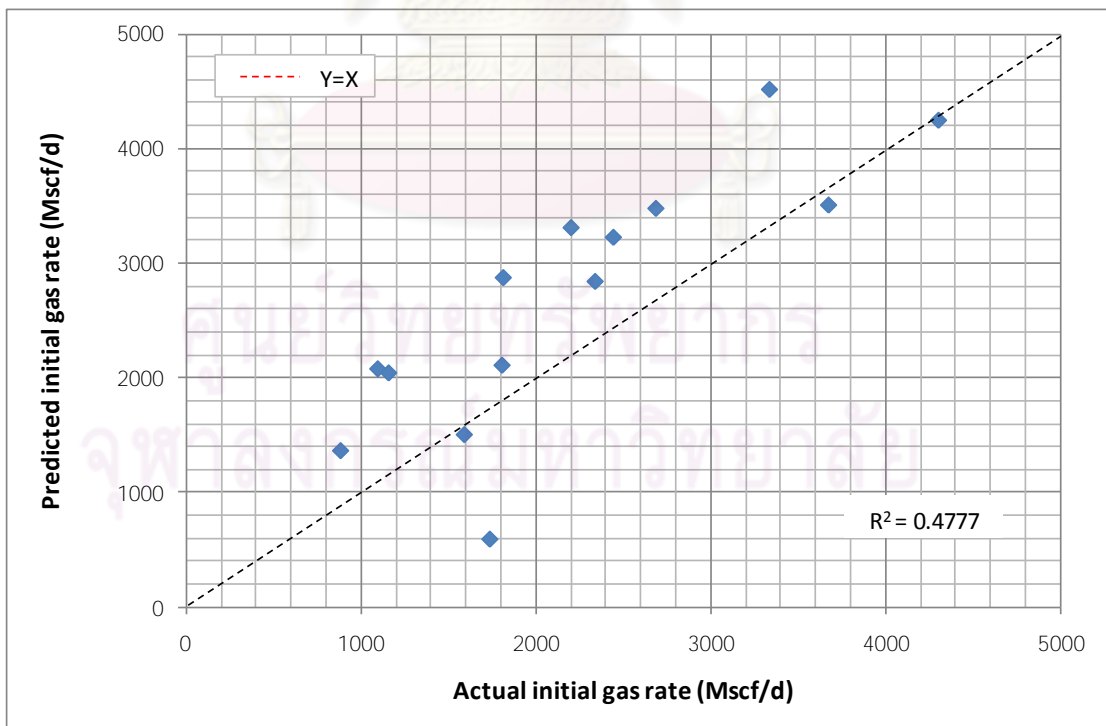


Figure 5.77: Cross plot of predicted vs actual initial gas rate for candidate well locations (positive value of Case 1-2-2)

From the Figures 5.76 and 5.77, R^2 of Case 1-2-1 and Case 1-2-2 are equal to 0.7368 and 0.4777, respectively. Therefore, the best performance model which produces the most accurate predicted output is Case 1-2-1.

In order to ensure this concept, we will try to predict the initial gas production rate at the time shorter than 1 year to see how accurate the prediction is. Figure 5.78 shows a prediction performance of Case 1-2-1 when used to predict the performance of well number 101 at different times to drill the next infill well: 1 year, 6 months, 3 months, and 1 month on the same graph.

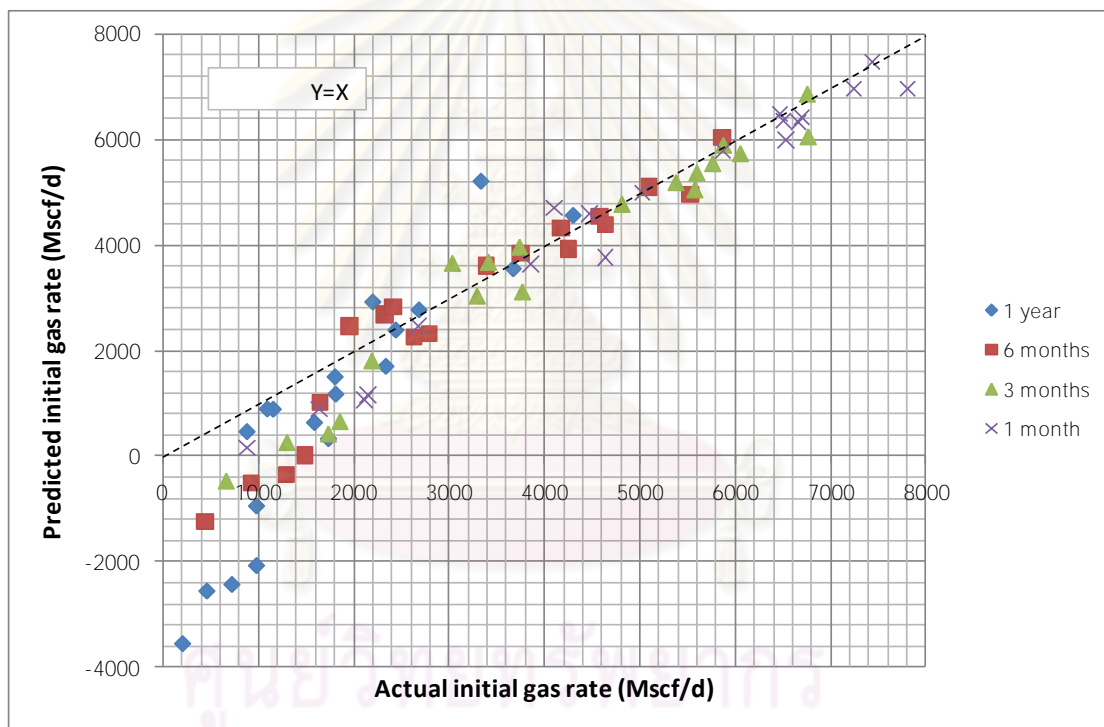


Figure 5.78: Cross plot of predicted vs actual initial gas rate for candidate well locations (Case 1-2-1) for different drill date

From Figure 5.78, the result is most accurate when using a trained ANN to predict the gas rate at 1 month after drilling well number 100. The worst case happens at 1 year. This is because the later the drill date is, the more the pressure decreases. The case of 1 month is the case with the shortest time. So, the pressure is not much

lower than the minimum value used in the training. Consequently, the ANN gives the highest accuracy when used to predict the initial gas rate.

Although, the prediction of ANN does not give us accurate result for all well locations, we only need one location to infill which is the location that gives the highest gas production rate. This location is associated with high pressure. Therefore, we can use the ANN to roughly estimate the gas rate for this location.

At this point, we are able to predict initial gas productions at candidate well locations as illustrated in Figure 5.79.

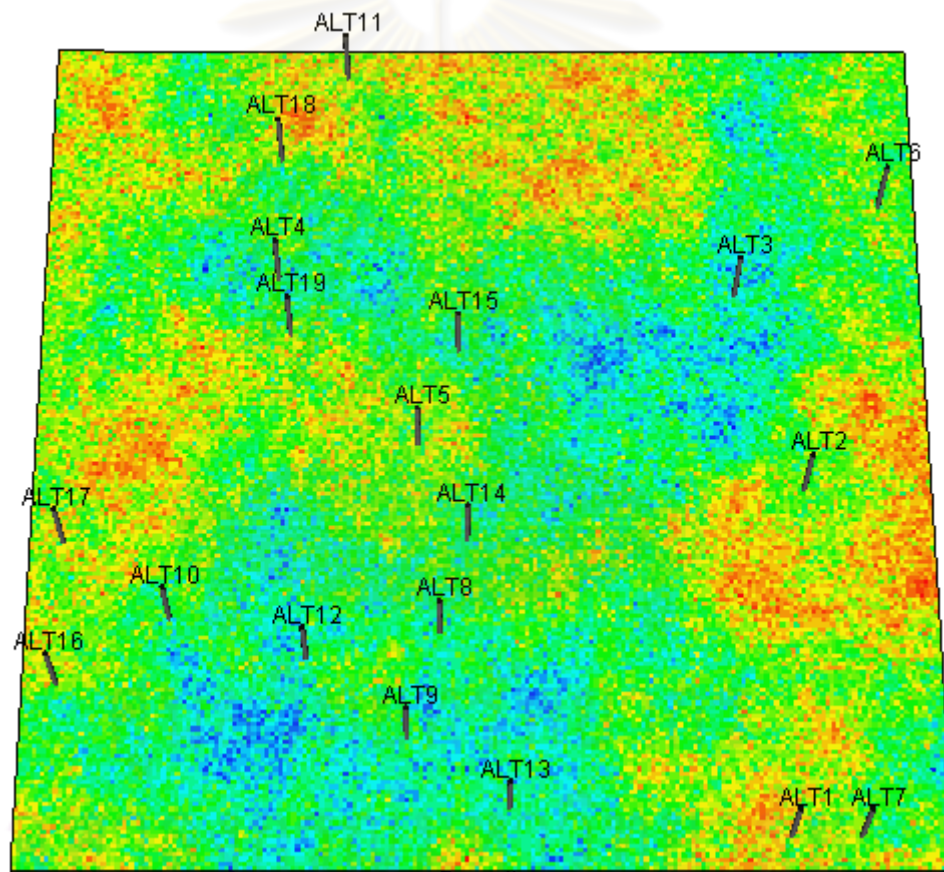


Figure 5.79: Candidate well locations on porosity distribution map

In order to determine the best location to drill well number 101, the predicted initial gas rate obtained from each ANN model is ranked and tabulated as shown in Table 5.11

Table 5.11: Order of candidate location from a higher to lower initial gas rate

Order	Candidate location for well number 101				
	Reality	Case 1-1	Case 1-2-1	Case 1-2-2	Case 1-3
1	location 9	location 9	location 1	location 1	location 11
2	location 13	location 1	location 9	location 9	location 17
3	location 1	location 13	location 13	location 13	location 7
4	location 5	location 5	location 19	location 5	location 5
5	location 17	location 19	location 5	location 19	location 1
6	location 7	location 17	location 17	location 17	location 19
7	location 19	location 7	location 7	location 11	location 16
8	location 11	location 11	location 16	location 7	location 6
9	location 16	location 16	location 11	location 16	location 18
10	location 8	location 6	location 18	location 18	location 9
11	location 6	location 4	location 4	location 4	location 13
12	location 4	location 18	location 6	location 6	location 2
13	location 18	location 8	location 2	location 2	location 4
14	location 14	location 2	location 8	location 8	location 14
15	location 12	location 12	location 14	location 14	location 8
16	location 2	location 14	location 12	location 12	location 12
17	location 3	location 3	location 3	location 3	location 3
18	location 15	location 15	location 15	location 15	location 15
19	location 10	location 10	location 10	location 10	location 10

In reality, location 9 is the one that yields the highest initial gas production rate and the top 3 candidate locations to infill a well are well number 9, 13, and 1, respectively. Case 1-1 which uses the actual pressure as input parameter can predict the same group of top 3 locations as in reality with the right best location at location 9. Case 1-2-1 and 1-2-2 which uses arithmetic and inverse-distance average pressure as input parameter can also predict the right group of top 3 locations as in reality except that the best location predicted by these cases are location 1 which is the third order in reality. Case 1-3 which does not use any pressure as input parameter cannot predict the right location. Top 3 locations of this case are location 11, 17, and 7 which are the 8th, 5th, and 6th order in reality, respectively. Because location 1 is predicted by both Case 1-2-1 and 1-2-2 while location 11 is predicted by only Case 1-3. Therefore, location 1 will be used as the best location to infill well even in fact this location is not the best location.

Table 5.12: Summary of error in initial gas rate predicted for location 1.

Case	Initial gas rate (Mscf/d)	Error (%)
Reality	3337.93	-
1-2-1	5215.98	56.26
1-2-2	4517.90	35.35
1-3	4397.76	31.75

Table 5.12 summarises the error of prediction when using each model to predict initial gas production rate of candidate location 1. From Table 5.12, the initial gas rate obtained from each model is quite high when comparing with the actual value from reality case. Even though the error may be high, we achieve the objective of being able to determine very good location for the next infill well even if it is not the very best.

5.4.2 Cumulative Gas Production Prediction

The purpose of this case study is to predict a 1 year cumulative gas production after drilling the well. The input parameters of ANN will be the same as the ones used to predict the initial gas production rate which are permeability and pressure at the drill date of the predicting well. However, after the wells were drilled, the pressure at that location will continue to decrease. Therefore, if we want to predict the cumulative gas production, we need to include parameters that affect the changing of pressure after wells are drilled as input of ANN.

Since porosity around the predicting well directly affects the amount of gas in place, it affects the change in pressure and needs to be included as an input parameter.

This section is divided into 4 case studies similar to Section 5.4.1. Case 2-1 has 2 input parameters as in Case 1-1 and the average porosity of each ring of the three rings. The next set of case studies uses average pressure from surrounding wells instead. There are 2 average methods that are 1) arithmetic average (Case 2-2-1) and 2) inverse-distance average (Case 2-2-2). The last case (Case 2-3) uses porosity, drill date, and the number of surrounding wells as a proxy of pressure at well location.

In each case study, the ANN is trained with training and validating sets by varying network configurations that are the number of hidden nodes, number of hidden layers, learning rate, and momentum on a trial and error basis. Only 2 models with give the lowest and next to lowest error of validating sets are used to test the accuracy with testing set. After that the only one model with lowest error of the testing set will be chosen to be the best performance model for each case study. Finally, best performance model will be used to predict the 1 year cumulative gas production of the well number 101 which is scheduled to be drilled one year after well number 100.

5.4.2.1 Case 2-1

In this case, the network consists of 5 input parameters. The first 2 input are the same as in Case 1-1 which are permeability and initial pressure at the drill date. To represent the effect of decline of initial pressure after starting the production, average porosity of 1st, 2nd, and 3rd rings are included as inputs. The number of hidden layers and number of neurons in each hidden layer are varied base on trial and error basis to get the best performance to predict the output which is 1-year cumulative gas production. Figure 5.80 illustrates the schematic diagram of ANN in this case.

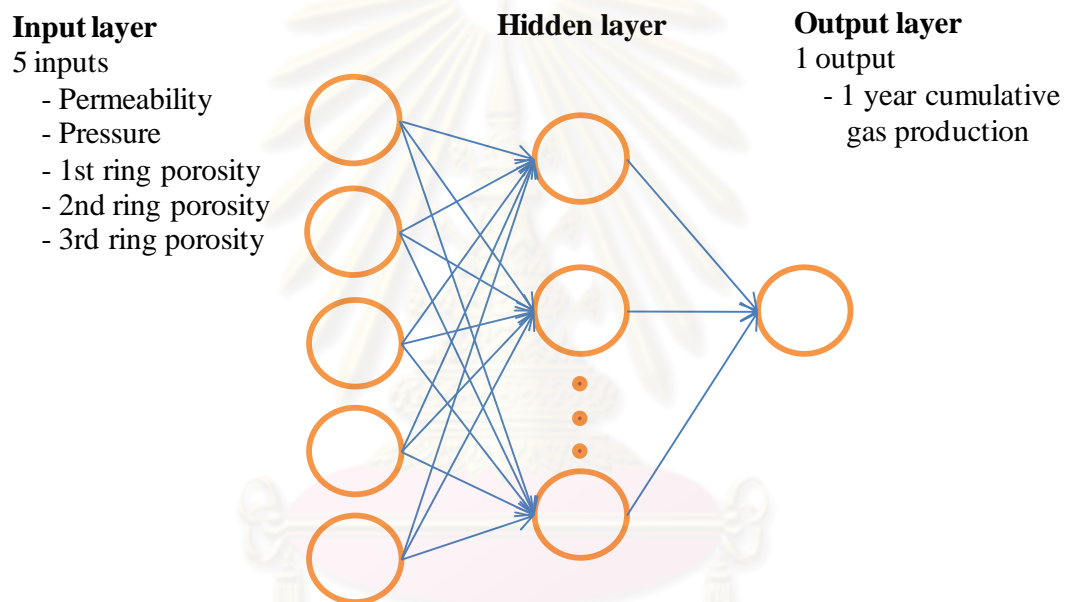


Figure 5.80: Schematic diagram of ANN for Case 2-1

5.4.2.1.1 Data Preprocessing

Similar to the previous case, a total of 75 data sets taken from well number 26 to 100 were divided into three main sets namely, training, validating, and testing sets with ratio of 4:1:1 (51:12:12 data sets). The wells in each data set are still the same as the ones in the previous case study. All input parameters from all data sets are plotted to observe the distributions which are the same as those in Case 1-1.

5.4.2.1.2 Model Training

The ANN model was trained with various network configurations based on a trial and error basis. There are 20 models that were run in this case. Each model was trained many times to obtain the lowest MSE possible. The network configurations and their MSE of validating set are summarized in Table 5.13.

Table 5.13: Model configuration for Case 2-1

Model No	Number of neurons		Learning rate	Momentum	MSE
	Hidden Layer 1	Hidden Layer 2			
1	5	0	0.1	0.1	13,649,834,695
2	5	0	0.5	0.1	13,778,152,454
3	5	0	0.1	0.5	13,649,834,695
4	5	0	0.5	0.5	13,778,152,454
5	10	0	0.1	0.1	20,478,539,230
6	10	0	0.5	0.1	23,036,574,398
7	10	0	0.1	0.5	19,365,278,493
8	10	0	0.5	0.5	21,035,647,362
9	20	0	0.1	0.1	25,478,453,672
10	20	0	0.5	0.1	22,304,384,304
11	20	0	0.1	0.5	26,478,394,045
12	20	0	0.5	0.5	24,857,494,455
13	5	5	0.1	0.1	14,699,731,417
14	5	5	0.5	0.1	22,175,493,419
15	5	5	0.1	0.5	17,108,192,582
16	5	5	0.5	0.5	20,101,001,784
17	10	10	0.1	0.1	21,349,589,473
18	10	10	0.5	0.1	24,789,304,463
19	10	10	0.1	0.5	26,433,748,953
20	10	10	0.5	0.5	27,634,735,484

From a total 20 model configurations, two models with the lowest and next to lowest MSE (model 1 and 2) which has the MSE of 13,649,834,695 and 13,778,152,454, respectively, are chosen. Model 1 consists of only one hidden layers with 5 neurons. The learning rate and momentum of 0.1 were used. Similar to model 1, model 2 consists of only one hidden layers with 5 neurons. However, the learning rate and momentum were set to be 0.5 and 0.1, respectively. The performance curves of

model 1 and 2 are shown in Figures 5.81 and 5.82, respectively. Model 1 was trained until epoch 31 but the weight and bias were updated until epoch 16 only because the MSE of validating set started to increase in this epoch. The training was continued for 15 more epochs for validation check. The lowest MSE of model 1 is 13,649,834,695. For model 2, the training was performed until epoch 143. However, the weight and bias were not updated after epoch 128 due to the same reason for model 1. The lowest MSE of model 2 is 13,778,152,454 in epoch 128.

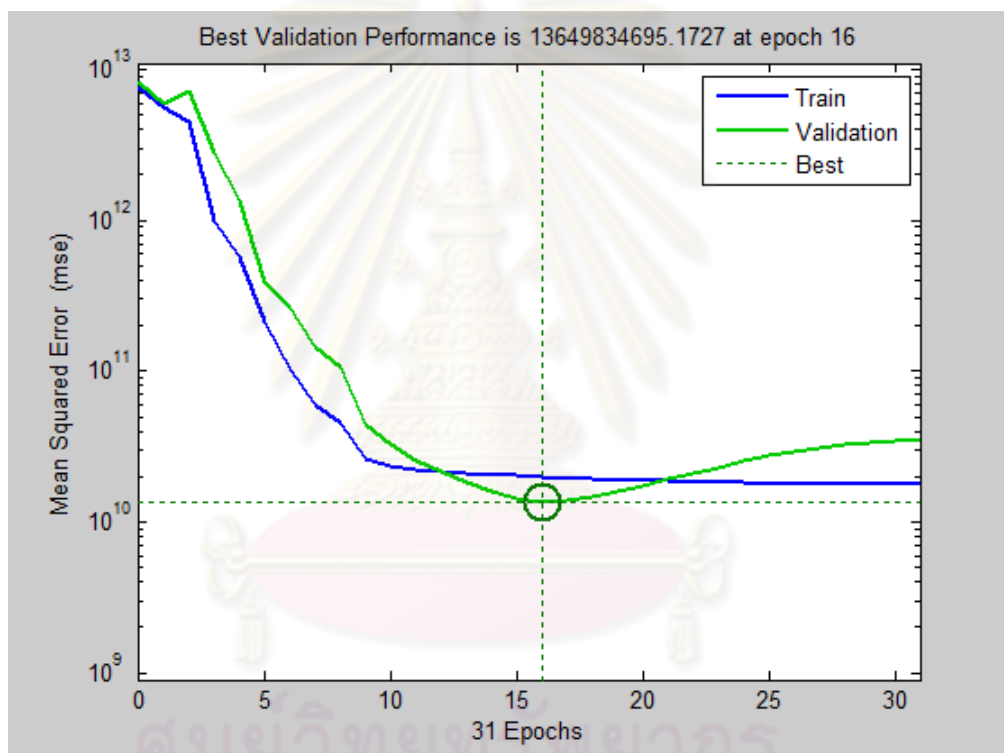


Figure 5.81: Performance curve of model 1 (Case 2-1)

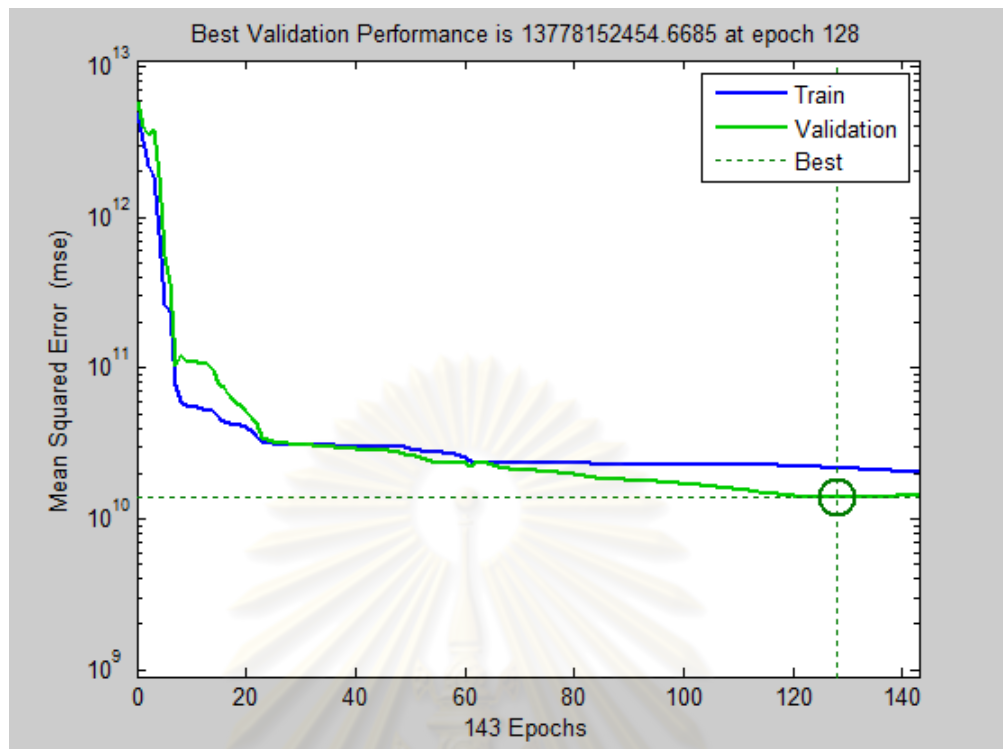


Figure 5.82: Performance curve of model 2 (Case 2-1)

Next, both models that produce the lowest MSE are further checked to ensure the accuracy of prediction. The outputs predicted by the ANN are compared with the target outputs of the training and validating sets by cross plotting them. Figures 5.83 and 5.84 represent the cross plots for the training and validating sets for model 1, respectively, while Figures 5.85 and 5.86 represent the cross plots for the training and validating sets for model 2, respectively. From the graph, the line $Y = X$ refers to correct prediction, i.e., each point on the 45-degree line is where predicted output is matched with the target output. So the closer the data points are located near the $Y = X$ line, the higher the accuracy of prediction. We can determine the accuracy of the prediction using regression coefficient of determination (R^2) as a criterion. R^2 equal to 1 represents a perfect fit to the $Y = X$ line. ANN will predict accurate output when R^2 is close to 1.

From Figures 5.83 to 5.86, we can see that the ANN can predict accurate output for both training and validating sets for both model 1 and 2. Model 1 gives R^2 of training and validating sets equal to 0.9928 and 0.9966, respectively, and model 2

gives R^2 of the training and validating sets equal to 0.9921 and 0.9965, respectively as well.

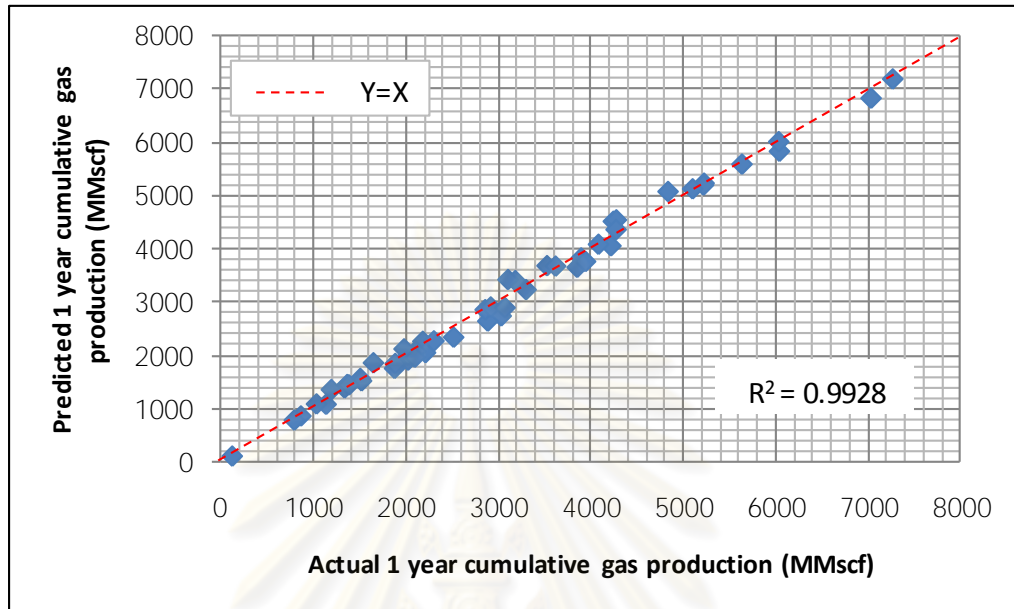


Figure 5.83: Cross plot of predicted vs actual 1 year cumulative gas of training sets of model 1 (Case 2-1)

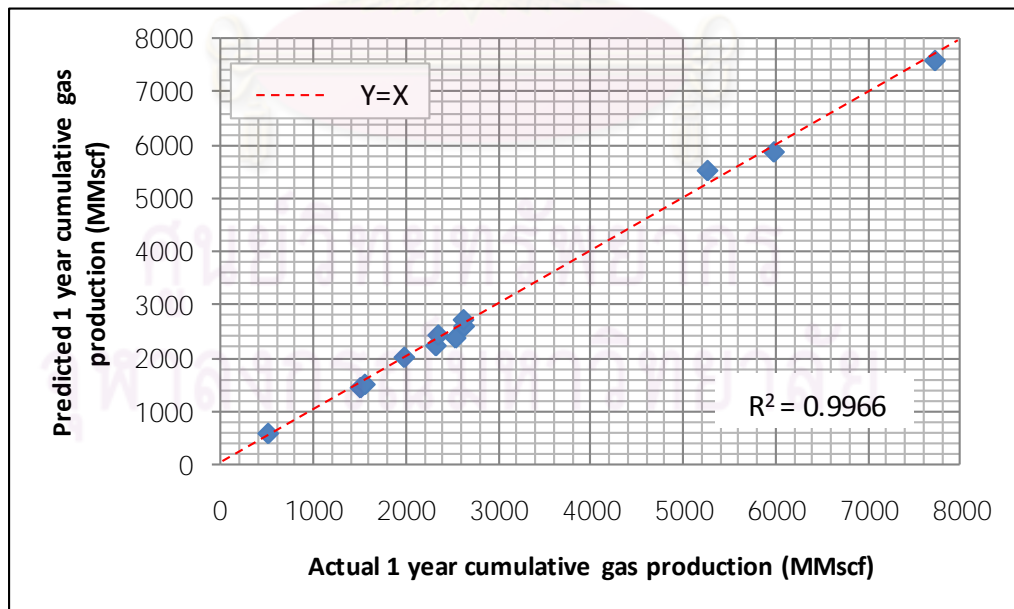


Figure 5.84: Cross plot of predicted vs actual 1 year cumulative gas of validating sets of model 1 (Case 2-1)

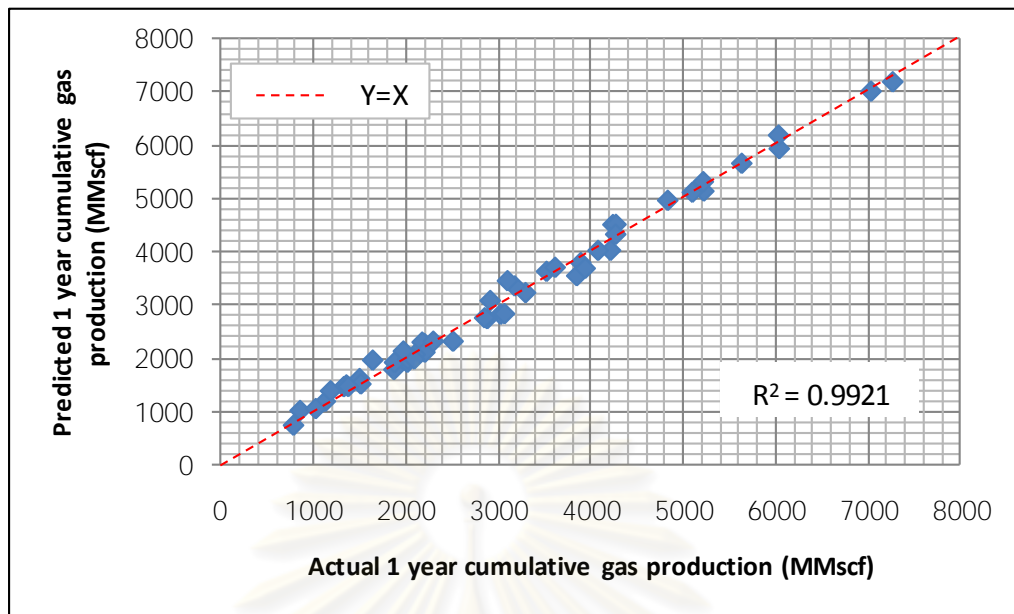


Figure 5.85: Cross plot of predicted vs actual 1 year cumulative gas of training sets of model 2 (Case 2-1)

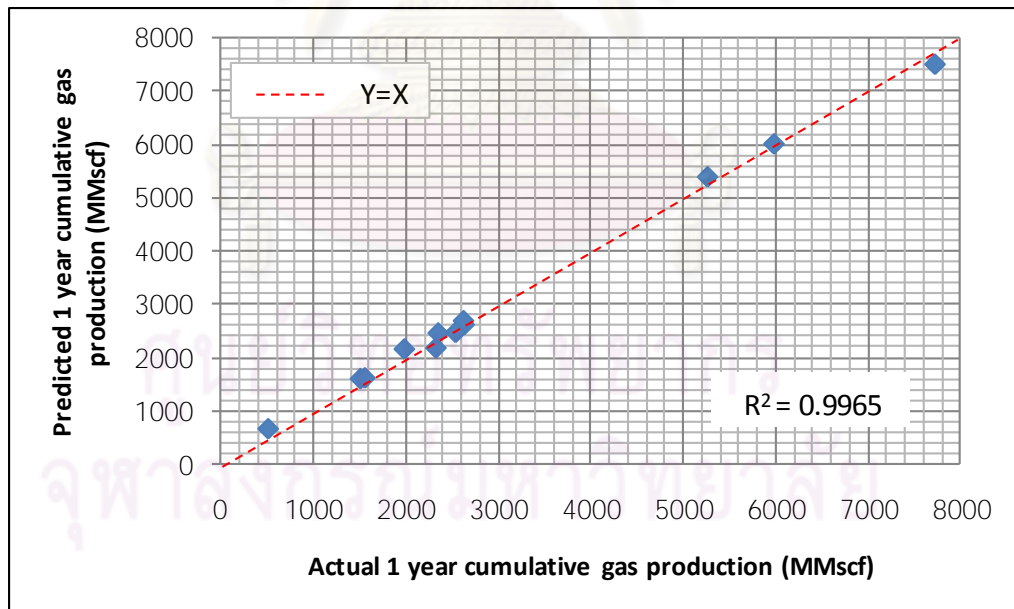


Figure 5.86: Cross plot of predicted vs actual 1 year cumulative gas of validating sets of model 2 (Case 2-1)

5.4.2.1.3 Model Testing Results and Discussion

In order to ensure the accuracy of ANN prediction when faced with unseen data sets, model 1 and 2 which yield the lowest MSE were then tested for accuracy by using testing data sets. After testing the ANN with testing sets, the outputs predicted by ANN are then compared with the target outputs by cross plotting them. Figures 5.87 and 5.88 represent the cross plots for model 1 and 2, respectively. From the graphs, R^2 of model 1 and 2 are equal to 0.9818 and 0.9863, respectively.

This means that both model 1 and 2 good performance of predicting the 1 year cumulative gas production. However, the coefficient of determination for model 2 is higher. Therefore, the best performance model which produces the most accurate predicted output is model 2.

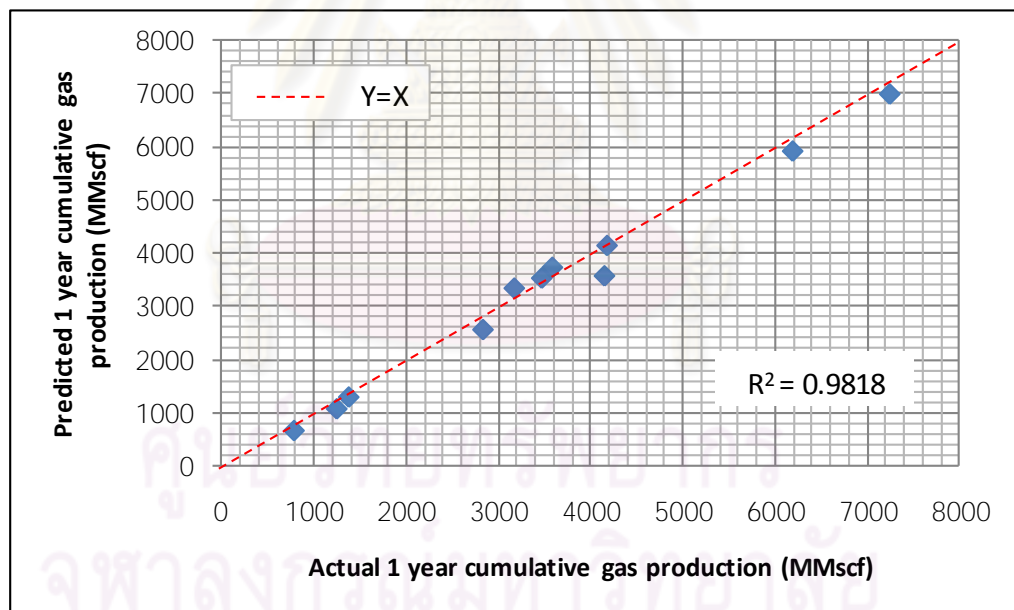


Figure 5.87: Cross plot of predicted vs actual 1 year cumulative gas of testing sets of model 1 (Case 2-1)

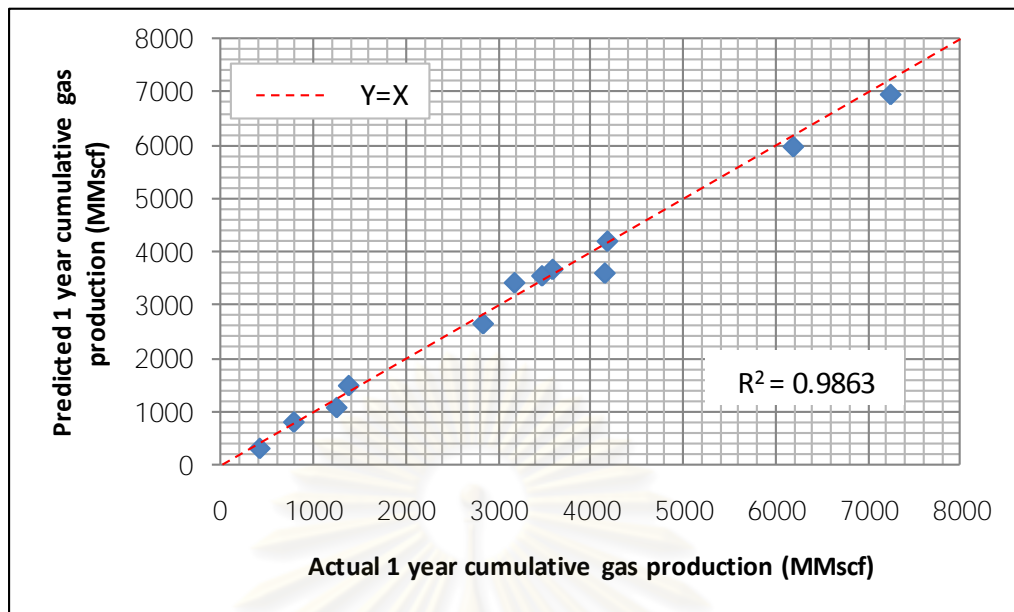


Figure 5.88: Cross plot of predicted vs actual 1 year cumulative gas of testing sets of model 2 (Case 2-1)

In summary, we have proved that the ANN can be used to predict 1-year cumulative gas production when necessary input information is available. However, in practice, even the best model from this study cannot be used to predict the output of well number 101 because it is impossible to know the pressure of well number 101 prior to drilling. In any case, this case study helps ensure that 1-year cumulative gas production is strongly related to permeability and pressure at the predicting well location.

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

5.4.2.2 Case 2-2-1

As described in Case 1-2-1, we do not know the pressure at the predicting well locations in real situations. So, this case study will use the average pressure of the surrounding wells instead. These pressures were taken from surrounding wells location at the time when drilling the predicting well. Figure 5.89 illustrates the schematic diagram of ANN in this case. The average pressure used in this study is the same as used in Case 1-2-1.

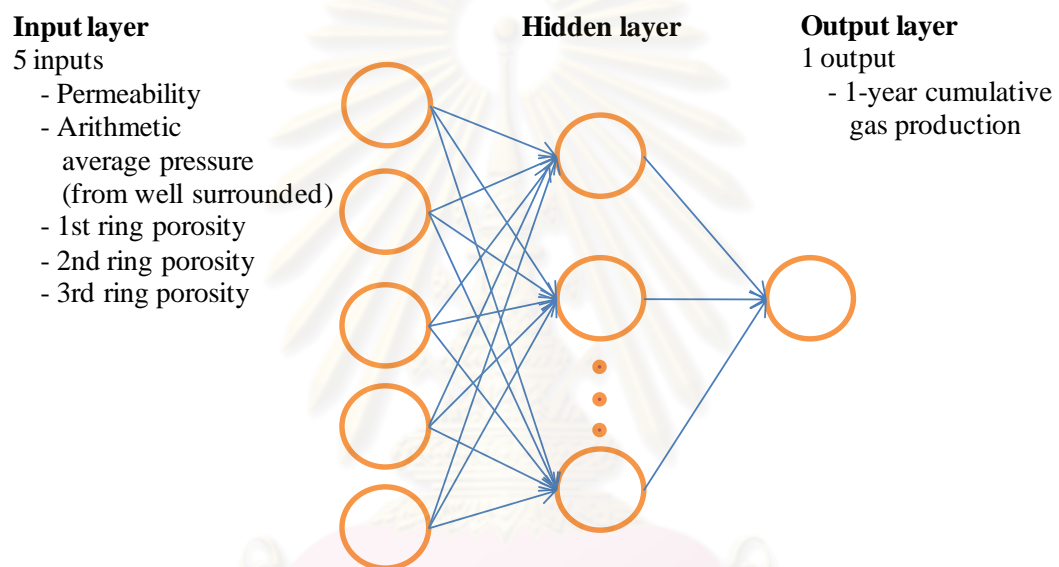


Figure 5.89: Schematic diagram of ANN for Case 2-2-1

5.4.2.2.1 Data Preprocessing

Similar to the previous case, a total of 75 data sets taken from well number 26 to 100 were divided into three main sets namely, training, validating, and testing sets with ratio of 4:1:1 (51:12:12 data sets). The wells in each data set are still the same as the ones in the previous case study. All input parameters from all data sets are plotted to observe the distributions which are the same as those in Case 1-2-1.

5.4.2.2.2 Model Training

The ANN model was trained with various network configurations based on a trial and error basis. There are 20 models that were run in this case. Each model was trained many times to obtain the lowest MSE possible. The network configurations and their MSE of validating set are summarized in Table 5.14.

Table 5.14: Model configuration for Case 2-2-1

Model No	Number of neurons		Learning rate	Momentum	MSE
	Hidden Layer 1	Hidden Layer 2			
1	5	0	0.1	0.1	79,471,345,810
2	5	0	0.5	0.1	138,090,469,358
3	5	0	0.1	0.5	103,478,498,463
4	5	0	0.5	0.5	98,473,678,436
5	10	0	0.1	0.1	154,892,615,335
6	10	0	0.5	0.1	147,689,823,465
7	10	0	0.1	0.5	165,489,345,243
8	10	0	0.5	0.5	142,453,645,891
9	20	0	0.1	0.1	165,243,152,345
10	20	0	0.5	0.1	174,532,652,435
11	20	0	0.1	0.5	164,273,524,543
12	20	0	0.5	0.5	182,638,845,637
13	5	5	0.1	0.1	97,706,258,262
14	5	5	0.5	0.1	89,627,857,993
15	5	5	0.1	0.5	105,985,209,239
16	5	5	0.5	0.5	108,884,086,527
17	10	10	0.1	0.1	120,457,389,946
18	10	10	0.5	0.1	135,462,878,463
19	10	10	0.1	0.5	126,745,367,843
20	10	10	0.5	0.5	145,678,845,231

From a total 20 model configurations, two models with the lowest and next to lowest MSE (model 1 and 14) which has the MSE of 79,471,345,810 and 89,627,857,993, respectively, are chosen. Model 1 consists of only one hidden layers with 5 neurons. The learning rate and momentum of 0.1 were used. But model 14 consists of 2 hidden layers with 5 neurons in each layer. However, the learning rate and momentum were set to be 0.5 and 0.1, respectively. The performance curves of

model 1 and 14 are shown in Figures 5.90 and 5.91, respectively. Model 1 was trained until epoch 20 but the weight and bias were updated until epoch 5 only because the MSE of validating set started to increase in this epoch. The training was continued for 15 more epochs for validation check. The lowest MSE of model 1 is 79,471,345,810. For model 14, the training was performed until epoch 25. However, the weight and bias were not updated after epoch 10 due to the same reason for model 1. The lowest MSE of model 14 is 89,627,857,993 in epoch 10.

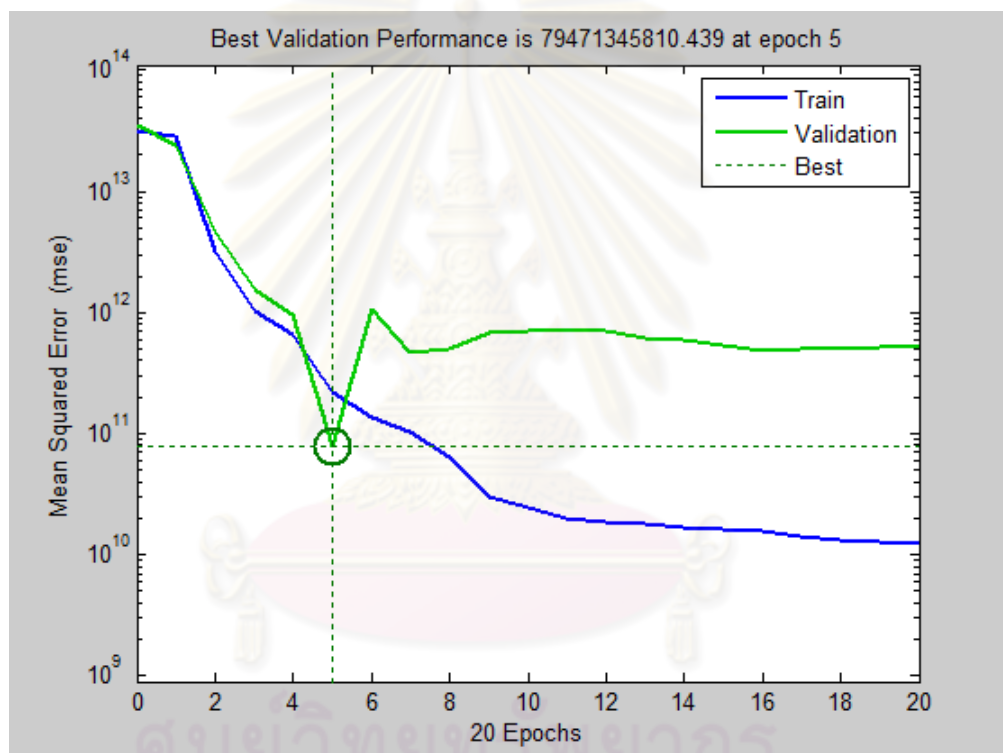


Figure 5.90: Performance curve of model 1 (Case 2-2-1)

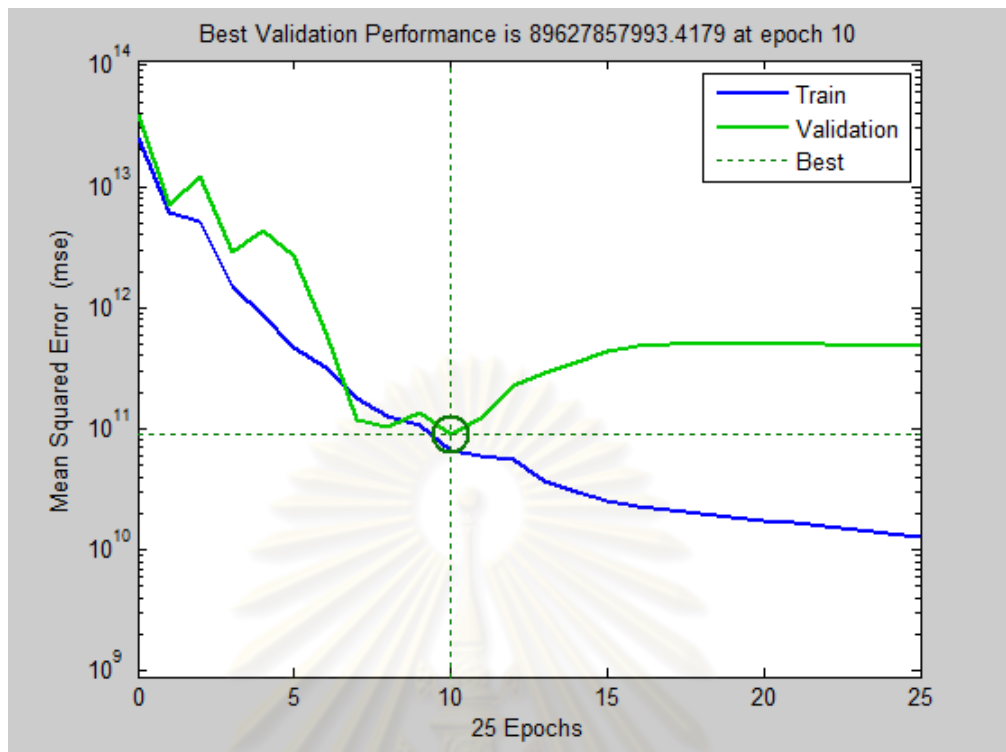


Figure 5.91: Performance curve of model 14 (Case 2-2-1)

Next, both models that produce the lowest MSE are further checked to ensure the accuracy of prediction. The outputs predicted by the ANN are compared with the target outputs of the training and validating sets by cross plotting them. Figures 5.92 and 5.93 represent the cross plots for the training and validating sets for model 1, respectively, while Figures 5.94 and 5.95 represent the cross plots for the training and validating sets for model 14, respectively. From the graph, the line $Y = X$ refers to correct prediction, i.e., each point on the 45-degree line is where predicted output is matched with the target output. So the closer the data points are located near the $Y = X$ line, the higher the accuracy of prediction. We can determine the accuracy of the prediction using regression coefficient of determination (R^2) as a criterion. R^2 equal to 1 represents a perfect fit to the $Y = X$ line. ANN will predict accurate output when R^2 is close to 1.

From Figures 5.92 to 5.95, we can see that the ANN can predict accurate output for both training and validating sets for both model 1 and 14. Model 1 gives R^2 of training and validating sets equal to 0.8989 and 0.9781, respectively, and model 14

gives R^2 of the training and validating sets equal to 0.9753 and 0.9785, respectively as well.

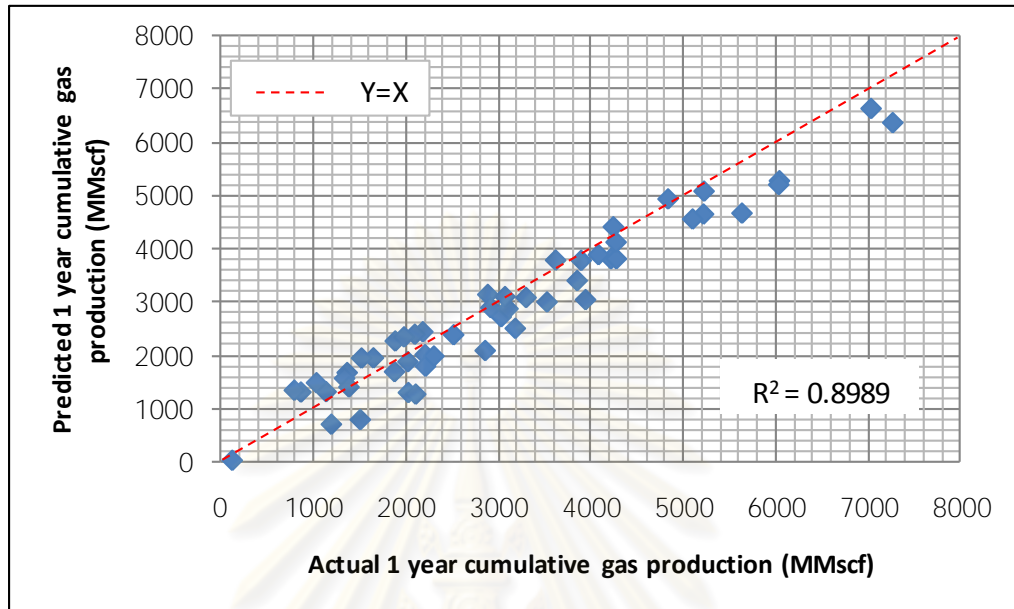


Figure 5.92: Cross plot of predicted vs actual 1 year cumulative gas of training sets of model 1 (Case 2-2-1)

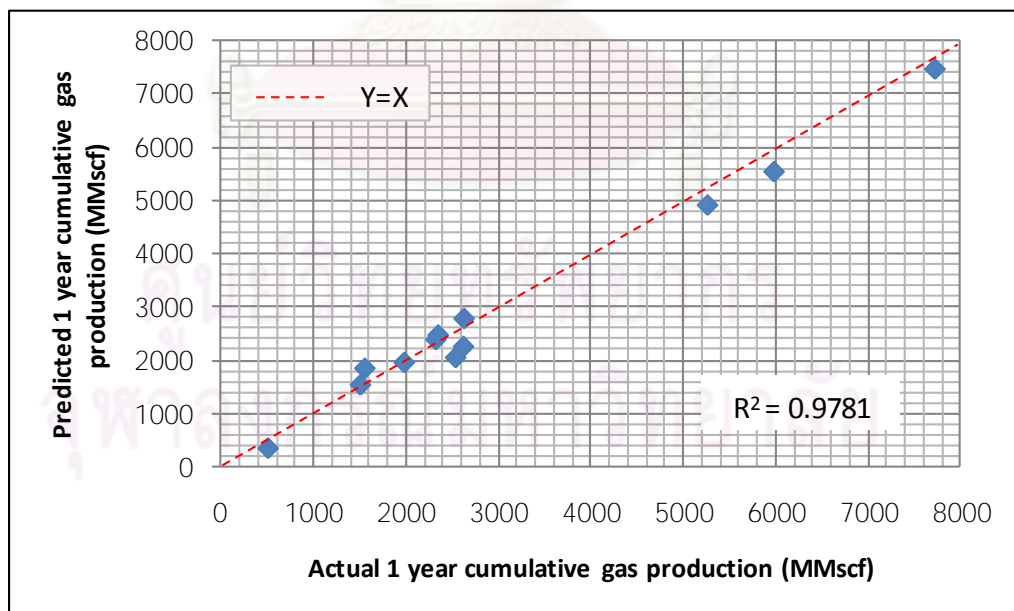


Figure 5.93: Cross plot of predicted vs actual 1 year cumulative gas of validating sets of model 1 (Case 2-2-1)

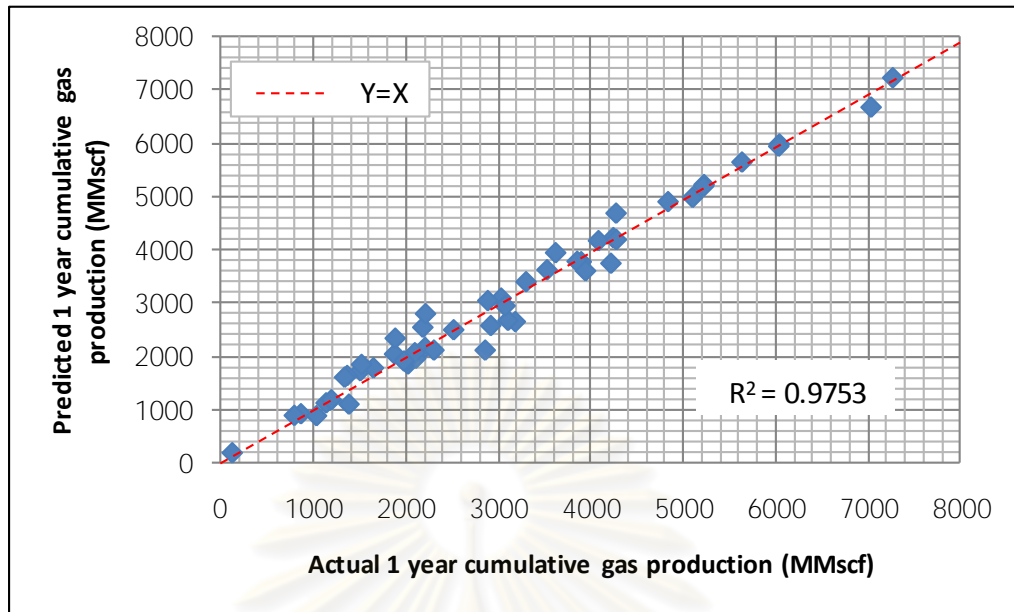


Figure 5.94: Cross plot of predicted vs actual 1 year cumulative gas of training sets of model 14 (Case 2-2-1)

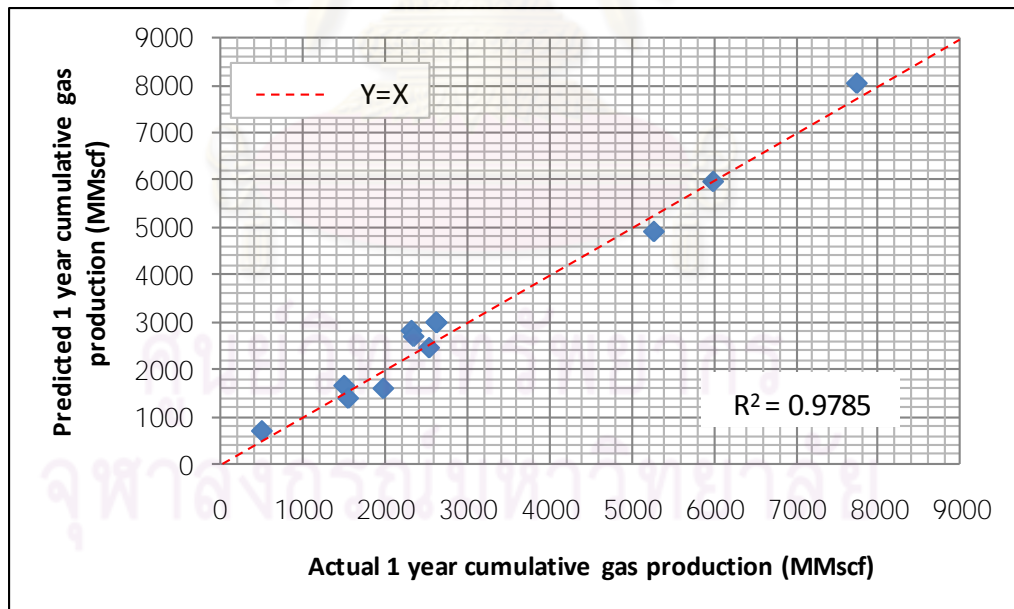


Figure 5.95: Cross plot of predicted vs actual 1 year cumulative gas of validating sets of model 14 (Case 2-2-1)

5.4.2.2.3 Model Testing Results and Discussion

In order to ensure the accuracy of ANN prediction when faced with unseen data sets, model 1 and 14 which yield the lowest MSE were then tested for accuracy by using testing data sets. After testing the ANN with testing sets, the outputs predicted by ANN are then compared with the target outputs by cross plotting them. Figures 5.96 and 5.97 represent the cross plots for model 1 and 14, respectively. From the graphs, R^2 of model 1 and 14 are equal to 0.7976 and 0.9100, respectively.

This means that both model 1 and 14 good performance of predicting the 1 year cumulative gas production. However, the coefficient of determination for model 14 is higher. Therefore, the best performance model which produces the most accurate predicted output is model 14. Consequently, this model will be used to predict 1-year cumulative gas production for well number 101 which is drilled 1 year after well number 100.

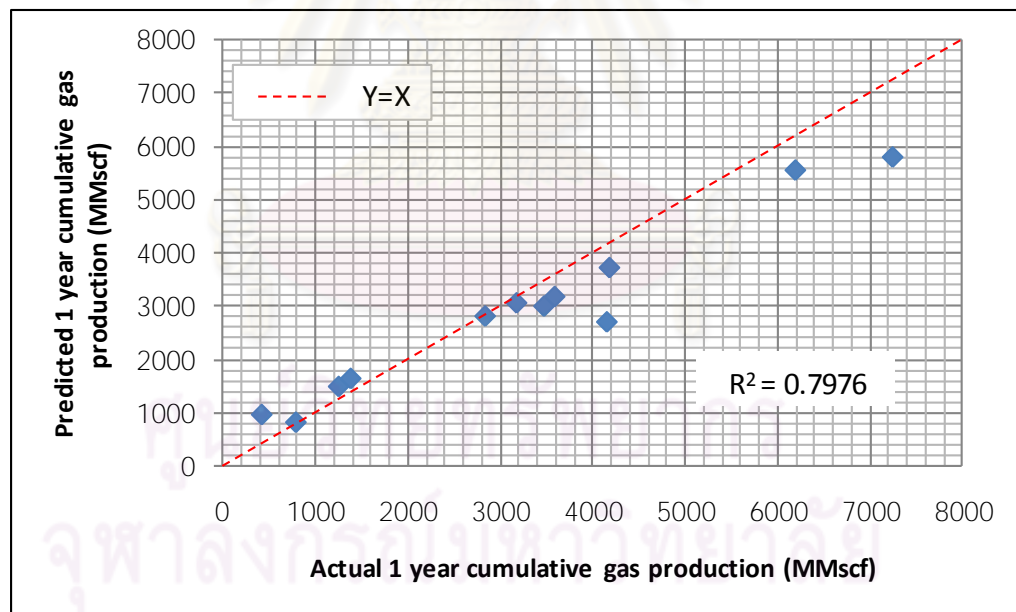


Figure 5.96: Cross plot of predicted vs actual 1 year cumulative gas of testing sets of model 1 (Case 2-2-1)

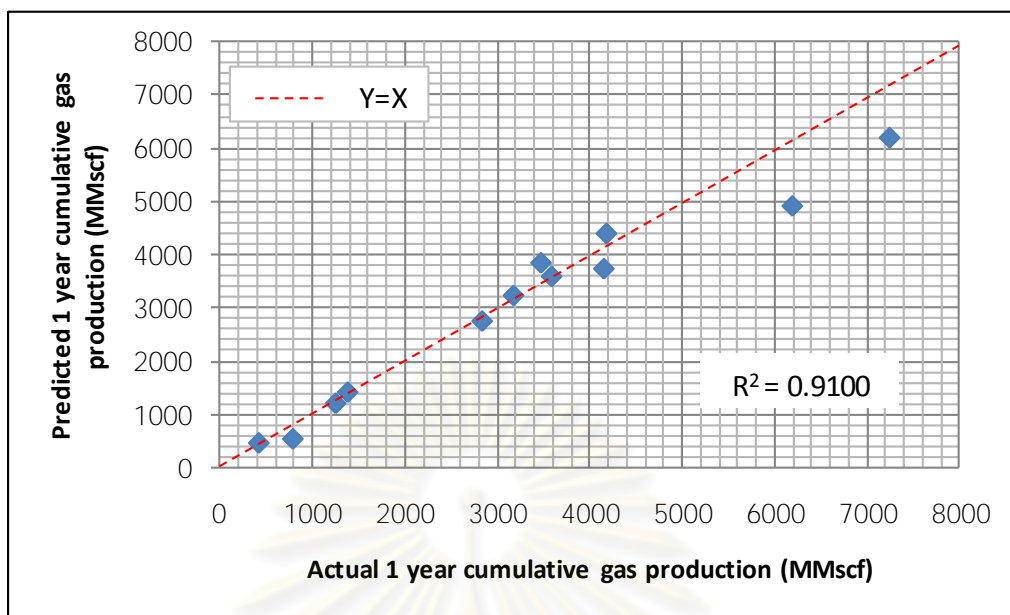


Figure 5.97: Cross plot of predicted vs actual 1 year cumulative gas of testing sets of model 14 (Case 2-2-1)

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

5.4.2.3 Case 2-2-2

This case study is the same as Case 2-2-1 except that the average pressure used in this study is inverse-distance average pressure. The calculation method has been discussed in Case 1-2-2 already. Figure 5.98 illustrates the schematic diagram of ANN in this case. The inverse-distance average pressure used in this study has the same values as the ones used in Case 1-2-2.

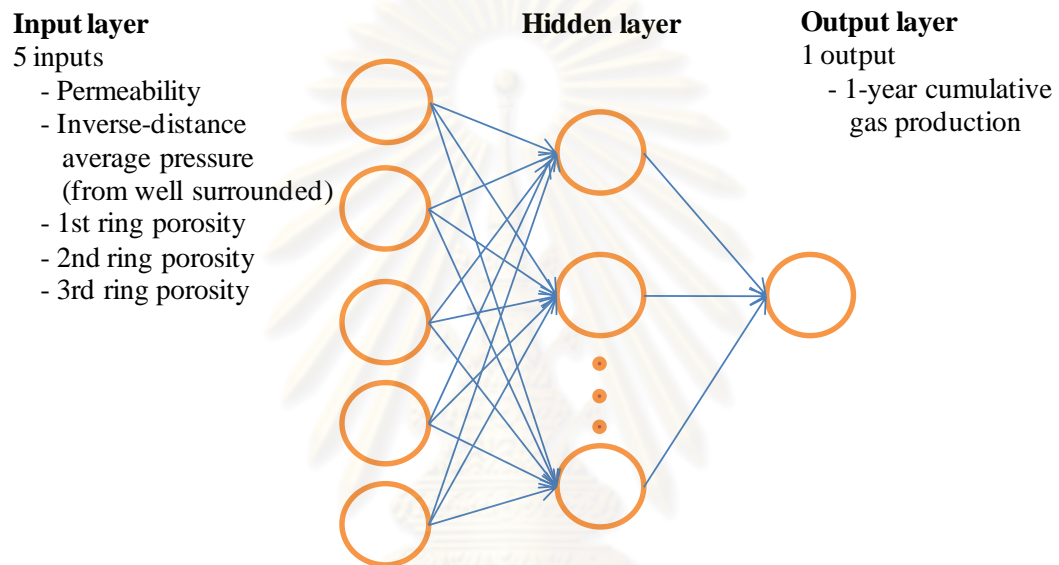


Figure 5.98: Schematic diagram of ANN Case 2-2-2

5.4.2.3.1 Data Preprocessing

Similar to the previous case, a total of 75 data sets taken from well number 26 to 100 were divided into three main sets namely, training, validating, and testing sets with ratio of 4:1:1 (51:12:12 data sets). The wells in each data set are still the same as the ones in the previous case study. All input parameters from all data sets are plotted to observe the distributions which are the same as those in Case 1-2-2.

5.4.2.3.2 Model Training

The ANN model was trained with various network configurations based on a trial and error basis. There are 20 models that were run in this case. Each model was trained many times to obtain the lowest MSE possible. The network configurations and their MSE of validating set are summarized in Table 5.15.

Table 5.15: Model configuration for Case 2-2-2

Model No	Number of neurons		Learning rate	Momentum	MSE
	Hidden Layer 1	Hidden Layer 2			
1	5	0	0.1	0.1	105,600,788,958
2	5	0	0.5	0.1	98,779,908,580
3	5	0	0.1	0.5	96,571,558,355
4	5	0	0.5	0.5	76,423,259,510
5	10	0	0.1	0.1	136,945,416,557
6	10	0	0.5	0.1	146,578,263,874
7	10	0	0.1	0.5	135,774,836,389
8	10	0	0.5	0.5	145,637,284,363
9	20	0	0.1	0.1	164,353,784,563
10	20	0	0.5	0.1	173,454,746,374
11	20	0	0.1	0.5	153,536,273,843
12	20	0	0.5	0.5	165,348,463,539
13	5	5	0.1	0.1	90,151,568,528
14	5	5	0.5	0.1	98,258,858,400
15	5	5	0.1	0.5	106,519,131,340
16	5	5	0.5	0.5	127,671,193,955
17	10	10	0.1	0.1	136,453,674,653
18	10	10	0.5	0.1	125,467,395,272
19	10	10	0.1	0.5	123,674,836,383
20	10	10	0.5	0.5	136,745,362,736

From a total 20 model configurations, two models with the lowest and next to lowest MSE (model 4 and 13) which has the MSE of 76,423,259,510 and 90,151,568,528, respectively, are chosen. Model 4 consists of only one hidden layers with 5 neurons. The learning rate and momentum of 0.5 were used. But model 13 consists of 2 hidden layers with 5 neurons in each layer. However, the learning rate and momentum were set to be 0.1. The performance curves of model 4 and 13 are

shown in Figures 5.99 and 5.100, respectively. Model 4 was trained until epoch 21 but the weight and bias were updated until epoch 6 only because the MSE of validating set started to increase in this epoch. The training was continued for 15 more epochs for validation check. The lowest MSE of model 4 is 76,423,259,510. For model 13, the training was performed until epoch 22. However, the weight and bias were not updated after epoch 7 due to the same reason for model 4. The lowest MSE of model 13 is 90,151,568,528 in epoch 7.

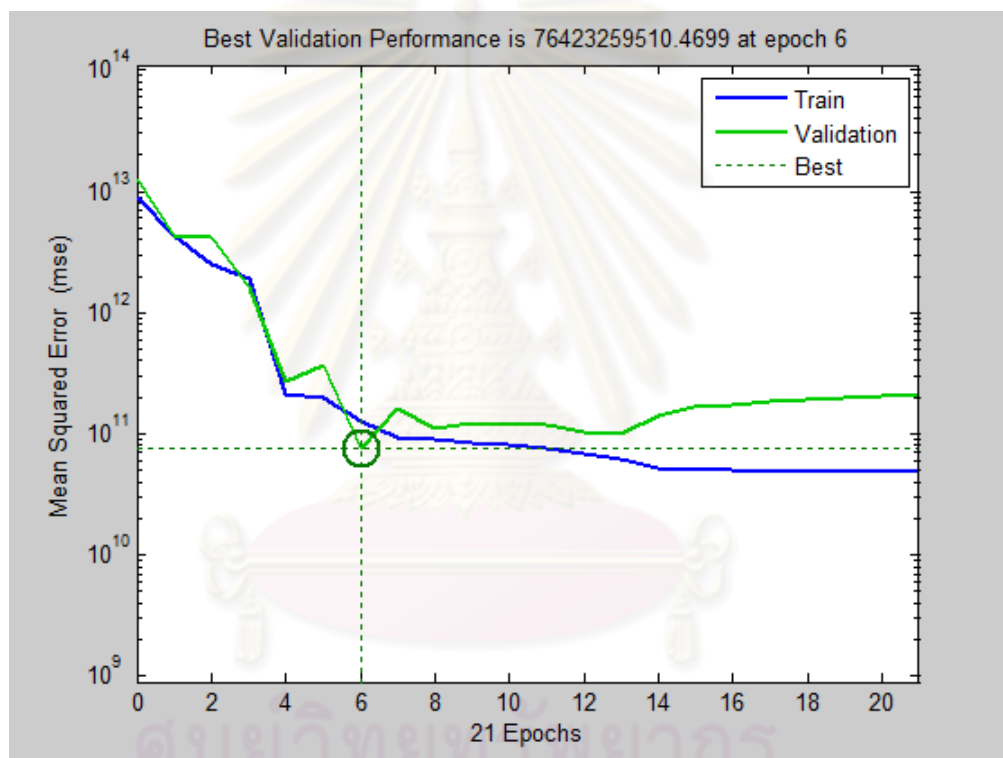


Figure 5.99: Performance curve of model 4 (Case 2-2-2)

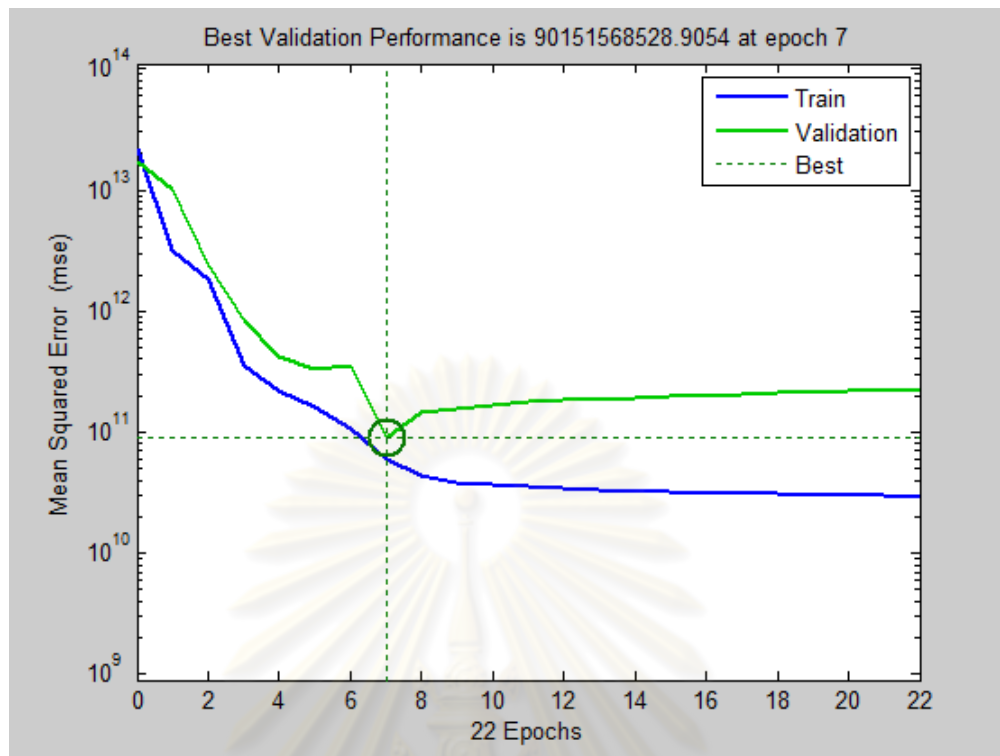


Figure 5.100: Performance curve of model 13 (Case 2-2-2)

Next, both models that produce the lowest MSE are further checked to ensure the accuracy of prediction. The outputs predicted by the ANN are compared with the target outputs of the training and validating sets by cross plotting them. Figures 5.101 and 5.102 represent the cross plots for the training and validating sets for model 4, respectively, while Figures 5.103 and 5.104 represent the cross plots for the training and validating sets for model 13, respectively. From the graph, the line $Y = X$ refers to correct prediction, i.e., each point on the 45-degree line is where predicted output is matched with the target output. So the closer the data points are located near the $Y = X$ line, the higher the accuracy of prediction. We can determine the accuracy of the prediction using regression coefficient of determination (R^2) as a criterion. R^2 equal to 1 represents a perfect fit to the $Y = X$ line. ANN will predict accurate output when R^2 is close to 1.

From Figures 5.101 to 5.104, we can see that the ANN can predict accurate output for both training and validating sets for both model 4 and 13. Model 4 gives R^2 of training and validating sets equal to 0.9455 and 0.9801, respectively, and model 13

gives R^2 of the training and validating sets equal to 0.9785 and 0.9775, respectively as well.

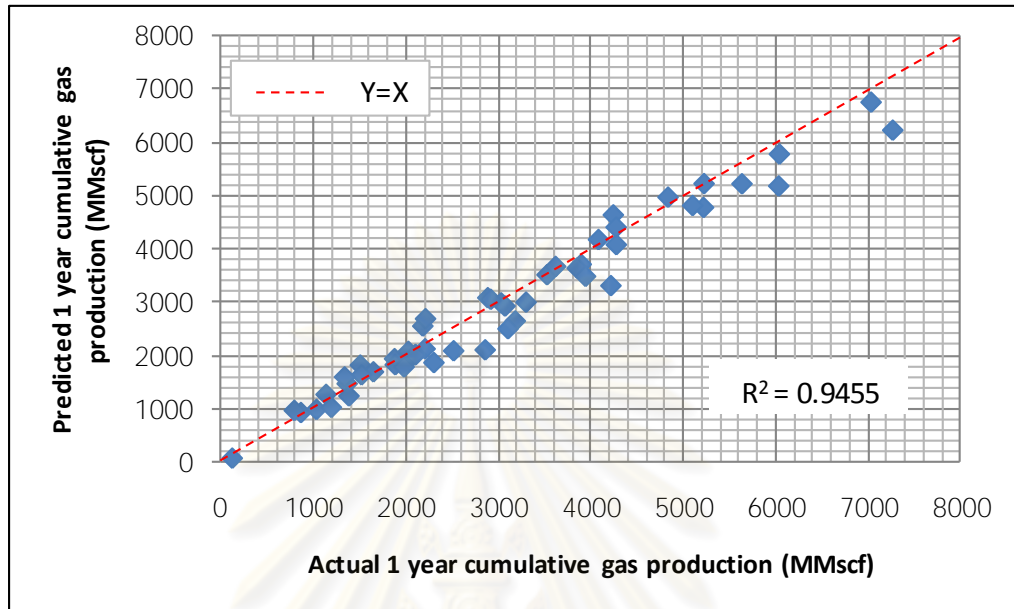


Figure 5.101: Cross plot of predicted vs actual 1 year cumulative gas of training sets of model 4 (Case 2-2-2)

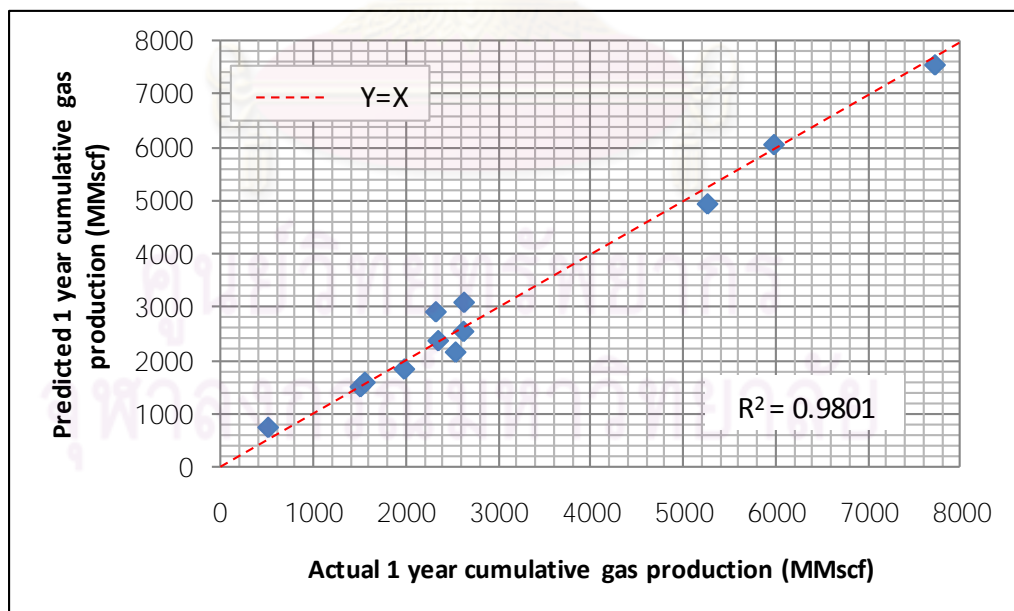


Figure 5.102: Cross plot of predicted vs actual 1 year cumulative gas of validating sets of model 4 (Case 2-2-2)

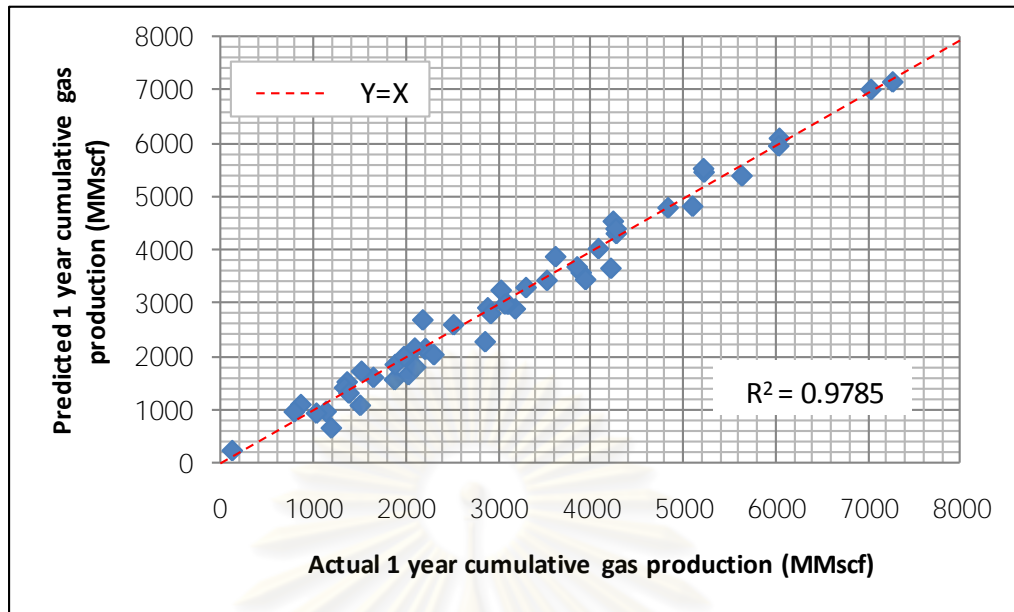


Figure 5.103: Cross plot of predicted vs actual 1 year cumulative gas of training sets of model 13 (Case 2-2-2)

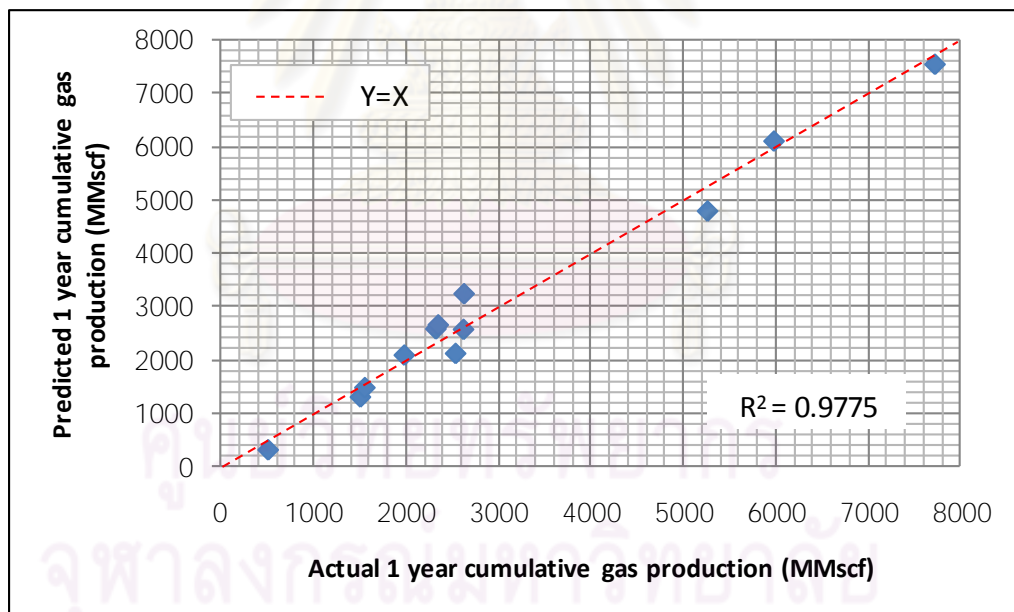


Figure 5.104: Cross plot of predicted vs actual 1 year cumulative gas of validating sets of model 13 (Case 2-2-2)

5.4.2.3.3 Model Testing Results and Discussion

In order to ensure the accuracy of ANN prediction when faced with unseen data sets, model 4 and 13 which yield the lowest MSE were then tested for accuracy by using testing data sets. After testing the ANN with testing sets, the outputs predicted by ANN are then compared with the target outputs by cross plotting them. Figures 5.105 and 5.106 represent the cross plots for model 4 and 13, respectively. From the graphs, R^2 of model 4 and 13 are equal to 0.8826 and 0.9524, respectively.

This means that both model 4 and 13 good performance of predicting the 1-year cumulative gas production. However, the coefficient of determination for model 13 is higher. Therefore, the best performance model which produces the most accurate predicted output is model 13. Consequently, this model will be used to predict 1-year cumulative gas production for well number 101 which is drilled 1 year after well number 100.

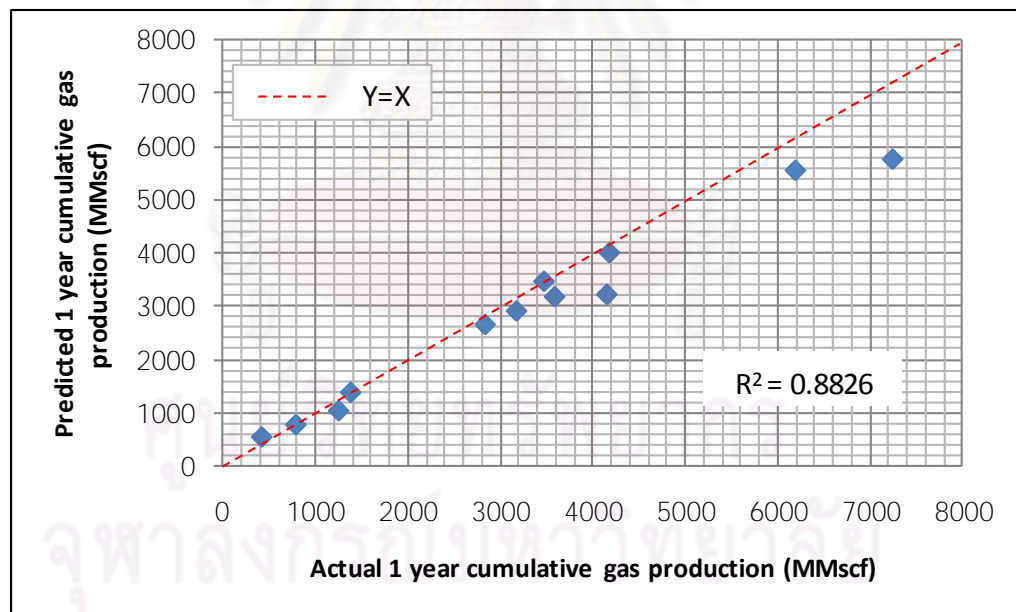


Figure 5.105: Cross plot of predicted vs actual 1 year cumulative gas of testing sets of model 4 (Case 2-2-2)

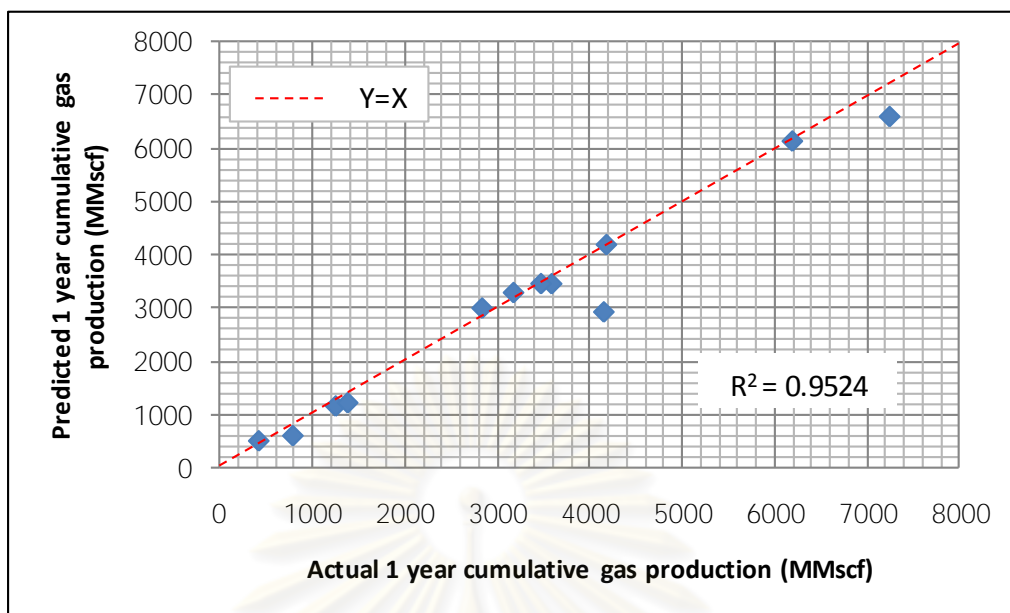


Figure 5.106: Cross plot of predicted vs actual 1 year cumulative gas of testing sets of model 13 (Case 2-2-2)

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

5.4.2.4 Case 2-3

In this case study, we use parameters which can refer to pressure at the predicting well location instead. These parameters are porosities of three different rings obtained by Geostatistics, drill date, and the number of surrounding wells. The porosities around the well are divided into 3 rings. The first ring covers an area of 100 x 100 ft at the center. The boundary of the second ring is located at 500 ft from the center in the x-and y-directions while the boundary of the third ring is 700 ft away from the center. Figure 5.107 illustrates the schematic diagram of ANN in this case.

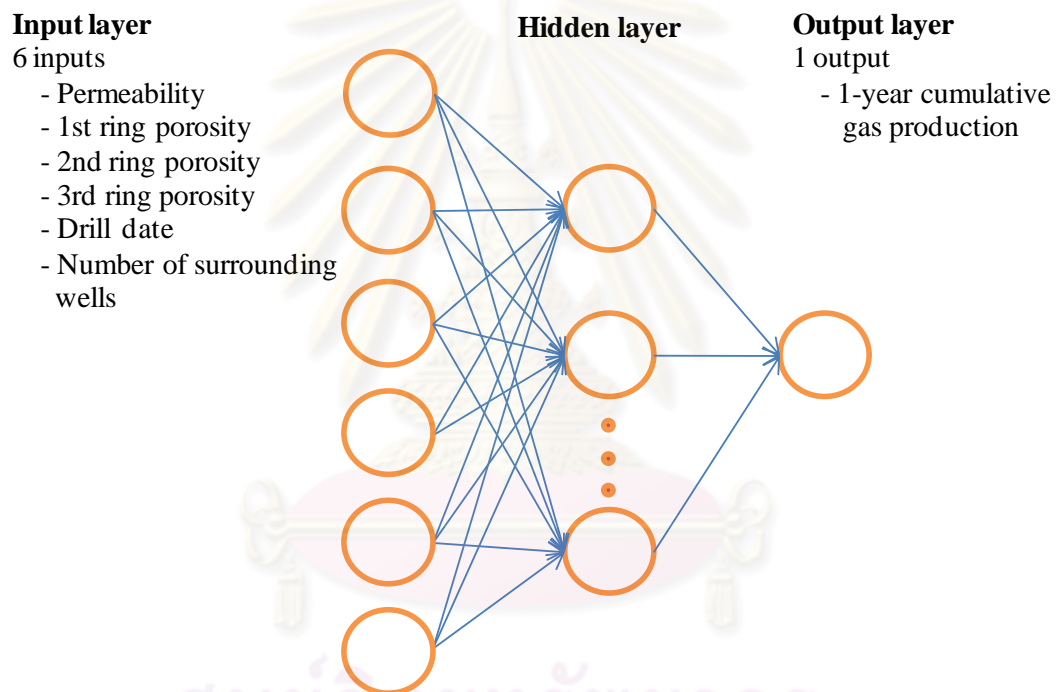


Figure 5.107: Schematic diagram of ANN for Case 2-3

5.4.2.4.1 Data Preprocessing

Similar to the previous case, a total of 75 data sets taken from well number 26 to 100 were divided into three main sets namely, training, validating, and testing sets with ratio of 4:1:1 (51:12:12 data sets). The wells in each data set are still the same as the ones in the previous case study. All input parameters from all data sets are plotted to observe the distributions which are the same as those in Case 1-3.

5.4.2.4.2 Model Training

The ANN model was trained with various network configurations based on a trial and error basis. There are 20 models that were run in this case. Each model was trained many times to obtain the lowest MSE possible. The network configurations and their MSE of validating set are summarized in Table 5.16.

Table 5.16: Model configuration for Case 2-3

Model No	Number of neurons		Learning rate	Momentum	MSE
	Hidden Layer 1	Hidden Layer 2			
1	5	0	0.1	0.1	32,028,257,296
2	5	0	0.5	0.1	24,933,493,346
3	5	0	0.1	0.5	30,035,281,420
4	5	0	0.5	0.5	39,114,195,592
5	10	0	0.1	0.1	40,453,647,436
6	10	0	0.5	0.1	41,234,353,647
7	10	0	0.1	0.5	42,647,463,648
8	10	0	0.5	0.5	39,454,363,738
9	20	0	0.1	0.1	42,356,474,874
10	20	0	0.5	0.1	41,267,483,738
11	20	0	0.1	0.5	50,674,849,476
12	20	0	0.5	0.5	49,837,345,637
13	5	5	0.1	0.1	51,414,438,904
14	5	5	0.5	0.1	38,253,881,774
15	5	5	0.1	0.5	27,416,977,226
16	5	5	0.5	0.5	50,884,150,400
17	10	10	0.1	0.1	52,678,493,467
18	10	10	0.5	0.1	51,324,647,657
19	10	10	0.1	0.5	52,345,363,467
20	10	10	0.5	0.5	50,896,356,738

From a total 20 model configurations, two models with the lowest and next to lowest MSE (model 2 and 15) which has the MSE of 24,933,493,346 and 27,416,977,226, respectively, are chosen. Model 2 consists of only one hidden layers with 5 neurons. The learning rate and momentum of 0.5 and 0.1, respectively, were used. But model 15 consists of 2 hidden layers with 5 neurons in each layer. However, the learning rate and momentum were set to be 0.1 and 0.5, respectively. The

performance curves of model 2 and 15 are shown in Figures 5.108 and 5.109, respectively. Model 2 was trained until epoch 19 but the weight and bias were updated until epoch 4 only because the MSE of validating set started to increase in this epoch. The training was continued for 15 more epochs for validation check. The lowest MSE of model 2 is 24,933,493,346. For model 15, the training was performed until epoch 20. However, the weight and bias were not updated after epoch 5 due to the same reason for model 2. The lowest MSE of model 15 is 27,416,977,226 in epoch 5.

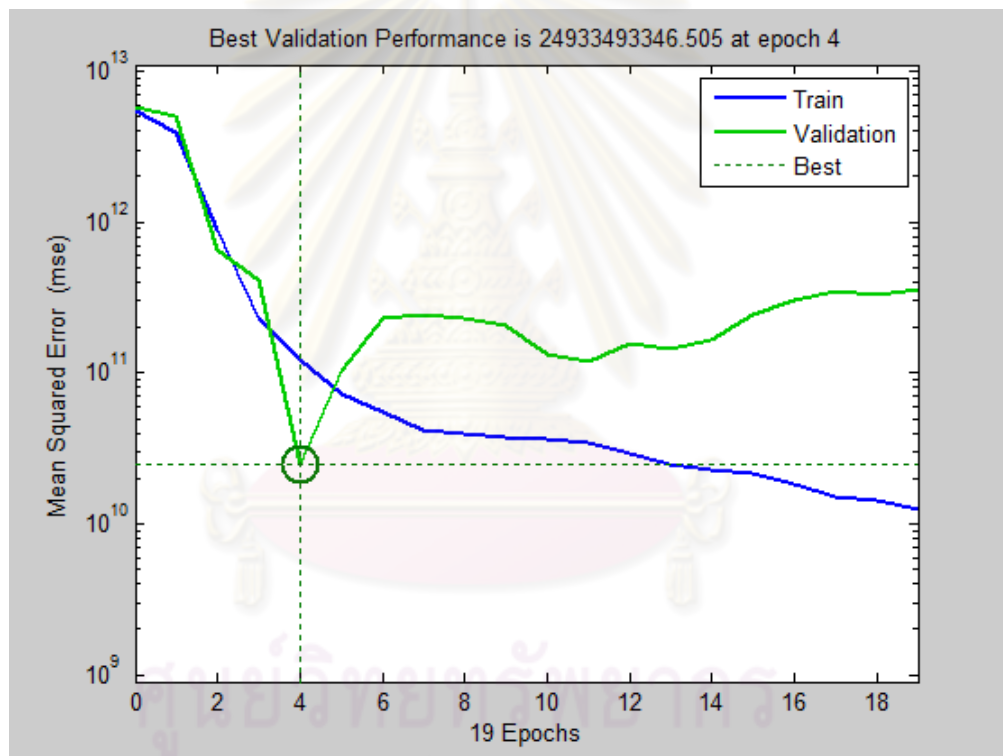


Figure 5.108: Performance curve of model 2 (Case 2-3)

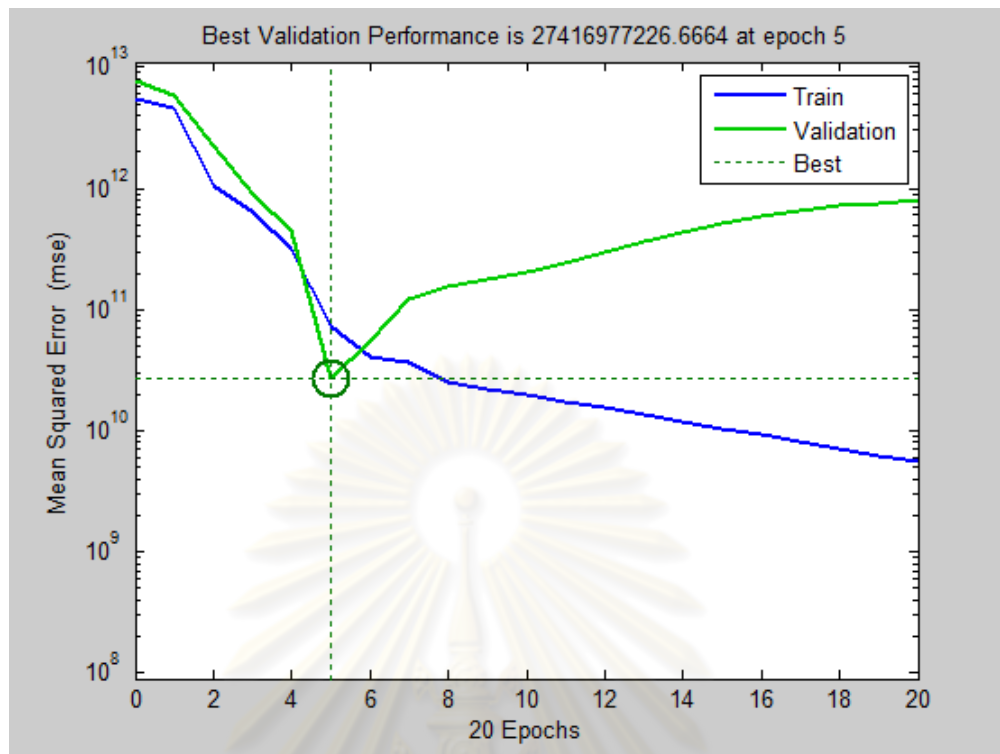


Figure 5.109: Performance curve of model 15 (Case 2-3)

Next, both models that produce the lowest MSE are further checked to ensure the accuracy of prediction. The outputs predicted by the ANN are compared with the target outputs of the training and validating sets by cross plotting them. Figures 5.110 and 5.111 represent the cross plots for the training and validating sets for model 2, respectively, while Figures 5.112 and 5.113 represent the cross plots for the training and validating sets for model 15, respectively. From the graph, the line $Y = X$ refers to correct prediction, i.e., each point on the 45-degree line is where predicted output is matched with the target output. So the closer the data points are located near the $Y = X$ line, the higher the accuracy of prediction. We can determine the accuracy of the prediction using regression coefficient of determination (R^2) as a criterion. R^2 equal to 1 represents a perfect fit to the $Y = X$ line. ANN will predict accurate output when R^2 is close to 1.

From Figures 5.110 to 5.113, we can see that the ANN can predict accurate output for both training and validating sets for both model 2 and 15. Model 2 gives R^2 of training and validating sets equal to 0.9549 and 0.9940, respectively, and model 15

gives R^2 of the training and validating sets equal to 0.9731 and 0.9936, respectively as well.

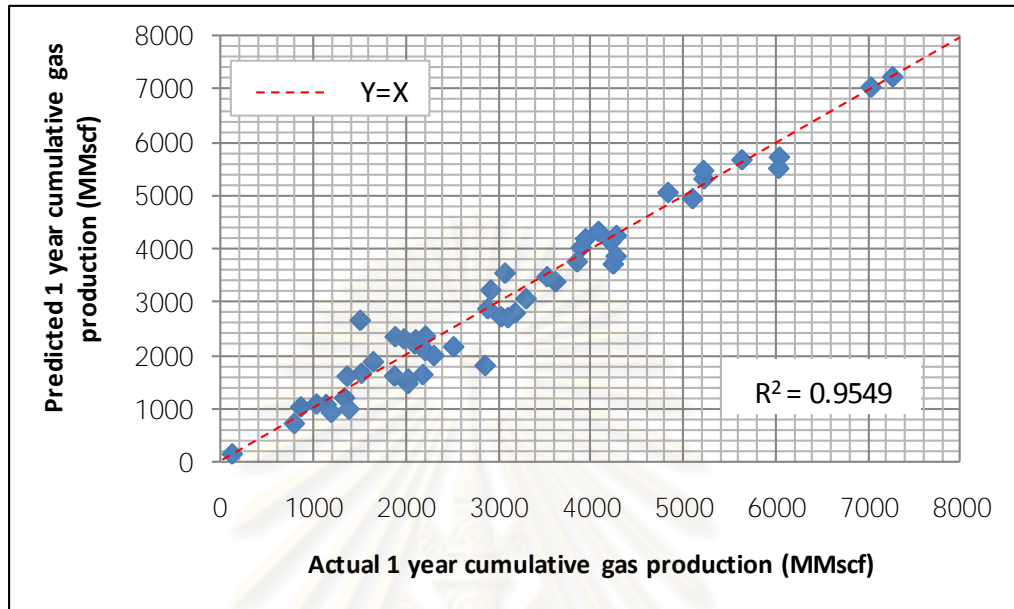


Figure 5.110: Cross plot of predicted vs actual 1 year cumulative gas of training sets of model 2 (Case 2-3)

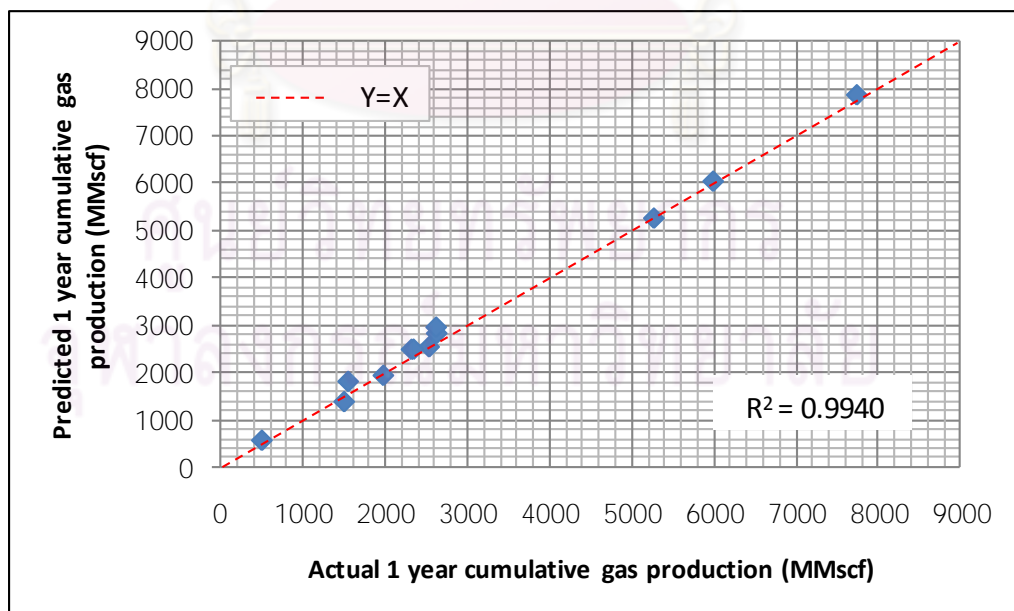


Figure 5.111: Cross plot of predicted vs actual 1 year cumulative gas of validating sets of model 2 (Case 2-3)

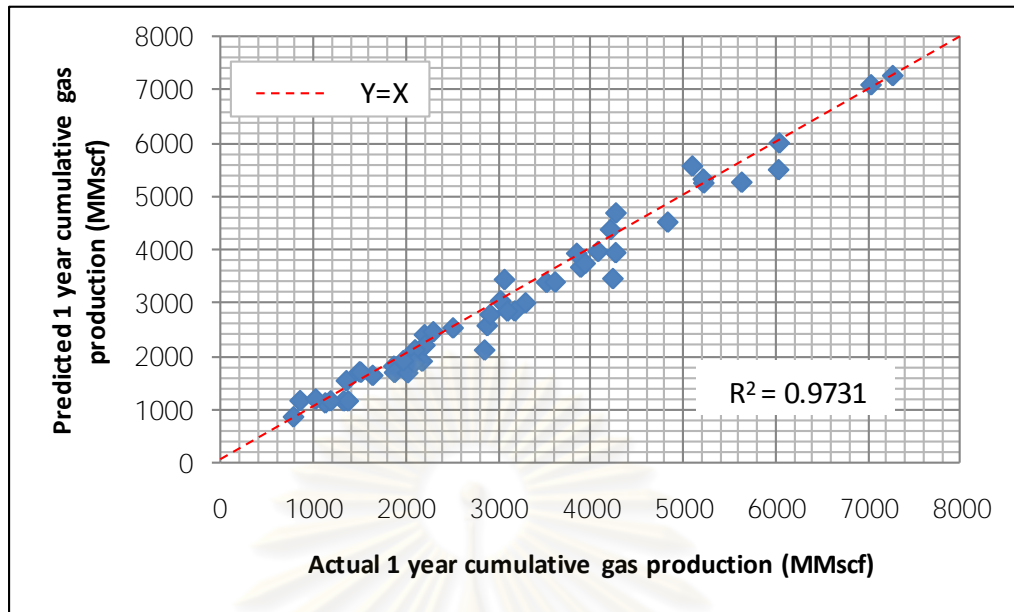


Figure 5.112: Cross plot of predicted vs actual 1 year cumulative gas of training sets of model 15 (Case 2-3)

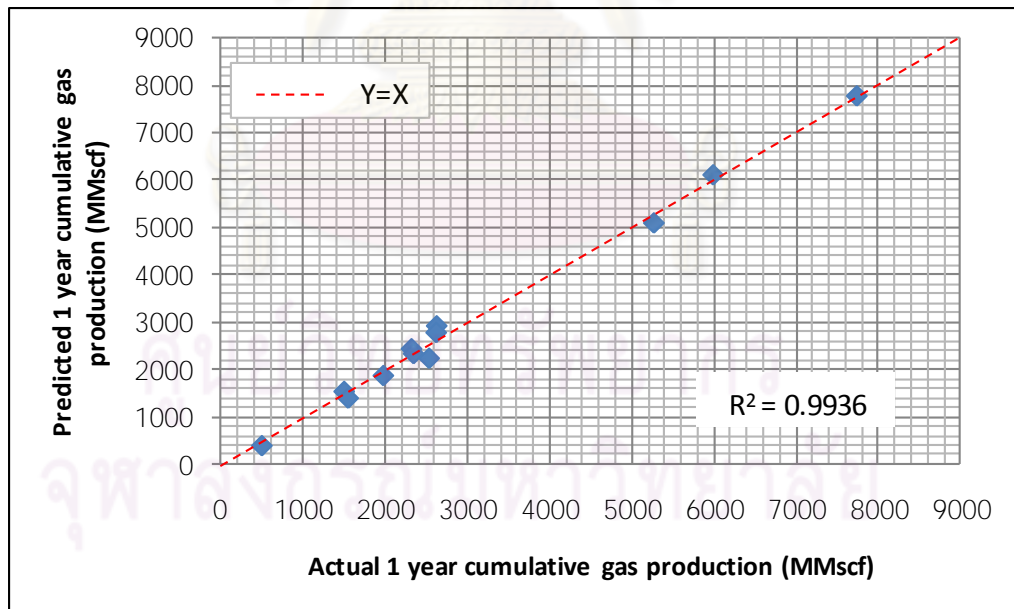


Figure 5.113: Cross plot of predicted vs actual 1 year cumulative gas of validating sets of model 15 (Case 2-3)

5.4.2.4.3 Model Testing Results and Discussion

In order to ensure the accuracy of ANN prediction when faced with unseen data sets, model 2 and 15 which yield the lowest MSE were then tested for accuracy by using testing data sets. After testing the ANN with testing sets, the outputs predicted by ANN are then compared with the target outputs by cross plotting them. Figures 5.114 and 5.115 represent the cross plots for model 2 and 15, respectively. From the graphs, R^2 of model 2 and 15 are equal to 0.9654 and 0.9750, respectively.

This means that both model 2 and 15 good performance of predicting the 1 year cumulative gas production. However, the coefficient of determination for model 15 is higher. Therefore, the best performance model which produces the most accurate predicted output is model 15. Consequently, this model will be used to predict 1-year cumulative gas production for well number 101 which is drilled 1 year after well number 100.

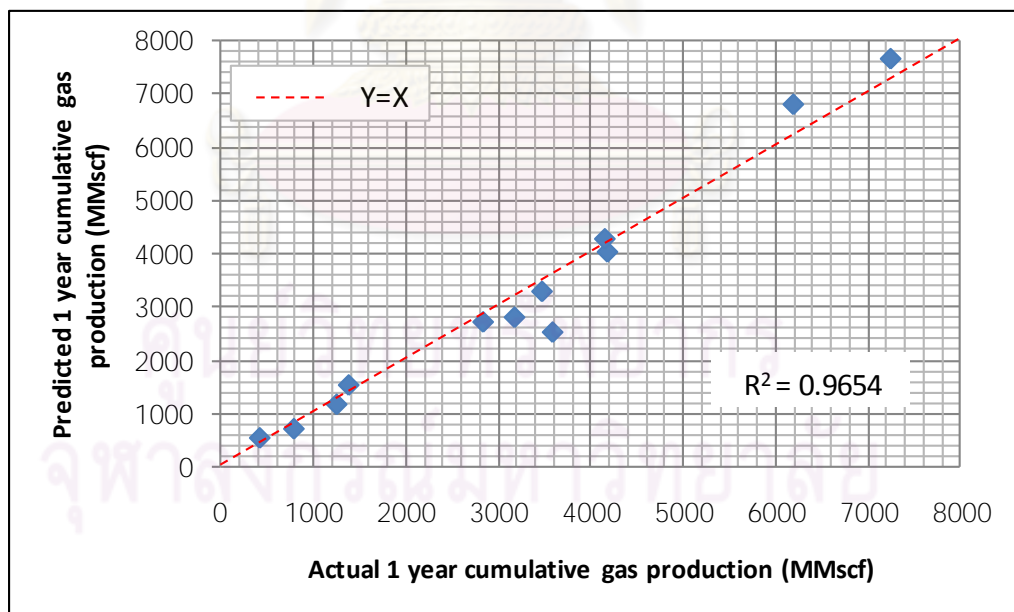


Figure 5.114: Cross plot of predicted vs actual 1 year cumulative gas of testing sets of model 2 (Case 2-3)

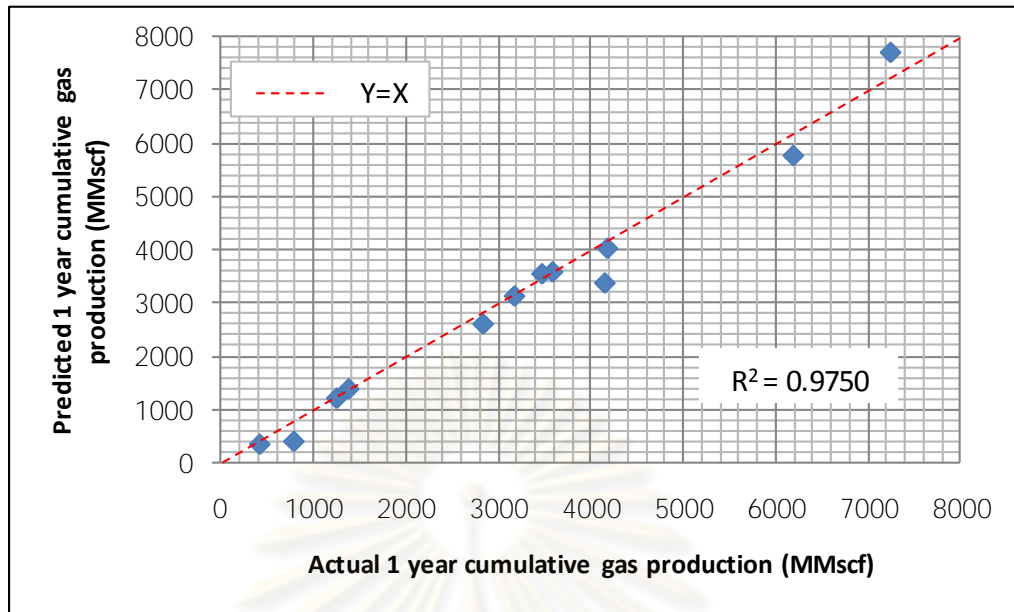


Figure 5.115: Cross plot of predicted vs actual 1 year cumulative gas of testing sets of model 15 (Case 2-3)

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

5.4.2.5 Performance of ANN Prediction

After several case studies with different input parameters were performed, we use each best performance model to predict the 1 year cumulative gas production of the new well planned to be drill (well number 101) which is drilled 1 year after well number 100. Table 5.17 summarize the best performance model for each case study.

Table 5.17: Summarized of best performance model for each case.

Case No.	Model No.	R ²		
		Training sets	Validating sets	Testing sets
2-1	2	0.9921	0.9965	0.9863
2-2-1	14	0.9753	0.9785	0.9100
2-2-2	13	0.9785	0.9775	0.9524
2-3	15	0.9731	0.9936	0.9750

The group of candidate locations to drill well number 101 is the same locations as described earlier in Case 1. After the predictions have been performed, the outputs predicted by the ANN are compared with the target output taken from reservoir simulation by cross plotting them. Figures 5.116, 5.117, 5.118, and 5.119 represent the results for Case 2-1, 2-2-1, 2-2-2, and 2-3, respectively.

In Figures 5.116 to 5.119, several points are located near the $Y=X$ line, representing good prediction. However, many points are located far away from this line (some points even have negative values). Predictions of negative flow rates occur when the actual cumulative gas productions are small (less than 100 MMscf for Case 2-1, around 300 MMscf for Case 2-2-1, and 200 MMscf for Case 2-2-2).

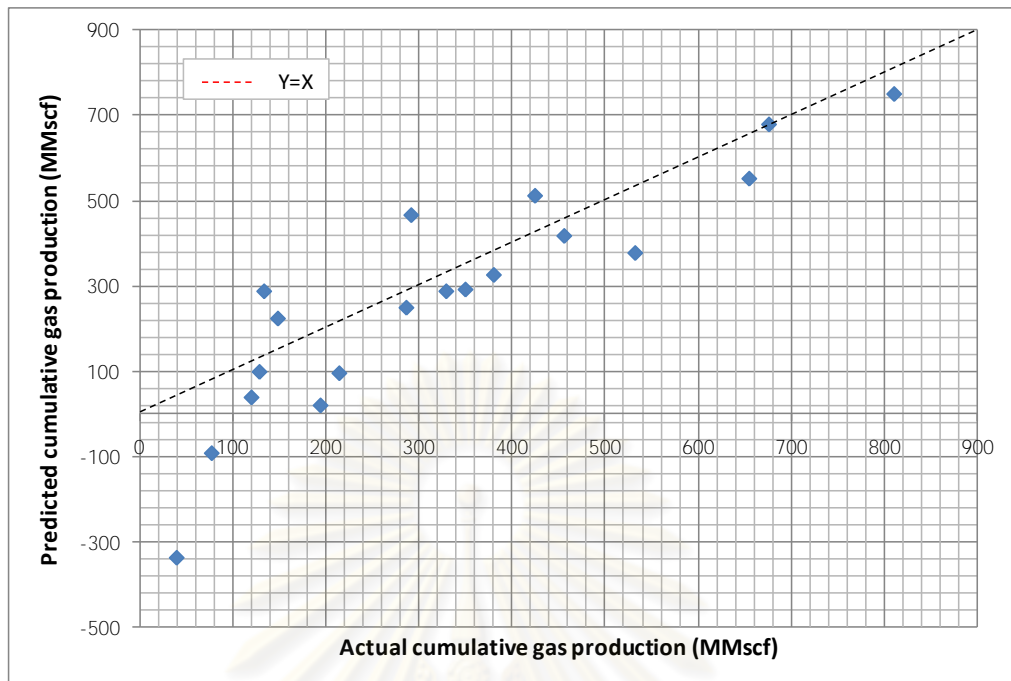


Figure 5.116: Cross plot of predicted vs actual cumulative gas production for candidate well location (Case 2-1)

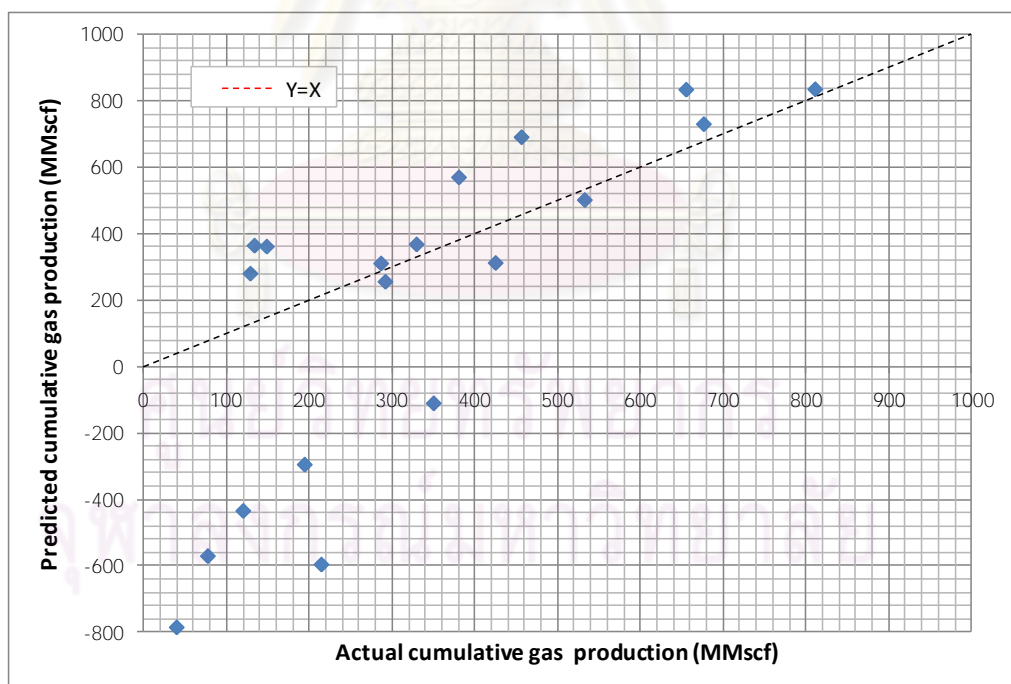


Figure 5.117: Cross plot of predicted vs actual cumulative gas production for candidate well location (Case 2-2-1)

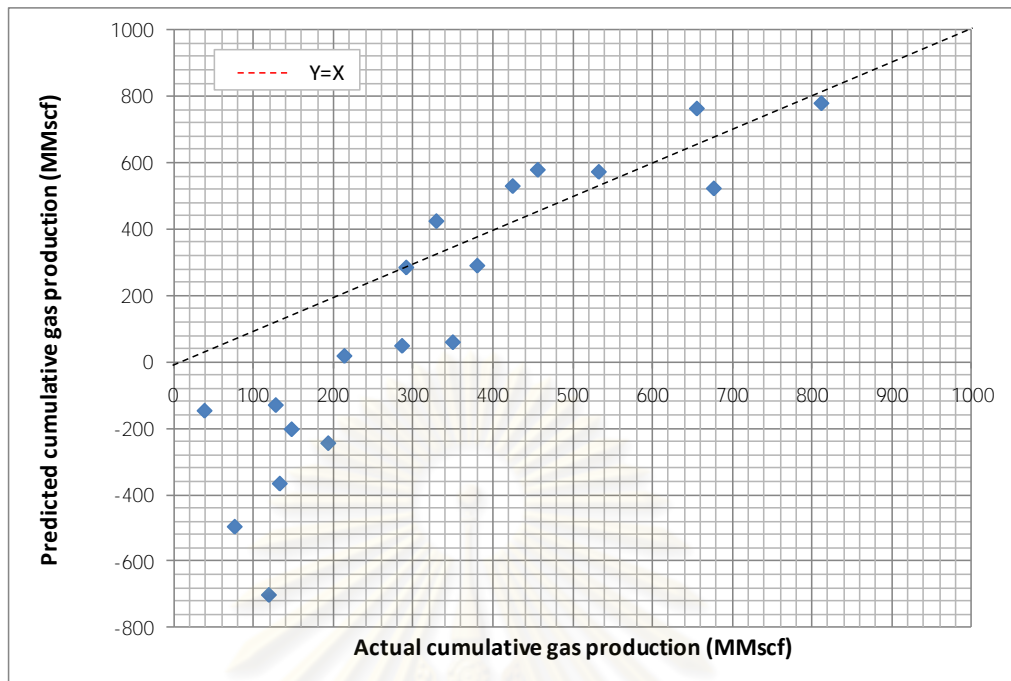


Figure 5.118: Cross plot of predicted vs actual cumulative gas production for candidate well location (Case 2-2-2)

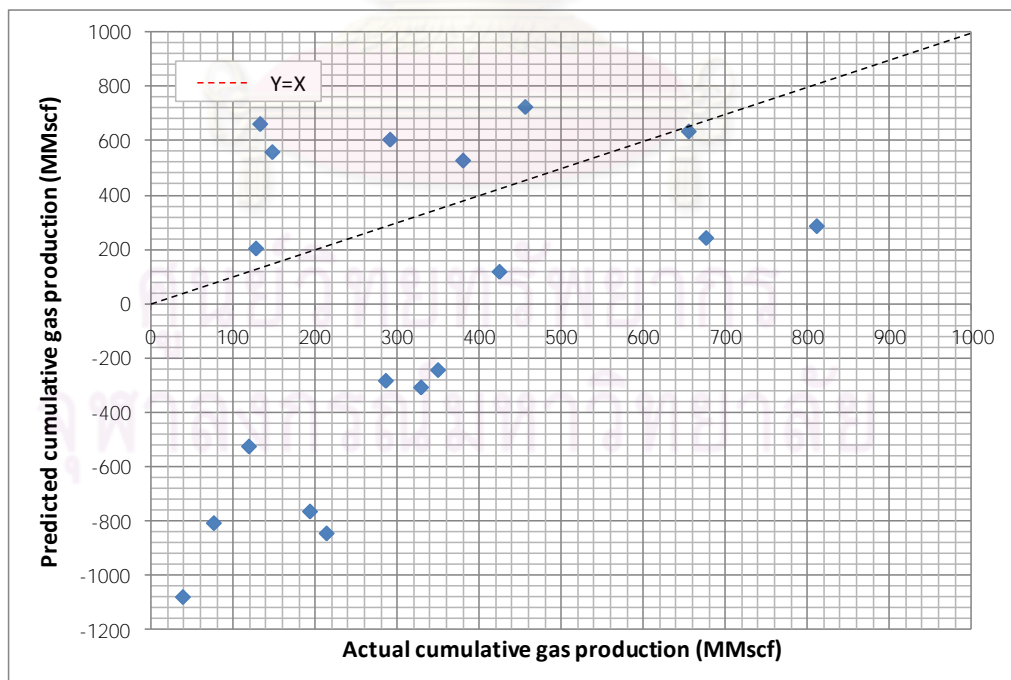


Figure 5.119: Cross plot of predicted vs actual cumulative gas production for candidate well location (Case 2-3)

From the study of Hettiarachchi et.al^[17], they found a problem when using the ANN to predict the relationship between rainfall and streamflow. A problem arises in extrapolation, i.e., the prediction is not accurate when the training set does not contain the maximum or minimum possible input and output values. Therefore, if we use the trained ANN to predict the output that is out of range or using the input which is out of range, the prediction is inaccurate.

In this study, the type of reservoir is closed boundary depletion drive gas reservoir. Both the pressure and cumulative gas production continues to decrease as a function of time. The pressure at well number 101 which is drilled 1 year afterward may be lower than the minimum value of input in training sets. Therefore, the prediction is based on extrapolation, causing inaccurate prediction. With this reason, the ANN model that uses pressure as an input parameter (Case 2-1, 2-2-1, 2-2-2) will show a good prediction only when the pressure at that location is not much lower than the minimum pressure in the training data set.

From 4 cases, Case 2-1 is the best predictive model for this study. But as described earlier that the pressure at the location to be drilled is not known prior to drilling. Therefore, the trained ANN of Case 2-1 cannot be used as a tool to predict the 1-year cumulative gas production. After comparing all remaining cases, Case 2-3 presents the most inaccurate model. So, we will choose the best prediction model between Case 2-2-1 and Case 2-2-2. Figures 5.120 and 5.121 show the cross plot (only positive value) for Case 2-2-1 and Case 2-2-2, respectively.

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

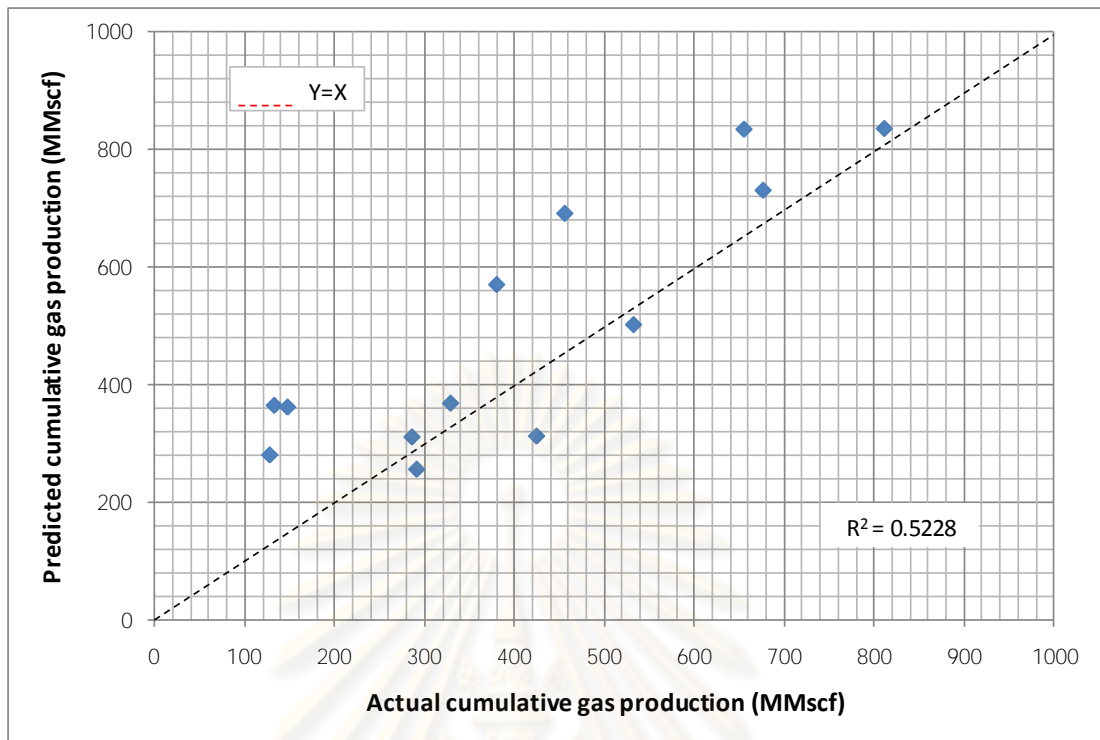


Figure 5.120: Cross plot of predicted vs actual initial gas rate for candidate well locations (positive value of Case 2-2-1)

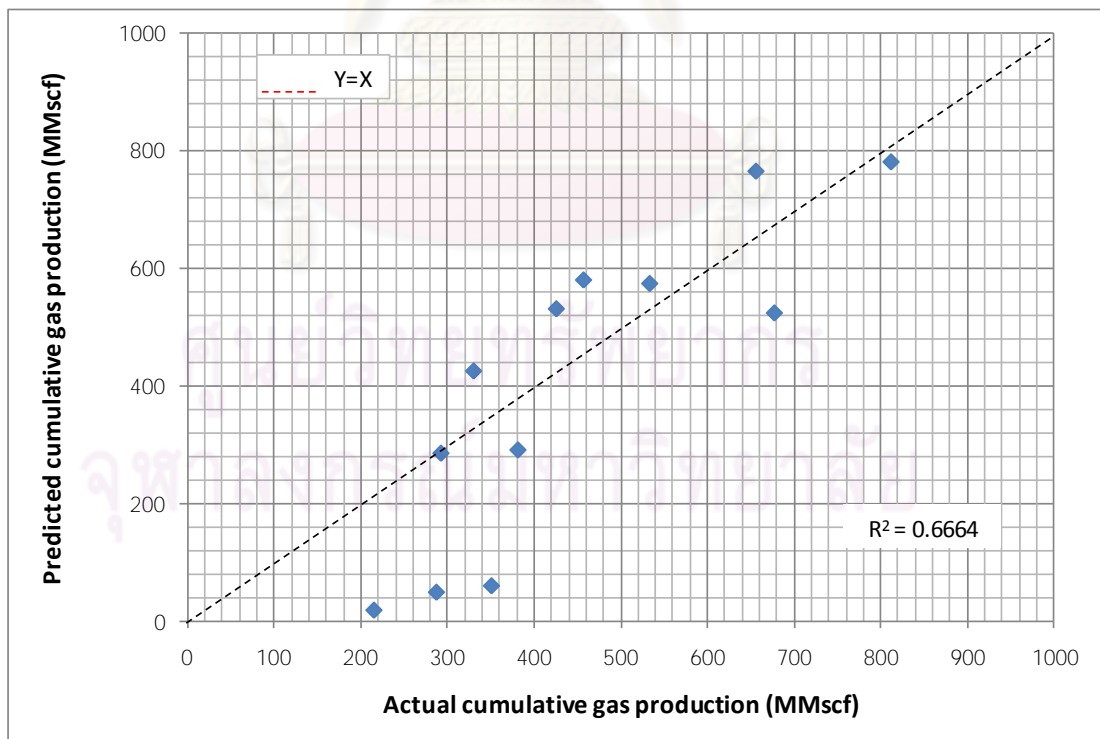


Figure 5.121: Cross plot of predicted vs actual initial gas rate for candidate well locations (positive value of Case 2-2-2)

From Figures 5.120 and 5.121, R^2 of Case 2-2-1 and Case 2-2-2 are equal to 0.5228 and 0.6664, respectively. Therefore, the best performance model which produces the most accurate predicted output is Case 2-2-2.

Although, the prediction of ANN does not give us accurate result for all well locations, we only need one location to infill which is the location that gives the highest 1 year cumulative gas production. This location is associated with high pressure. Therefore, we can use the ANN to roughly estimate the 1 year cumulative gas production for this location.

At this point, we are able to predict 1 year cumulative gas productions at candidate well locations as illustrated in Figure 5.79.

In order to determine the best location to drill well number 101, the predicted 1 year cumulative gas production obtained from each ANN model is ranked and tabulated as shown in Table 5.18.

Table 5.18: Order of candidate location from a higher to lower 1 year cumulative gas production

Order	Candidate location for well number 101				
	Reality	Case 2-1	Case 2-2-1	Case 2-2-2	Case 2-3
1	location 9	location 9	location 9	location 9	location 5
2	location 13	location 13	location 1	location 1	location 17
3	location 1	location 1	location 13	location 17	location 4
4	location 5	location 7	location 17	location 5	location 1
5	location 17	location 11	location 19	location 7	location 11
6	location 7	location 17	location 5	location 13	location 18
7	location 19	location 5	location 16	location 16	location 19
8	location 8	location 19	location 4	location 19	location 9
9	location 16	location 8	location 18	location 11	location 13
10	location 11	location 16	location 7	location 8	location 2
11	location 6	location 4	location 6	location 6	location 7
12	location 12	location 6	location 2	location 12	location 8
13	location 14	location 18	location 11	location 2	location 6
14	location 18	location 2	location 8	location 10	location 16
15	location 4	location 12	location 14	location 18	location 3
16	location 2	location 3	location 3	location 14	location 14
17	location 3	location 14	location 15	location 4	location 15
18	location 15	location 15	location 15	location 15	location 15
19	location 10	location 10	location 10	location 3	location 10

In reality, location 9 is the one that yields the highest 1 year cumulative gas production and the top 3 candidate locations to infill a well are location 9, 13, and 1, respectively. Case 2-1 and 2-2-1 which use the actual and arithmetic average pressure as input parameter can predict the same group of top 3 locations as in reality with the right best location at location 9. Case 2-2-2 which uses inverse-distance average pressure as input parameter can also predict the right best location as in reality but different in group of top 3 locations. Case 2-3 which does not use any pressure as input parameter cannot predict the right location. Top 3 locations of this case are location 5, 17, and 4 which are the 4th, 5th, and 15th order in reality, respectively. Because location 9 is predicted by both Case 2-2-1 and 2-2-2 while location 5 is predicted by only Case 2-3. Therefore, location 9 will be used as the best location to infill the well that matched with the best location in reality.

Table 5.19: Summary of error in 1 year cumulative gas production predicted for location 9.

Case	1 year cumulative gas production (MMscf)	Error (%)
Reality	811.10	-
2-2-1	834.95	2.94
2-2-2	780.47	-3.78
2-3	285.08	-64.85

Table 5.19 summarises the error of prediction when using each model to predict 1 year cumulative gas production of candidate location 9. From Table 5.19, the 1 year cumulative gas production obtained from Case 2-2-1 is quite high when comparing with the reality case. In the other hand, the value is quite low when using Case 2-2-2 and 2-3. Even though the error may be high, we achieve the objective of being able to determine very good location for the next infill.

5.4.3 Location for Infill Well Number 101

From Section 5.4.1 and 5.4.2 which use ANN to predict the initial gas production rate and 1 year cumulative gas production, respectively, we obtain a location to infill the well number 101 for each case as follows:

Table 5.20: Summary of best candidate well location for each case

Case	Predicted output	Best well location
1	Initial gas production rate	location 1
Reality		location 9
2	1 year cumulative gas production	location 9
Reality		location 9

From Table 5.20, based a different predicted output, different best well location is obtained. As described in Section 5.1 about output of ANN, initial gas production rate indicates short term performance while 1 year cumulative gas production indicates long term performance. Normally, gas companies always prefer long term performance rather than short term performance. Consequently, we should drill the infill well at candidate location 9 which is predicted by ANN model that uses 1 year cumulative gas production as the output. Finally, we can choose the best location for an infill well.

CHAPTER VI

CONCLUSIONS AND RECOMMENDATIONS

6.1 Conclusions

The purpose of this study is to use Artificial Neural Network (ANN) as a tool to predict the gas production of infill wells. The predicted results will be used to determine the best location for a new well. The data sets used to train the network were prepared from a reservoir simulation, generated by reservoir and fluid properties referred from a gas field in Gulf of Thailand. There are 5 main steps needed to be performed. First, specify input and output parameters for ANN by reviewing related literature and theories and obtain data for these parameters. Second, partition and rearrange the data sets to ensure that the training, validating, and testing sets have similar distribution. Third, develop an ANN model with various kinds of network configuration. Fourth, choose 2 best models to test with the testing set. Fifth, use the best performance model to predict the performance of an infill well. The conclusions from these model developments are summarized as follows:

1) For a closed boundary depletion drive gas reservoir, the main factors which directly affect the prediction of initial gas production rate and cumulative gas production are pressure and permeability at the predicting well location.

2) Although the pressure at the predicting well location may be unknown prior to drilling but we can use arithmetic average and inverse-distance average pressure of surrounding wells, porosity, drill date, and the number of surrounding wells as input. The result shows accurate prediction when tested with the testing data sets and when used to predict the initial gas rates and cumulative gas production at locations where the pressure is not much lower than the pressure used in the training data sets.

3) A problem arises when using the ANN to predict the output based on extrapolation. Since the pressure continues to decrease as a function of time, the pressure at the infill location is smaller than the ones used in the training, validating, and testing processes. This causes error in the prediction of initial gas rate and cumulative gas production.

6.2 Recommendations

The recommendations for future work are summarized as follow:

1) This study focused on many fixed parameters such as well configuration, layer thickness etc. Therefore, the study may be extended to include variations in these parameters in order to obtain data sets which are more representative of an actual field as much as possible.

2) Closed boundary depletion drive gas reservoir seems to have problem with extrapolation due to the decline in reservoir pressure along with time. Therefore, using this methodology in the other type of reservoir which has smaller change in pressure such as water drive gas reservoir may provide better results due to less extrapolation problem.

3) The larger the number of input parameters, the larger the number of training data sets. The ANN cannot predict good results if it lacks a sufficient number of training data sets. This must be taken into consideration when training an ANN.

REFERENCES

- [1] Coats, K.H. An Approach to Locating New Wells in Heterogeneous, Gas Producing Fields. Paper SPE 2264 presented at 43rd Annual SPE Fall Meeting, Houston Sept. 29-October. 2, 1968.
- [2] McCain, W.D. and Voneiff, G.W. A Tight Gas Field Study; Carthage (Cotton Valley) Field. Paper SPE 26141 presented at the SPE Gas Technology Symposium held in Calgary, Alberta, Canada 28-30 June 1993.
- [3] Voneiff, G.W and Cipolla, C. A New Approach to Large-Scale Infill Evaluations Applied to the Ozona (Canyon) Gas Sands. Paper SPE 35203, presented at the Permian Basin Oil & Gas Recovery Conference held in Midland, Texas 27-29 March 1996.
- [4] Guan, L., McVay, D.A., Jenson, J.L., and Voneiff, G.W. Evaluation of Statistical Infill Candidate Selection Technique. Paper SPE 75718, presented at the SPE Gas Technology Symposium held in Calgary, Alberta, Canada 30 April-2 May 2002.
- [5] Gao, H. and McVay, D.A. Gas Infill Well Selection Using Rapid Inversion Methods. Paper SPE 90545, presented at the Annual Technical Conference and Exhibition, Houston, TX, USA. 26-29 September 2004.
- [6] Guan, L., Gao, H., Wang, Z., and Du, Y. New Methods for Determining Infill Drilling Potentials in Large Tight Gas Basins. Paper PETSOC -2005-059-P, presented at the Petroleum Society's 6th Canadian International Petroleum Conference (56th Annual Technical Meeting), Calgary, Alberta, Canada. 7-9 June 2005.
- [7] Robert J. B. Predicting Production Using a Neural Network (Artificial Intelligence Beats Human Intelligence). Paper SPE 30202, presented at the Petroleum Computer Conference held in Houston, TX, USA. 11-14 June 1995
- [8] Fattah, S.M. and Startzman, R.A. Predicting Natural Gas Production Using Artificial Neural Network. Paper SPE 68593, presented at the SPE

Hydrocarbon Economics and Evaluation Symposium, Dallas, Texas. 2-3 April 2001.

- [9] Sampaio, T.P., Ferreira, V.J.M., De Sa Neto, A. An Application of Feed Forward Neural Network as Nonlinear Proxies for the Use During the History Matching Phase. Paper SPE 122148, presented at the SPE Latin American and Caribbean Petroleum Engineering Conference held in Cartagena, Colombia. 31 May-3 June 2009.
- [10] Doraisamy, H. Methods of Neuro-Simulation for Field Development. Paper SPE 39962, presented at the Rocky Mountain Regional/Low-Permeability Reservoirs symposium held in Denver, Colorado. 5-8 April 1998.
- [11] Soto, R., Wu, C.H. and Bubela, A.M. Infill Drilling Recovery Models for Carbonate Reservoirs – A Multiple Statistical, Non-Parametric Regression, And Neural Network Approach. Paper SPE 57458, presented at the SPE Eastern Regional Conference and Exhibition held in Charleston, WV. 21-22 October 1999.
- [12] Jalali, J. and Mohaghegh, S.D. Reservoir Simulation and Uncertainty Analysis of Enhanced CBM Production Using Artificial Neural Networks. Paper SPE 125959, presented at SPE Eastern Regional Meeting held in Charleston, West Virginia, USA. 23-25 September 2009.
- [13] The MathWorks, inc. MATLAB software manual, 2008.
- [14] Uarnganun, A. Artificial Neural Network class teaching media.
- [15] Lawrence, S., Lee Giles, C., Ah Chung Tsoi. Lessons in Neural Network Training: Overfitting May be Harder than Expected. Paper AAAI-97, presented at the Fourteenth National Conference on Artificial Intelligence, Menlo Park, California. 1997.
- [16] Amyx, Bass, Whiting. Petroleum Reservoir Engineering: Physical Properties
- [17] Hettiarachchi, P., Hall, M.J., and Minns, A.W. The extrapolation of artificial neural networks for the modeling of rainfall-runoff relationships. Journal of Hydroinformatics. 2005



APPENDICES

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

APPENDIX A

“MATLAB” source code for ANN development

```

1  %Input & Target Output
2  - input=[];
3  - target=[];
4
5  %Construct a network
6  - net = newff(input, target,[5],{'logsig'},'trainlm');
7
8  - net.inputweights(1,1).initFcn = 'rands';
9  - net.biases(1).initFcn = 'rands';
10 - net = init(net);
11 - net.divideFcn = 'divideblock';
12
13 - [trainV,valV,testV,trainInd,valInd,testInd] = divideblock(input,4,1,0);
14 - [trainT,valT,testT] = divideind(target,trainInd,valInd,testInd);
15
16 %Network configuration (Learning rate & Momentum)
17 - net.trainParam.lr = 0.1;
18 - net.trainParam.mc = 0.5;
19 %Criteria to stop training
20 - net.trainParam.min_grad = 1e-20;
21 - net.trainParam.epochs = 1000;
22 - net.trainParam.mu_max = 1e2000;
23 - net.trainParam.max_fail = 15;
24
25 %Partitioning ratio for training, validating, and testing sets
26 - net.divideParam.trainRatio = 4;
27 - net.divideParam.valratio = 1;
28 - net.divideParam.testratio = 0;
29
30 %Start to train a network
31 - [net,tr]=train(net,input,target);
32
33 %Simulate a network
34 - z=sim(net,simulate);

```

APPENDIX B

Details of 100 wells taken from reservoir simulation model

Well No	Coordinate		Input parameters									Output parameters	
	X	Y	k (mD)	P _{actual} (psia)	P _{normal average} (psia)	P _{distance average} (psia)	Φ _{1st ring} (%)	Φ _{2nd ring} (%)	Φ _{3rd ring} (%)	Start drill date (day)	Surrounding well Qty (wells)	Initial gas rate (Mscf/d)	1 year cumulative (MMscf)
1	98	163	3.58	1753.65	0	0	17	17.88	17.63	10	0	14113.04	4347.82
2	71	46	5.80	1901.59	0	0	18	18.96	18.42	20	0	18864.63	6034.71
3	184	179	15.17	1985.39	0	0	20	20.67	21.88	30	0	24601.32	7948.75
4	40	170	1.37	1544.64	0	0	15	14.17	15.88	40	0	6540.95	1936.11
5	149	170	24.54	1716.52	0	0	21	18.63	20.21	50	0	22093.86	6988.91
6	155	133	439.74	2146.46	0	0	27	26.25	25.79	60	0	31061.39	10025.64
7	145	21	271.83	2124.47	0	0	26	24.33	24.54	70	0	30623.43	9999.64
8	136	90	3.58	1722.47	0	0	17	17.08	17.29	80	0	13780.60	3859.96
9	22	127	39.69	2110.45	0	0	22	24.13	23.33	90	0	28896.55	9379.45
10	100	93	15.17	1961.01	0	0	20	20.25	20.04	100	0	24251.21	7522.94
11	178	73	39.69	1921.22	0	0	22	19.33	18.75	110	0	26059.06	7918.61
12	111	125	5.80	1857.18	0	0	18	18.88	19.75	120	0	18302.05	5533.10
13	111	48	9.38	1959.00	0	0	19	21.13	21.54	130	0	22218.46	6923.75
14	25	57	39.69	1992.29	0	0	22	20.96	19.96	140	0	27123.39	8224.26
15	25	89	39.69	2086.40	0	0	22	24.58	25.08	150	0	28534.72	8897.69
16	21	19	64.21	2090.35	0	0	23	23.38	23.83	160	0	29251.34	9038.58
17	58	105	271.83	2039.54	0	0	26	23.00	21.38	170	0	29327.96	8852.42
18	186	147	271.83	2018.64	0	0	26	24.13	24.92	180	0	29006.43	8774.05
19	142	52	3.58	1799.70	0	0	17	19.71	19.79	190	0	14639.85	4309.13
20	52	16	39.69	2034.80	0	0	22	24.00	23.83	200	0	27759.31	8416.61
21	178	29	39.69	1782.11	0	0	22	18.83	19.17	210	0	23979.26	7063.76
22	62	139	5.80	1741.49	0	0	18	17.88	17.58	220	0	16850.36	4447.67

Well No	Coordinate		Input parameters									Output parameters	
	X	Y	k (mD)	P _{actual} (psia)	P _{normal average} (psia)	P _{distance average} (psia)	Φ _{1st ring} (%)	Φ _{2nd ring} (%)	Φ _{3rd ring} (%)	Start drill date (day)	Surrounding well Qty (wells)	Initial gas rate (Mscf/d)	1year cumulative (MMscf)
23	83	15	15.17	1894.32	0	0	20	21.25	21.88	230	0	23293.31	6886.72
24	186	113	64.21	1889.27	0	0	23	22.54	22.92	240	0	26204.47	7579.51
25	113	17	39.69	1933.63	0	0	22	22.67	23.42	250	0	26240.75	7895.70
26	158	99	1.37	1513.80	1587.56	1565.40	15	15.38	16.04	260	2	6308.64	1502.47
27	75	116	9.38	1817.56	1736.49	1749.17	19	21.58	21.17	270	3	20286.91	5644.03
28	177	131	271.83	1858.05	1857.74	1856.83	26	23.38	23.25	280	3	26545.14	7736.50
29	42	114	15.17	1732.06	1798.89	1813.66	20	21.17	20.63	290	4	20980.79	6050.16
30	77	157	3.58	1811.70	1518.72	1518.40	17	21.33	19.58	300	2	14773.53	4156.48
31	139	113	5.80	1593.67	1541.88	1533.19	18	18.17	18.17	310	3	14950.38	3946.66
32	182	56	15.17	1635.20	1630.08	1630.08	20	19.42	19.04	320	1	19581.50	5226.94
33	177	90	439.74	1787.66	1557.53	1553.34	27	24.46	23.96	330	3	25564.72	7280.05
34	80	88	24.54	1696.90	1705.59	1698.21	21	20.75	22.67	340	2	21797.88	5991.87
35	147	192	271.83	1781.77	1420.21	1420.21	26	22.46	22.63	350	1	25376.95	7248.13
36	43	76	64.21	1769.44	1721.10	1726.85	23	23.42	23.21	360	2	24390.87	7045.00
37	116	173	2.21	1513.05	1445.90	1445.90	16	16.50	16.79	370	1	8715.76	2103.71
38	110	72	5.80	1577.03	1517.80	1534.44	18	18.75	18.21	380	3	14731.58	3528.58
39	132	150	2.21	1538.82	1493.22	1488.90	16	18.88	18.13	390	4	8966.72	2210.24
40	71	72	9.38	1534.22	1533.35	1538.43	19	19.71	18.92	400	2	16377.69	4184.42
41	40	50	3.58	1364.58	1615.40	1604.91	17	16.50	17.33	410	2	9696.27	2327.42
42	60	124	5.80	1323.76	1498.17	1490.15	18	17.42	17.42	420	4	11486.92	2633.32
43	95	65	5.80	1552.76	1472.31	1467.10	18	21.00	19.88	430	5	14412.38	3476.80
44	154	79	3.58	1182.76	1343.02	1331.76	17	14.83	14.58	440	4	7609.77	1557.71
45	120	192	5.80	1511.85	1324.96	1324.96	18	17.79	17.75	450	1	13882.77	3183.98
46	14	175	1.37	1612.33	1280.36	1280.36	15	18.00	19.58	460	1	7055.65	1874.25
47	54	187	0.85	1581.34	1274.64	1274.64	14	15.17	16.33	470	1	4707.06	1188.51
48	70	185	2.21	1476.92	1530.05	1530.05	16	15.75	16.38	480	1	8366.80	2024.06

Well No	Coordinate		Input parameters									Output parameters	
	X	Y	k (mD)	P _{actual} (psia)	P _{normal average} (psia)	P _{distance average} (psia)	Φ _{1st ring} (%)	Φ _{2nd ring} (%)	Φ _{3rd ring} (%)	Start drill date (day)	Surrounding well Qty (wells)	Initial gas rate (Mscf/d)	1year cumulative (MMscf)
49	25	36	168.04	1589.27	1405.53	1415.04	25	24.75	24.25	490	3	22273.64	6195.89
50	91	186	3.58	1542.99	1313.40	1319.40	17	17.08	17.25	500	3	11706.68	2857.58
51	89	50	64.21	1464.49	1331.67	1340.15	23	21.71	21.83	510	5	19759.82	5108.48
52	103	32	39.69	1522.23	1446.27	1448.42	22	24.83	24.79	520	4	20060.59	5271.67
53	83	132	9.38	1252.84	1262.27	1263.49	19	17.92	19.71	530	4	12481.98	3030.90
54	46	147	64.21	1365.89	1248.01	1231.04	23	19.13	18.75	540	4	18261.02	4280.57
55	165	51	2.21	1282.27	1237.71	1235.46	16	17.13	16.88	550	4	6557.72	1332.03
56	130	16	39.69	1452.91	1447.22	1448.04	22	24.50	24.67	560	2	19010.64	4841.67
57	145	36	439.74	1443.10	1336.34	1343.26	27	24.46	24.58	570	4	20274.15	5232.52
58	166	178	24.54	1328.20	1215.09	1206.53	21	23.63	23.46	580	3	16304.98	4222.07
59	160	35	5.80	1236.37	1249.65	1260.08	18	18.71	17.92	590	6	10379.35	2180.78
60	27	190	271.83	1685.69	1312.25	1324.44	26	20.04	21.33	600	2	23908.06	6042.83
61	184	9	9.38	1378.12	1162.61	1163.22	19	21.21	22.25	610	2	14214.28	3105.09
62	106	143	3.58	1280.86	1143.41	1146.89	17	20.29	20.25	620	4	8737.45	2019.22
63	8	97	15.17	1323.99	1340.15	1340.15	20	23.33	23.46	630	1	15069.58	3621.90
64	193	26	15.17	1276.91	1193.62	1182.79	20	21.13	20.83	640	2	14385.34	3179.19
65	134	167	168.04	1172.97	1081.42	1067.22	25	19.75	20.42	650	5	15875.10	3855.64
66	18	142	39.69	1352.88	1328.65	1328.65	22	21.83	21.79	660	1	17499.83	4277.82
67	40	129	39.69	1222.93	1116.88	1114.66	22	21.13	19.63	670	7	15534.35	3591.65
68	59	167	0.85	1184.64	1152.69	1156.21	14	13.33	14.67	680	5	2563.83	509.71
69	10	39	103.87	1224.11	1185.13	1191.16	24	23.79	23.46	690	3	16436.92	4086.82
70	116	93	15.17	1045.10	994.87	998.71	20	19.13	18.71	700	4	10968.08	2205.65
71	68	24	103.87	1181.27	1078.18	1085.93	24	23.83	23.96	710	4	15777.42	3899.14
72	170	158	15.17	1041.48	1032.03	1033.63	20	22.29	21.71	720	5	10912.42	2515.35
73	130	130	64.21	1086.04	1033.09	1025.35	23	21.54	20.79	730	5	13948.05	3300.78
74	193	71	168.04	995.97	963.63	950.84	25	21.04	21.13	740	3	13094.65	3072.73

Well No	Coordinate		Input parameters									Output parameters	
	X	Y	k (mD)	P _{actual} (psia)	P _{normal average} (psia)	P _{distance average} (psia)	Φ _{1st ring} (%)	Φ _{2nd ring} (%)	Φ _{3rd ring} (%)	Start drill date (day)	Surrounding well Qty (wells)	Initial gas rate (Mscf/d)	1year cumulative (MMscf)
75	61	89	39.69	1028.10	1001.64	1002.92	22	20.92	21.79	750	5	12521.13	2839.91
76	11	192	168.04	1352.31	1221.89	1223.71	25	22.00	22.54	760	2	18638.13	4247.86
77	130	69	2.21	1026.25	874.67	876.54	16	18.17	17.42	770	6	4201.88	792.22
78	121	110	5.80	924.03	903.43	901.38	18	18.75	17.58	780	6	6333.93	1131.20
79	22	159	3.58	1132.19	990.87	1008.34	17	20.83	20.63	790	4	7032.28	1379.33
80	84	30	3.58	911.65	933.96	935.71	17	21.08	22.17	800	5	4659.69	858.04
81	130	31	103.87	954.07	923.20	928.76	24	23.67	24.38	810	6	12214.97	2885.97
82	43	92	39.69	964.02	925.51	928.96	22	23.38	23.04	820	5	11531.65	2352.60
83	26	104	439.74	995.64	929.48	931.46	27	27.00	26.42	830	6	13314.05	2918.36
84	100	108	64.21	877.78	813.18	812.98	23	21.00	19.88	840	7	10693.65	2301.33
85	163	20	3.58	730.55	815.64	808.43	17	16.21	15.75	850	5	2696.37	422.81
86	10	112	439.74	924.96	909.22	908.49	27	24.17	23.63	860	4	12192.26	2627.13
87	26	73	24.54	823.08	833.12	836.49	21	20.67	21.83	870	6	8560.46	1516.23
88	55	47	168.04	839.08	780.85	773.09	25	20.29	21.42	880	4	10616.74	1977.39
89	185	163	168.04	785.21	808.67	807.34	25	20.88	21.08	890	4	9734.81	2094.43
90	157	114	439.74	859.38	791.61	784.77	27	25.88	25.38	900	6	11151.58	2541.14
91	152	152	24.54	724.28	760.04	754.24	21	18.67	19.13	910	8	6955.64	1254.97
92	10	81	39.69	747.08	754.57	753.84	22	21.46	21.08	920	5	8008.23	1359.88
93	36	11	9.38	719.22	803.87	803.08	19	19.25	19.71	930	3	5003.50	786.62
94	40	33	271.83	787.06	725.71	726.67	26	24.50	22.71	940	7	9912.53	1985.29
95	8	60	168.04	758.90	710.57	706.22	25	22.67	22.21	950	5	9267.45	1644.71
96	193	86	39.69	704.13	669.20	671.99	22	22.33	22.92	960	3	7284.34	1385.10
97	98	10	15.17	693.40	691.15	689.99	20	22.88	23.08	970	4	5631.59	1027.21
98	192	129	439.74	717.78	690.99	690.88	27	26.25	26.04	980	3	8765.67	1881.67
99	114	158	0.52	793.10	719.80	725.00	13	14.54	15.75	990	5	593.61	112.91
100	126	50	64.21	676.13	663.15	662.21	23	23.71	22.33	1000	7	7275.37	1509.41

APPENDIX C

Details of 19 candidate wells of well number 101 and result from all cases of ANN

Well No 101	Coordinate		Target Output		Predicted Output							
	X	Y	Initial gas rate (Mscf/d)	1year cumulative (MMscf)	Case 1-1	Case 1-2-1	Case 1-2-2	Case 1-3	Case 2-1	Case 2-2-1	Case 2-2-2	Case 2-3
					Initial gas rate (Mscf/d)	Initial gas rate (Mscf/d)	Initial gas rate (Mscf/d)	Initial gas rate (Mscf/d)	1year cumulative (MMscf)	1year cumulative (MMscf)	1year cumulative (MMscf)	1year cumulative (MMscf)
Alt 1	166	193	3337.93	655.22	4033.45	5215.98	4517.90	4397.76	551.25	833.68	764.57	633.50
Alt 2	172	113	885.09	128.41	931.00	461.19	1360.72	3107.38	99.33	279.21	-130.37	202.95
Alt 3	158	65	726.94	119.79	-589.00	-2441.13	-1497.70	192.68	39.51	-435.93	-703.35	-525.26
Alt 4	54	61	1159.88	133.41	1589.35	886.34	2040.51	1196.46	287.77	363.65	-367.35	660.77
Alt 5	86	102	2688.44	532.56	3190.72	2773.22	3478.06	4464.25	377.30	501.06	573.56	1231.68
Alt 6	192	42	1592.64	286.43	1793.25	633.25	1501.18	3409.22	249.41	309.83	48.46	-283.72
Alt 7	181	193	2340.95	424.82	2754.21	1701.60	2840.06	4473.68	511.14	311.35	530.57	117.27
Alt 8	91	147	1738.66	349.97	1299.77	327.74	587.34	279.38	291.90	-111.75	59.38	-244.33
Alt 9	84	171	4304.76	811.10	4457.71	4566.72	4248.34	3216.16	749.23	834.95	780.46	285.08
Alt 10	32	144	210.13	39.30	-2124.62	-3565.01	-2556.93	48.56	-335.52	-786.85	-147.38	-1079.20
Alt 11	69	8	1815.59	291.66	2191.10	1172.30	2873.28	4659.74	465.92	254.86	284.89	603.32
Alt 12	62	153	984.59	214.28	150.39	-2082.59	-1302.59	233.04	96.13	-597.48	17.89	-845.05
Alt 13	106	187	3675.97	676.46	3913.02	3554.43	3507.41	3214.99	678.11	729.78	523.47	241.62
Alt 14	97	125	984.60	194.05	45.61	-950.97	-292.89	346.11	20.83	-296.36	-245.38	-764.44
Alt 15	95	79	464.27	77.01	-1220.98	-2566.08	-1569.64	170.03	-90.96	-572.02	-497.38	-807.27
Alt 16	8	159	1808.56	329.33	2045.58	1501.73	2107.40	3412.40	287.79	367.30	424.74	-308.02
Alt 17	8	126	2445.37	456.20	3001.85	2389.47	3225.60	4498.04	417.25	690.43	579.57	723.50
Alt 18	54	30	1097.70	148.24	1518.82	891.79	2076.30	3358.11	223.99	360.71	-203.34	557.54
Alt 19	57	75	2203.56	380.51	3149.52	2923.66	3310.16	4309.41	326.01	569.30	290.48	526.31

VITAE

Nattaphon Temkiatvises was born on November 5, 1982 in Bangkok, Thailand. He received his B.Eng in Mechanical Engineering from the Faculty of Engineering, Chulalongkorn University in 2004. After graduating, he continues his studies in the Master of Petroleum Engineering program at the Department of Mining and Petroleum Engineering, Faculty of Engineering, Chulalongkorn University



ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย