PERMEABILITY PREDICTION USING NEURAL NETWORK MODEL WITH POST-PROCESSING

A Thesis

Submitted to the Department of Petroleum & Mineral Resources Engineering

in partial fulfillment of the requirements for the degree of Master of Science in Engineering (Petroleum).

By

MD. JAKARIA

DEPARTMENT OF PETROLEUM & MINERAL RESOURCES ENGINEERING
BANGLADESH UNIVERSITY OF ENGINEERING & TECHNOLOGY,
DHAKA
BANGLADESH

JANUARY, 2004.

RECOMMENDATION OF THE BOARD OF EXAMINERS

The undersigned certify that they have read and recommended to the Department of Petroleum and Mineral Resources Engineering, for acceptance, a thesis entitled PERMEABILITY PREDICTION USING NEURAL NETWORK MODEL WITH POST-PROCESSING submitted by MD. JAKARIA in partial fulfillment of the requirements for the degree of MASTER OF SCIENCE IN ENGINEERING in PETROLEUM AND MINERAL RESOURCES.

Chairman (Supervisor):

Dr Zulfiquar Ali Reza
Assistant Professor
Dept. of Petroleum & Mineral Resources
Eng. BUET

Member

Dr. Mohammad Tamim
Head & Professor
Dept. of Petroleum & Mineral Resources
Eng. BUET

Member :

Dr. Ijaz Hossain
Professor
Dept. of Chemical Eng
BUET

Member (External) :

Dr. Harendra Nath Mondal
Associate Professor
Dept. of Chemical Eng.
BUET

Date: January 17, 2004.

ABSTRACT

Practicable and realistic reservoir characterization is essential for optimal reservoir management. In this study, a randomized back-propagation neural network model is developed for formation permeability prediction. The model has only one hidden-layer, and the inputs to the model are core porosity, facies identifier, sample thickness, and well sample location. A number of sensitivity studies for permeability prediction are performed. Prediction errors from the model are analyzed and a post-processing scheme for error mitigation is investigated. Neural network responses were compared with those using conventional methods for permeability determination. There are some specific advantages of using the developed model. Characterization of prediction space is observed to be better. However, the limitations of the study were also highlighted. A variety of applications of artificial neural networks in reservoir engineering problems are reviewed in this study.

ACKNOWLEDGEMENTS

I am grateful to Dr. Zulfiquar Ali Reza for his valuable and relentless guidance, encouragement and supervision throughout the entire work. As a result, the work has become fruitful.

I would like to thank Dr. M. Tamim, the head of the Petroleum and Mineral Resources Department, for his support and cooperation for this work.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

The demand for energy is ever increasing. For today's world, oil and gas is one of the main sources of energy for the industrial development and maintaining living standard However, the finite sources of known reserves are depleting. As a consequence, there is growing endeavor for search for unknown or unexploited reserves and optimal reservoir management of the developed fields. Fierce competition and the urge for environmental friendly exploitation of natural resources drive most oil and gas companies to have greater and better control over their plays Characterization of the reservoir is thus deemed an essential aspect in reservoir management.

Reservoir characterization is an important domain of petroleum engineering. It is the precondition for reservoir simulation as well as effective reservoir management. It entails various aspects of reservoir and rock-fluid modeling, namely formation property like porosity, permeability, fluid saturation, lithology prediction, and so forth. For proper and effective reservoir characterization, the directional and spatial variation of rock property should be determined. This study focuses on one of the processes involved in permeability characterization.

The traditional methods for permeability determination are often empirical in nature. Carman-Kozeny equation for permeability determination is a common approach. A number of limitations exist in the applicability of this approach. It is valid for packs of uniformly sized spheres. Another drawback is that surface area can be determined only by core analysis, and only with special equipment. Variation of Kozeny constant for consolidated porous media creates another problem to work with Carman-Kozeny equation. Multiple linear regression method is also used for the estimation of permeability. Multiple regression methods often have an inherent averaging tendency. Nonetheless, the assessment of uncertainty in estimation is precluded in these methods.

The influence and application of virtual intelligence have grown in numerous fields of science and engineering. Virtual intelligence tool such as artificial neural network is being widely used. Artificial neural networks have emerged as useful tools in petroleum engineering, particularly reservoir characterization. In this study, a neural network model based on back propagation algorithm has been developed for permeability determination.

The neural network is trained with available data. During training, the neural network gathers knowledge about the system and stores them as memory matrix. This memory matrix is then used for prediction. Neural network is error tolerant and a data driven model. Very detail knowledge of the system is not necessary to work with neural network. It requires less input parameters for reservoir characterization with respect to conventional mathematical model. So, neural network can be used for reservoir characterization in an inexpensive way.

CHAPTER 2

STATEMENT OF THE PROBLEM

Determining permeability from well logs is an essential element in reservoir modeling. The task of deriving permeability values is however not a simple exercise. Present day technologies reduce the complexity of the problem to some extent. The general convention is to use simple permeability-porosity relationships These simple relationships are only valid for unconsolidated sand and homogeneous lithology. In fact, permeability is a complex function of several interrelated factors such as lithology, pore fluid composition and porosity. Well logs respond directly to these factors. There have been attempts to establish their relationships. As an alternative to conventional methods such as Carman-Kozeny equation, multiple linear regression analysis, one can implement an artificial neural network technique to predict permeability more accurately from well logs.

The artificial neural network models are data driven that is entailed in the "training" stage of model development. The trained neural network models are able to capture information and knowledge about the system. They can extract nonlinear relationships that are immensely difficult to model from the first principles.

Primarily the objectives of this work are:

- To develop a neural network model to predict well permeability
- To optimize the prediction process and development of a post-processing scheme to minimize estimation error
- To efficiently estimate permeabilities at "uncored" wells
- To compare neural network model responses with conventional methods.

Outline of the work is given below.

- Review of application of neural network in Petroleum Field
- Development of the computer code to implement a neural network model using core-data
- Prediction of well permeability using neural network model. This entails a case study with real core data.
- Optimal parameter selection of the neural network model
- Sensitivity analysis using neural network model
- Error analysis of the model predictions
- Development of a post-processing scheme for error correction
- Permeability estimation using conventional methods.

CHAPTER 3

LITERATURE REVIEW

This chapter reviews the available literature of artificial neural network applied to petroleum engineering. Only selective works have been considered. The review should not be considered an exhaustive account of the applications of neural network in petroleum engineering domain.

Artificial neural networks (ANN) are collections of computational techniques that mimic the functions of human brain These techniques can recognize patterns, approximate functions, feature detection and other scientific and engineering problems. A discussion of the basics of artificial neural networks will be given in Chapter 4 In a general sense, applications of neural networks are most effective when there are no established relations among the parameters. Neural networks attempt to capture the non-linear relationship through some data driven algorithms. However, the application of artificial neural networks should be made with caution; it is strongly recommended not to use these algorithms in a "black-box" manner.

Applications of neural networks in petroleum engineering have been around since the mid-nineties. Amongst the first uses of neural network in petroleum engineering are in the field of reservoir characterization. In the following sections, some of these applications will be briefly discussed here. These applications are categorized into two broad sections: one for reservoir characterization, and the other for reservoir engineering.

3.1 Applications of Neural Networks in Reservoir Characterization

Reservoir characterization entails building reservoir models and characterizing reservoir properties. Information from different sources is obtained for reservoir characterization. Different sources of information include the domains of geology, geophysics, geochemistry, stratigraphy, sedimentology, engineering, and relevant geosciences. Data from these sources can vary in resolution, information-content, and scale. The challenge in reservoir characterization is to integrate these distinctly

different pieces of information into building plausible reservoir models Applications of neural networks in reservoir characterization are briefly discussed here.

Estimation of permeability is very difficult in uncored well as there is no direct relationship between permeability and lithology. In fact, the relationship between permeability, porosity and other data such as grain size, type and others are very complex, and area specific. Log responses depend on lithology, fluid composition and porosity. A predictive equation for uncored intervals can be developed using multiple linear regression analysis. But fundamental assumption for linearity may not be valid Huang and Shimel (1994) developed a back-propagation neural network model to predict permeability in uncored well. Input parameters were latitude, longitude, depth, spontaneous potential, gamma ray, density, sonic, neutron porosity, density correction. Permeability was output parameter. The networks contain one hidden layer. Number of nodes in input, hidden and output layers were 9, 12, and 1, respectively.

Mohaghegh and Arefi (1994) developed neural networks model to predict rock property such as porosity, permeability as well as water, gas and oil saturation using geophysical logs. Mohaghegh and Arefi (1995) developed a neural network model to predict permeability using log data.

Mohaghegh and Popa (1995a) developed a virtual intelligence software tool based on detail analysis of a group of well logs. The software was then used to analyze the remaining well logs at a reduced time and cost.

Balan, Mohaghegh and Ameri (1995) developed methods for empirical, statistical (multiple regression) and a back-propagation neural network model to determine permeability of heterogeneous oil bearing formation from well log data. It was found that empirical modeling requires individual equation for individual field. Further, fluid saturation, porosity and cementation factors are necessary to start modeling through empirical method. A neural network model of three layers with 18 hidden neurons was developed and trained. The neural network outperformed the other techniques. The major advantage of these models is that they do not require fluid saturation, porosity data, and they are not affected by the cementation factor.

Mohaghegh et al. (1995b) compared the prediction capability of multiple regression and back-propagation neural network model. Multiple regression systematically underestimated the permeability. Also, it was not able to predict permeability for the entire domain. Neural network models predicted the target permeability values closer to standard core data. Furthermore, the entire domain permeability could be predicted.

White and Molnar (1995) developed neural networks for zone identification in a complex reservoir. The neural networks were trained using geophysical data along with previously defined various zones. Then the developed neural networks were used to identify zones, previously known by core analysis, for different wells.

Walls and Taner (1995) developed a model for reservoir classification based on seismic attributes and borehole data. Core, well log, and post-stack seismic data were used to predict lithology. The method was based on combination of core, well logs and pattern recognition via neural networks. Borehole parameters to the network were density, primary velocity, secondary velocity, clay volume, and water saturation. Output was reservoir classification indicators. Seismic data are found most effective to predict inter-well lithology. To tie well log derived attributes and seismic attributes, synthetic seismograms were generated and used to train the neural networks. A suite of attributes was derived. At last five attributes were found sufficient enough to generate acceptable results. These five attributes were input to the neural network and output was lithology column. Using a neural network trained to log lithology in the time domain gives lithology from synthetic seismic with reasonable accuracy at the well locations. The neural network weights derived synthetic seismograms were applied to attributes from real seismic revealing the producing wells to be inside the indicated oil sand area and the non-producing wells to be outside.

Mohaghegh and Richardson (1998) used well logs to predict effective porosity and fluid saturation. Magnetic resonance imaging is used to measure free fluid, irreducible water, and effective porosity accurately. Permeability is then calculated using mathematical function. Magnetic resonance images are also capable of

calculating recoverable reserves. Recent studies show that neural network have potential to produce synthetic magnetic resonance images from conventional well logs. Virtual magnetic resonance imaging logs were produced by a neural network model and used to measure free fluid, irreducible water, and effective porosity for training and verification data. Recoverable reserves were calculated for different reserves using the virtual logs.

Mohaghegh (1999) developed a neural network model to predict permeability. The neural network model was trained using geophysical well log data (bulk density, gamma ray, and induction logs) as inputs, and core data as output. Neural network performance was satisfactory to predict permeability for verification data.

Mohaghegh and Goddar (2000) used neural network models producing virtual magnetic resonance imaging logs from conventional logs (spontaneous potential, gamma ray, caliper, and resistivity, density and induction logs) to predict free fluid, effective porosity, irreducible water saturation.

Modeling of the lithofacies distribution is an important aspect in reservoir characterization. Siripitayananon and Hui-Chuan (2001) developed a neural network model to predict lithofacies distribution The relation between seismic attributes and lithofacies is complicated and nonlinear and this relationship cannot be developed based on first principle. A back-propagation neural network was developed to capture the nonlinear relationship between seismic attributes and lithofacies. The model could also be used for petrophysical modeling. The numbers of nodes for the input, hidden and output layers are 11, 7, and 4, respectively. The input parameters used are seismic attributes (amplitude, instantaneous frequency, perigram, cosine of phase, instantaneous phase, reflection strength, and response phase), the location of the seismic trace, two way time to the reflector, and one bias node to impose an extra degree of freedom. Output layer nodes represent 4 categories of lithofacies. Data were prepared using a k-nearest neighbors classification algorithm.

Reeves and Mohaghegh (2002) showed that a high-resolution reservoir characterization is possible through the integration of different scale and type of data using artificial intelligence (such as neural networks). Relationships between data of

different scales and type had been established, including conventional well logs, reservoir imaging logs, cross-well seismic, and surface seismic based on artificial intelligence. Through integration of multiscale data a 3D reservoir image was established which is required for reservoir flow simulation.

3.2 Applications of Neural Networks in Reservoir Engineering

This section discusses the use of neural networks in reservoir and production engineering operations. Again, this is not an exhaustible account of such applications.

Post-fracture well performance prediction is a challenging task. Mohaghegh and McVey (1995) developed a back-propagation neural network model to predict post-fracture deliverability of wells. Unlike the conventional simulator based on mathematical modeling, neural networks do not require a lot of reservoir data. Thus, it is cost effective with respect to large cost for data collection. This process is being currently used to select candidates well economically feasible for stimulation treatment Input parameters were well number, year and date the wells were fractured, number of fracture jobs, type of fracture, fluid viscosity, total water used, nitrogen used per barrel of water, total sand used, sand concentration, sand type, acid volume and type, chemicals, treatment injection rate, occurrence of new screen-out, contractor, hole size, completion type, well type, date of completion, date converted to storage, well group number, sand thickness, minimum 20 year flow test value, maximum 20 year flow test value, and flow test before refracturing. Output was maximum flow test after fracture. Networks predicted results were compared with actual data. This model is found cost effective with respect to conventional fracture simulation.

Ternyik and Bilgesu (1995a) developed a virtual intelligence tool using neural networks to predict flowing bottomhole pressure under multiphase flow and inclined wellbore conditions. The input parameters were oil, gas and water flow rates, temperature, oil and gas gravity, pipe length, surface pressure and inclination angles of the pipe. The developed virtual measurement tool was compared with published data.

A new method, virtual measurement in pipes was developed by Ternyik and Bilgesu (1995b) using neural networks to predict liquid holdup and flow regime in pipelines and well bores. The predicted output was tested with published data for validation. The method proved to be an accurate measuring tool.

Nikravesh and Kovsech (1996) developed a neural network to predict water injection rate as a function of wellhead pressure and vice versa in a fractured, low permeability oil reservoir. The result was satisfactory though water flood behavior is complex. Neural networks were efficient in finding the major input parameters. In the same reservoir, neural networks were applied to correlate the injection pressure and rates, and temperature responses in different wells using data from a dual injector steam drive pilot. Assuming a future pressure policy, neural network was used to predict injection rate and growth of reservoir heated volume

Mohaghegh and Hefner (1996) developed a new software tool, Fracture Optimization eXpert (FOX), based on neural network and genetic algorithm. It can predict post fracture deliverability and the best possible combination of fracture parameters using production history data and completion data. It can do without using reservoir data. This tool was applied in North Eastern Ohio (Clinton Sand) and the result was satisfactory. Mcvey and Mohaghegh (1994, 1995) identified major parameters for hydraulic fracture based on neural networks Mohaghegh and Balan (1996) designed and optimized hydraulic fracture treatment based on neuro-genetic approach.

Bilgesu and Tetric (1997) developed a neural network model to predict rate of penetration and bit wear condition under various formation types and parameters. Drilling data was generated using a simulator. The data generated was used to develop relationship between complex patterns such as weight on bit, rotary speed, pump rates, formation hardness, and bit type. The validity of the model was demonstrated with data from an existing field. Bilgesu and Altmis (1998) developed a three-layer back-propagation neural network model to predict bit wear and life. Six measured parameters (weight on bit, rotary speed, pump rate, formatiou hardness, bit type, and torque) were used. The validity of the model was demonstrated.

Mohaghegh and Platon (1998) developed neural networks and genetic algorithm for suitable fracture candidate selection. Neural networks provide realistic model for successful fracturing jobs and chemical treatments; genetic algorithms were developed to design optimization economic analysis. Different neural networks were used for different chemical treatment. With each neural network model the first series of genetic algorithm was used for optimum treatment design. A separate genetic algorithm was used for economic analysis

Mohaghegh and Popa (1999b) developed intelligent software to design fracture jobs in an inverse manner. The inputs to the software were fracture geometry and reservoir characteristics and outputs were fluid, proppant and treatment schedule. This new tool was tested and shown to be capable of designing hydraulic fractures.

Mohaghegh and Reeves (2000) developed a methodology incorporating artificial intelligence techniques (neural network, genetic algorithm and fuzzy logic) to select wells for restimulation. Neural networks were used to develop a representative model of completion and hydraulic fracturing process for a specific field. Genetic algorithm as search and optimization tool was used to identify missed production based on the neural network model. Finally, fuzzy logic was used to capture field experiences as well as detrimental parameter and incorporated them in decision-making process. This methodology was applied in a tight sand field and the results were satisfactory.

Reeves, Bastian and Flumerfelt (2000) identified restimulation candidate wells in an effective manner. Production statistics, virtual intelligence and type curve matching were investigated to select restimulation candidates. Restimulating a number of simulated reservoir models and observing the incremental production responses established restimulation potential. Simple production data could not effectively select restimulation candidates. Virtual intelligence techniques were found to be most effective.

Stundner and Al-Thuwani (2001) used back-propagation neural network to develop models for injection production ratio optimization, well interaction. The key concept for this optimization is to keep well pressure above saturation pressure and avoid pressure sinks in the reservoir with the change in production injection patterns.

Reservoir pressure response should be modeled with the change in production and injection rates. Fluid production and water injection were model inputs. The model outputs were aquifer pressure and reservoir pressure. Inputs to the neural network were cumulative fluid production and water injection, fluid production and water injection rates. Average reservoir aquifer pressures were the model outputs. When production data were not available, choke size, tubing head flowing pressure and/or bottom hole flowing pressure were used for pressure response modeling. Wells were grouped based on connectivity. They developed another back-propagation neural network model for shut in pressure of a well. Fluid production and water injection rates were inputs and shut in pressure was the output of this model.

During the process of hydraulic fracturing of gas wells over the years, companies usually record the relevant data on methods and materials such as date of the job, fluid type and amount, proppant type and pump rate, breakers, additives, amount of nitrogen, etc. These data are of little use in 3D hydraulic fracture simulations. Mohaghegh and Gaskari (2002) processed these data coupled with general well information, well log data and production data using virtual intelligence to select the fracture candidate wells in a sand formation with low permeability. The job was successful in candidate selection. Mohaghegh and Platon (1998) performed similar exercise.

Mohaghegh and Hutchins (2002) developed a neural network model and trained the pressure and flow rate data from separation facilities with corresponding pressure and flow rate data at the inlet of central compressor. It was used as a tool to provide maximum oil production from a field to find out optimum discharge pressure and rate from separation unit.

There are more applications of neural networks in reservoir related studies. Due to lack of accessibility of their account, these could not be included in this review.

CHAPTER 4

BASICS OF NEURAL NETWORK

Artificial neural networks are computational techniques inspired by the mechanism, structure and functions of human brain. The literature on neural networks has grown exponentially over the recent years. The sheer number of applications of neural networks in various fields of science, engineering, and social science are monumental. In this chapter, it has been attempted to relate a brief discussion on the basics of the neural networks.

Essentially, the neural network uses a set of linear and nonlinear activation function. The neural networks are trained with a set of input and output data. Neural networks can store knowledge obtained from this training in weights. After training the neural network can be used for prediction, estimation and characterization. Neural network is a data driven model. Prior knowledge about the first principles of the system is not necessary to perform above tasks.

The outline of the chapter follows. Section 4.1 discusses the basic attributes of neural network. Section 4.2 depicts how a neuron produces an output signal corresponding to an input signal. Section 4 3 overviews the fundamental learning paradigms of a neural network Section 4.4 deals about optimization technique in order to minimize cost function. Section 4.5 introduces the basic principle of back propagation neural network under multilayer perceptron.

4.1 Neural Attributes

Architectures and functional properties of neurodynamics are the basic attributes of neural networks. Architecture deals with number of neurons and their interconnectivity. Neural networks consist of many interconnected neurons, or processing elements, with familiar characteristics, such as inputs, synaptic strengths, activation, outputs, and bias. Neurodynamics deal with training and learning, recalling, association, continuous comparison of information with existing knowledge, and classification. Neural networks process information based on parallel

decomposition of complex information into basic elements; an analogue is spectral decomposition of color and its reconstruction.

4.2 Basic Model of a Neuron

Neurons are the building blocks of an artificial neural network. Neurons are also referred as preprocessing elements. Each neuron has set of inputs ($x_1, x_2, \ldots, x_n$). Each input element (signal) is weighted (multiplied by weights (w_i)) and reached processing element. In addition, a neuron has a bias term, a threshold value (Θ) that has to be reached or exceeded to produce signal, a nonlinear function (f) that acts on activation (v) (weighted signal), and finally produce an output O. As a neuron becomes a part of a number of neurons in a network is referred to as a node. Inputs, weights, activation signals, output, threshold, and nonlinear function are written as $x_{ij}, w_{ij}, v_i, O_i, \Theta_i, f_i$, respectively. The basic architecture of a neural network is shown in Figure 4.1.

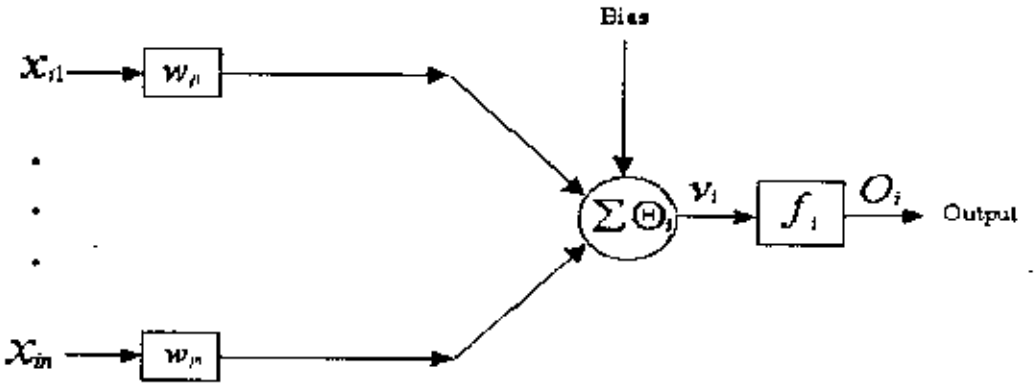

Figure 4.1: Basic architecture of a neural network model.

The transfer function of the basic model is described by the relation

$$O_i = f_i\left(\sum_{j=1}^{n} w_{ij} x_{ij}\right). \tag{4.1}$$

The neuron's activation condition is

$$\left(\sum_{j=1}^{n} w_{ij} x_{ij}\right) \geq \Theta_i, \tag{4.2}$$

where the index i represents the neuron in question and j represents the inputs neurons. The response of the neuron is bounded using a nonlinear function, known as activation function. Most common activation function used in neural networks is the Sigmoidal function that is monotonic, bounded and easily differentiable.

4.3 Learning in Artificial Neural Networks

Learning is the process by which neural network adjusts itself according to input stimuli to produce a desired output. It is a continuous classification process. When a set of input signal is represented to neural network it either recognizes or produces a new class. During learning, neural networks changes its synaptic weights so that its outputs converge to the desired outputs. Neural network completes its learning when the outputs are same as the desired outputs. At the end of learning, neural networks acquire knowledge and store them in the current weights. Some of the learning rules and types of learning are discussed below.

4.3.1 Delta Rule

It is also known as Widrow-Hoff rule. It essentially says that the adjustment made to a synaptic weight of a neuron will be proportional to the product of the error signal and the input signal. Let Δw_{kj}^{n} denotes the adjustment to the synaptic weight w_{kj} of neuron k excited by element x_j^n of the signal vector x^n at time step n. The delta rule says

$$\Delta w_{kj}^{n} = \eta e_k^n x_j^n, \qquad (4.3)$$

where η is a positive constant determining the rate of learning.

4.3.2 Supervised Learning

During the process of training, neural network produces an output response with respect to input stimuli. The network compares this output with the desired output and produces error signal. This error signal is a measure of network performance. All the error signals are summed and averaged to yield a cost function, a function of free parameters such as weights, bias. Then the weights are adjusted to minimize this cost function. As the error minimization process is directed according to the mismatch

from desired outputs, hence the name supervised learning. The supervised learning maps inputs to corresponding output patterns.

4.3.3 Unsupervised Learning

Unlike the supervised learning, unsupervised learning does not require the knowledge of the desired output. Unsupervised learning is suited for data classification by the neural network on its own. The network categorizes the data based on the interdependencies detected within them. For the implementation of unsupervised learning one may use competitive learning. For a network of two layers, input layer and competitive layer, as all input data are presented all neurons in the competitive layer compete with one another to respond to the input feature. The network learns based on "winner-takes-all" strategy in which a neuron with the greatest total input "wins" the competition and turns on, all other neurons are then switched off.

In the learning process, irrespective to types and rules of the process one requires a minimization technique. Following section discusses the problem of unconstrained minimization and some of the techniques.

4.4 Unconstrained Optimization

Consider a continuously differentiable cost function $\varepsilon(w)$ with respect to weight vector w. Cost function projects the weight vector into real number. One needs to find weight vector w^* for which cost function will be minimum. The necessary condition is

$$\varepsilon(w^*) \leq \varepsilon(w) \tag{4.4}$$

This is an unconstrained optimization problem. The optimum condition for the problem is $\nabla \varepsilon(w^*) = 0,$ (4.5)

where ∇ is the gradient operator

$$\nabla = \left[\frac{\partial}{\partial w_1}, \frac{\partial}{\partial w_2}, \ldots, \frac{\partial}{\partial w_m}\right]^T. \tag{4.6}$$

A class of unconstrained optimization algorithms is suitable for adaptive learning. Usually, a local iterative descent algorithm is used.

4.4.1 Method of Steepest Descent

In this algorithm, weights are adjusted according to the gradient of cost function. The goal is to find a weight vector in a direction so that the gradient vector $\nabla\varepsilon(w)$ is reduced at every step of iteration. For convenience, $g = \nabla\varepsilon(w)$. The formal description of steepest descent algorithm is

$$\Delta w^n = w^{n+1} - w^n = -\eta g^n \tag{4.7}$$

Equation (4.7) is in fact a formal statement of the error-correction rule.

The gradient vector is approximated applying first order Taylor series expansion around weight vector w^{n+1} and it is as follows

$$\begin{aligned}\varepsilon\left(w^{n+1}\right) &\approx \varepsilon\left(w^n\right) - \eta g^{n^T}\Delta w^n \\ &= \varepsilon\left(w^n\right) - \eta g^{n^T} g^n = \varepsilon\left(w^n\right) - \eta\left\|g^n\right\|\end{aligned}$$

This ensures the reduction of cost function in each step. However, this is applicable for small enough learning rates.

4.4.2 Least-Mean-Square (Delta Rule) Algorithm

The Least-Mean-Square (LMS) algorithm is developed using instantaneous values of cost function. The cost function is given by

$$\varepsilon\left(w^n\right) = \frac{1}{2}e^{n^2}, \tag{4.8}$$

where e^n is the error signal measured at time n. Differentiation of Equation 4.8 with respect to the weight vector w, yields

$$\frac{\partial\varepsilon\left(w^n\right)}{\partial w^n} = e^n\frac{\partial e^n}{\partial w^n} \tag{4.9}$$

LMS algorithm deals with linear neuron and the error signal is

$$e^n = d^n - x^{n^T}w^n. \tag{4.10}$$

Hence, $\frac{\partial e^n}{\partial w^n} = -x^n$, and $\frac{\partial \varepsilon(w^n)}{\partial w^n} = -x^n e^n$. Using an iterative descent approach, the LMS algorithm can be formulated as follows

$$\hat{w}^{n+1} - \hat{w}^n = -\eta\, x^n e^n \,, \qquad (4\ 11)$$

where η is the learning rate parameter. LMS algorithm provides an estimate of weight vectors. The important point is that unlike the steepest descent algorithm, LMS algorithm does not require knowledge of the statistics of the environment.

For strongly nonlinear problems, simple network architecture discussed so far cannot yield plausible response. One has to employ, more complex models like multilayer perceptron models. Following section gives the basic outline of a multilayer perceptron.

4.5 Multilayer Perceptron

Multilayer feed forward networks are an important class of neural networks. Typically it consists of input units that constitute the input layer, one or more hidden layers of computational nodes and one output layer of computational nodes. Input signals are passed in a forward direction from layer to layer. These neural networks are called multilayer perceptrons. Multilayer perceptrons are trained in a supervised learning manner based on error-correction learning. Error back-propagation learning consists of forward pass and backward pass. During forward pass input signals are passed through hidden layers to output layers. Weights are retained unchanged. In output layer, actual responses are subtracted from target outputs to produce error signals. During backward pass, these error signals are passed backward to adjust weights according to error correction rule to move the actual outputs closer to the target outputs.

4.5.1 Characteristics of a Multilayer Perceptron

A multilayer perceptron has three distinctive characteristics. They are as follows.

- Each neuron behaves according to a smooth, nonlinear activation function. Sigmoidal function is generally used for this purpose

$$y_j = \frac{1}{1+\exp(-v_j)},$$

where v_j is the local field (that is, the weighted sum of all synapse inputs plus the bias), and y_j the output of the neuron

- The networks consist of layers of hidden neurons. These hidden neurons lie between input and output layers and not part of the input or output neurons. They enable the networks to learn nonlinear relationship by extracting feature from input patterns.
- Each neuron in a layer is connected to the all neurons in the previous layer through synapses of the network. The weights are adjusted to change the connectivity of the network.

4.5.2 Some Preliminaries

Figure 4.2 shows the architecture of a multilayer perceptron with two hidden layers Any node in any layer is fully connected to all nodes in the previous layer. Thus, they are fully connected. Signal flow progresses in a forward direction from left to right and from one layer to another.

Two kinds of signals are identified in this network. They are function signals and error signals. Function signal comes in input layer, propagates through the hidden layers and emerges at the output layer. It performs a useful function at the output layer. At each neuron, through which it propagates, it is calculated as a function of inputs and associated weights. It is also referred to as input signal. While, an error signal is generated at an output node and passed backward. During this pass, weights are modified based on error signal. Figure 4.3 shows schematically the two signals The backward propagation algorithm is described hereafter. This is one of the most common algorithms for multiple layer perceptron models.

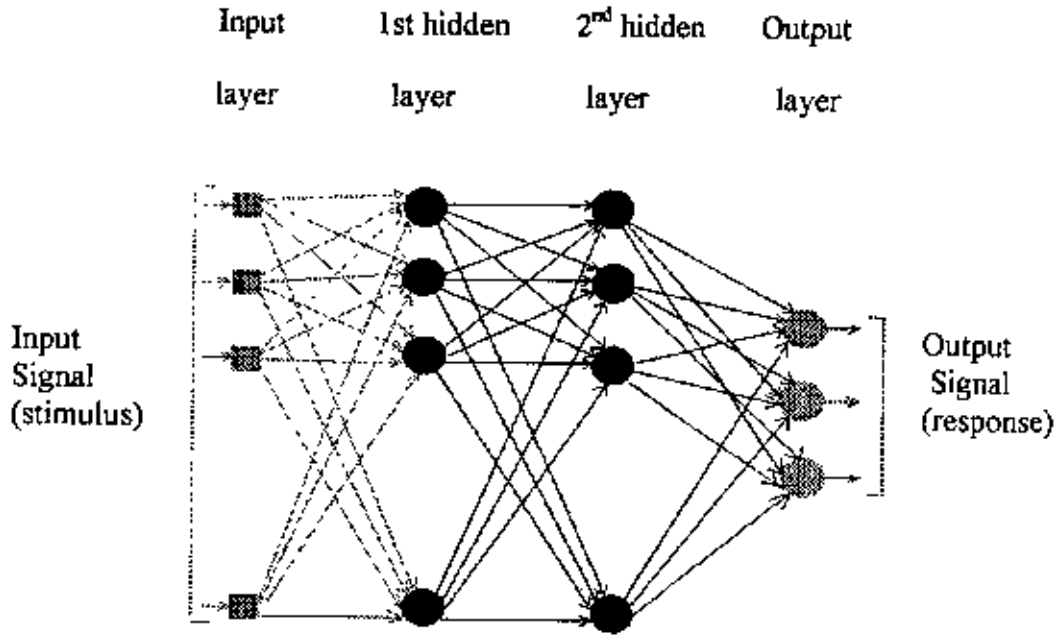

Figure 4.2: Architecture of a multilayer perceptron with two hidden layers.

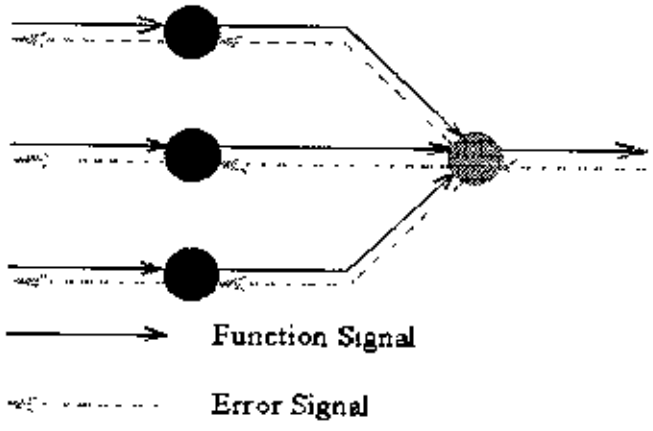

Figure 4.3 Illustration of the directions of two basic signal flows in a multilayer perceptron: forward propagation of function signals and back-propagation of error signals.

4.5.3 Back Propagation Algorithm

The back propagation algorithm is one of the most popular algorithms based on supervised learning. The network output is compared and the mismatch or error is propagated backward through the network. During this back propagation of error the weights are adjusted. This process is continued in an iterative manner The algorithm is briefly described here.

The error signal of neuron j at iteration n (that is, presentation of the n^{th} training pattern) is defined by

$$e_j^n = d_j^n - y_j^n. \tag{4.12}$$

Here, neuron j is an output node. By definition, the instantaneous value of total error energy is

$$\varepsilon^n = \frac{1}{2}\sum_{j \in C} e_j^{n^2}, \tag{4.13}$$

where C is all the neurons in the output layer of the network. The average squared error energy is

$$\varepsilon_{av}^n = \frac{1}{N}\sum_{n=1}^{N} \varepsilon^n, \tag{4.14}$$

where N is the total number of patterns in the training set. Average error energy is a function of all free parameters (weights, bias). It represents the performance of learning process. Learning process minimizes ε_{av}^n. In a simple training method, weights are updated on a pattern-by-pattern basis until one epoch is dealt with The weights are modified in accordance with respective errors computed for each pattern presented to the network. Figure 4.4 shows the transformation of function signals, produced by a set of neurons situated at the previous layer of neuron i layer to neuron j.

The induced local field is

$$v_j^n = \sum_{i=0}^{m} w_{ij}^n y_i^n , \tag{4.15}$$

which is the input to the activation function associated with neuron j. For bias, synaptic weight is given by $w_{j0}^n = b_j^n$. The output of neuron j is

$$y_j^n = \varphi_j\left(v_j^n\right) \tag{4.16}$$

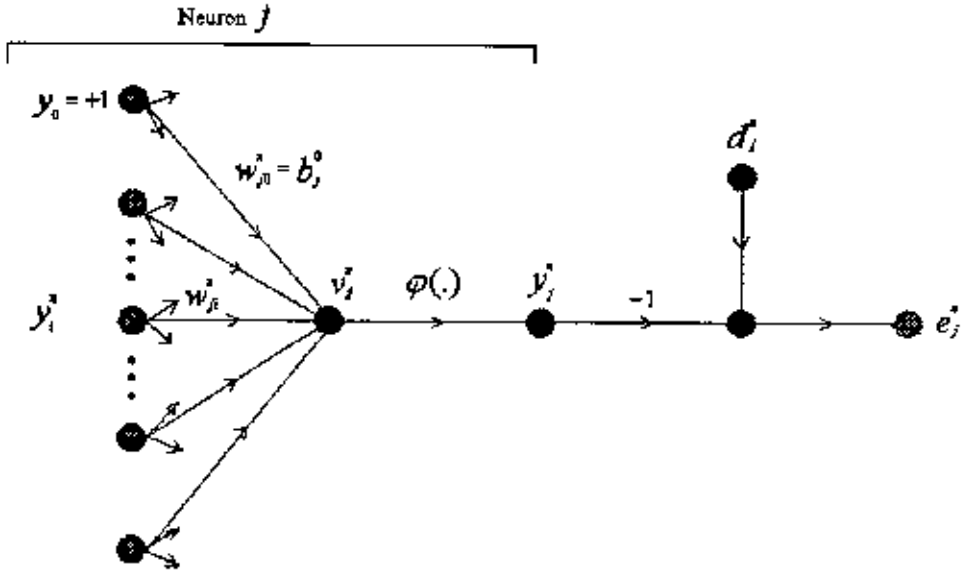

Figure 4.4: Signal-flow graph highlighting the details of output neuron j.

Back propagation algorithm applies weight correction principle just like LMS algorithm. According to this principle, weight correction Δw_{ji}^n is proportional to the partial derivative of $\frac{\partial \varepsilon^n}{\partial w_{ji}^n}$. According to the chain rule, one has

$$\frac{\partial \varepsilon^n}{\partial w_{ji}^n} = \frac{\partial \varepsilon^n}{\partial e_j^n} \cdot \frac{\partial e_j^n}{\partial y_j^n} \cdot \frac{\partial y_j^n}{\partial v_j^n} \cdot \frac{\partial v_j^n}{\partial w_{ji}^n} \tag{4.17}$$

The partial derivative represents a sensitivity factor, determining the direction of search in weight space for the synaptic weight w_{ji}.

Differentiating both sides of Equation (4.13) with respect to e_j^n, one obtains

$$\frac{\partial \varepsilon^n}{\partial e_j^n} = e_j^n. \tag{4.18}$$

While, differentiation of Equation (4.12) with respect to y_j^n, yields

$$\frac{\partial e_j^n}{\partial y_j^n} = -1 \tag{4.19}$$

Further, differentiating Equation (4.16) with respect to v_j^n, one gets

$$\frac{\partial y_j^n}{\partial v_j^n} = \varphi'\left(v_j^n\right) \tag{4.20}$$

Differentiation of Equation (4.15) with respect to w_{ji}^n yields

$$\frac{\partial v_j^n}{\partial w_{ji}^n} = y_i^n \tag{4.21}$$

Finally, one derives the form

$$\frac{\partial \varepsilon^n}{\partial w_{ji}^n} = -e_j^n \cdot \varphi'\left(v_j^n\right) \cdot y_i^n. \tag{4.22}$$

Now, the correction Δw_{ji}^n is defined by the delta rule:

$$\Delta w_{ji}^n = -\eta \cdot \frac{\partial \varepsilon^n}{\partial w_{ji}^n}, \tag{4.23}$$

where η is the learning -rate parameter of the back-propagation algorithm. The minus sign in Equation (4.23) is to force the weight change in the descent direction. From Equations (4.22) and (4.23) one obtains

$$\Delta w_{ji}^n = \eta \cdot \delta_j^n \cdot y_i^n, \tag{4.24}$$

where the local gradient δ_j^n is defined by

$$\delta_j^n = \frac{\partial \varepsilon^n}{\partial v_j^n} = \frac{\partial \varepsilon^n}{\partial y_j^n} \cdot \frac{\partial y_j^n}{\partial v_j^n} = -e_j^n \cdot \varphi'\left(v_j^n\right). \tag{4.25}$$

The local gradient is equal to the product of error signal and derivative of activation function $\varphi'(v_j^n)$ Weights are changed according to local gradients. Desired response is applied for each neuron at the output layer. For output layer neuron j, error signal can be calculated using Equation (4.12), and the local gradient using Equation (4 25).

Now when neuron is located in hidden layer, it has no specified desired response. Figure 4.5 shows hidden neuron signal flow graph for a neuron in a hidden layer. Error signal for neuron j is calculated from error signals of all neurons to which neuron j is connected. According to Equation (4.25), the local gradient δ_j^n hidden neuron j can be written as

$$\delta_j^n = \frac{\partial \varepsilon^n}{\partial y_j^n} \cdot \frac{\partial y_j^n}{\partial v_j^n} = -\frac{\partial \varepsilon^n}{\partial y_j^n} \cdot \varphi'(v_j^n). \tag{4.25}$$

To calculate the partial derivative of $\frac{\partial \varepsilon^n}{\partial y_j^n}$, one may proceed as follows. From Figure 4.5, one has

$$\varepsilon^n = \frac{1}{2}\sum_{k \in C} e_k^{n^2}. \tag{4.26}$$

Differentiating Equation (4.26) with respect to the function signal y_j^n, one obtains

$$\frac{\partial \varepsilon^n}{\partial y_j^n} = \sum e_k^n \cdot \frac{\partial e_k^n}{\partial y_j^n}. \tag{4.27}$$

Applying chain rule in Equation (4.27), one can write

$$\frac{\partial \varepsilon^n}{\partial y_j^n} = \sum e_k^n \cdot \frac{\partial e_k^n}{\partial v_k^n} \cdot \frac{\partial v_k^n}{\partial y_j^n}. \tag{4 28}$$

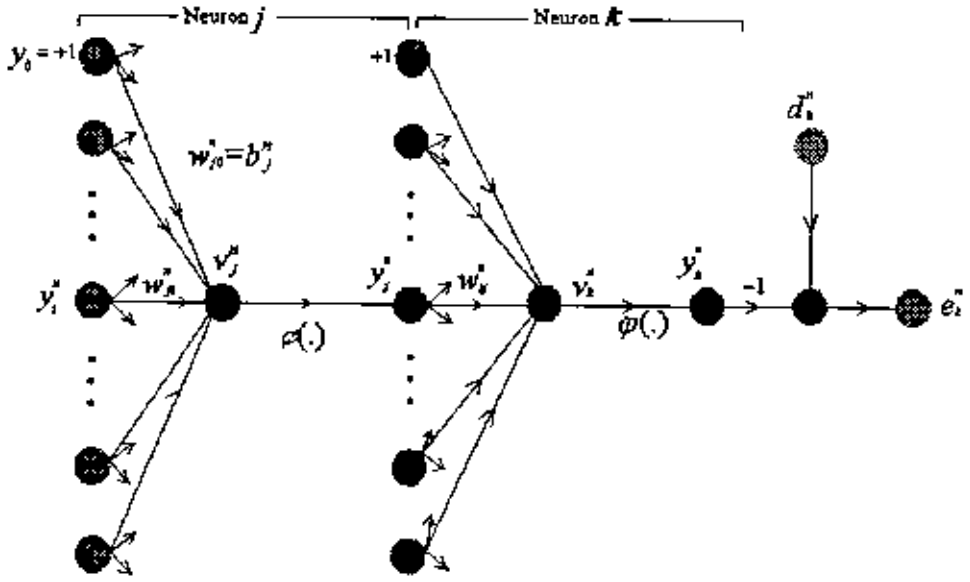

Figure 4.5: Flow graph highlighting the details of output neuron k connected to hidden neuron j.

However from Figure 4.5, it can be noted that when neuron k is output node

$$e_k^n = d_k^n - y_k^n = d_k^n - \varphi_k(v_k^n). \tag{4.29}$$

Differentiating the Equation (4.29) with respect to v_k^n, one obtains

$$\frac{\partial e_k^n}{\partial v_k^n} = -\varphi'_k(v_k^n). \tag{4.30}$$

From Figure 4.5, for neuron k the induced local field is

$$v_k^n = \sum_{j=0}^{m} w_{kj}^n \cdot y_j^n, \tag{4.31}$$

where m is the total number of inputs (excluding the bias) applied to neuron k. The synaptic weight is given by $w_{k0}^n = b_k^n$. Differentiating Equation (4.31) with respect to y_j^n, one obtains

$$\frac{\partial v_k^n}{\partial y_j^n} = w_{kj}^n \tag{4.32}$$

Equation (4.28) now has the form

$$\frac{\partial \varepsilon^n}{\partial y_j^n} = -\sum_k e_k^n \cdot \varphi'_k\left(v_k^n\right) \cdot w_{kj}^n = -\sum_k \delta_k^n \cdot w_{kj}^n . \tag{4.33}$$

The local gradient for hidden neuron now assumes the form

$$\delta_j^n = \varphi'_j\left(v_j^n\right) \sum_k \delta_k^n \cdot w_{kj}^n . \tag{4.34}$$

The back-propagation algorithm performs forward and backward passes computation. The weights are remained constant during the forward pass and neurons produce function signals. The back computation starts by passing the error signal from output layer to the hidden layer direction. During this pass, each neuron computes its local gradient and changes its associated weights according to Delta rule.

Options for the activation function are discussed in the following subsection.

4.5.4 Activation Function

Local gradient computation needs the derivatives of activation function. Thus, a requirement for the activation function is its continuous differentiability. Sigmoidal activation function is commonly used as an activation function. The general form of sigmoidal function (logistic function) is

$$\varphi_j\left(v_j^n\right) = \frac{1}{1+\exp\left(-av_j^n\right)}, \tag{4.35}$$

where $a > 0$ and $-\infty < v_j^n < \infty$. Here, v_j^n is the induced local field of neuron j. The derivative is given by

$$\varphi'_j\left(v_j^n\right) = \frac{a \cdot \exp\left(-av_j^n\right)}{\left[1+\exp\left(-av_j^n\right)\right]^2} . \tag{4.36}$$

Using $y_j^n = \varphi_j\left(v_j^n\right)$, one can eliminate the exponential term reducing to

$$\varphi'_j\left(v_j^n\right) = a \cdot y_j^n \cdot \left[1 - y_j^n\right] \tag{4.37}$$

Another common activation function is hyperbolic tangent function.

$$\varphi_j(v_j^n) = a \cdot \tanh(bv_j^n) \tag{4.38}$$

where a and b are constants. Hyperbolic tangent function is just the logistic function rescaled and biased. Differentiation of Equation (4.38) with respect to v_j^n yields

$$\varphi'_j(v_j^n) = ab \cdot \operatorname{sech}^2(bv_j^n) = ab(1 - \tanh^2(bv_j^n)) = \frac{b}{a}[a - y_j^n][a + y_j^n] \tag{4.39}$$

Following is a short note on the rate of learning.

4.5.5 Rate of Learning

Learning rate parameter, η, has a significant effect on network performance. Back propagation algorithm provides an approximate trajectory in weight space according to steepest descent rule. A smaller learning rate parameter yields smoother trajectory, smaller changes in weight, and network settles to a acceptable solution. However, more time is needed to train the network. Large learning rate parameter adjusts the weights faster. But the network may become unstable and oscillate around solution. To speed up the learning rate without oscillation, one can add momentum term to account a fraction of previous time step weight change. This additional term tends to keep the weight changes going in the same direction. The weight correction equation with the momentum term is

$$\Delta w_{ji}^n = \alpha \cdot \Delta w_{ji}^{n-1} + \eta \delta_j^n \cdot y_i^n, \tag{4.40}$$

where α is usually a positive member called momentum constant. It varies from 0 to 1.

Some implementation issues like mode of training and stopping criteria are discussed below.

4.5.6 Mode of Training

An epoch consists of entire training patterns, that is, a complete training set. The learning process is performed on an epoch-by-epoch basis. It continues until the

weights and bias level of the network attain stability and average error goes to acceptable value. For s given training set, back propagation learning may proceed in one of two basic ways: sequential mode and batch mode.

Sequential mode of back propagation learning is also referred to as on-line, pattern, stochastic mode. According to this mode, back propagation algorithm theories are derived Learning is performed on pattern by pattern by basis. Let us consider an epoch with N number of training examples such as $(x^1,d^1),(x^2,d^2),...,(x^N,d^N)$. The first pattern, (x^1,d^1), is presented to the network, forward and backward computations are performed, and weights and bias level updated. Then the second pattern (x^2,d^2) is presented to the network, forward and backward computations are performed again, weights and bias level are updated further and so on up to the last pattern (x^N,d^N).

In batch mode of learning, weights are updated after the presentation of all patterns in an epoch. Cost function is calculated based on Equations (4.18) and (4 19) as follows

$$\varepsilon_{av}=\frac{1}{2N}\sum_{n=1}^{N}\sum_{j\in C}e_j^{n2}. \qquad (4.41)$$

Error signal e_j^n is produced by neuron j for pattern n In Equation (4.41), the inner summation with respect to j is performed for each pattern over all neurons in the output layer. The outer summation with respect to n is performed for all patterns under an epoch. Weights are updated according to delta rule as follows

$$\Delta w_{ji}^n=-\eta\cdot\frac{\partial\varepsilon_{av}}{\partial w_{ji}^n}=\frac{\eta}{N}\sum_{n=1}^{N}e_j^n\frac{\partial e_j^n}{\partial w_{ji}^n}. \qquad (4.42)$$

Sequential mode of learning requires less storage for synaptic connections than batch mode. Pattern-by-pattern updating of weights makes the network search stochastic in weight space and helps the network to avoid local minimum problem. When the data are redundant, batch mode of learning is quite capable of taking the advantage of redundancy. Sequential mode of learning is highly popular for two reasons. The

algorithm is simple to implement, and it provides solution to large and difficult problems.

4.5.7 Stopping Criteria

Back propagation algorithm cannot converge and there is no well-defined stopping criterion. One can consider some reasonable criteria for termination of learning process. Local or global mininum error surface is one of them. The back-propagation algorithm can be considered to have converged when the Euclidean norm of the gradient vector reaches a sufficiently small gradient threshold. This criterion requires computation of derivative of error surface and its training time is high. These make the criterion disadvantageous.

Another stopping criterion of the back-propagation algorithm could be when the absolute rate of change in the average squared error per epoch is sufficiently small. Generally, the rate of change of error surface ranges from 0.01 to 1 per epoch. Unfortunately, this criterion may lead to premature termination of the learning process.

4.6 Discussion

This chapter explains the basics of a neural network. A reasonably good account of the back-propagation algorithm is given In this work, similar back-propagation model has been adapted. The neural network devised here has three layers – input, output and one hidden layer. In terms of activation function, both sigmoidal and hyperbolic tangent functions are considered. A number of implementation issues and sensitivity exercises are performed to optimize the performance of the model.

CHAPTER 5

EXPLORATORY DATA ANALYSIS

This chapter discusses the exploratory data analysis performed on petrophysical data available from core measurements. The data include well locations, permeability, porosity, facies, thickness, and relative stratigraphic depth. There are data from five wells. Of these four wells have permeability and facies measurements. Histograms, and scatter plots are obtained and analyzed. Essential statistics of data such as mean, median, maximum, minimum, standard deviation, correlation coefficients between variables of interest, and some derived statistics are obtained

The outline of this chapter is as follows. Section 5.1 gives a short description on the well locations. Section 5.2 explores the essential statistics through histograms and scatter plots of the core data. Section 5.3 concludes the chapter highlighting the summarized information from exploratory data analysis.

5.1 Well Locations

Short description on well locations and the spatial distribution of some of the variables are given in this section. The reservoir domain of interest extends from about coordinate location of 731500 ft to 740500 ft in the Easting, while 5704000 ft to 5711000 ft in the Northing direction. In the figures, X direction implies Easting, while Y implies Northing directions. The coordinates (in ft) of the five wells are Well 1. (732227.4, 5710441.0), Well 2: (735445.5, 5710225.0), Well 3: (734877.5, 5705548.0), Well 4: (734877.5, 5705548.0), and Well 5. (739558.8, 5710616.0), respectively.

Figure 5.1 shows vertically averaged porosity of five wells. Interesting to note that there is a presence of high porosity zone in the northeastern part of the reservoir Figures 5.2 and 5.3 show vertically averaged maximum permeability (K_{max}), and vertically averaged vertical permeability (K_{vert}). Evidently, there are only four wells with permeability measurements. It is clear from the figures that there is a presence of permeability zone in the northeastern part of the reservoir. Figure 5.4 shows the

spatial distribution of well thickness. Figure 5.4 reveals a trend in the well thickness. The formation appears to be thinning in the eastward direction.

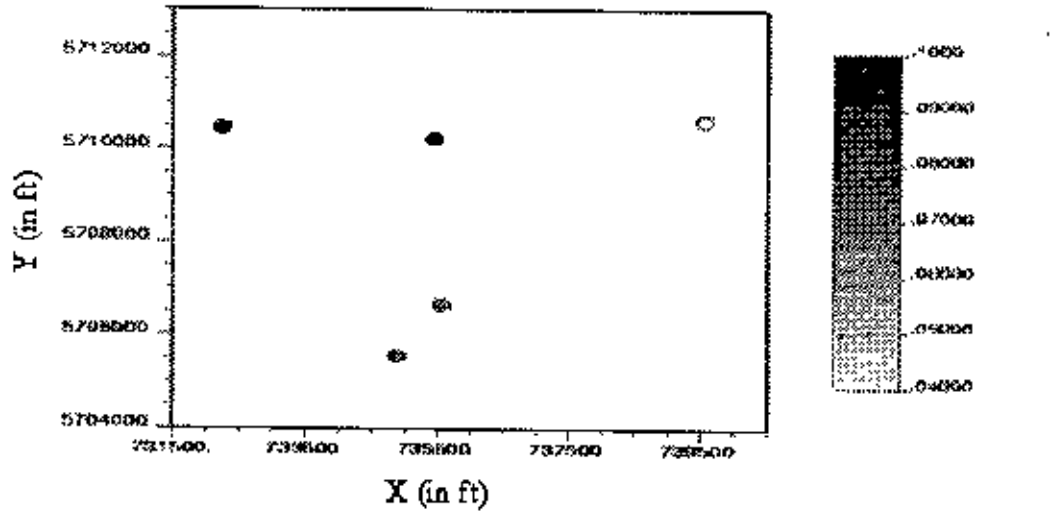

Figure 5.1: Location map of vertically averaged core porosity at well locations.

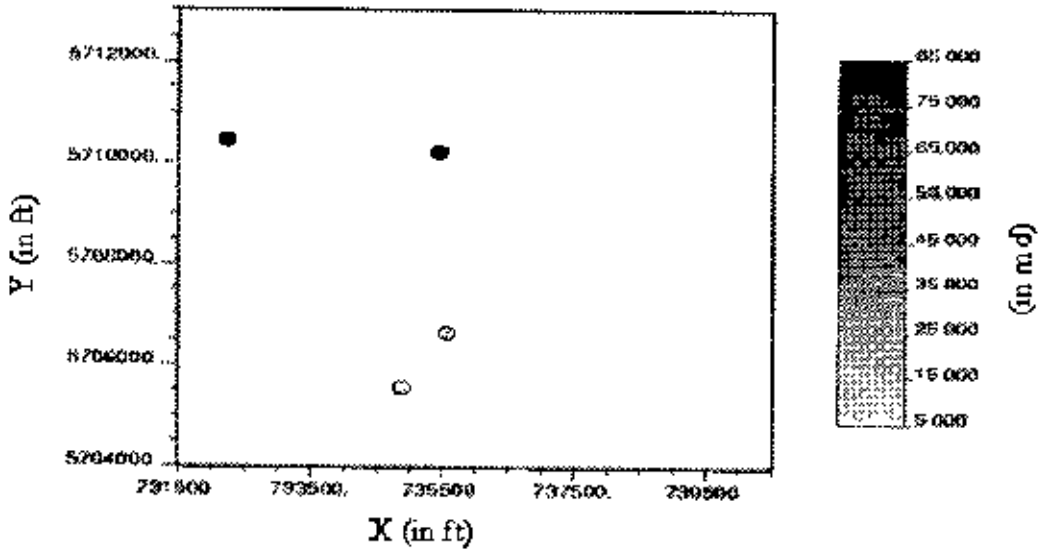

Figure 5.2: Location map of vertically averaged maximum core permeability at well locations.

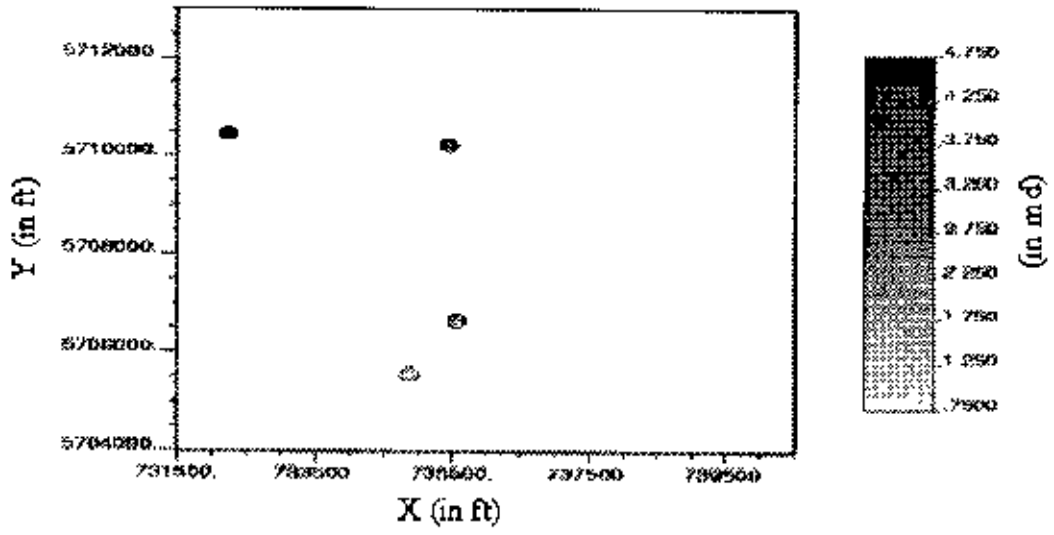

Figure 5.3: Location map of vertically averaged vertical core permeability at well locations.

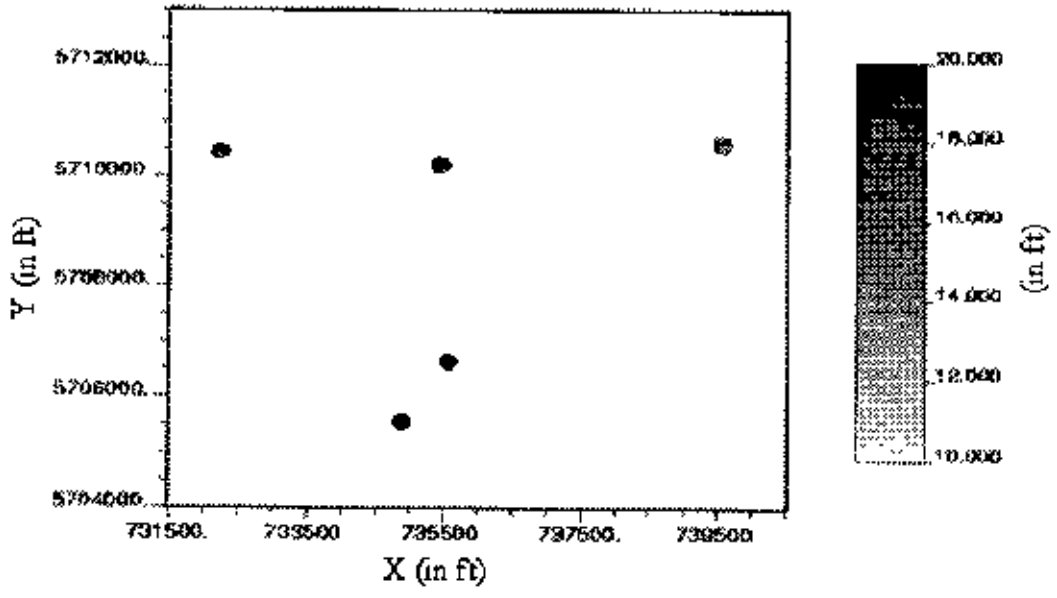

Figure 5.4: Location map of well thickness.

In the following section, some basic statistics of the core data are explored. Outlier analysis is also performed.

5.2 Univariate and Bivariate Statistics of Core Data

Histograms and scatter plots of core data are obtained. Histograms are basically plots for univariate statistics. Univariate statistics include mean, standard deviation, median, range of data. Scatter plots are for bivariate statistics between two variables Bivariate statistics include the correlation coefficient and rank correlation coefficients. Outlier information may also reveal with both histograms and scatter plots. This section also discusses correlation between the variables before and after the removal of outliers. Well-wise descriptions are given below.

5.2.1 Well 1 Statistics

Figure 5.5 shows histogram of maximum permeability (K_{max}). Mean and standard deviation of K_{max} are 62.93 and 106.36, respectively. K_{max} value ranges from 0.62 to 62.59. Figure 5 6 shows histogram of vertical permeability (K_{vert}) Mean and standard deviation of K_{vert} are 4 134 and 4.32, respectively. The maximum and minimum K_{vert} values are 0.38 and 24.4. Histogram of porosity is given in Figure 5 7. Mean and standard deviation are 0.08 and 0.04, while the values range from 0.02 to 0.16. There are 6 facies numbered 0 to 5. Figure 5.8 shows the facies histogram. Clearly, Facies 1 is the dominant facies in Well 1.

Porosity variation with facies in Well 1 is explored. Figure 5.9 shows the porosity histogram for Facies 0. Facies 0 have low porosity with mean of 0.031, and ranges from 0.017 to 0.045. Figure 5.10 shows the porosity histogram for Facies 1. Facies 1 have moderate porosity with mean 0.105, and a range from 0.062 to 0.144. Figure 5.11 shows porosity histogram for Facies 2. Facies 2 porosities are even higher than Facies 1. The mean is about 0.131, and it ranges from 0.099 to 0.161. Figure 5 12 shows the porosity histogram for Facies 3. The porosity mean is 0.111 ranging from 0.072 to 0.155.

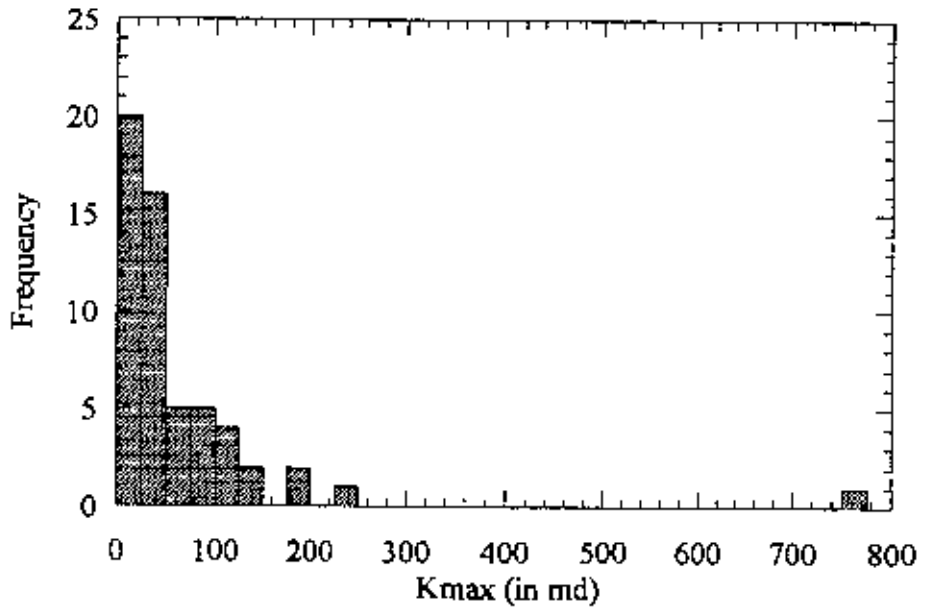

Figure 5.5: Maximum permeability histogram for Well 1.

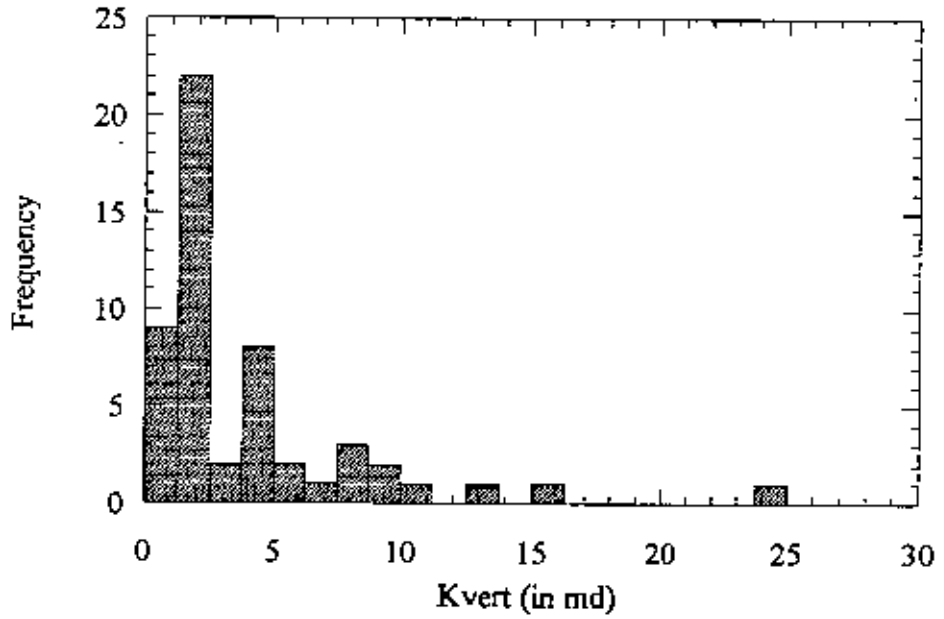

Figure 5.6: Vertical permeability histogram for Well 1.

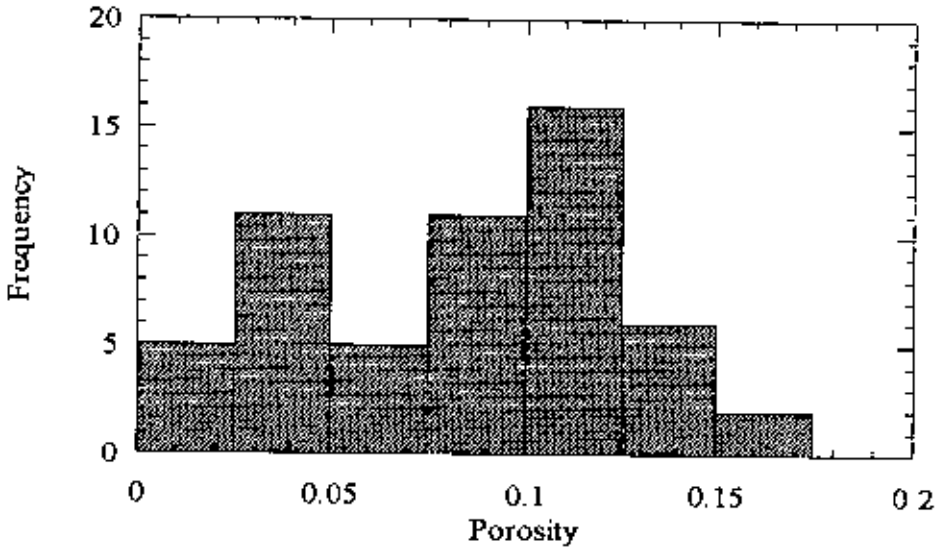

Figure 5.7: Porosity histogram for Well 1.

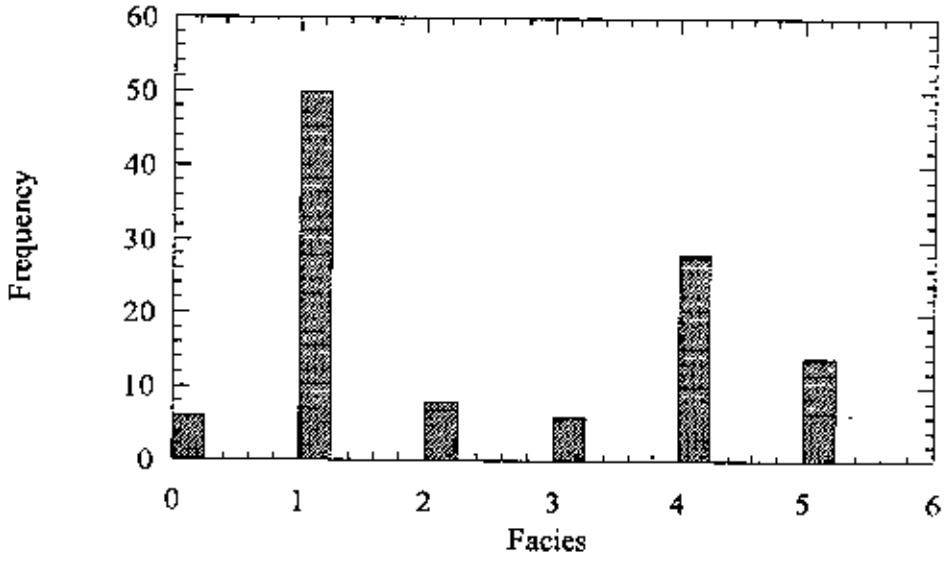

Figure 5.8: Facies histogram for Well 1.

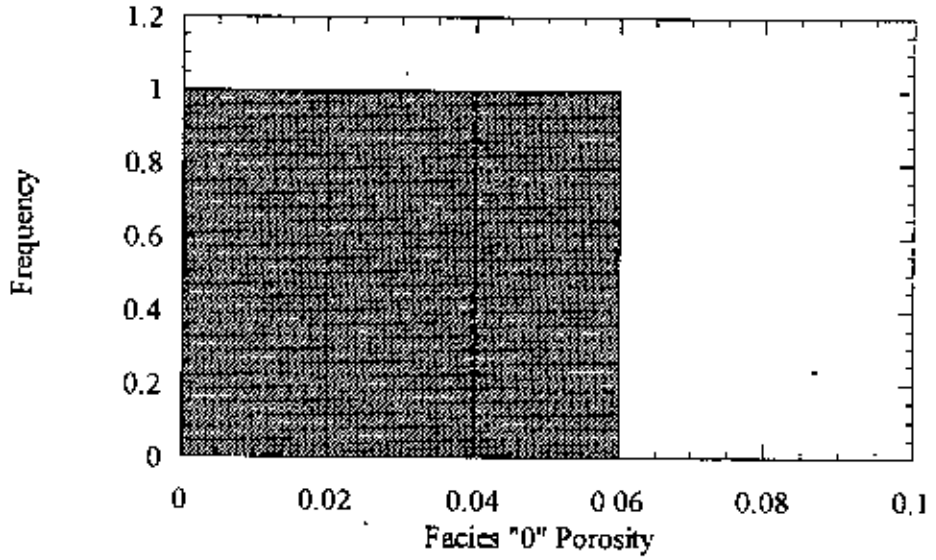

Figure 5.9: Facies 0 porosity histogram for Well 1.

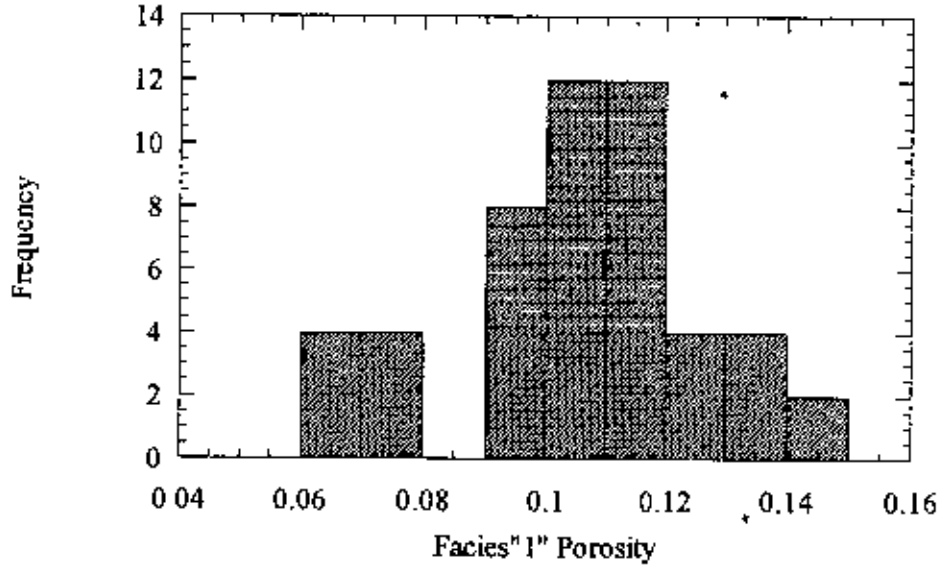

Figure 5.10: Facies 1 porosity histogram for Well 1.

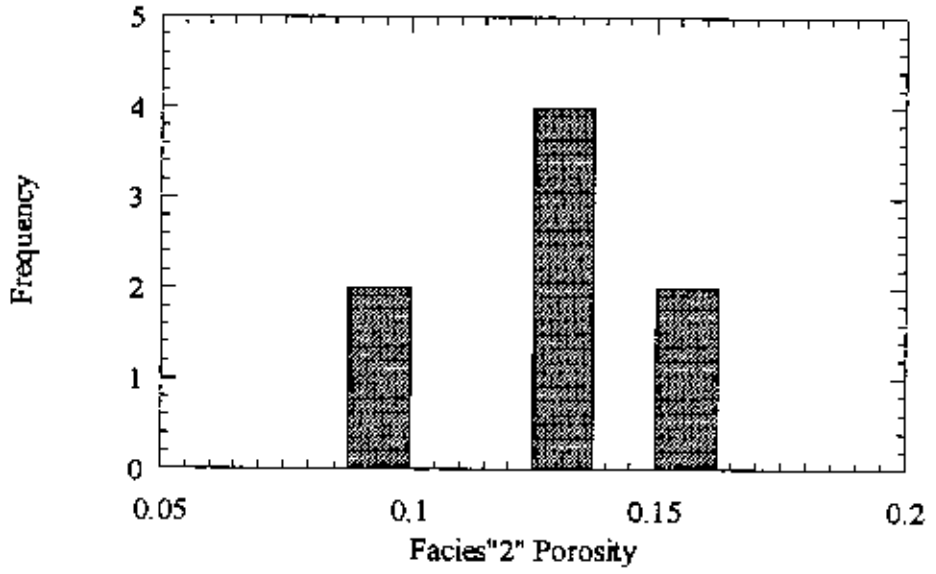

Figure 5.11: Facies 2 porosity histogram for Well 1.

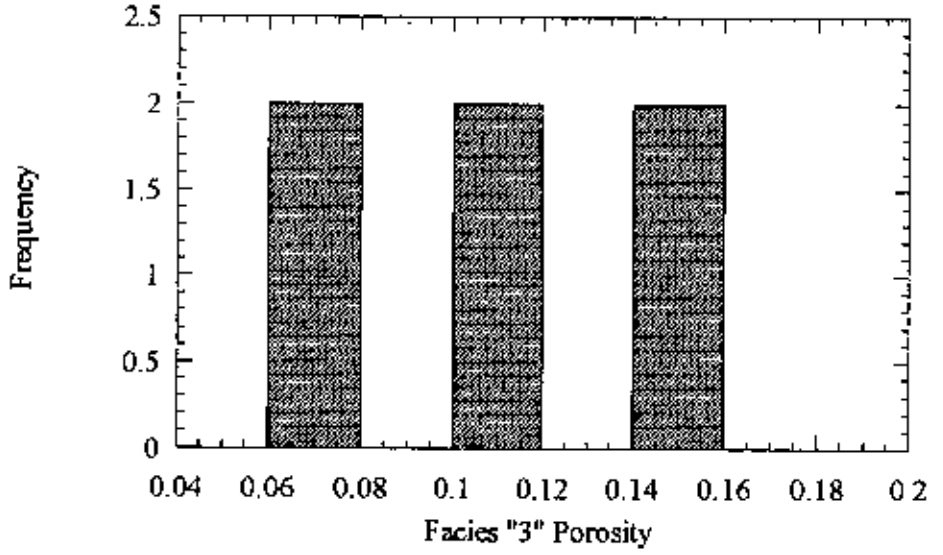

Figure 5.12: Facies 3 porosity histogram for Well 1.

Figure 5.13 shows porosity histogram for Facies 4. Mean porosity is found to be 0.048, and its range is from 0.018 to 0.111. Figure 5.14 shows the porosity histogram for Facies 5. Facies 5 have again low porosity with a mean 0.059, and range from 0.017 to 0.096

Bivariate statistics between the variables for Well 1 are explored. Figure 5.15 is the scatter plot of maximum permeability and vertical permeability. The correlation coefficient is 0.347. The poor correlation between the two permeabilities may be attributed to some outlier values. After removal of outlier, the correlation improves from 0.347 to 0.411. Figure 5.16 is the maximum permeability versus porosity scatter plot with outliers. Removal of outliers leads to improved correlation coefficient from 0.282 to 0.437. Figure 5.17 is the vertical permeability versus porosity scatter plot. Removing the outliers, improves the correlation significantly from 0.331 to 0.598. Figure 5.18 shows facies versus relative strata scatter plot. It is plotted to find a trend between facies and depth. This plot shows Faces 0 generally lies at top and Facies 4 and 5 at the bottom of the formation. Correlation is found 0 75. Figure 5.19 shows support versus porosity scatter plot with correlation 0.504. This plot indicates higher porosity values have lower support.

5.2.2 Well 2 Statistics

Similar analysis is performed for Well 2. Figure 5.20 shows maximum permeability (K_{max}) histogram. The mean and standard deviation of K_{max} are found to be 73.8 and 175.884 respectively, while the values range from 0.06 to132. Figure 5.21 shows histogram of vertical permeability (K_{vert}). The mean and standard deviation of K_{vert} are 3.109 and 4.088. The values range from 0.03 to 21.2 Figure 5.22 shows histogram of porosity It has a mean of 0 083, and standard deviation 0.036. Porosity value ranges from 0.016 to 0.193. Figure 5.23 shows facies histogram. Only Facies 0 to 4 are present. Again as in Well 1, Facies 1 is the dominant facies Remarkably, Facies 5 is absent in Well 2. Variation of porosity with facies is determined. Figure 5 24 is porosity histogram for Facies 0. Facies 0 porosity in Well 2 is slightly higher than that of Well 1, with a mean of 0.068 and range from 0.016 to 0.107.

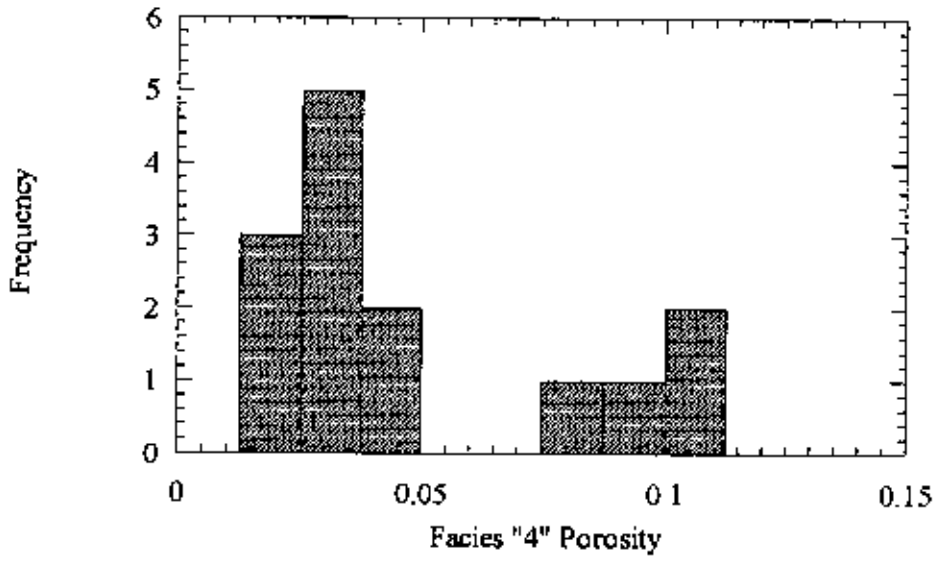

Figure 5.13: Facies 4 porosity histogram for Well 1

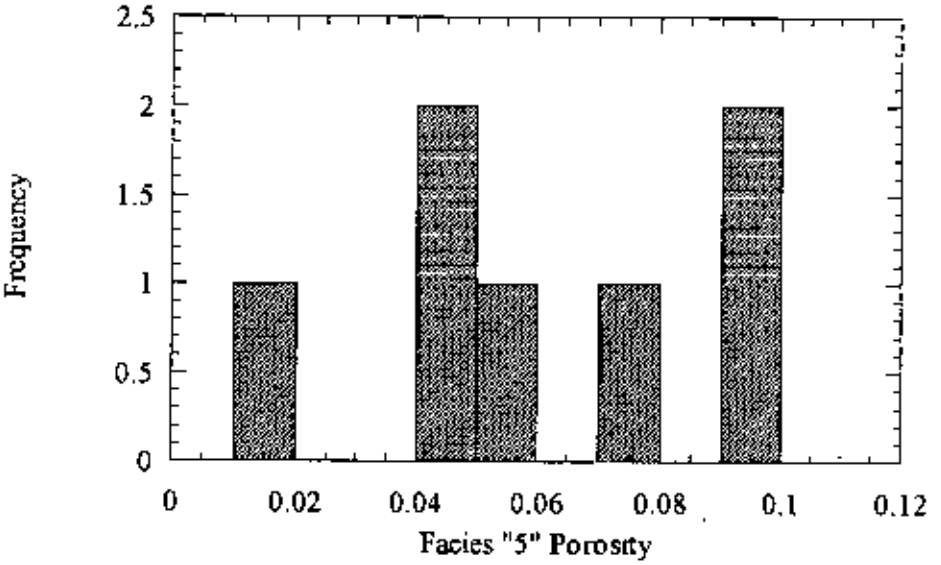

Figure 5.14: Facies 5 porosity histogram for Well 1.

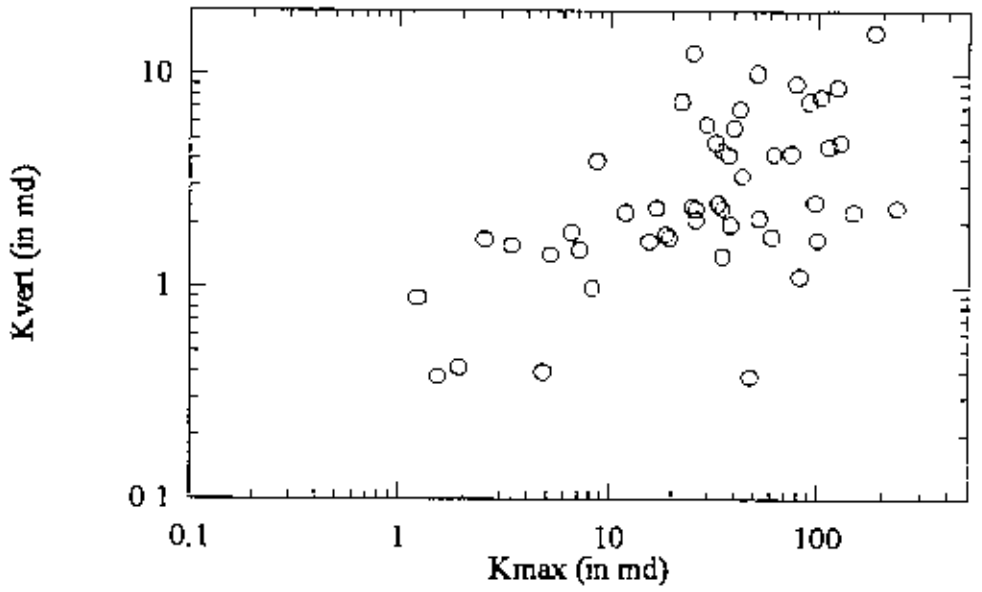

Figure 5.15: Vertical permeability versus maximum permeability scatter plot for Well 1.

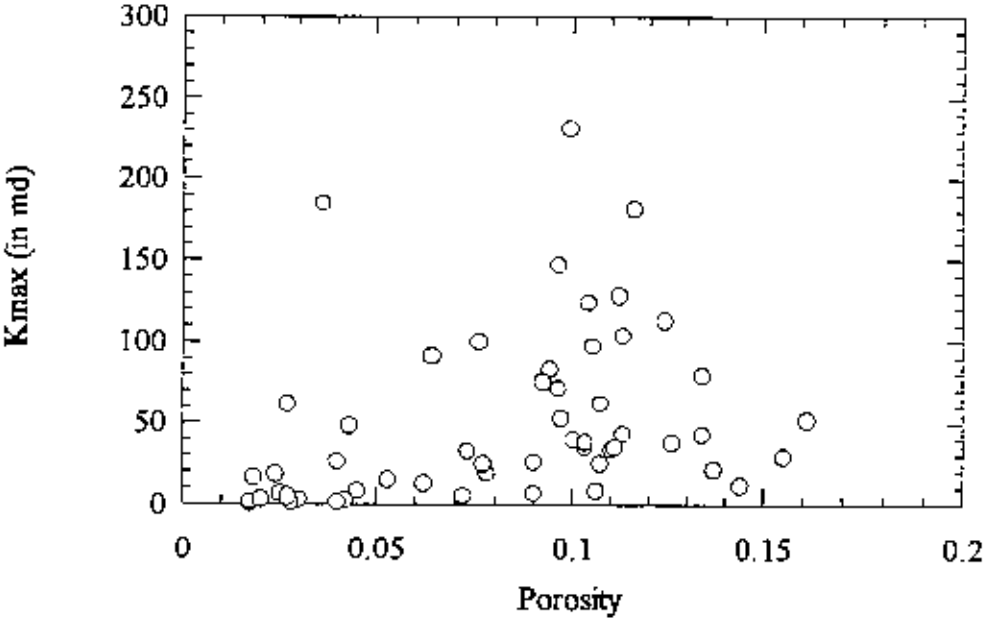

Figure 5.16: Maximum permeability versus porosity scatter plot for Well 1.

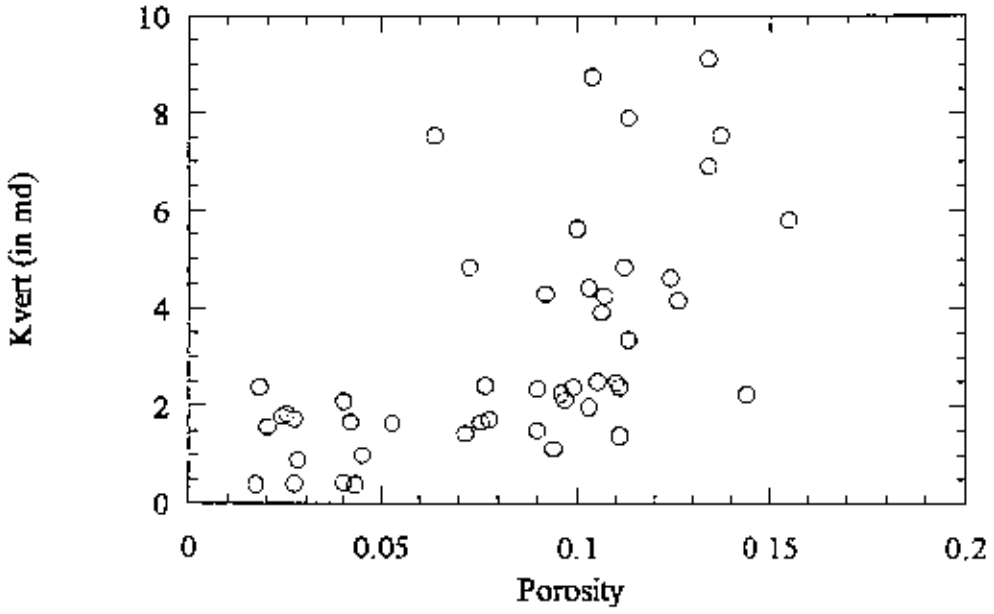

Figure 5.17: Vertical permeability versus porosity scatter plot for Well 1.

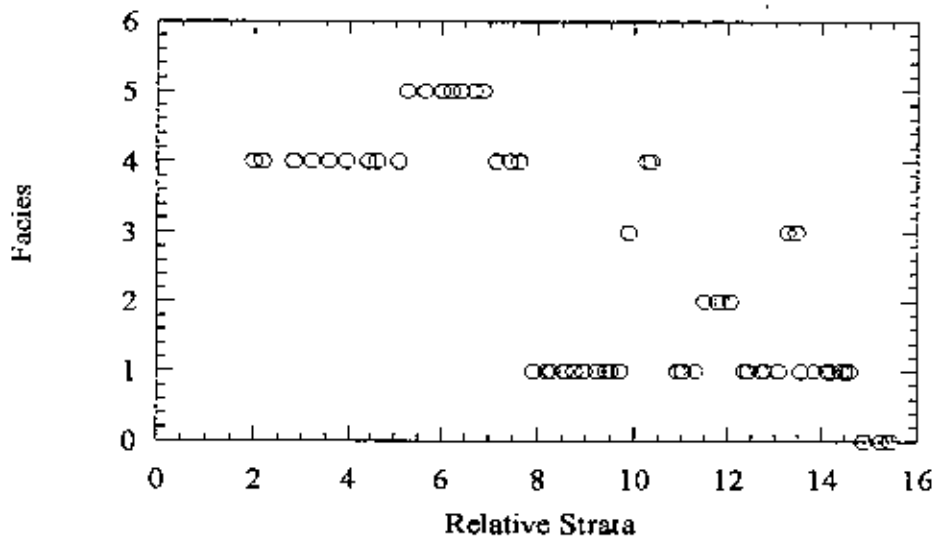

Figure 5.18: Facies versus relative strata scatter plot for Well 1.

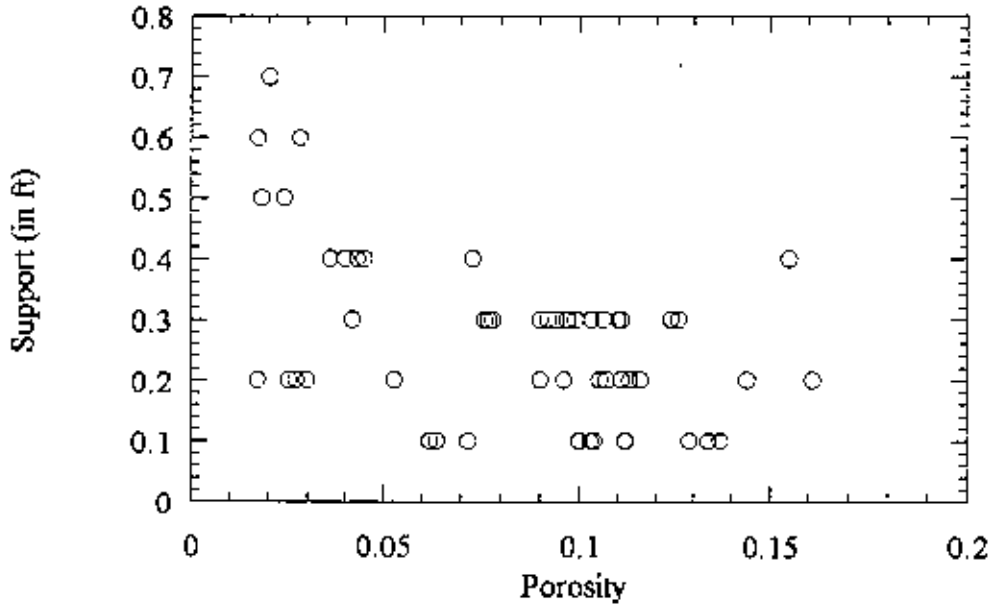

Figure 5.19: Measurement support versus porosity scatter plot for Well 1.

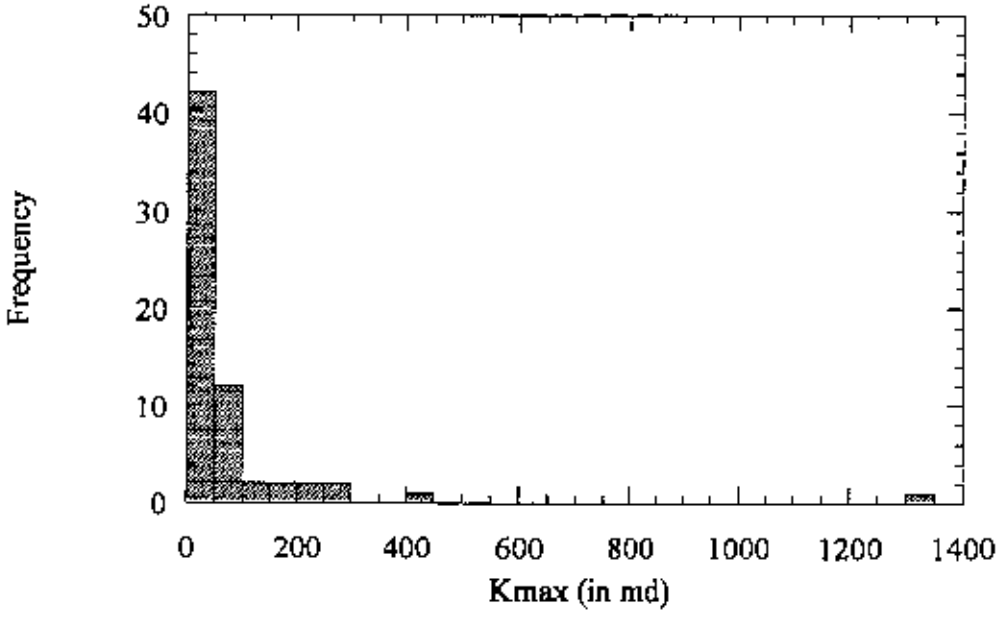

Figure 5.20: Maximum permeability histogram for Well 2.

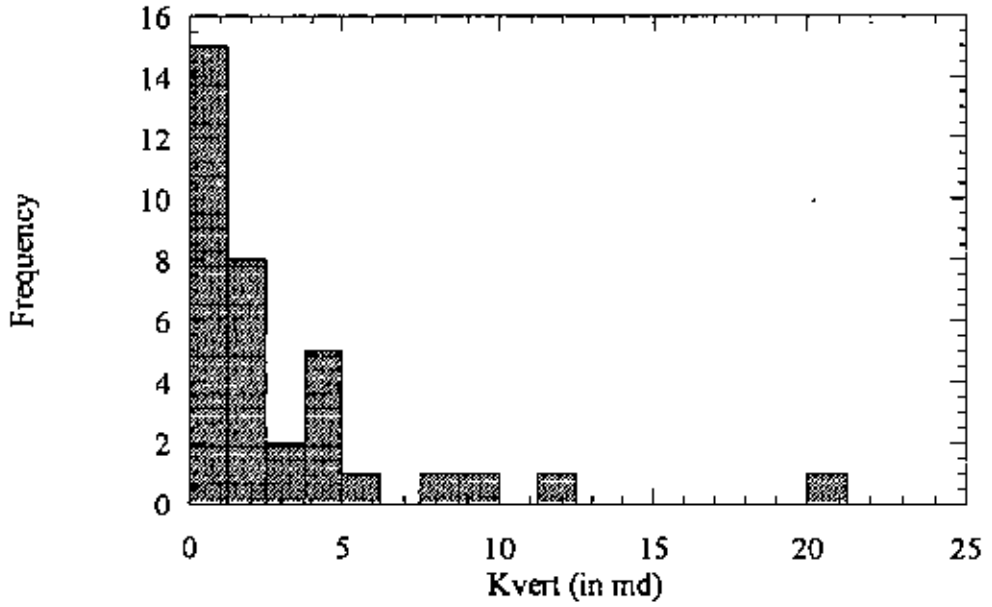

Figure 5.21: Vertical permeability histogram for Well 2

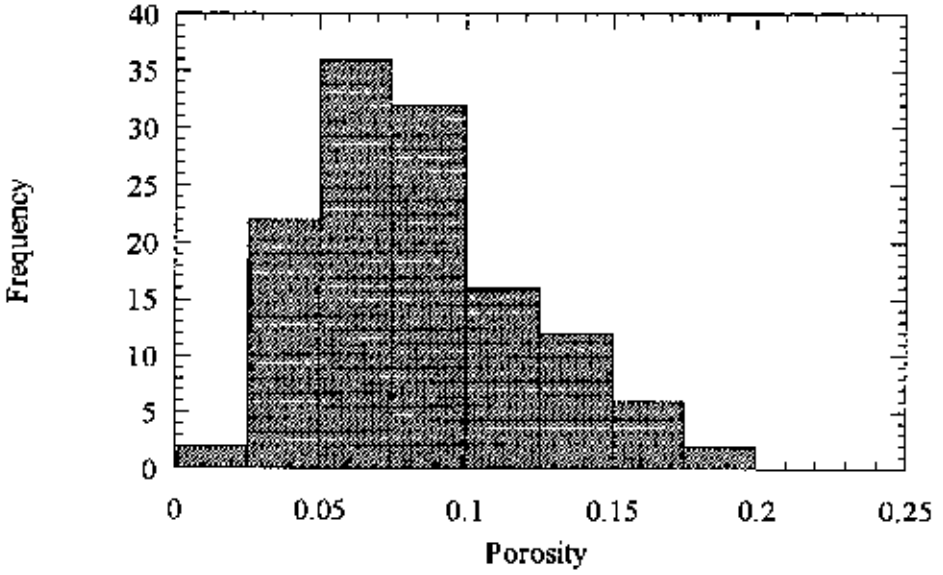

Figure 5.22: Porosity histogram for Well 2.

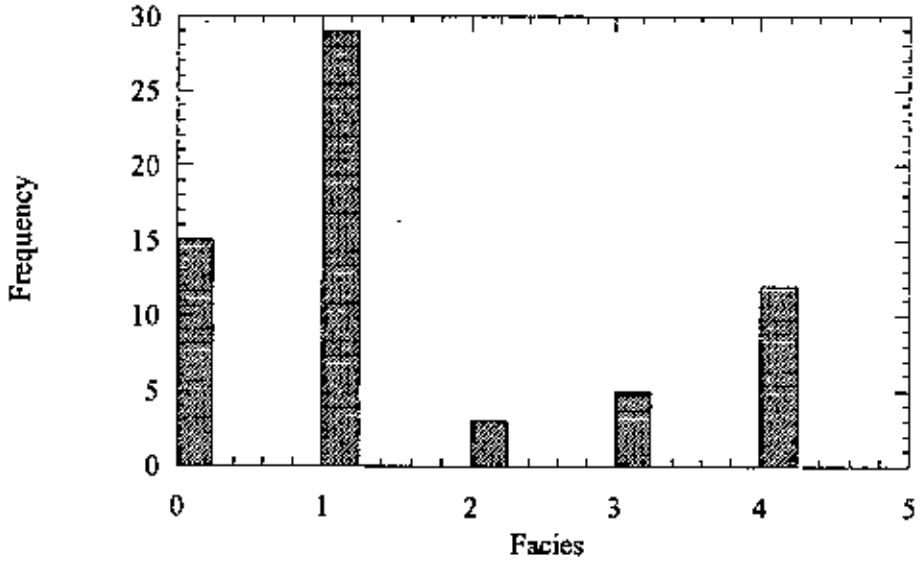

Figure 5.23: Facies histogram for Well 2.

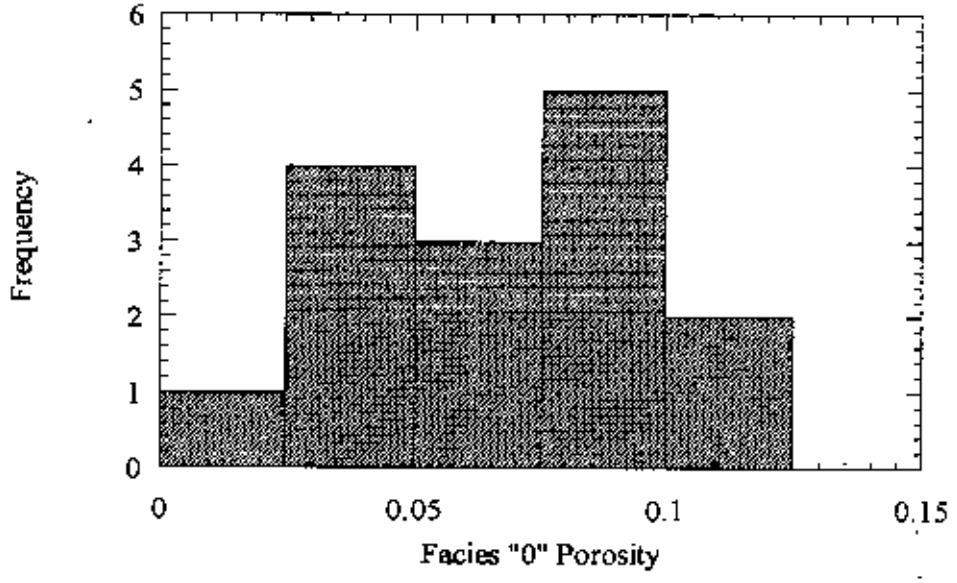

Figure 5.24: Facies 0 porosity histogram for Well 2.

Figure 5.25 shows porosity variation for Facies 1. Its mean is 0.087 (higher than Facies 0 porosity), and the range is from 0.029 to 0.167. Figure 5.26 shows the porosity histogram for Facies 2. The mean is found to be 0.068 The values range from 0.035 to 0.11. Figure 5.27 shows the porosity histogram for Facies 3. The mean is 0.068, and the range is between 0.035 and 0.11. Figure 5.28 shows porosity histogram for Facies 4. Notably, Facies 4 has the highest mean porosity of 0.095, with a range from 0.06 to 0.193.

Bivariate statistics of Well 2 are discussed here. Figures 5.29 is the vertical permeability versus maximum permeability scatter plot. Removing the outliers, correlation improves from 0.265 to 0.485. Maximum Permeability versus porosity scatter plot is shown in Figures 5.30 showing a correlation 0.172. Figure 5.31 shows vertical permeability versus porosity scatter plot. The correlation coefficient is 0.478 Figure 5.32 shows facies versus relative strata scatter plot. This plot shows a very good correlation (0.82) between facies and strata. At the top Facies 0 is found, while on the bottom Facies 4 is dominant.

5.2.3 Well 3 Statistics

Well 3 statistics are briefly described here. Figure 5.33 is the histogram of maximum permeability. Its mean and standard deviation are 22.111 and 740.463, while its range is from 0 04 to 227. Vertical permeability histogram is shown in Figure 5 34. Its mean and standard deviation are 2.195 and 1.731, and the values range from 0.01 to 32.2. Figure 5 35 shows histogram of porosity having mean and standard deviation of 0.061 and 0.023, respectively. The minimum and the maximum porosity values are 0.018 and 0.111. Figure 5.36 shows facies histogram. Facies 1 is again the dominant facies in this well alike Wells 1 and 2. Only Facies 1, 3, and 4 are present.

For facies porosity variation, the figures are not included here for convenience. However, the statistics are described. Facies 1, 3, and 4 porosity means are found to be 0.061, 0.078, and 0 04, respectively The corresponding porosity ranges for the three facies are (0.033, 0.096), (0.042, 0.111), and (0.018, 0.096), respectively.

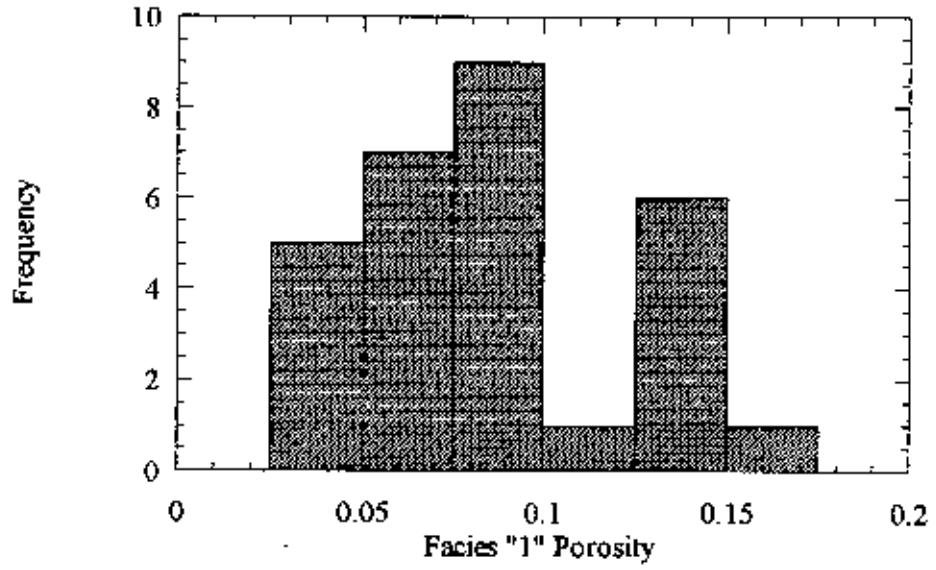

Figure 5.25: Facies 1 porosity histogram for Well 2.

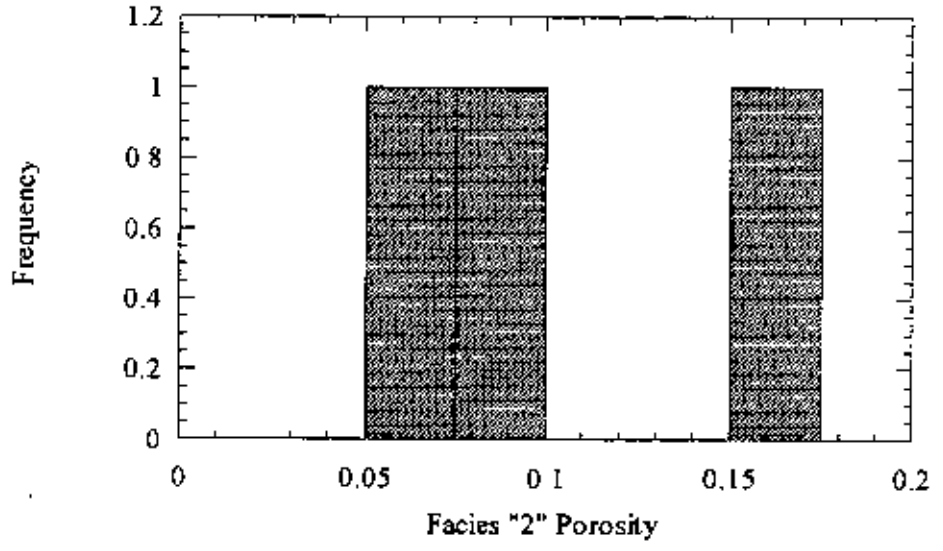

Figure 5.26: Facies 2 porosity histogram for Well 2.

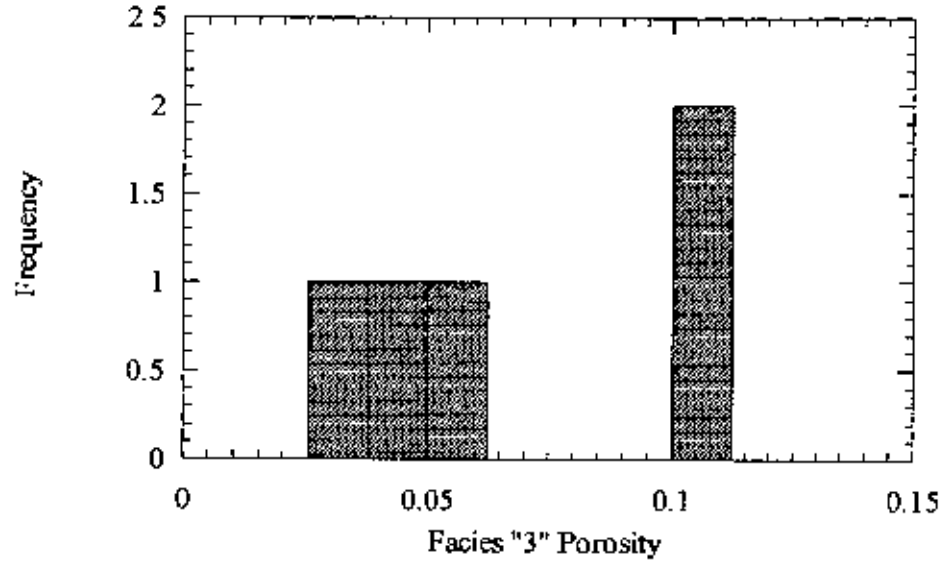

Figure 5.27: Facies 3 porosity histogram for Well 2.

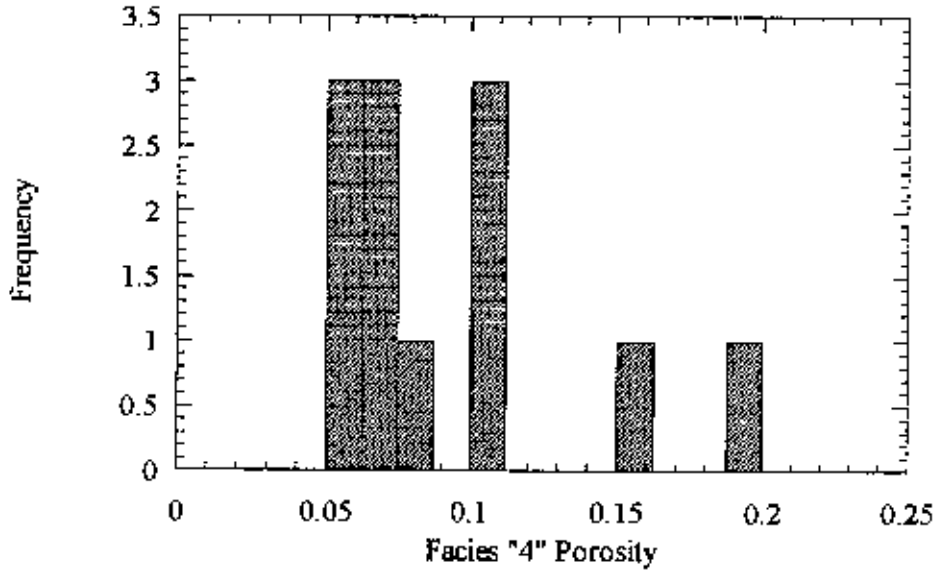

Figure 5.28: Facies 4 porosity histogram for Well 2.

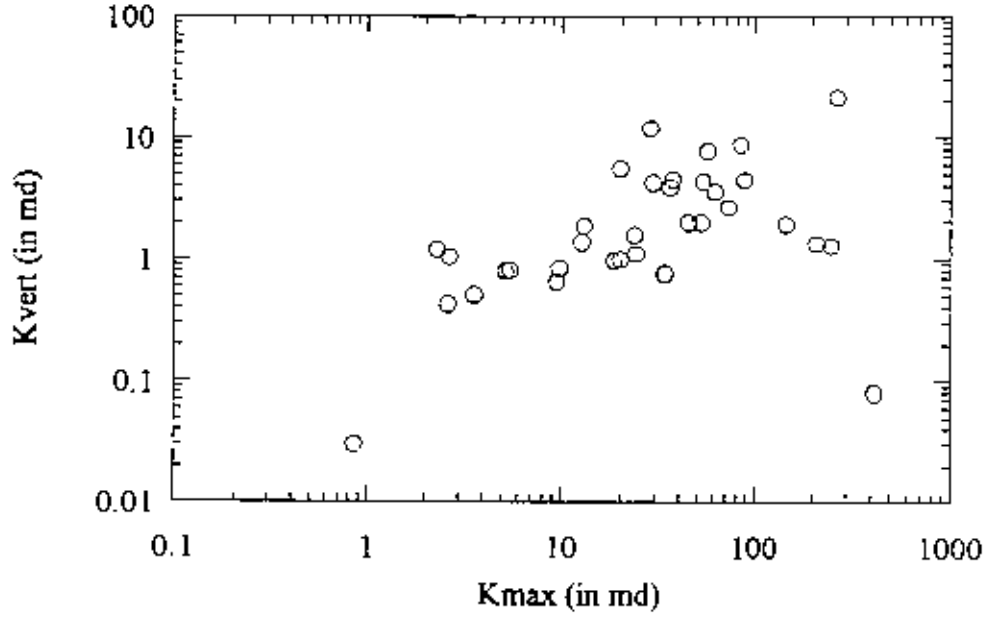

Figure 5.29: Vertical permeability versus maximum permeability scatter plot for Well 2.

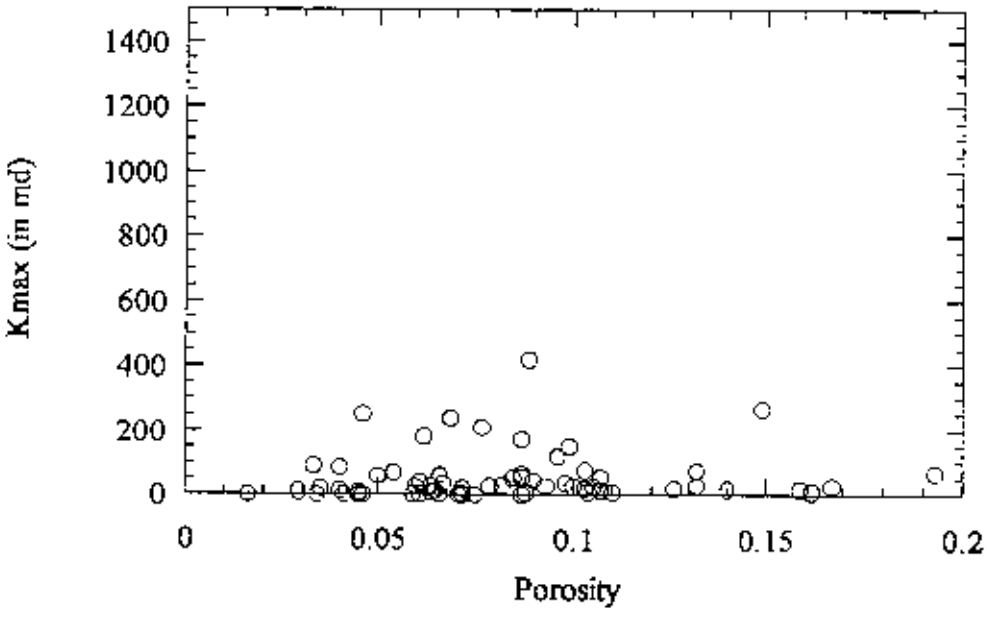

Figure 5.30: Maximum permeability versus porosity scatter plot for Well 2.

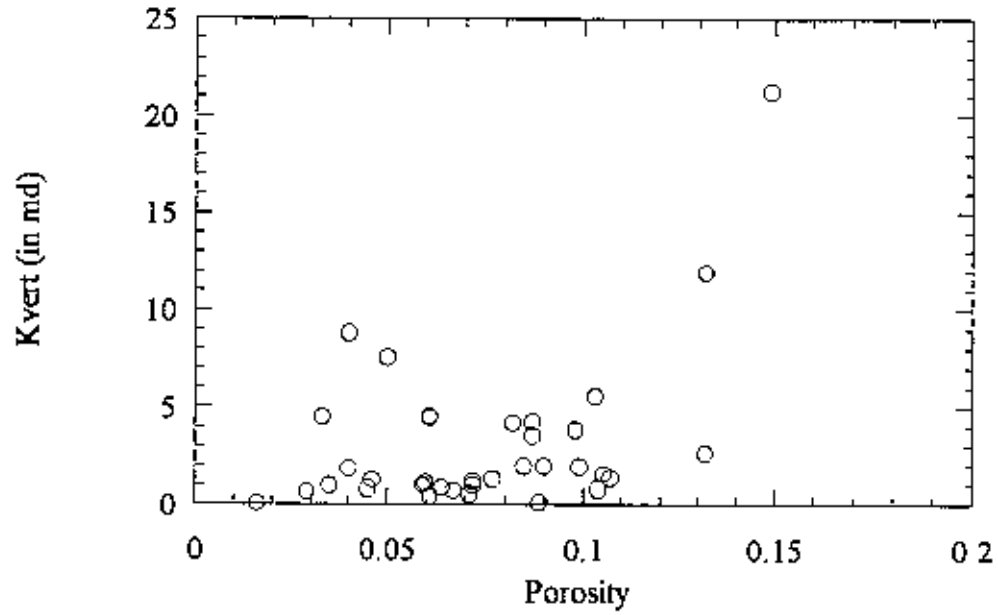

Figure 5.31: Vertical permeability versus porosity scatter plot for Well 2.

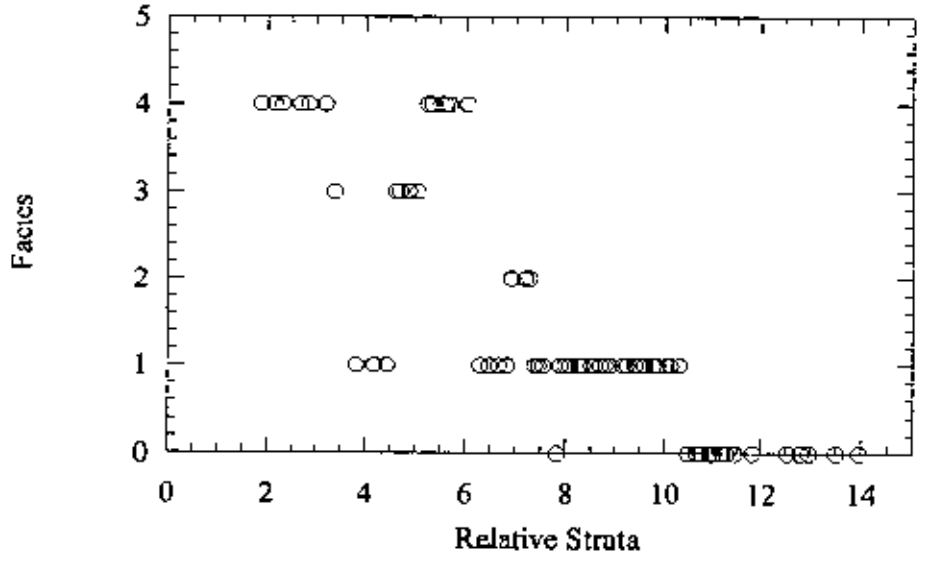

Figure 5.32: Facies versus relative strata scatter plot for Well 2.

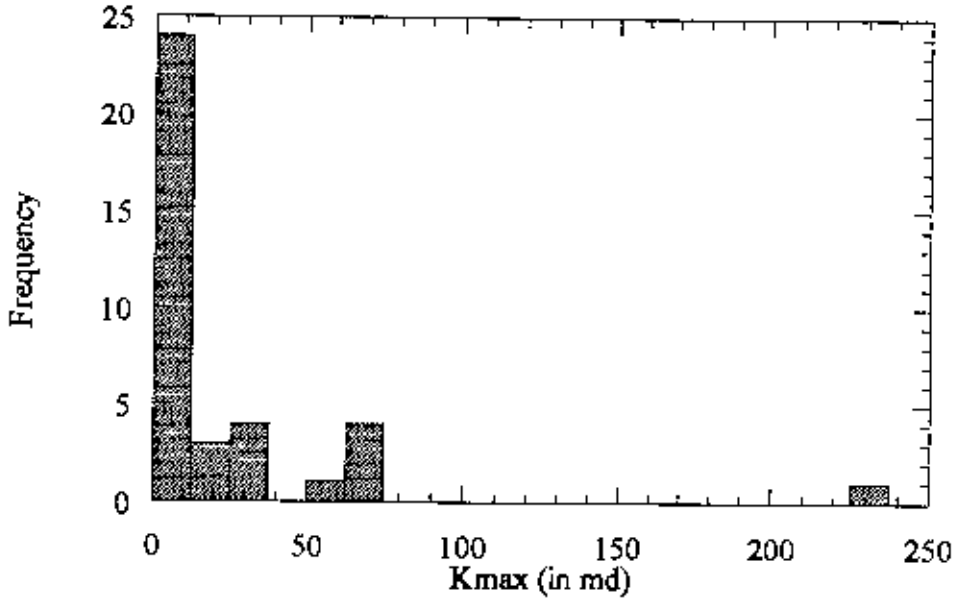

Figure 5.33: Maximum permeability histogram for Well 3.

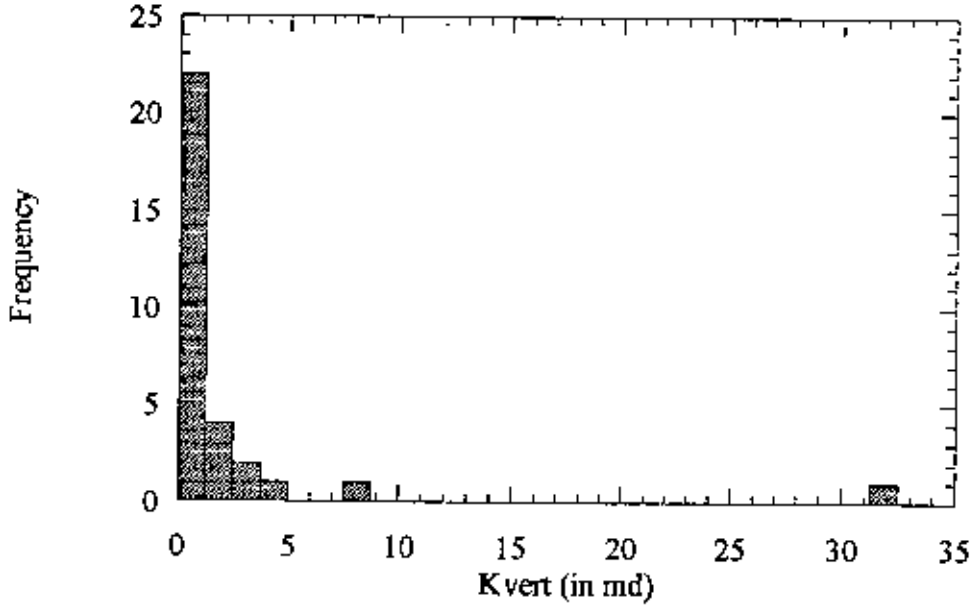

Figure 5.34: Vertical permeability histogram for Well 3.

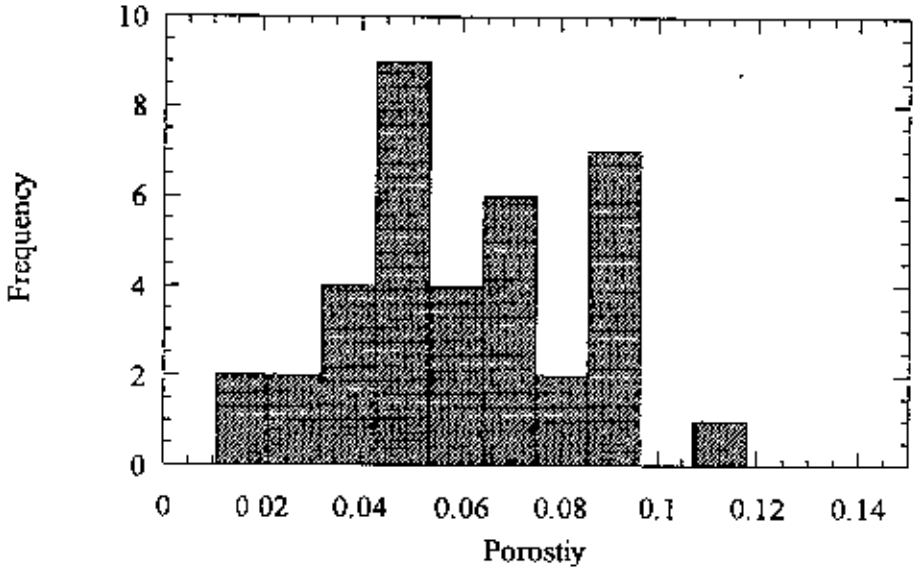

Figure 5.35: Porosity histogram for Well 3.

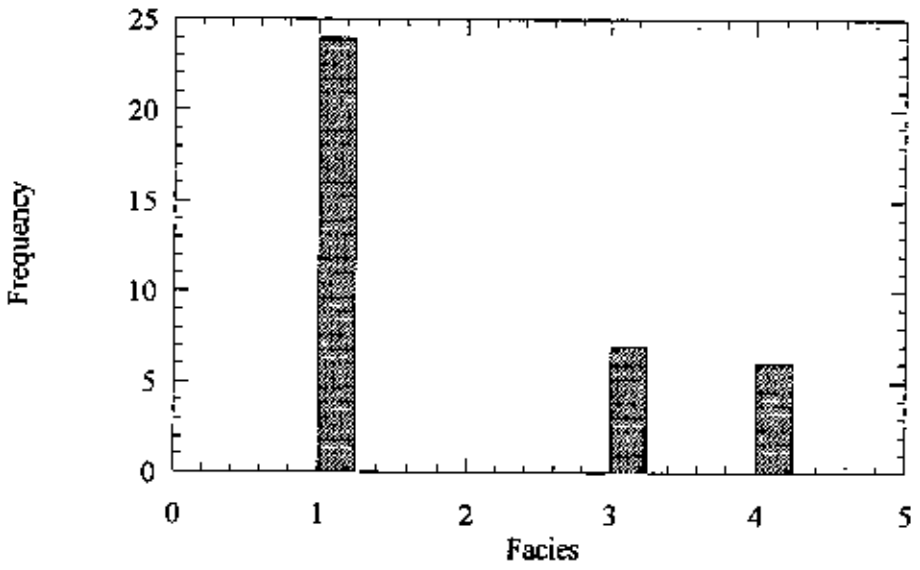

Figure 5.36: Facies histogram for Well 3.

Well 3 bivariate statistics are discussed here. Figure 5.37 is the vertical permeability versus maximum permeability scatter plot. Removing the outliers, correlation improves significantly from 0 029 to 0.83. Figure 5.38 shows maximum permeability versus porosity scatter plot. Removing outliers, correlation improves from 0.014 to 0.34. Figure 5.39 shows vertical permeability versus porosity scatter plot Correlation improves from 0.005 to 0.341 after removing the outlier. Figure 5.40 shows facies versus relative strata scatter plot. This plot shows correlation 0.871 between facies and relative strata.

5.2.4 Well 4 Statistics

Well 4 statistics are described below. Figures 5.41, 5.42, and 5.43 are the histograms of maximum permeability, vertical permeability, and porosity. The corresponding means are found to be 5.926, 0.947, and 0.065; while the standard deviations are 7.292, 1.765, and 0.02, respectively. And the corresponding ranges are (0.04, 25.9), (0.07, 8.95), and (0.019, 0.104), respectively. Figure 5.44 shows facies histogram. Only Facies 1 and 2 are present, Facies 1 being the major one. For porosity variation with facies, Facies 1, 2 porosity means are found to be 0.74 and 0.042. The corresponding porosity ranges are (0.05, 0.104) and (0.019, 0.061)

As for bivariate statistics, Figure 5.45 shows vertical permeability versus maximum permeability scatter plot. Having outliers removed, correlation improves from 0.032 to 0.498. Maximum permeability versus porosity scatter plot is given in Figure 5.46. With and without outliers, the correlation coefficients are 0.345 and 0.411. Figure 5.47 is the vertical permeability versus porosity scatter plot. In this case, the correlation coefficients with and without outliers are 0.158 and 0.445. Figure 5.48 shows facies versus relative strata scatter plot. This plot also shows very good correlation (0.837) between facies and relative strata.

5.2.5 Well 5 Statistics

It should be noted that Well 5 does not have permeability measurements. Figure 5.49 gives the porosity histogram. Porosity value ranges from 0.025 to 0.067 with a mean and standard deviation of 0.044 and 0.014. Figure 5.50 is the facies histogram. Only Facies 0

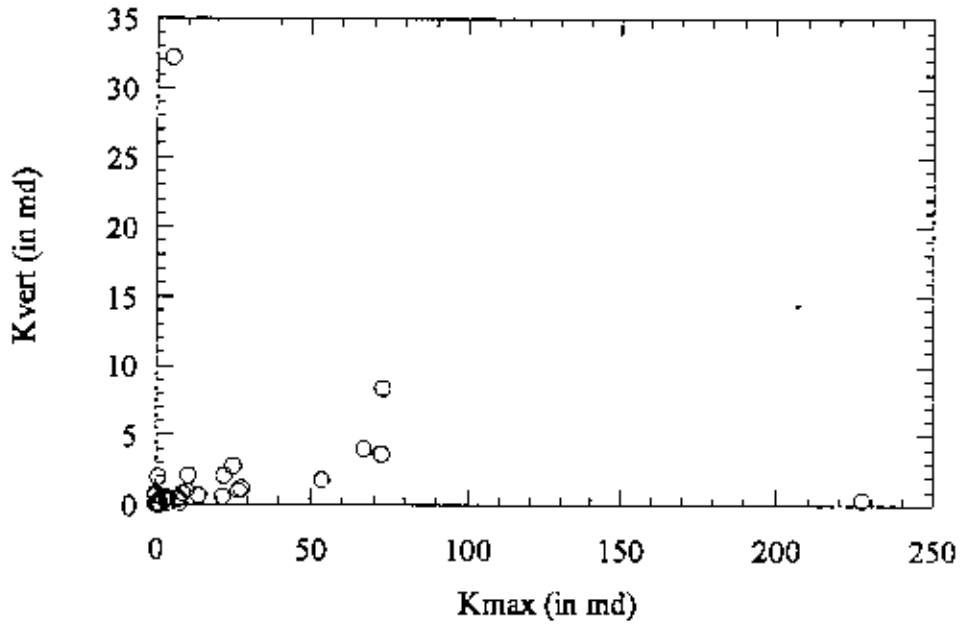

Figure 5.37: Vertical permeability versus maximum permeability for Well 3.

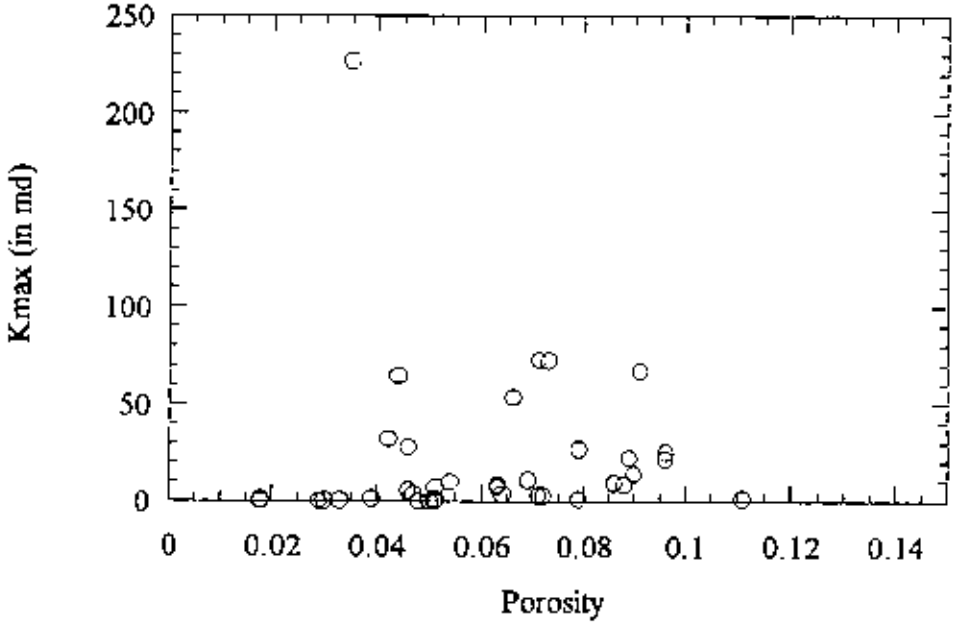

Figure 5.38: Maximum permeability versus porosity scatter plot for Well 3.

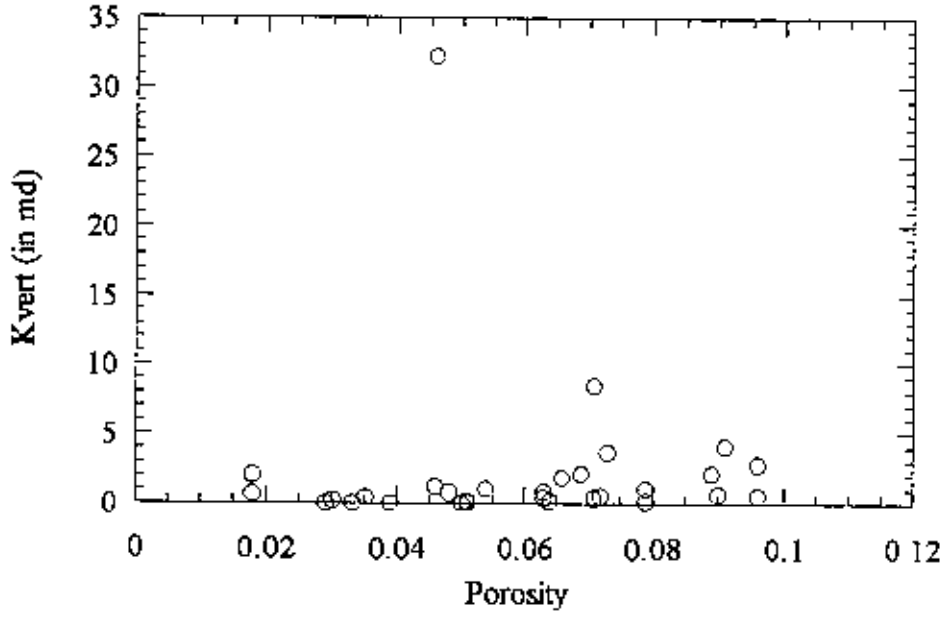

Figure 5.39: Vertical permeability versus porosity scatter plot for Well 3.

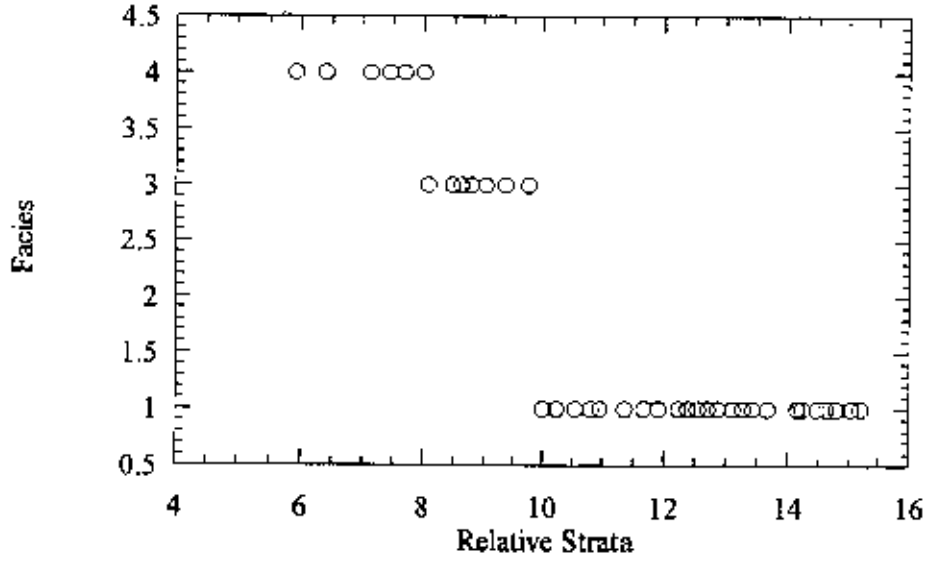

Figure 5.40: Facies versus relative strata scatter plot for Well 3.

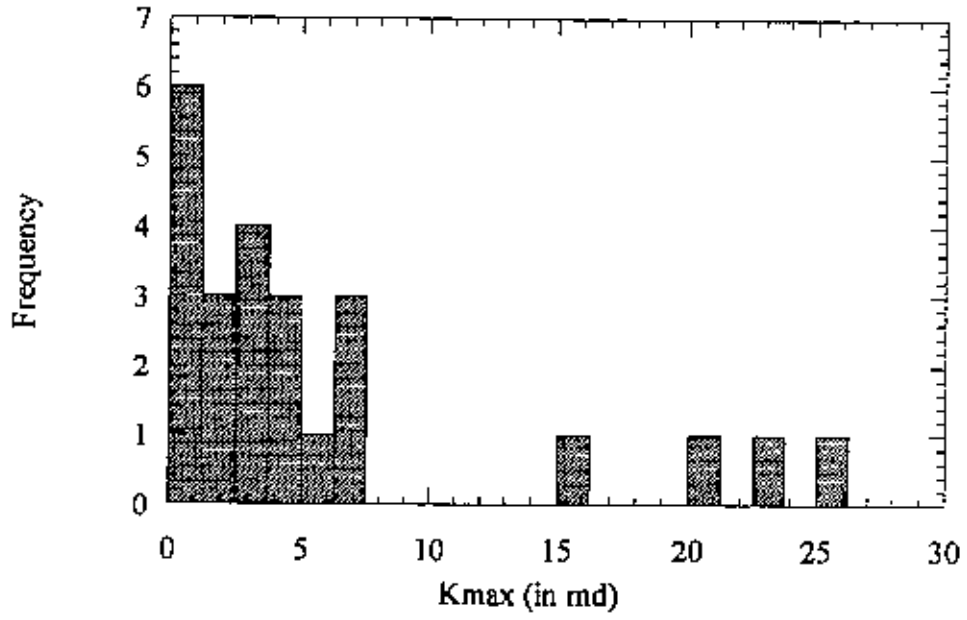

Figure 5.41: Maximum permeability histogram for Well 4.

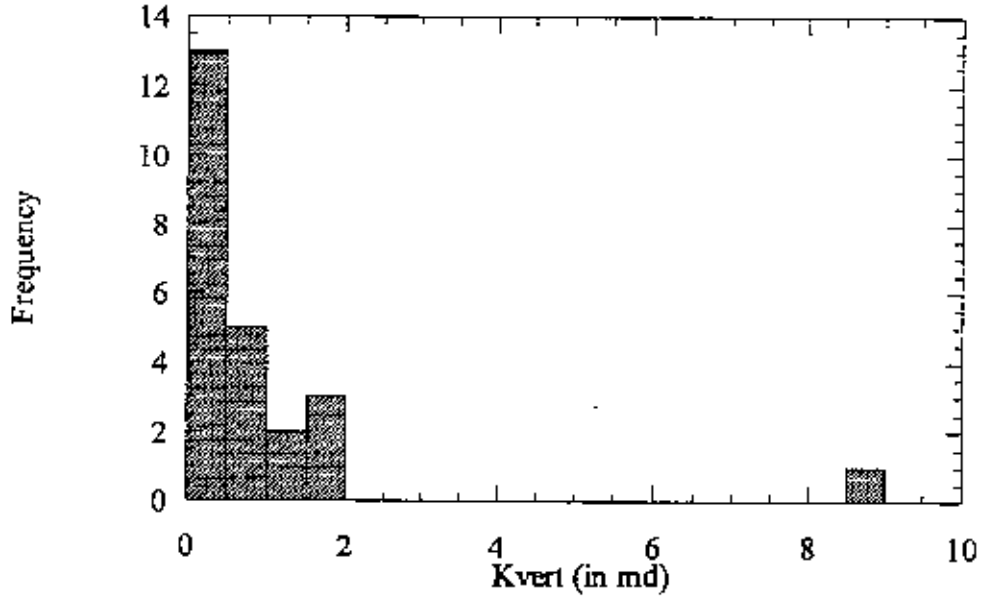

Figure 5.42: Vertical permeability histogram for Well 4.

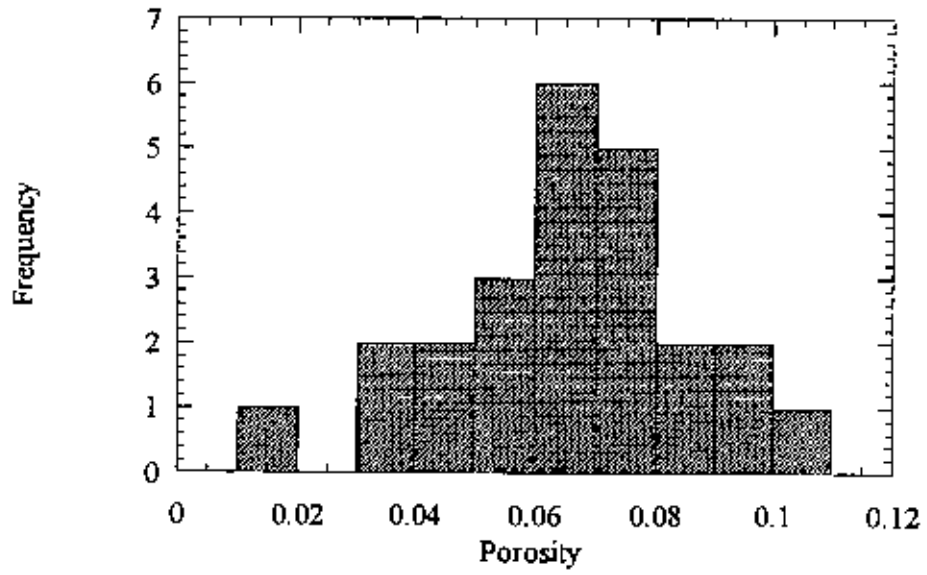

Figure 5.43: Porosity histogram for Well 4.

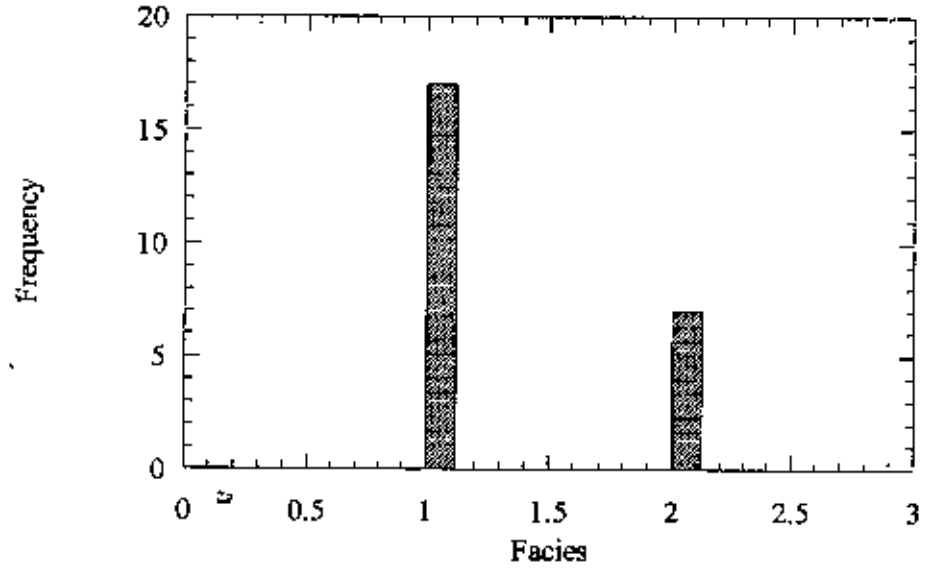

Figure 5.44: Facies histogram for Well 4.

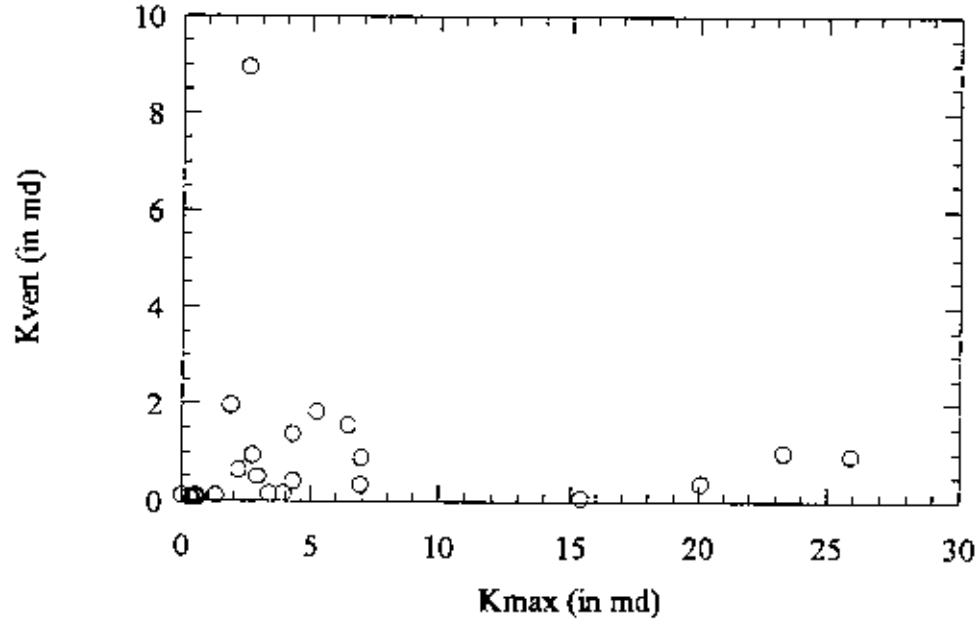

Figure 5.45: Vertical permeability versus maximum permeability scatter plot for Well 4.

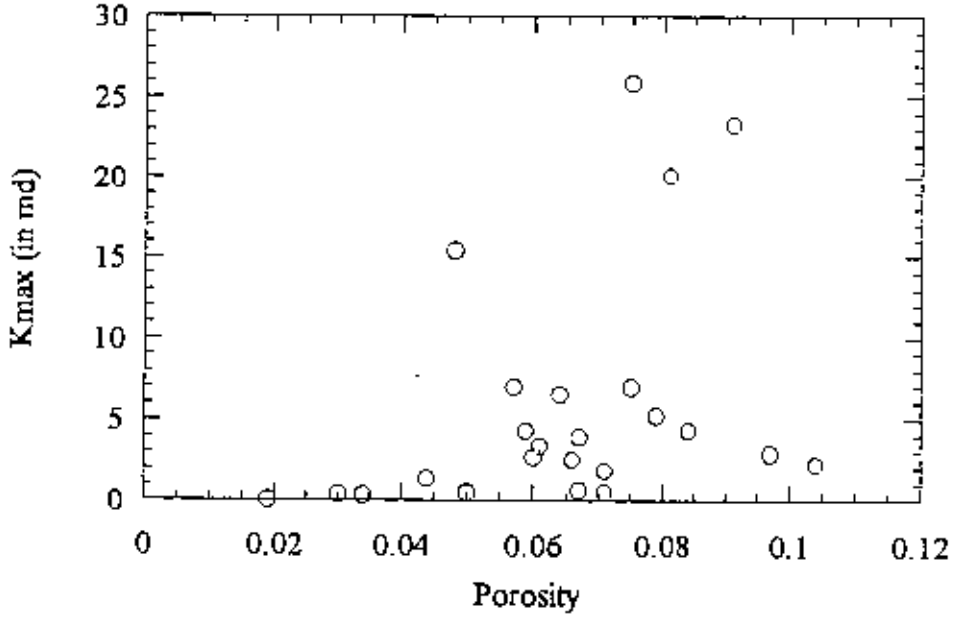

Figure 5.46: Maximum permeability versus porosity scatter plot for Well 4.

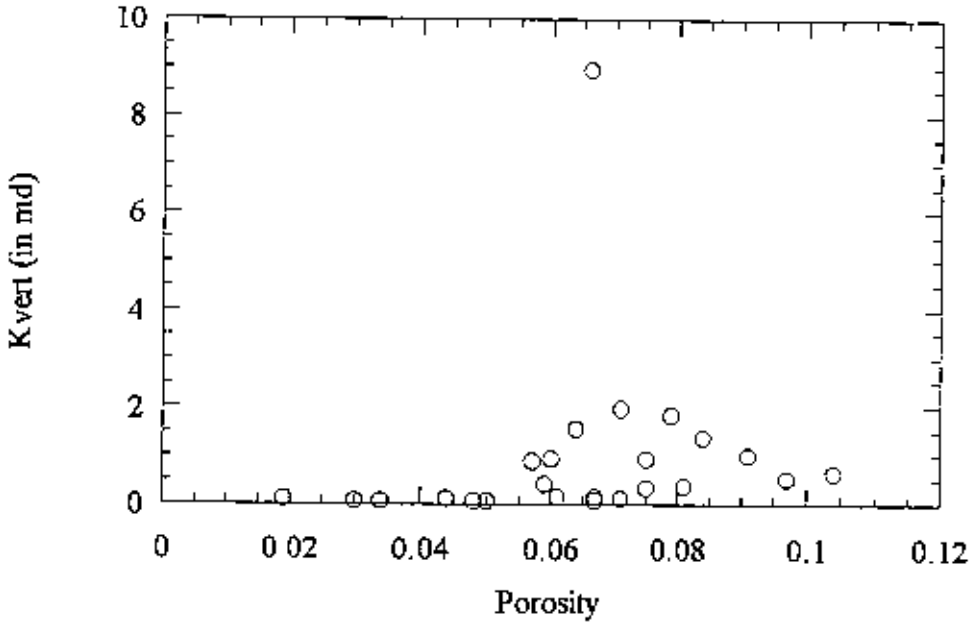

Figure 5.47: Vertical permeability versus porosity scatter plot for Well 4

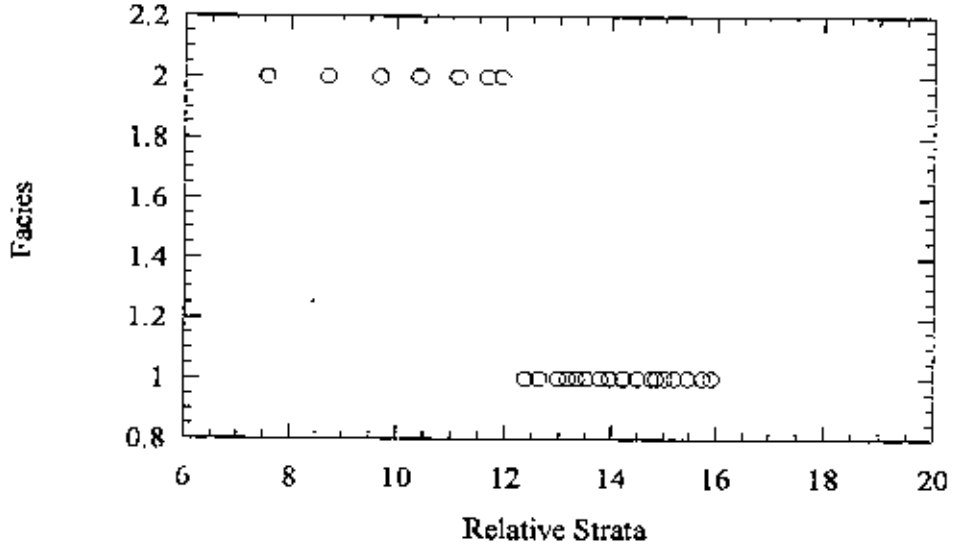

Figure 5.48: Facies versus relative strata scatter plot for Well 4.

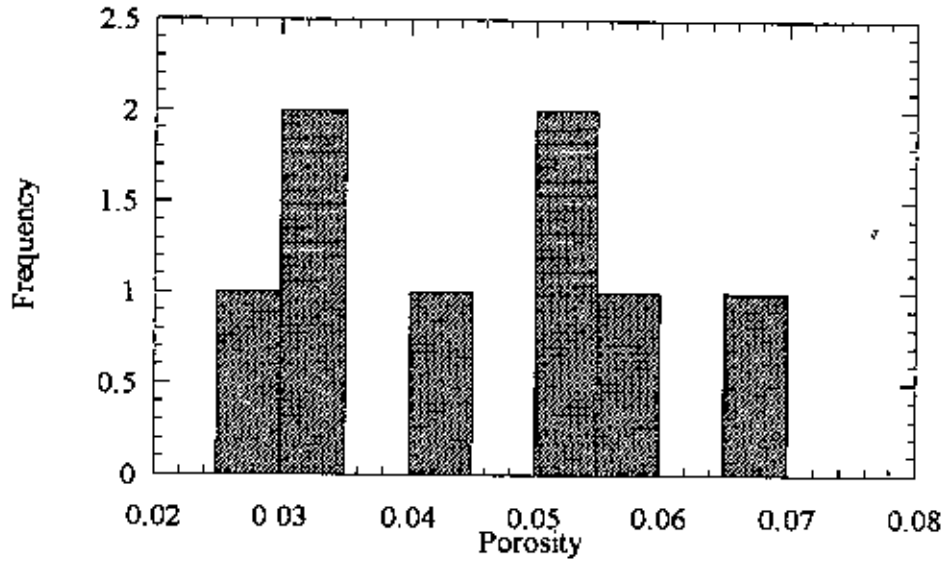

Figure 5.49: Porosity Histogram for Well 5.

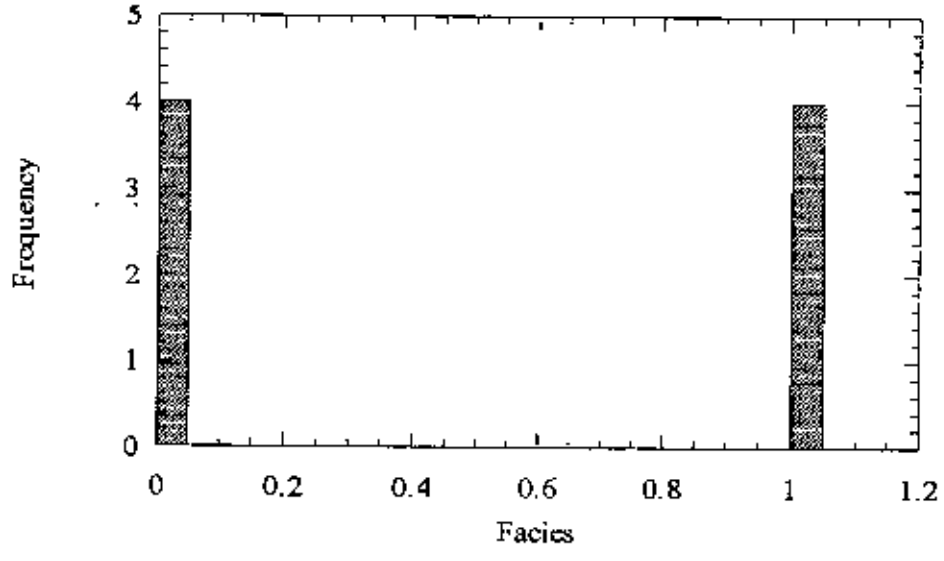

Figure 5.50: Facies histogram for Well 5.

and 1 are present. Interestingly, both facies proportions are equal. Facies 0 porosity values slightly higher mean of 0.049 than Facies 1 mean of 0.04. However, the corresponding porosity ranges are (0.025, 0 067) and (0.031, 0.056), respectively. Figure 5 51 is the facies versus relative strata scatter plot. Correlation is found 0.08.

5.2.6 Global Statistics

Only well statistics were considered till this point. The information from these statistics is localized. To retrieve the global information about the reservoir, combined statistics need to be considered. Figure 5.52 is the maximum permeability histogram. Maximum permeability value ranges from 0.04 to 1320 with a global mean and standard deviation 50.861 and 124.137. Vertical permeability histogram is shown in Figure 5.53. The global mean and standard deviation are 2.95 and 4.49. The values range from 0.01 to 32.2. Figure 5.54 shows porosity histogram It has global mean and standard deviation 0 075 and 0.035. Facies histogram is given in Figure 5.55. Not surprisingly, Facies 1 is the dominant one.

For porosity variation with facies, porosity histograms for Facies 0, 1, 2, 3, 4, and 5 are shown in Figures 5.56, 5.57, 5.58, 5.59, 5.60, and 5.61 respectively. The corresponding means are 0.059, 0.081, 0.08, 0.82, 0.064, and 0.059; while the corresponding ranges are (0 016, 0.107), (0.029, 0.167), (0.019, 0.162), (0.035, 0 155), (0.018, 0.193), and (0.017, 0.096), respectively. Figure 5 62 shows facies versus relative strata scatter plot. Correlation coefficient between the parameters is 0.671.

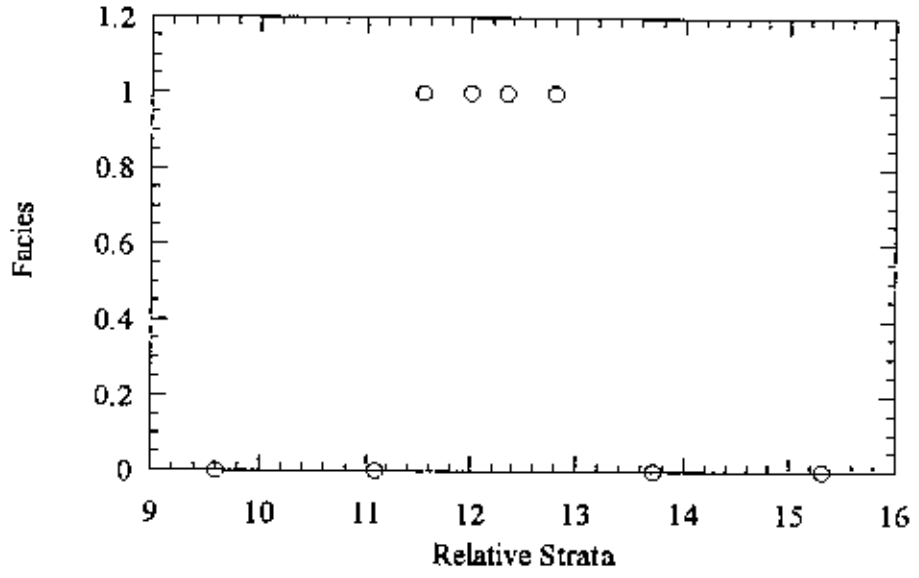

Figure 5.51: Facies versus relative strata scatter plot for Well 5.

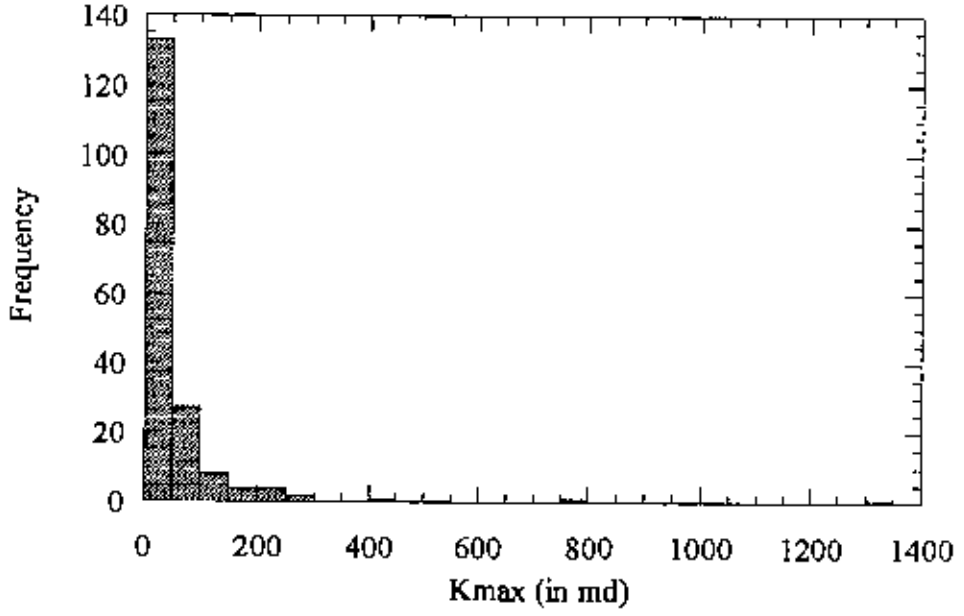

Figure 5.52: Global maximum permeability histogram.

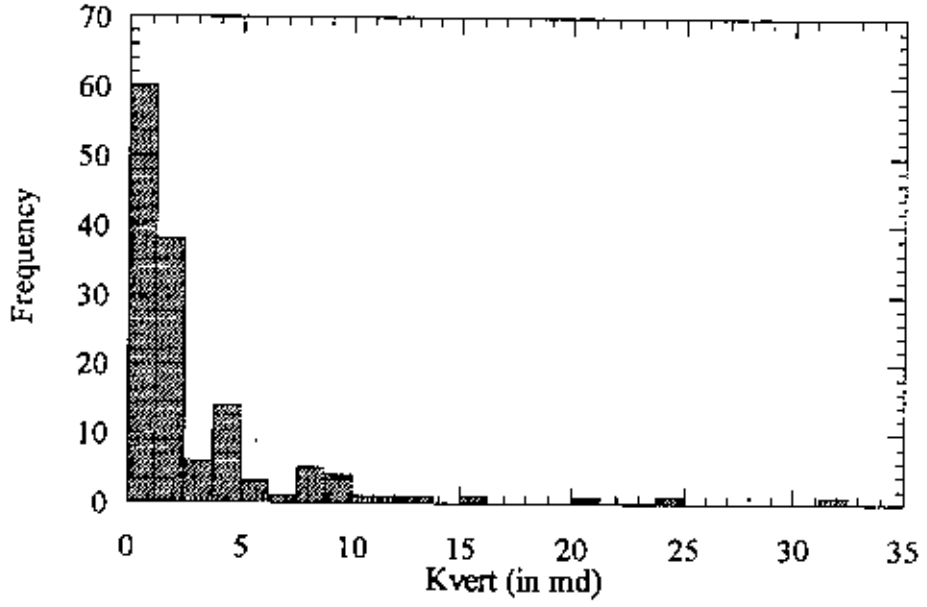

Figure 5.53: Global vertical permeability histogram.

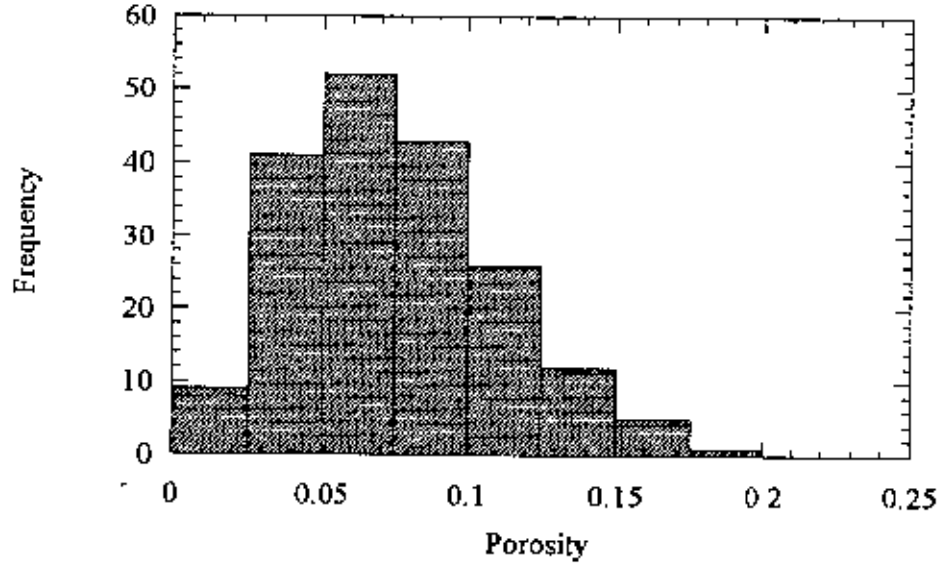

Figure 5.54: Global porosity histogram.

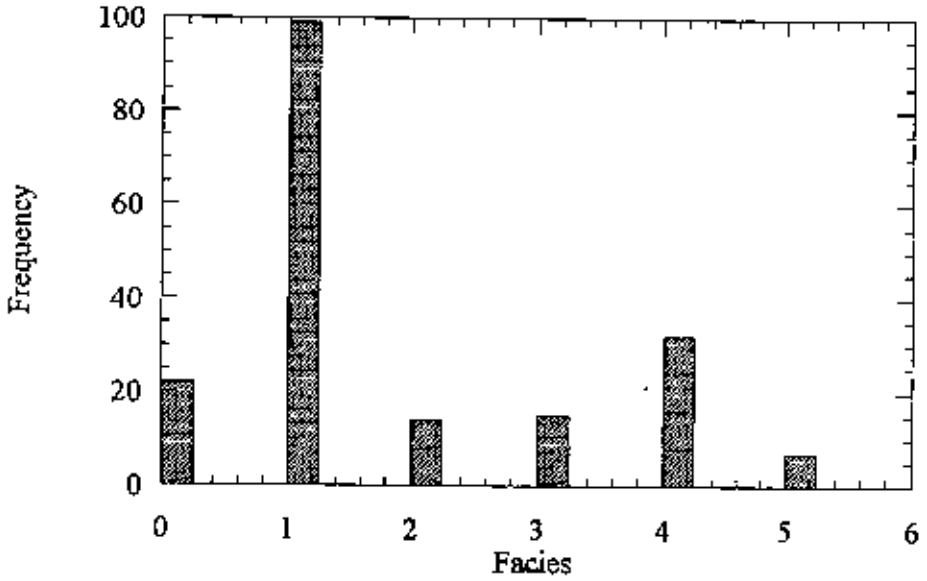

Figure 5.55: Global facies histogram.

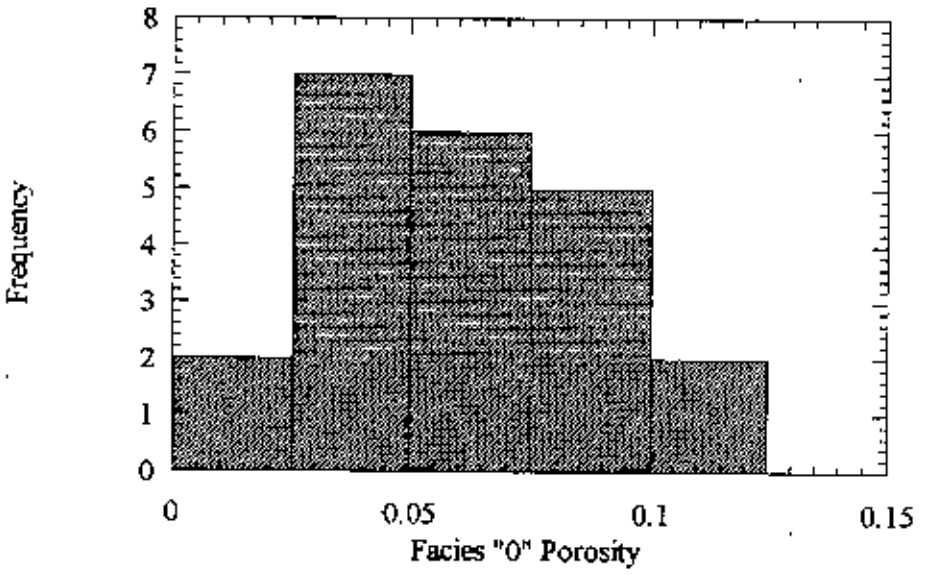

Figure 5.56: Facies 0 global porosity histogram.

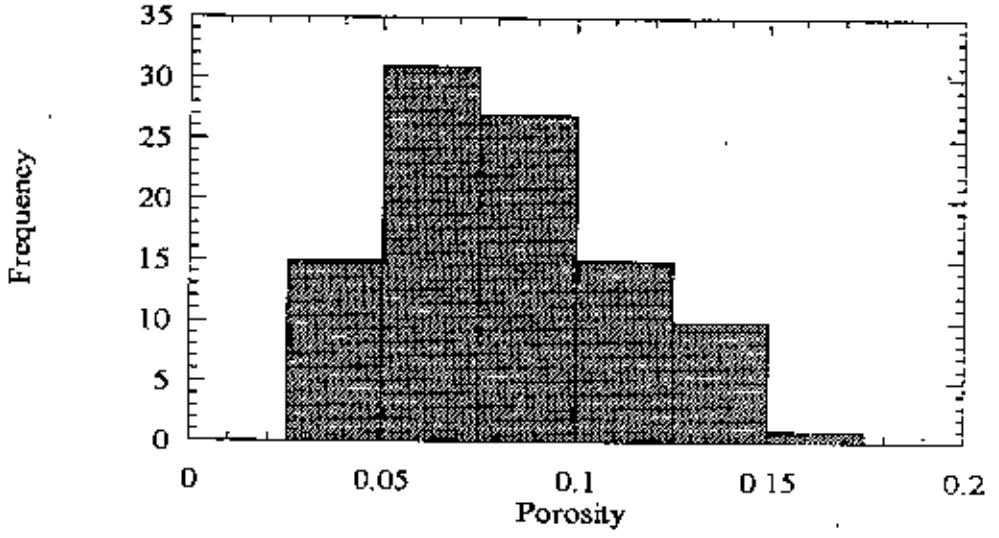

Figure 5.57: Facies 1 global porosity histogram.

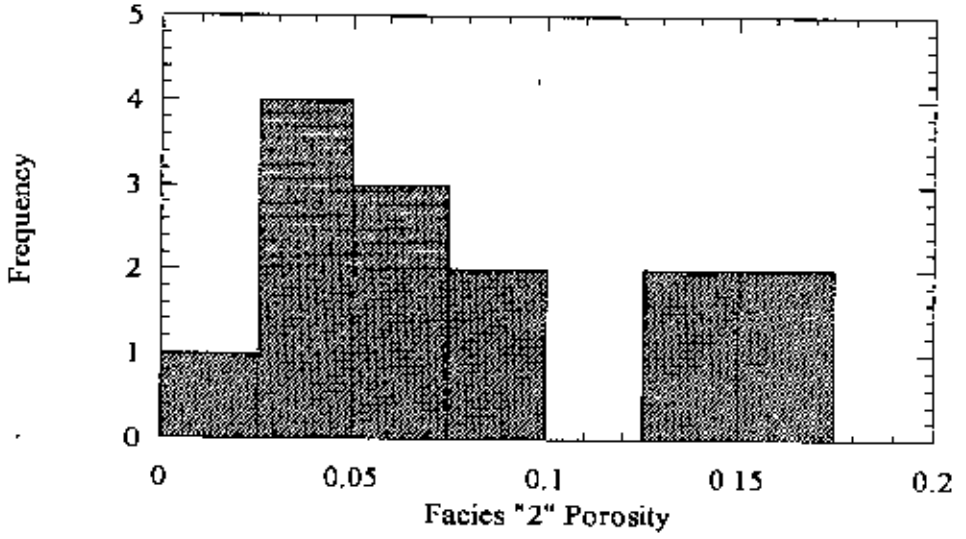

Figure 5.58: Facies 2 global porosity histogram.

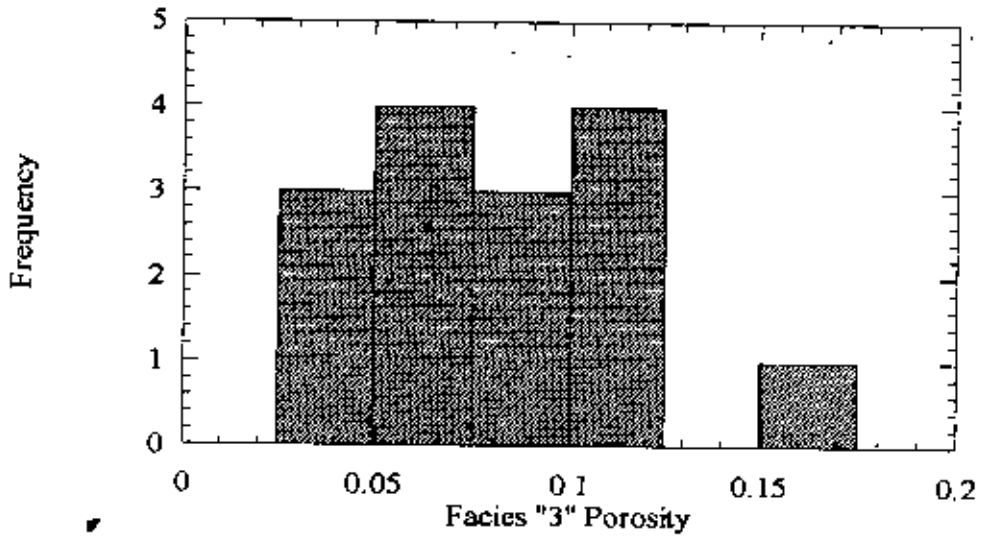

Figure 5.59: Facies 3 global porosity histogram.

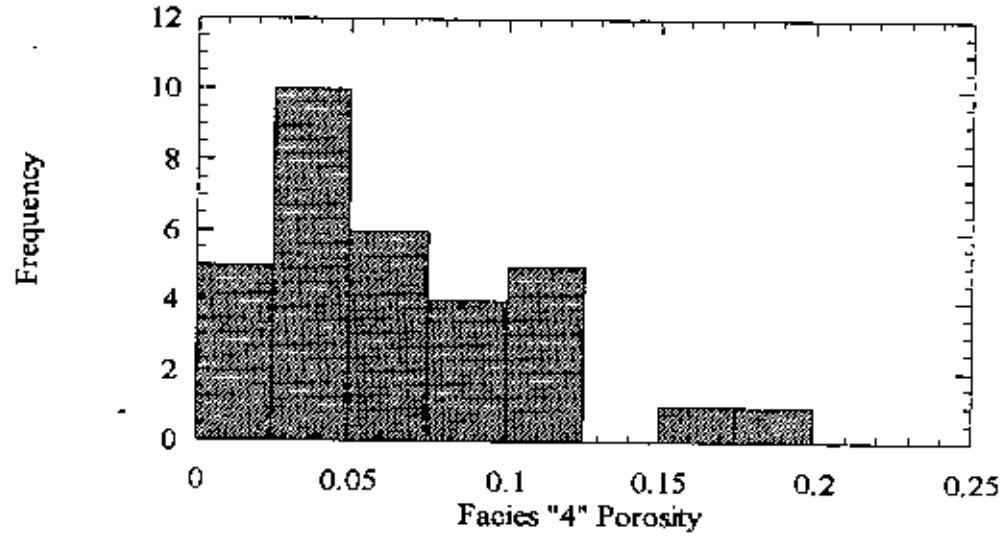

Figure 5.60: Facies 4 global porosity histogram.

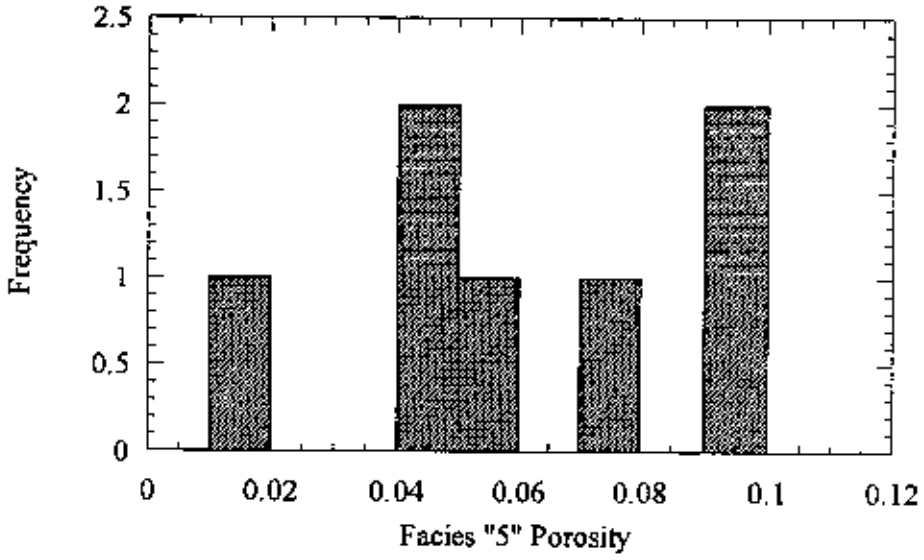

Figure 5.61: Facies 5 global porosity histogram.

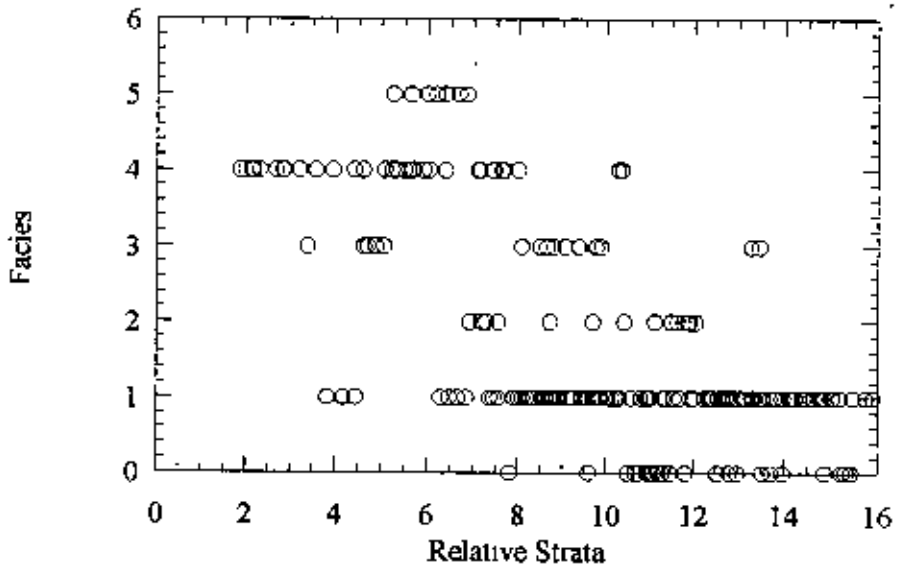

Figure 5.62: Global facies versus relative strata scatter plot.

5.3 Conclusion

Exploratory data analysis is performed on the available core data. Both localized and global statistics are retrieved. Some of the conclusions from the analysis are given below.

- Available data are well locations, depth, porosity, facies, maximum permeability, and vertical permeability. Most of these variables are derived properties except for the locations and depth.
- Local and global statistics can be significantly different at places.
- Correlations between the variables are quite poor. The raw petrophysical and primary response signals of the wells are not available. Thus, it almost prohibits the use of models based on first principles or some simple regressive models to predict permeability. This warrants the use of highly non-linear, effective tools like neural network. This rationalizes the scope of the work of this thesis.
- Notwithstanding the use of neural networks, the poor correlations and the absence of primary signals substantiate that there will be significant prediction error Some post-processing of the estimates could be considered.

CHAPTER 6

PERMEABILITY PREDICTION USING NEURAL NETWORK

This chapter discusses permeability prediction using the neural network model. Previously, exploratory data analysis was performed (discussed in Chapter 5). It revealed some directions on how to set up the framework of the neural network model. It was identified that spatial data (Easting, Northing, relative stratigraphy), porosity, facies and thickness could be used as the argument to the model. It should be pointed out that permeability depends on various other factors like grain size and distribution, angularity, connectivity, geomechanical properties, diagenesis effect and so forth. Absence of the information of these parameters is a handicap and it will influence the quality of the prediction using any model. Notwithstanding the fact, we devise the neural network model for the prediction.

The outline of the chapter is as the following. Section 6.1 gives a brief discussion on the architecture and training of the neural network. Sections 6.2, 6.3, and 6.4 deal with optimal parameter selections, that is, the selection of bias, learning rates, and number of hidden nodes, respectively. Section 6.6 discusses the learning profile. In section 6.7 it has been tried to find out the major input parameter for output parameter, permeability. Section 6.7 discusses about the randomness effect. Section 6.8 predicts the permeability of Well 5 while the Section 6.9 gives conclusion of the developed model.

6.1 Training of the Neural Network Model

A back propagation neural network model with one hidden layer is developed to predict permeability. The input parameters to the model are Easting (X), Northing (Y), relative stratigraphic coordinates (Z_{strata}), porosity, facies, and thickness. The output is the permeability. There is only one hidden layer present. All the input and output data are normalized according to affine scaling. In order to consider the entire population (reservoir), respective maximum and minimum values of different

parameters for normalization are taken beyond their maximum and minimum values found in the sample (well) data. Table 6.1 shows maximum and minimum values. Maximum permeability (K_{max}) is converted to log scale and then normalized. Thus, the model consists of six plus one (bias) input nodes, five plus one (bias) hidden nodes and one output node.

Table 6.1: Maximum and minimum of the input and output parameters.

Parameter	X	Y	Z_{grid}	Facies	Porosity	Thickness	Log Permeability
Maximum	739551	5710616	15.91	0	0.25	20	3.12
Minimum	732227	5705548	1.867	5	0.016	14	-1.398
Normalized Maximum	750000	5715000	16	0	0.25	25	3.176
Normalized Minimum	730000	5703000	0	5	0.001	5	-3

The optimum parameter selection is an important aspect for the effectiveness of the neural network. The parameters optimized are the bias, the learning rates, and the number of hidden node. Each optimal parameter is selected based on minimum value of objective function, Root Mean Square Error (RMSE), keeping all other parameters constant. A range of the parameter is considered and the RMSE values determined. In order to minimize the random effect due to in-built randomness, RMSE values are obtained by averaging the same from 3 runs. The reduction factor is also considered. The number of epoch is kept 10^5 in each case. The following sections discuss the optimum parameter selection.

6.2 Bias Selection

Input layer and hidden layer bias values are incorporated in the model. Because of the use of normalized variables, input and hidden layer bias values range from 0 to 1. In order to obtain the optimum values in the bias space, 25 pairs of bias values

are considered. The pairs considered are (1, 1), (0.75, 1), (0.5, 1), (0.25, 1), (0.05, 1), (1, 0.75), ..., and (0.05, 0.05), and the RMSE values are observed. The optimum pair of bias values selected (0.5, 0.05). The low bias value 0.05 (for hidden layer) indicates that the effect of bias is negligible to the output layer. This optimum pair of bias values is used subsequently.

6.3 Learning Rates Selection

Learning rates for both hidden-to-output-layer and input-to-hidden-layer back propagation of weight correction should be optimally selected. For convenience, hidden-to-output-layer learning rate will be invoked by BLR, while input-to-hidden-layer learning rate by ALR.

6.3.1 Hidden-to-Output-Layer Learning Rate Selection

The optimum hidden-to-output-layer learning rate was determined through a sensitivity study. Keeping the number of hidden layer node and the number of epochs constant, the input-to-hidden-layer learning rate is fixed at some values and the hidden-to-output-layer learning rate was varied from 0.0001 to 1. The number of hidden layer node is fixed at 5, while the number of epoch considered is 10^5 for the sensitivity study. A number of plots of RMSE valnes are generated. The reduction factors of the RMSE are also plotted. Figures 6.1 and 6.2 show variation of RMSE and reduction factor with BLR variation for ALR = 0.01. Figure 6.1 shows very little change of RMSE values with respect to BLR. It is evident from the figures that the optimum range of BLR is between 0.0002 and 0.002. Figures 6 3 and 6.4 show similar plots for ALR = 0.02. Figure 6.3 shows greater RMSE variation than that evident in Figure 6 1. The optimum range of BLR is found to be from 0.0002 to 0.005. The RMSE variations for BLR for various values of ALR are shown in Figures 6.5, 6.7, 6.9, 6.11 and 6.13. The corresponding ALR values in these figures are 0.05, 0.1, 0.2, 0.5, and 1, respectively. The corresponding plots showing reduction factor of RMSE with BLR are given in Figures 6.6, 6.8, 6.10, 6.12 and 6.14. Examining the respective figures, the corresponding optimum ranges of BLR

are found (0.0005, 0.01), (0.0005, 0.005), (0.0001, 0.02), (0.0002, 0.02), and (0.0002, 0.005), respectively.

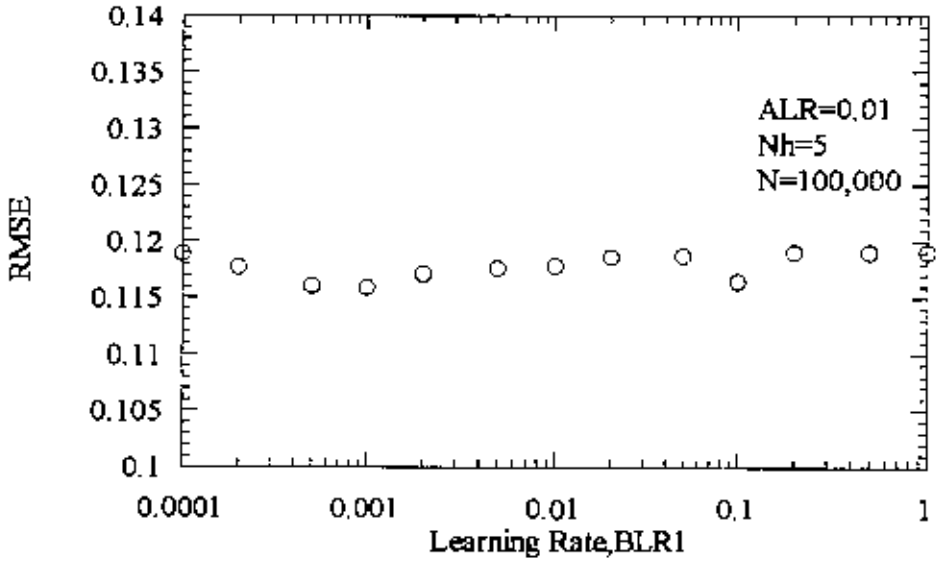

Figure 6.1: RMSE versus BLR for ALR=0.01.

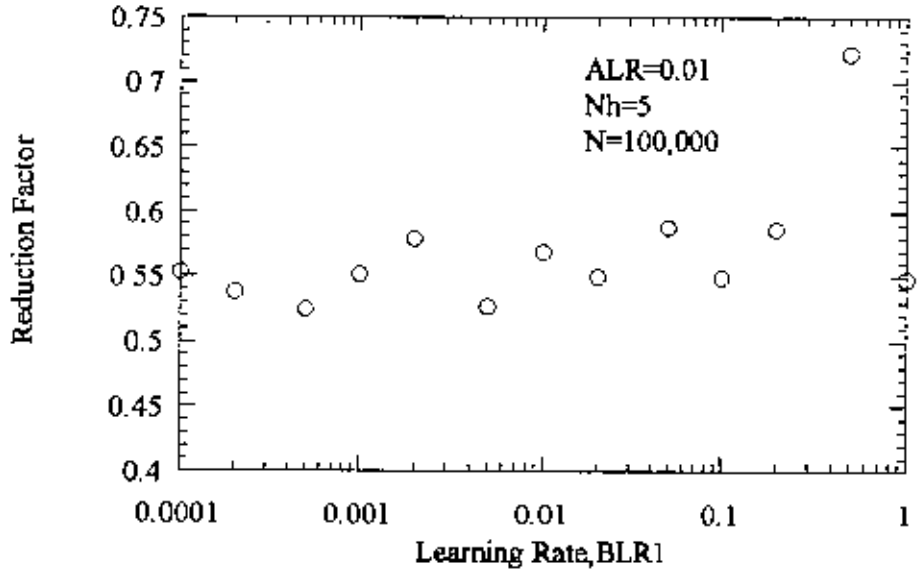

Figure 6.2: Reduction factor of RMSE versus BLR for ALR=0.01.

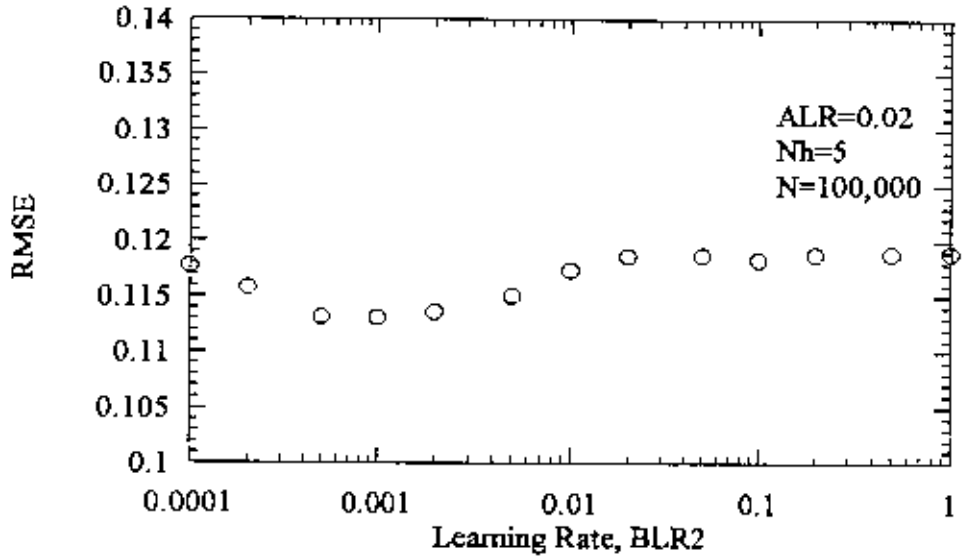

Figure 6.3: RMSE versus BLR for ALR=0.02.

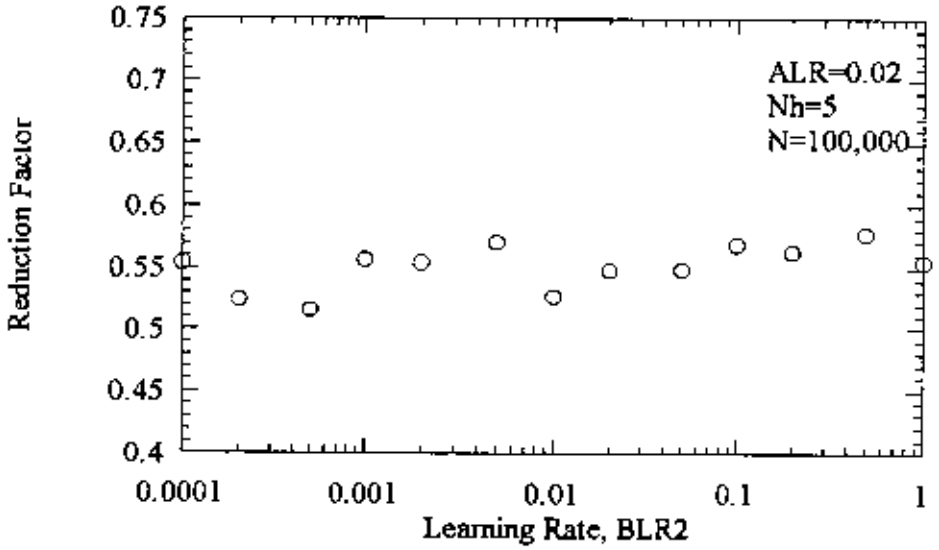

Figure 6.4: Reduction factor of RMSE versus BLR for ALR=0.02.

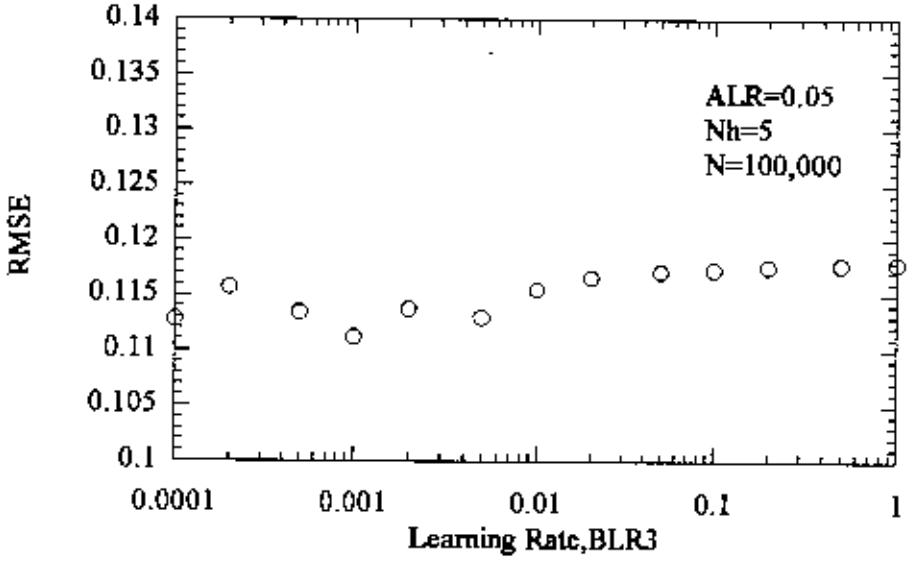

Figure 6.5: RMSE versus BLR for ALR=0.05.

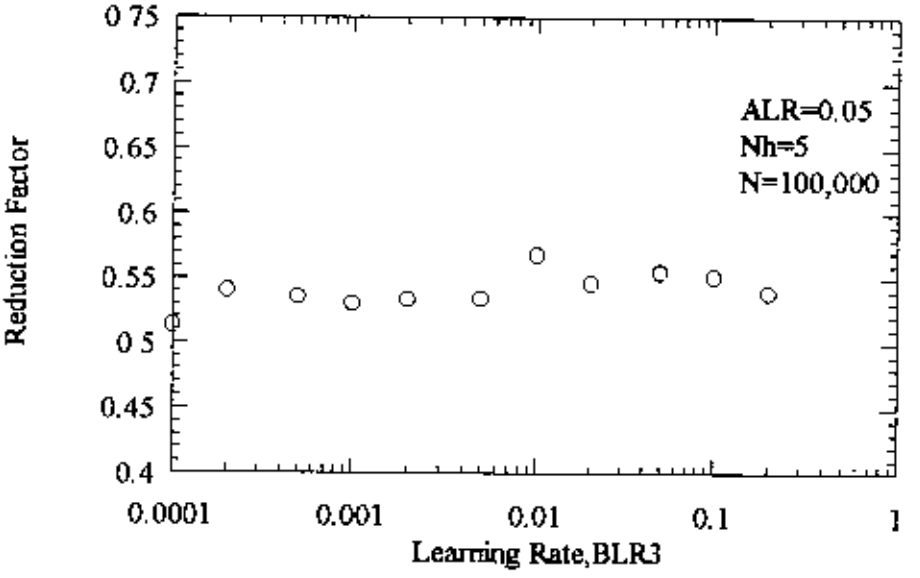

Figure 6.6: Reduction factor of RMSE versus BLR for ALR=0.05.

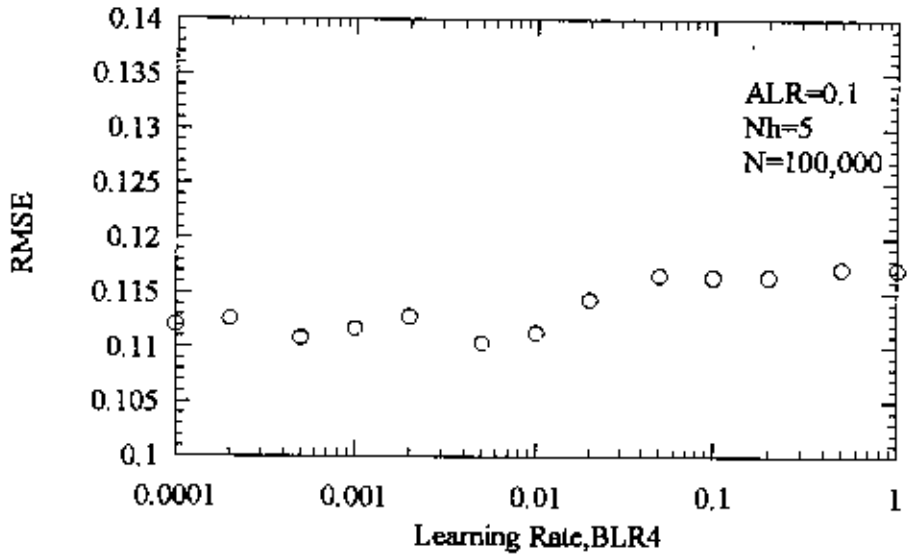

Figure 6.7: RMSE versus BLR for ALR=0.1.

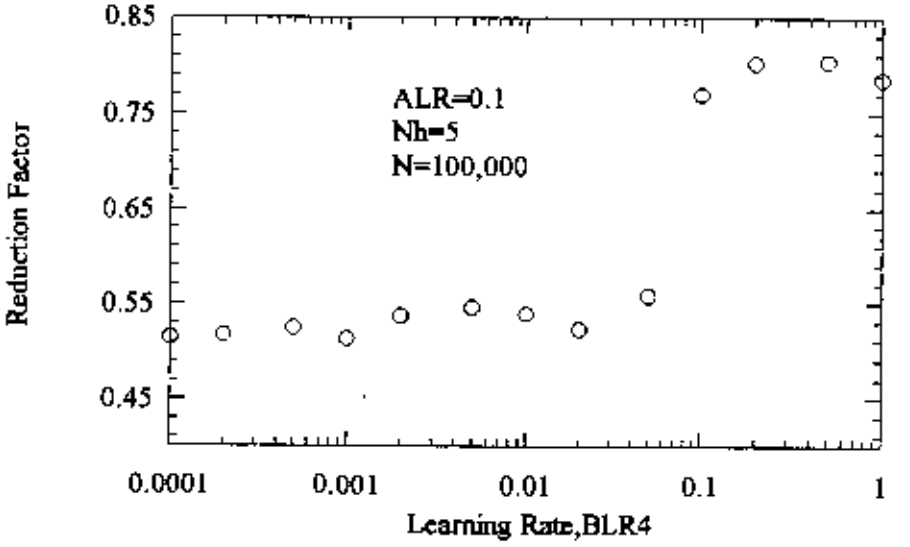

Figure 6.8: Reduction factor of RMSE versus BLR for ALR=0.1.

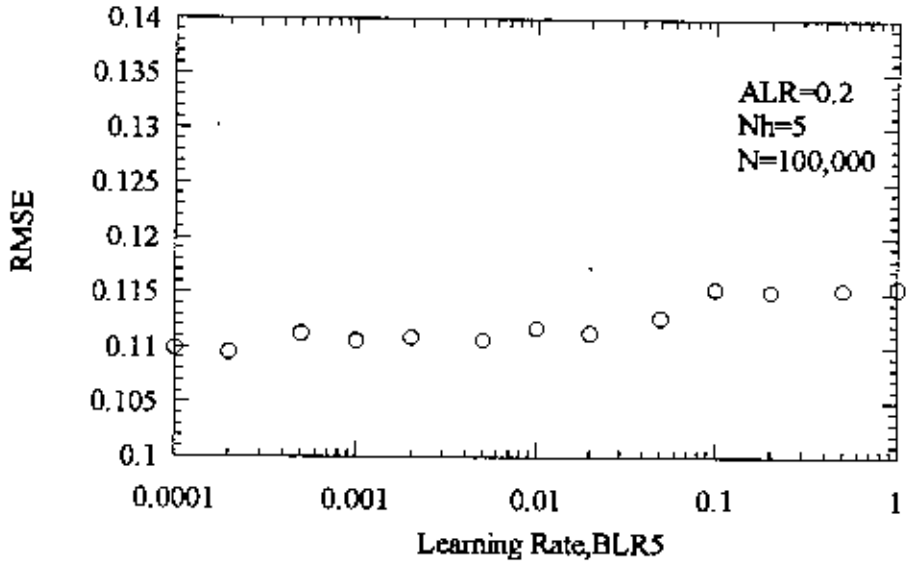

Figure 6.9: RMSE versus BLR for ALR=0.2.

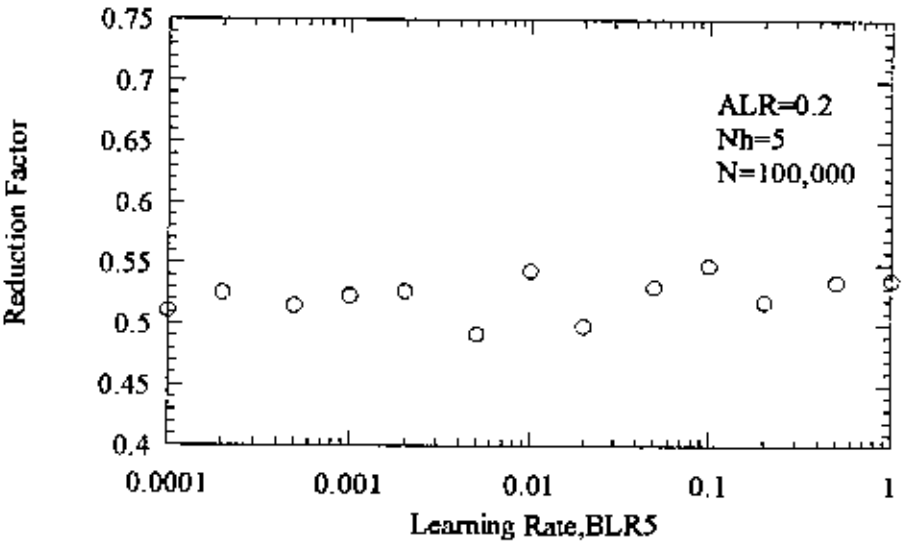

Figure 6.10: Reduction factor of RMSE versus BLR for ALR=0.2.

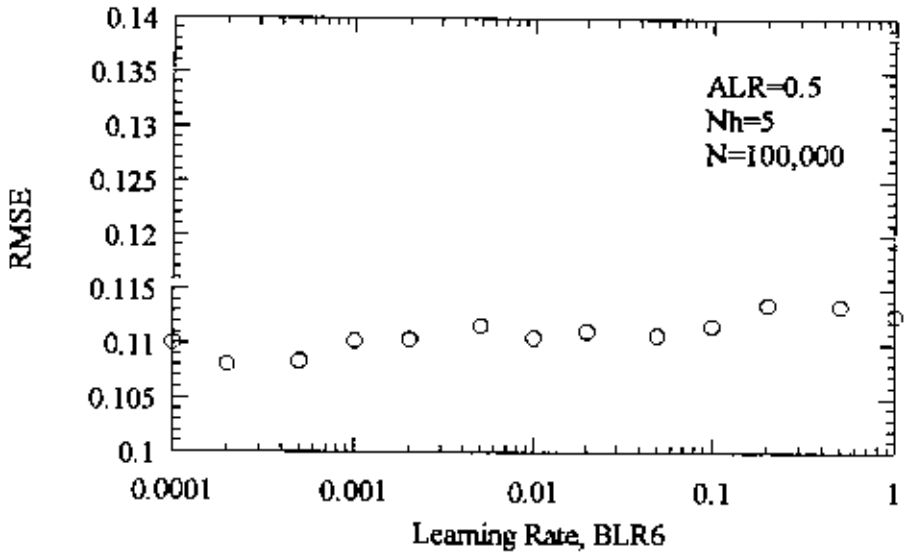

Figure 6.11: RMSE versus BLR for ALR=0.5.

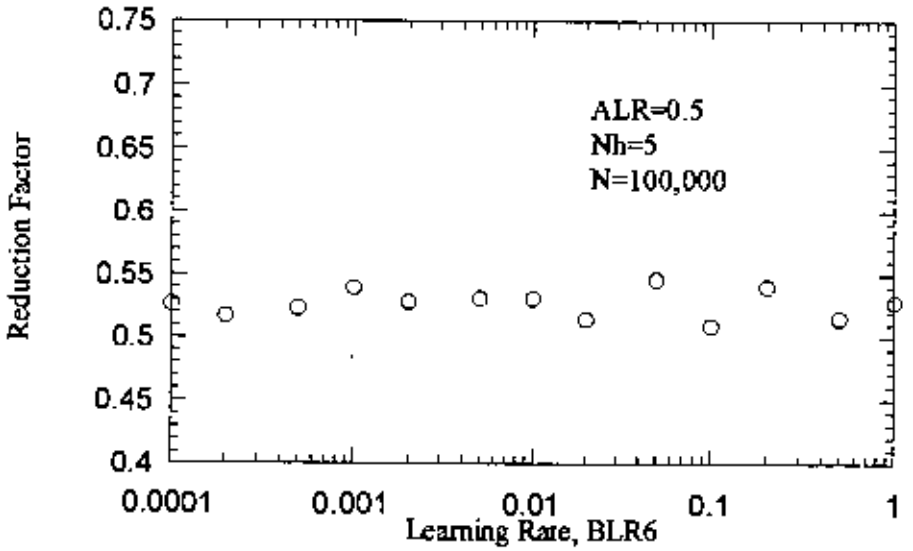

Figure 6.12: Reduction factor of RMSE versus BLR for ALR=0.5.

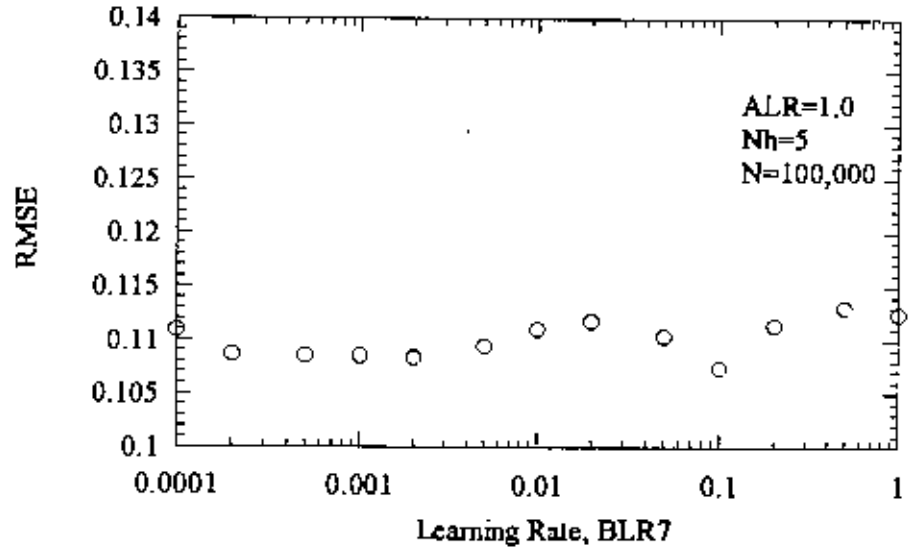

Figure 6.13: RMSE versus BLR for ALR=1.

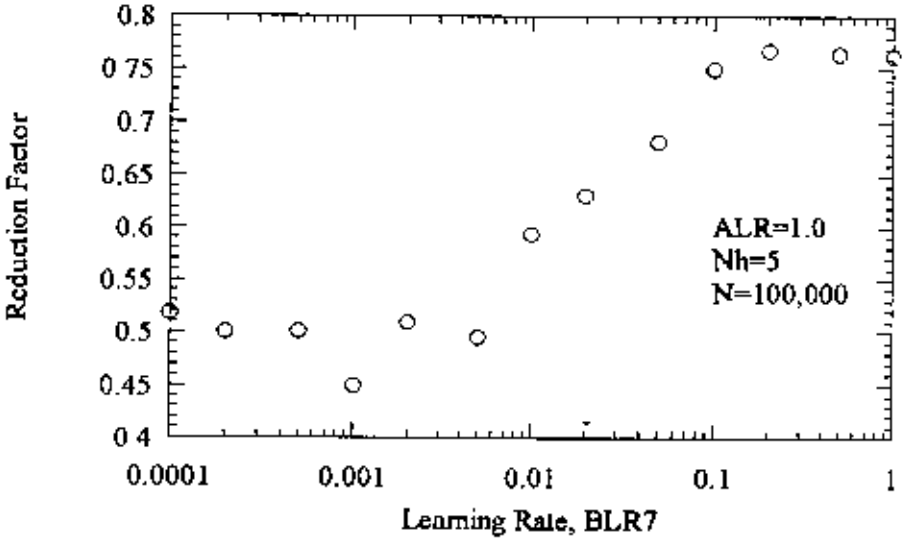

Figure 6.14: Reduction factor of RMSE versus BLR for ALR=1.0.

Considering the RMSE variation with BLR in the above cases, an overall optimum range of BLR is found from 0.0002 to 0.005. This range appears to yield good response for all the cases considered. Although it should be noted the change in RMSE with respect to BLR is not so high. Thus, it can be concluded that BLR value may not be very significant.

6.3.2 Input-to-Hidden-Layer Learning Rate Selection

Similar to hidden-to-output-layer learning rate (BLR) selection procedure, the optimum input-to-hidden-layer learning rate (ALR) was determined through a sensitivity study. Keeping the number of hidden layer node and the number of epochs constant, BLR is fixed at some values within the optimum range found in the previous section, and ALR is varied from 0.0001 to 1. Again, the number of hidden layer node is fixed at 5, while the number of epoch considered is 10^5 for the sensitivity study. Plots of RMSE values and their reduction factors against ALR are generated.

The BLR values are fixed at 0.0002, 0.0005, 0.001, 0.002, and 0.05. The corresponding RMSE variations versus ALR are shown in Figures 6.15, 6.17, 6.19, 6.21, and 6.23, respectively. The variations in reduction factor of RMSE versus ALR plots are given in Figures 6.16, 6.18, 6.20, 6.22, and 6.24, respectively. The corresponding optimum ranges for ALR are found to be (0.5, 5), (1, 10), (0.1, 5), (0.1, 5), and (0.1, 10), respectively. Figure 6.15 shows significant reduction of RMSE particularly at low valne of ALR and subsequent regular reduction of RMSE over the range. Figure 6.21 shows very steep and optimum range of ALR in the range of 5 to 20. Figure 6.23 indicates the variation of RMSE for BLR=0.005 is not significant.

The variation of RMSE values with respect to ALR is evidently more significant than that with respect to BLR values. From the sensitivity study, the optimum values of ALR and BLR are selected at 2 and 0.0005, respectively.

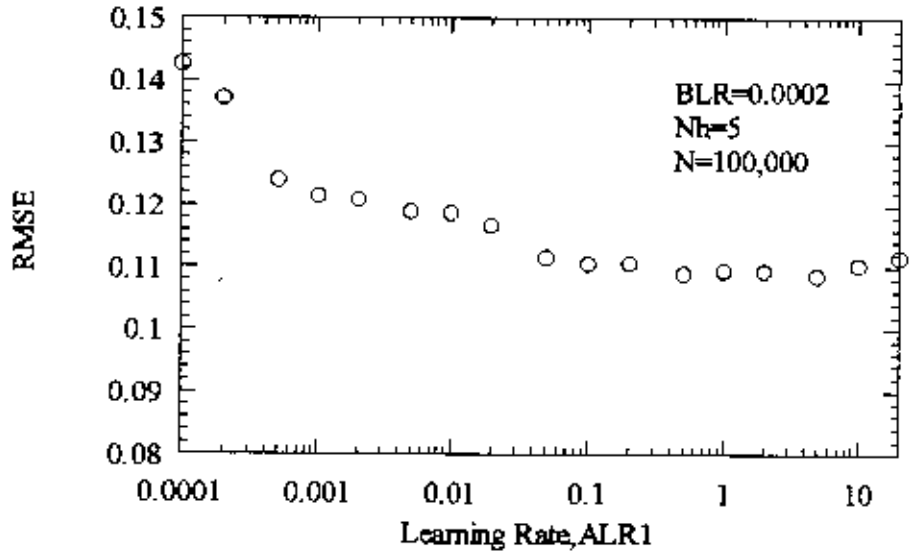

Figure 6.15: RMSE versus ALR for BLR=0.0002.

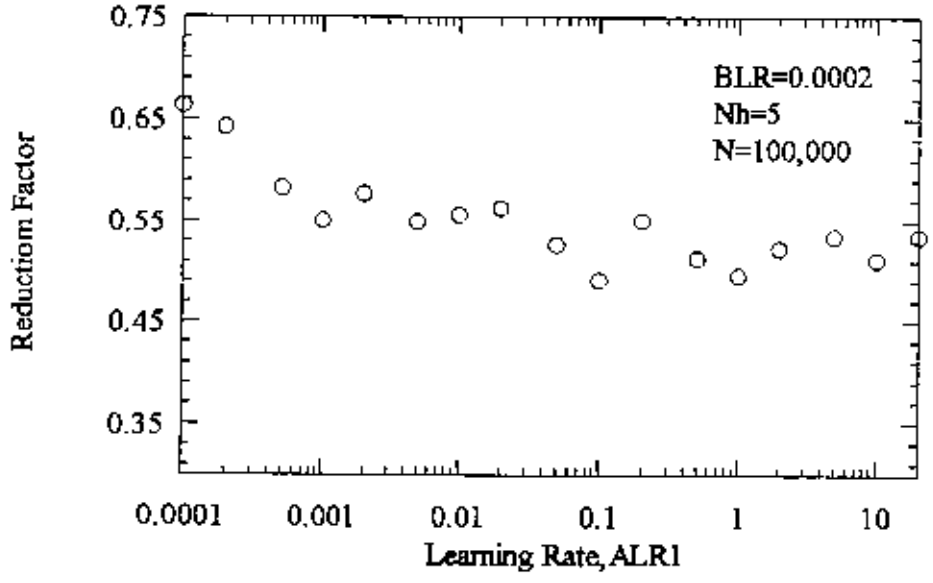

Figure 6.16: Reduction factor of RMSE versus ALR for BLR=0.0002.

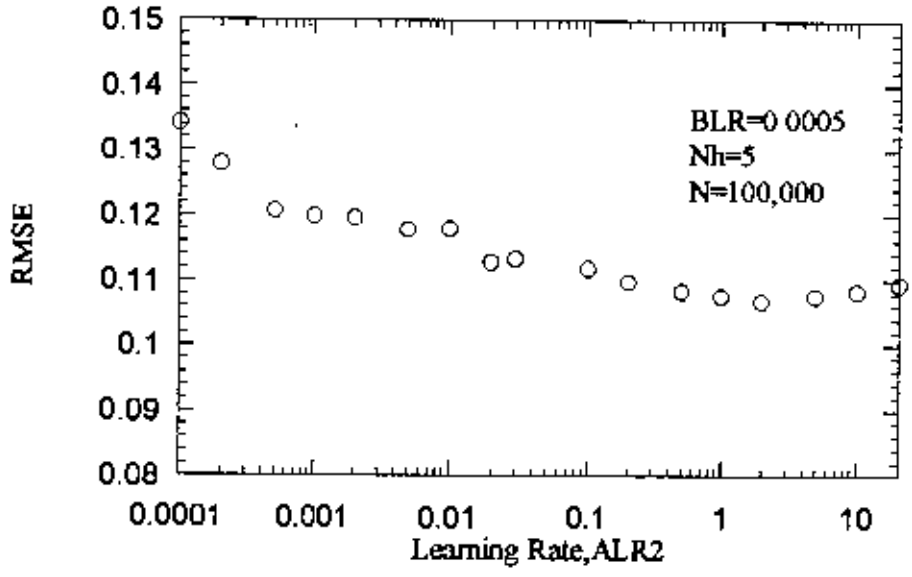

Figure 6.17: RMSE versus ALR for BLR=0.0005.

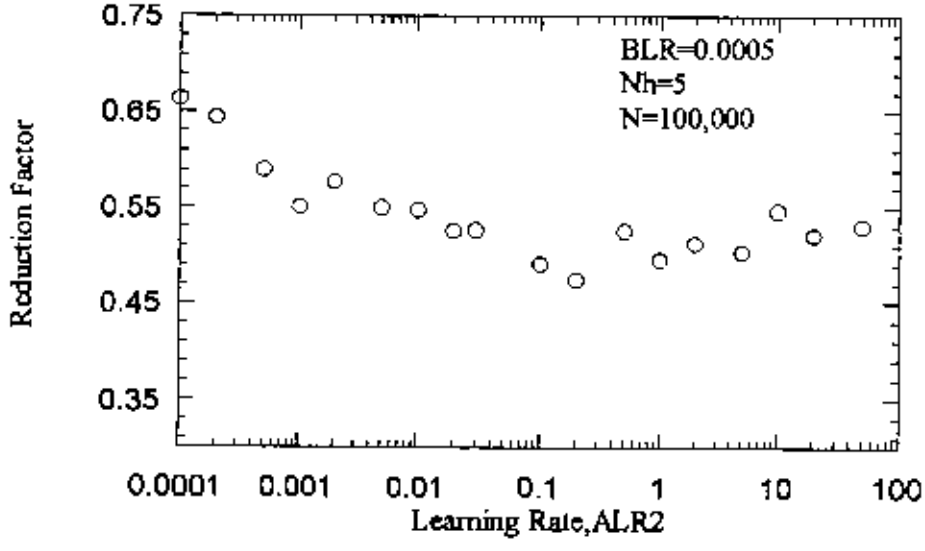

Figure 6.18: Reduction factor of RMSE versus ALR for BLR=0.0005.

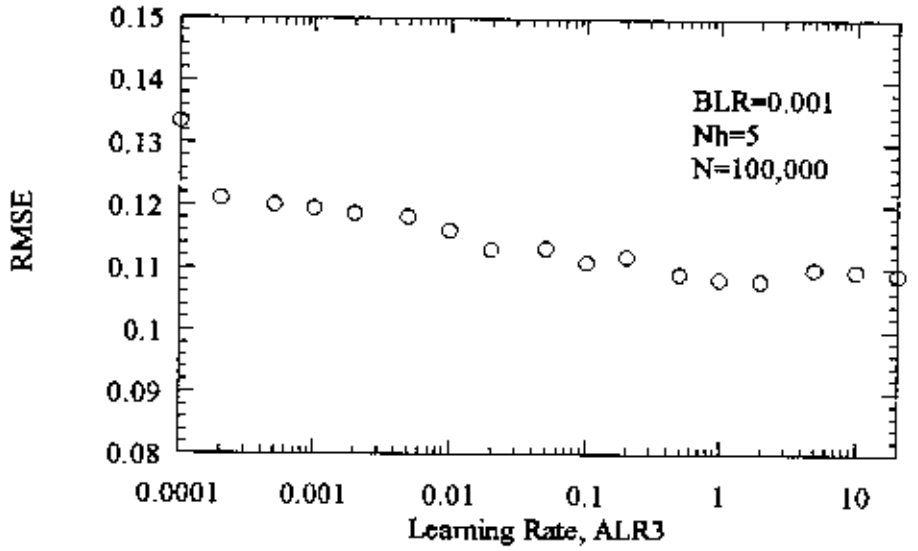

Figure 6.19: RMSE versus ALR for BLR=0.001.

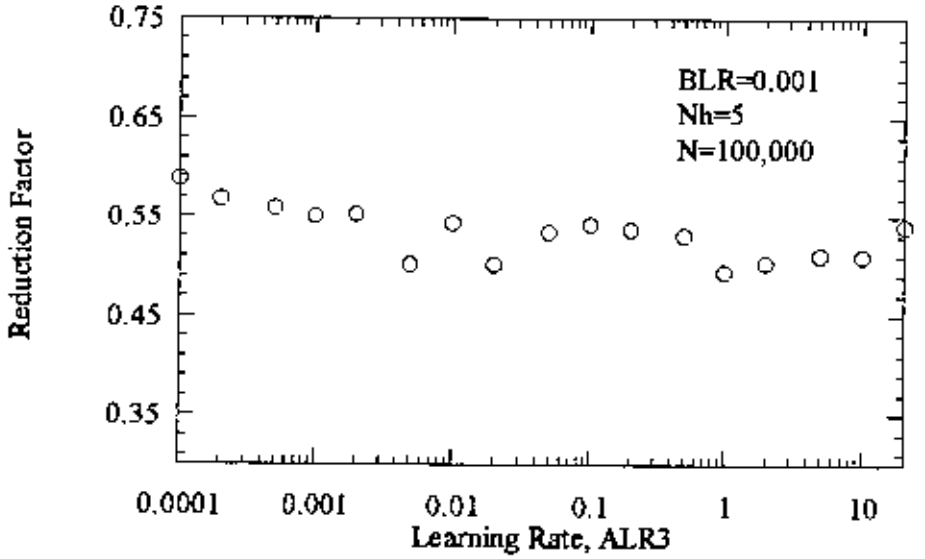

Figure 6.20: Reduction factor of RMSE versus ALR for BLR=0.001.

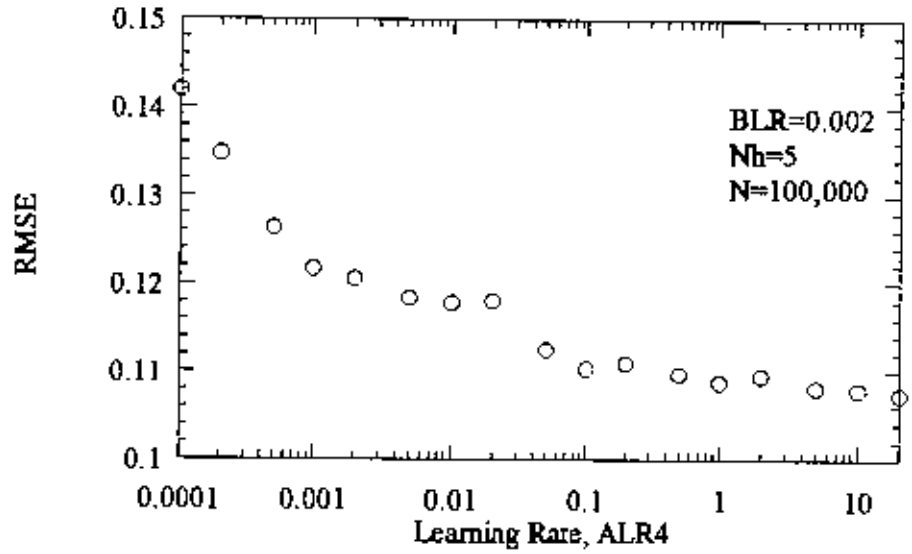

Figure 6.21: RMSE versus ALR for BLR=0.002.

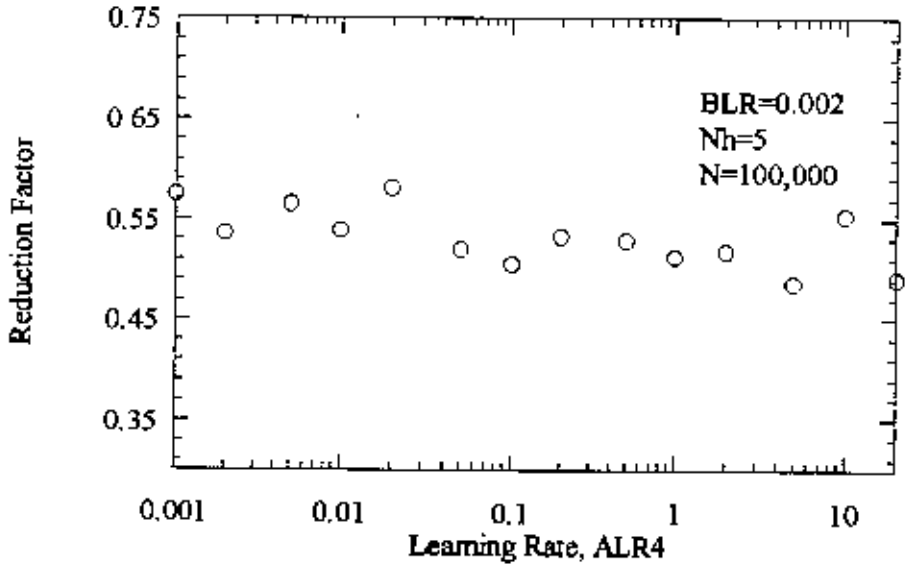

Figure 6.22: Reduction factor of RMSE versus ALR for BLR=0.002.

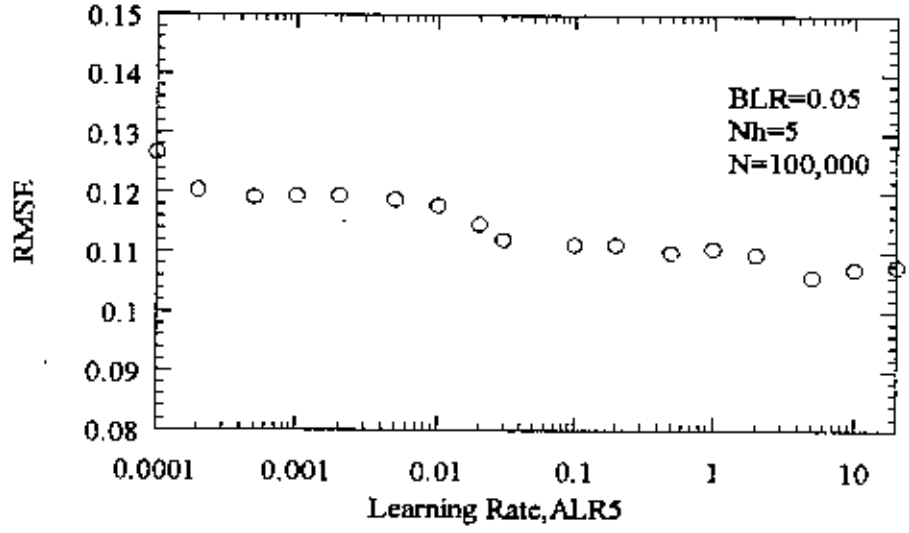

Figure 6.23: RMSE versus ALR for BLR=0.05.

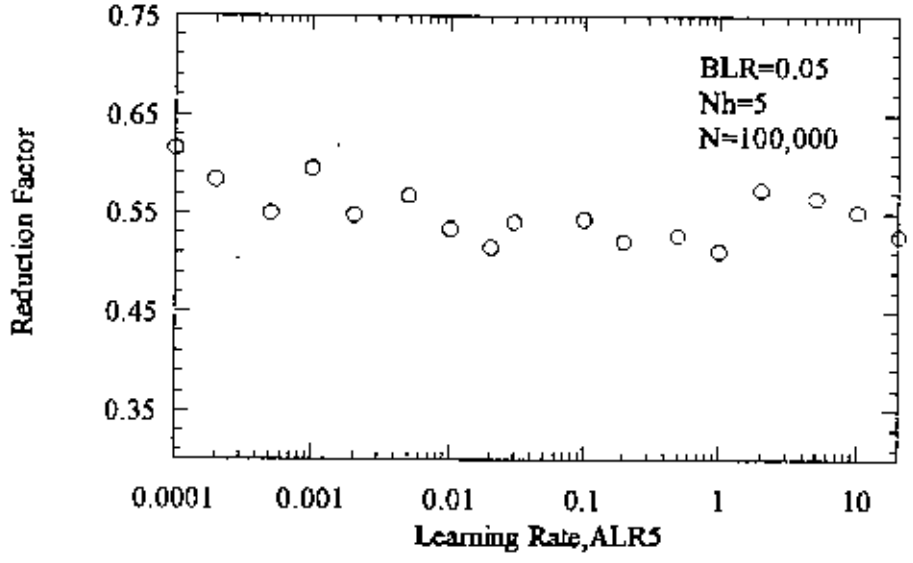

Figure 6.24: Reduction factor of RMSE versus ALR for BLR=0.05.

6.4 Number of Hidden Nodes Selection

The optimum number of hidden node is determined using the previously obtained optimum values of bias and learning rates. The number of hidden layer nodes is varied from 1 to 15 and the RMSE variation is examined. Figure 6.25 shows initial increase of RMSE with just one hidden node. However, a subsequent rapid reduction of RMSE with the number of hidden node is evident in the figure with RMSE eventually becoming stabilized. Stabilization occurs when the number of node is about 6. The optimum number of hidden node is fixed at 6 considering the thumb rule (number of connections should be at least three times higher than the number of patterns) and the computational load (computation increases with number of hidden node).

The following section discusses the learning profile of the neural networks. It mainly discusses how fast it learns or converges to expected local minimum error.

6.5 Learning Profile

The variation of RMSE with number of epochs gives the learning profile. Figure 6.26 shows RMSE versus number of epochs for the optimum parameters. It is apparent from the figures that training starts with large error, but rapid reduction of error takes place at moderate number of epochs, with eventually the RMSE flattening out Most of the training is completed within epoch number of 1000. Subsequent training rate is slow, but RMSE continuously decreases along the path of convergence.

During learning neural network changes its weights with respect to epoch number. Change of weights with respect to epoch number 1, 2, 5, 10, 20, 50,100, 200, 500, 1,000, 2,000, 5,000, 10,000, 20,000, 50,000 and 100,000 has been shown in a combined form in Figure 6.27. This figure shows that at the beginning of training, weights are near about to zero. Though the change in RMSE after epoch number 1,000 is very slow, the significant change in weights is continued up to epoch number 20,000. Final weights have been shown in the last one (for epoch number 10^5) and neural networks perform its prediction based on this set of weights.

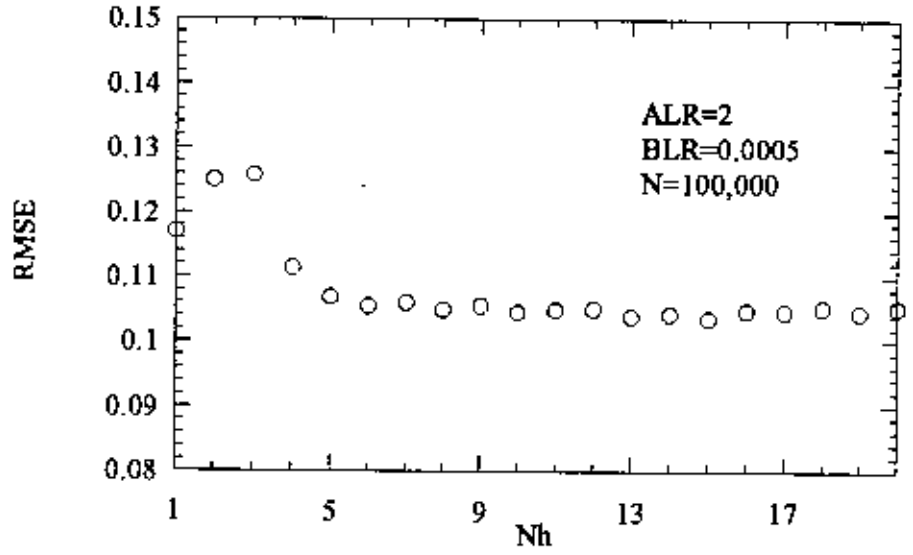

Figure 6.25: RMSE versus number of hidden node.

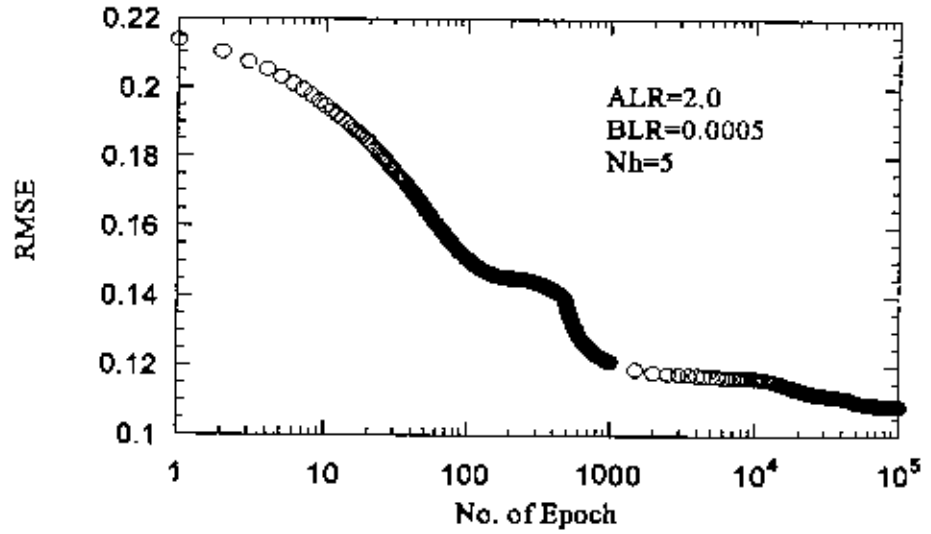

Figure 6.26: RMSE versus number of epoch.

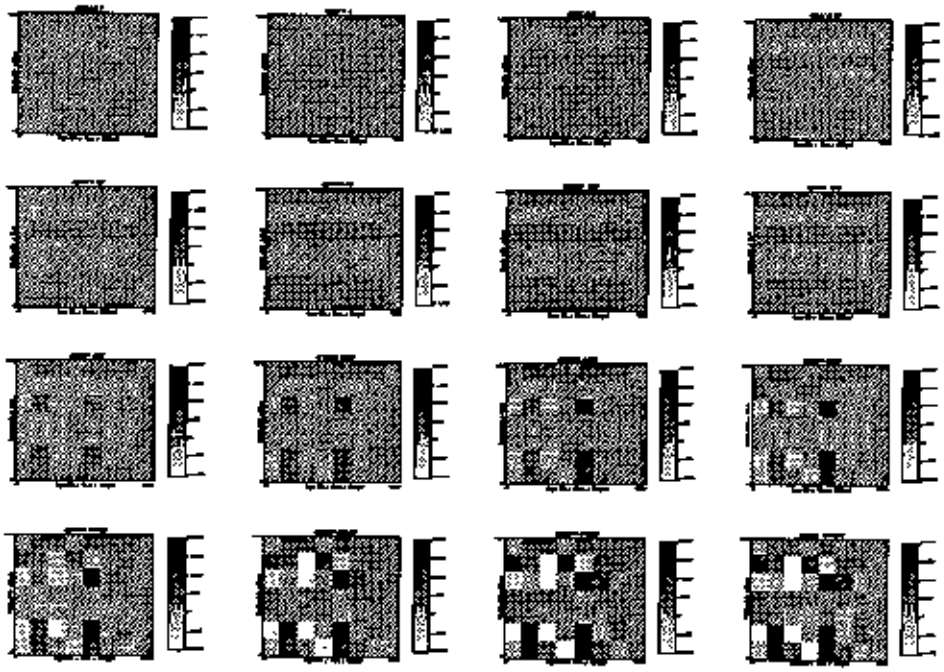

Figure 6.27: Changes in weights with respect to epoch number from 1 to 10^5.

In the following Section 6.6, it has been tried to show the major input parameters for permeability prediction.

6.6 Impact of Input Parameter

The Figure 6.29 is the final absolute weights after epoch number 10^5. The Figure consists of 7(row) × 8(column) blocks. The first top row is the bias from hidden to output layer. Subsequent rows represent the first, second, third, fourth, fifth and sixth hidden node respectively. Similarly, columns (from right side) represent input parameters x, y, relative strata, facies, porosity and thickness. The seventh column is the bias from input to hidden layer. Output is represented by eighth column From this Figure, it is clear that the major input parameters are porosity and relative strata. Y location is also important input parameter. The effect of biases is negligible.

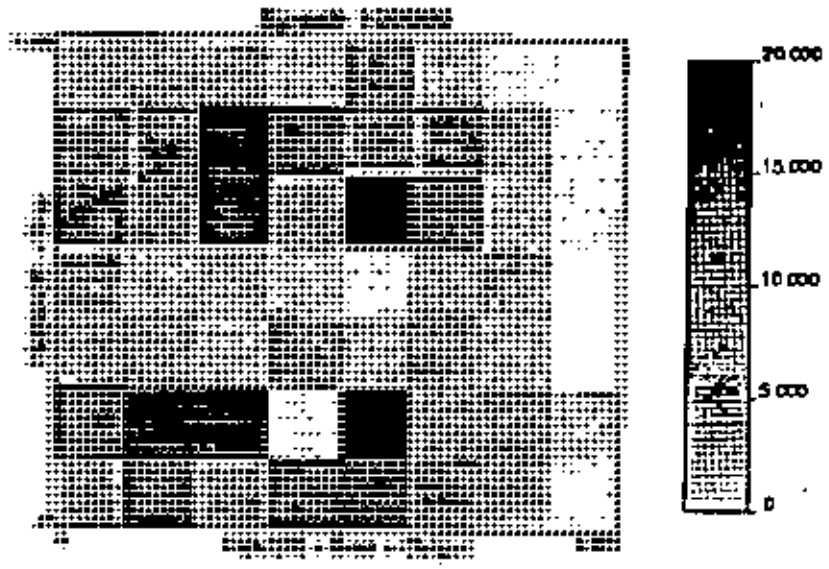

Figure 6.28: Significant of input parameter to output through absolute weights.

6.7 Randomness Effect

Figure 6.27 shows randomness plot. The figure shows RMSE variation from 0.10503 to 0.1107 due to randomness effect. The reduction factor varies from 0.46 to 0.56 due to randomness effect. An average value of RMSE is previously considered to reduce the randomness effect during the optimal parameter selection.

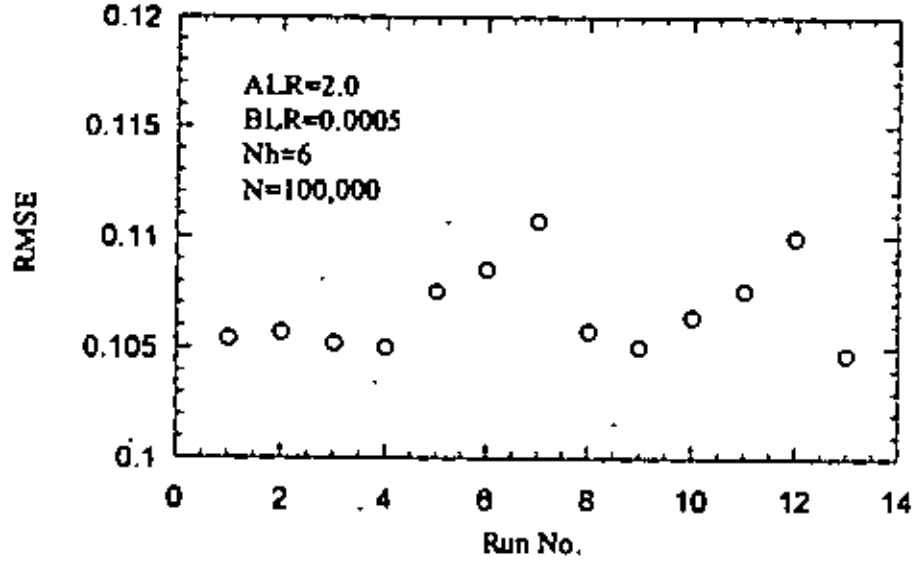

Figure 6.29: RMSE variation at different runs signifying randomness.

6.8 Well 5 Permeability Prediction

The developed neural network model has been trained upto epoch number 25×10^5 using the data of well 1, 2, 3, and 4. During the training, the selected optimum parameters were input to hidden layer bias = 0.5, hidden to output layer bias = 0 05, ALR=2.0, BLR = 0.0005, number of hidden node = 6. RMSE was 0.10041. Table 6.2 shows predicted permeability of Well 5.

6.9 Conclusion

- Development of a neural network modeling mainly deals with the optimum parameter selection. To find each optimum parameters, the neural net work with one hidden layer consists of 6 hidden node were trained upto epoch number 10^5. In most cases the RMSE was near about to 0.1060. The optimum parameters were input to hidden layer bias = 0.5, hidden to output layer bias =0.05, ALR=2 0, BLR = 0.0005, number of hidden node = 6.

- The higher value of ALR indicates that the input to hidden layer learning rate is more significant than the hidden to output layer learning rate.

- After training the model upto epoch number 10^5, the RMSE is quite high (near about to 0.11) and rate of change of RMSE with respect to epoch number is very slow after epoch number 10^5. This is due to the lack of information (other input parameters on which permeability depends) in the input layer. So, the model prediction error will be high.

- The change in RMSE is high upto epoch number 1000 and then it gets slow. But the change in weights is significant upto epoch number 20×10^3.

- Porosity and relative strata show their more impact to permeability prediction. The effect of biases is negligible.

- Average of RMSE should always be considered during the optimum parameter selection to reduce the randomness effect.

Table 6.2 Well 5 permeability prediction training Well 1, 2, 3 and 4 (Epoch No. 25×105).

Input						Output
X	Y	Relative Strata	Facies	Porosity	Thickness	Permeability
0.79657	0.63467	0.95714	0.00	0.096386	0.45	0.51216
0.79657	0.63467	0.85714	0.00	0.196790	0.45	0.51110
0.79657	0.63467	0.80000	0.20	0.120480	0.45	0.51103
0.79657	0.63467	0.77143	0.20	0.220880	0.45	0.51103
0.79657	0.63467	0.75000	0.20	0.164660	0.45	0.51102
0.79657	0.63467	0.72143	0.20	0.120480	0.45	0.51101
0.79657	0.63467	0.69285	0.00	0.265060	0.45	0.51136
0.79657	0.63467	0.60000	0.00	0.208840	0.45	0.51217

CHAPTER 7

CROSS VALIDATION

Neural network model responses are verified using cross validation method. This chapter discusses the analyses obtained from the cross validation exercise. As stated earlier, the present study develops a neural network model (Chapter 6) to predict well permeability using available petrophysical and spatial information including porosity, facies, well location, relative stratigraphic depth and formation thickness. Cross validation is a statistical analysis in which model responses are obtained for some known subset of input parameter space without using prior information of this subset and analyzing some statistical measures of the difference between the known subset and model response. In this study for all the wells that already have permeability information, we predict the permeabilities twice. First, these values are predicted using prior information. Second time, we do not use the prior information. The difference between the two predictions provides us information regarding the quality of the model responses for any particular well or area of the reservoir. It is important to note that all other parameters were kept unchanged during the training phase. We used an epoch number of 10^5 for the study.

The outline of the chapter follows. Section 7.1 discusses the cross validation analyses for Well 1. Similarly, Sections 7.2, 7.3 and 7.4 discuss the same for Wells 2, 3 and 4. Section 7.5 relates the uncertainty characterization of the permeability prediction space. Some conclusions are drawn in Section 7.6.

7.1 Well 1 Response Validation

Figure 7.1 is the permeability prediction versus pattern plot of Well 1 for the case in which the training set includes prior information of Well 1. The training root mean square error (RMSE) was 0.1063. The predicted and training patterns means are very close and they are 0.717 and 0.718. Figure 7.2 shows pattern-wise prediction errors. The absolute prediction error ranges from 0 to 0.15 with a mean value 0.05.

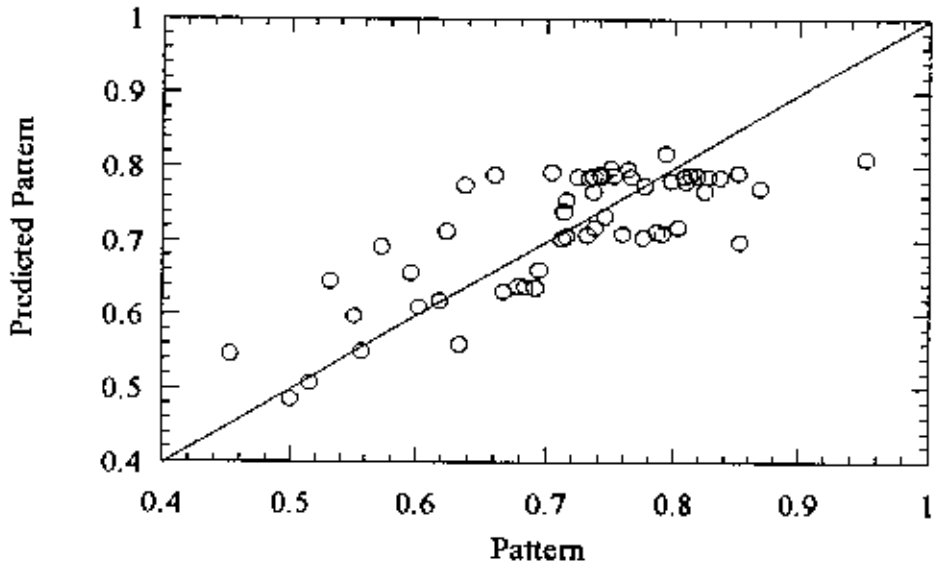

Figure 7.1 Permeability prediction versus pattern for Well 1 (Training set Wells 1, 2, 3, and 4).

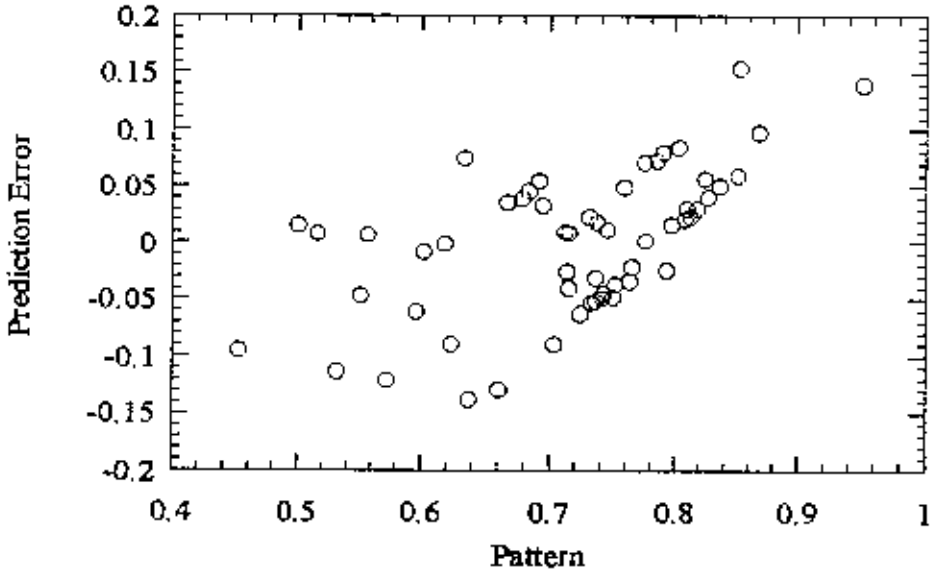

Figure 7.2: Permeability prediction error versus pattern for Well 1 (Training set: Wells 1, 2, 3, and 4).

For the model responses that do not use prior Well 1 information, the prediction versus pattern plot is given in Figure 7.3. Training RMSE in this case is 0.1084. The predicted pattern mean is 0.721. Prediction errors are shown in Figure 7.4. Mean absolute error is 0.05 while the absolute error values range from 0 to 0.16. It is apparent that the prediction performances for Well 1 are more or less same in both cases.

7.2 Well 2 Response Validation

Permeability prediction for Well 2 using prior information in training set is shown in Figure 7.5. Training RMSE is 0.1063. Predicted mean is 0.683 and while pattern mean 0.681. The predicted patterns and patterns means are near about to equal. Pattern wise prediction errors are shown in Figure 7.6. Mean value of absolute error is 0.096 with a range from 0 to 0.40. For model responses without Well 2 prior information, permeability predictions versus patterns are shown in Figure 7.7. Prediction errors are shown in Figure 7.8. RMSE is 0.0882. Large prediction errors are clearly evident. The difference between prediction mean (0.721) and pattern mean (0.716) is significant. Absolute error for prediction ranges from 0 to 0.44 while the mean is 0.111. So, the responses are less reliable for Well 2 compared to that for Well 1.

7.3 Well 3 Response Validation

Well 3 permeability prediction and prediction errors are shown in Figures 7.9 and 7.10 for the case when the Well 3 data is included in the training patterns. RMSE is 0.1066. The prediction and pattern means are 0.574 and 0.597. Mean absolute error for prediction is 0.094, and it ranges from 0 to 0.29. While, permeability prediction and prediction errors are shown in Figure 7.11 and 7.12 for the case when Well 3 prior information are not included in the training phase. Bias in the prediction worsens. The mean absolute error for prediction is 0.117. Pattern wise absolute prediction error varies from 0 to 0.35.

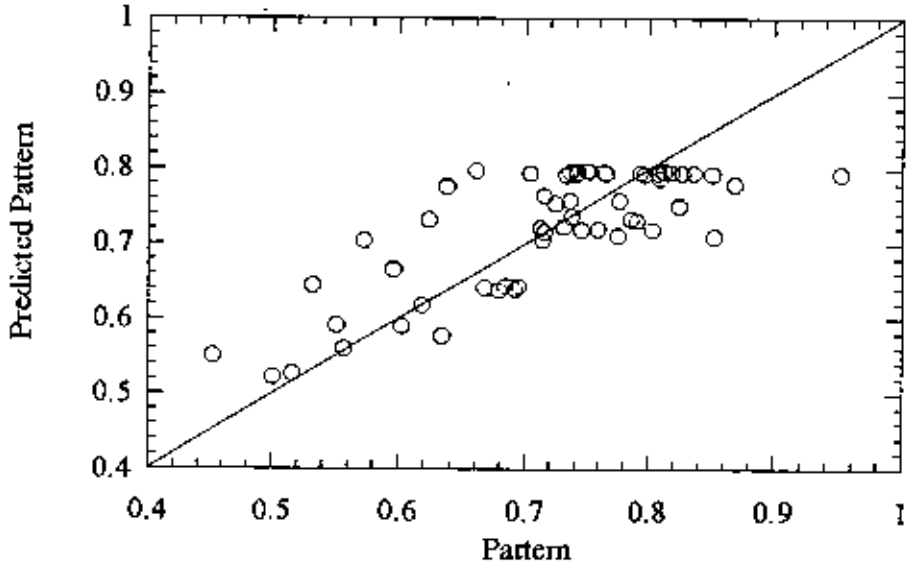

Figure 7.3: Permeability prediction versus pattern for Well 1 (Training set: Wells 2, 3, and 4).

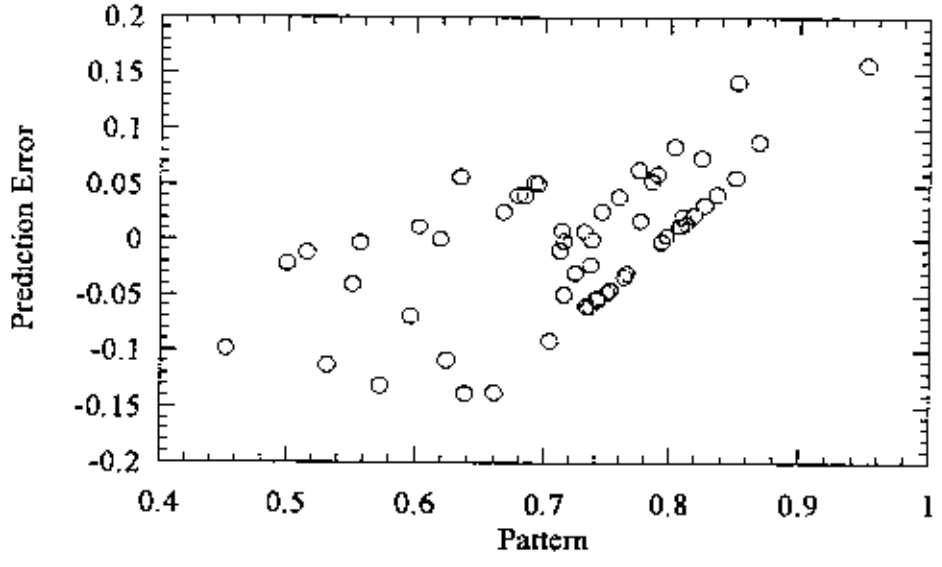

Figure 7.4: Permeability prediction error versus pattern for Well 1 (Training set: Wells 2, 3, and 4).

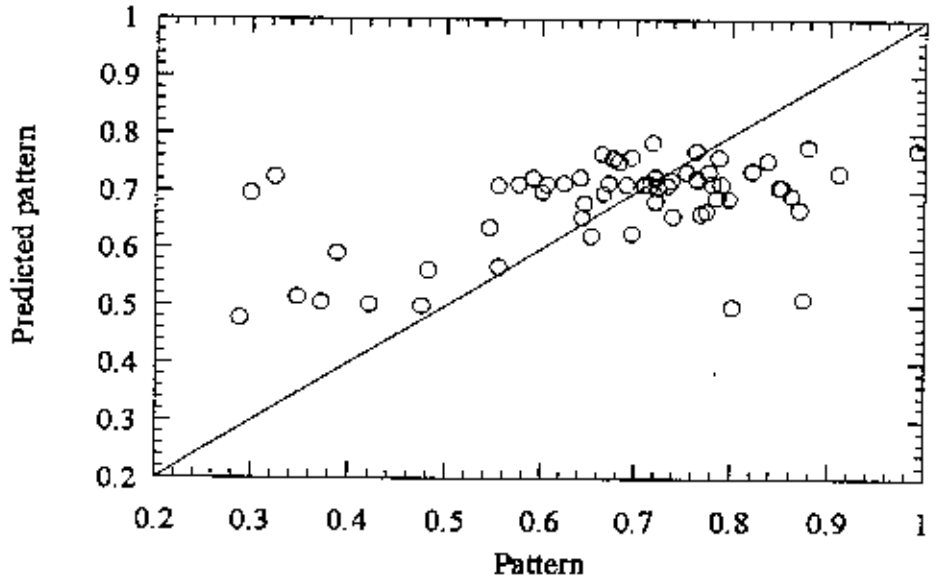

Figure 7.5: Permeability prediction versus pattern for Well 2 (Training set: Wells 1, 2, 3, and 4).

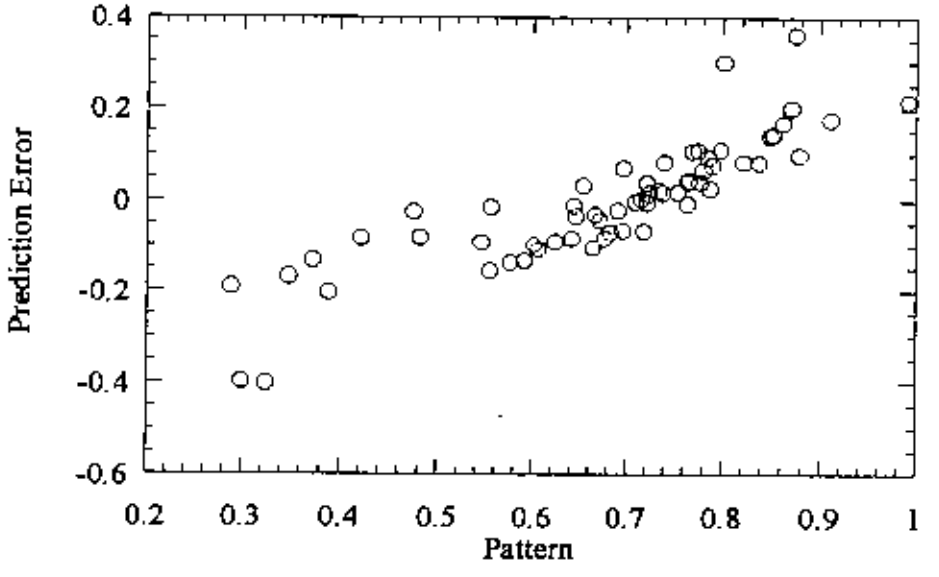

Figure 7.6: Permeability prediction error versus pattern for Well 2 (Training set: Wells 1, 2, 3, and 4).

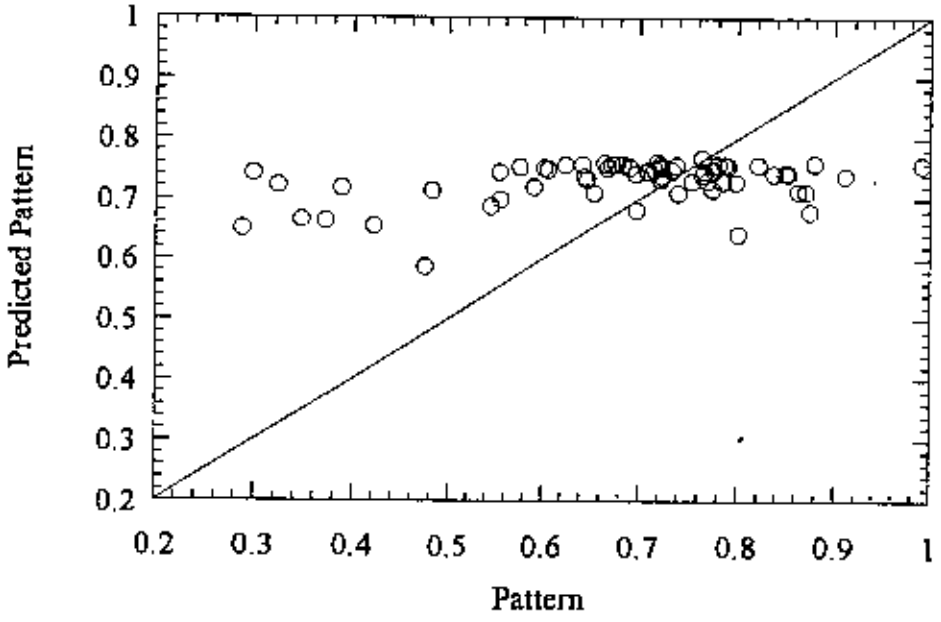

Figure 7.7: Permeability prediction versus pattern for Well 2 (Training set: Wells 1, 3, and 4).

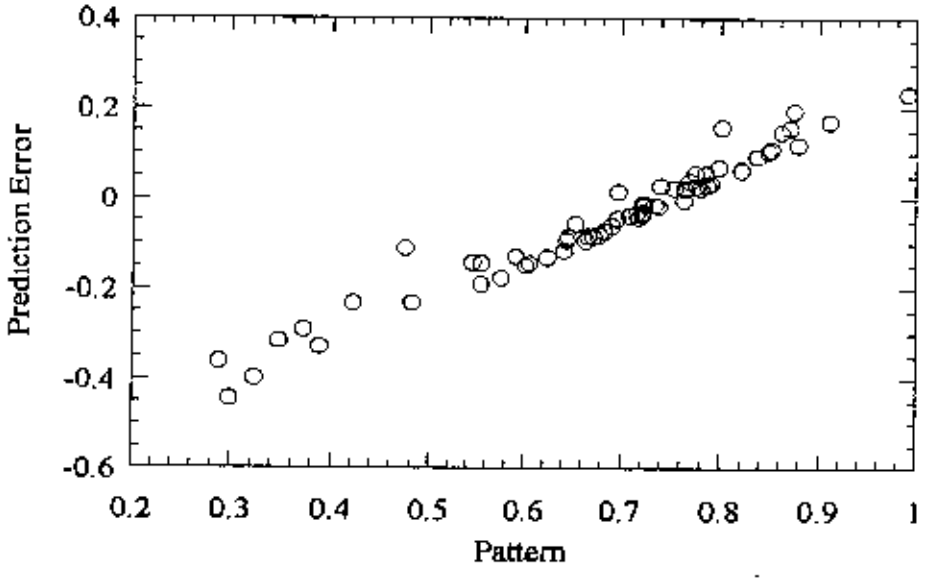

Figure 7.8: Permeability prediction error versus pattern for Well 2 (Training set: Wells 1, 2, 3, and 4).

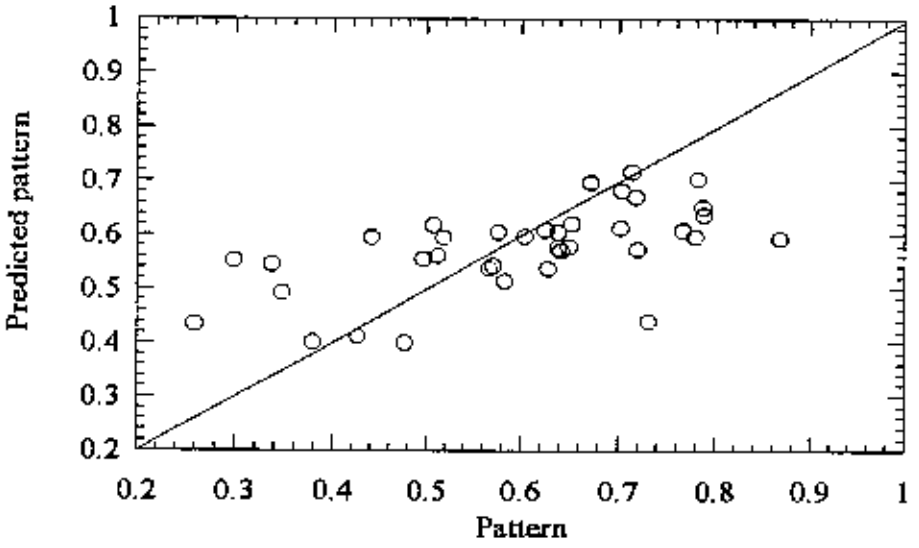

Figure 7.9: Permeability prediction versus pattern for Well 3 (Training set: Wells 1, 2, 3, and 4).

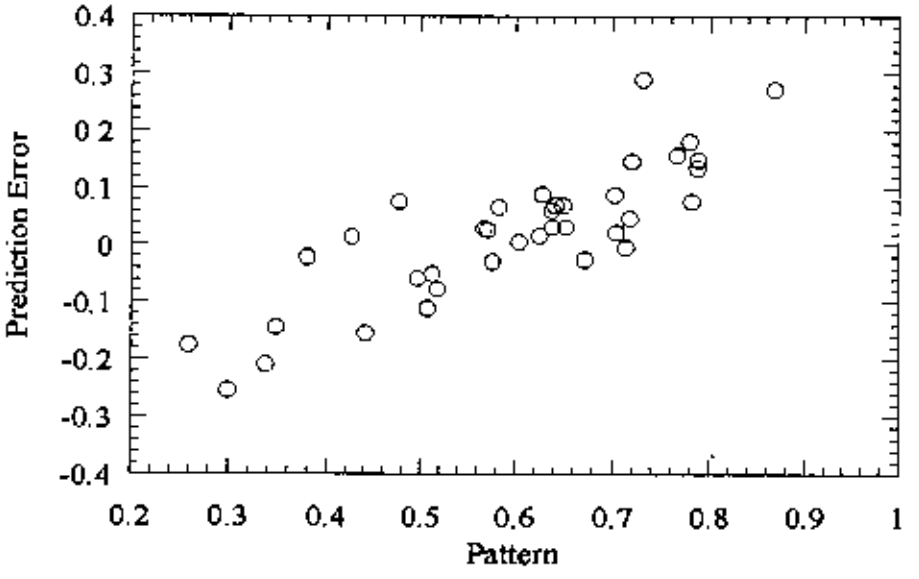

Figure 7.10: Permeability prediction error versus pattern for Well 3 (Training set: Wells 1, 2, 3, and 4).

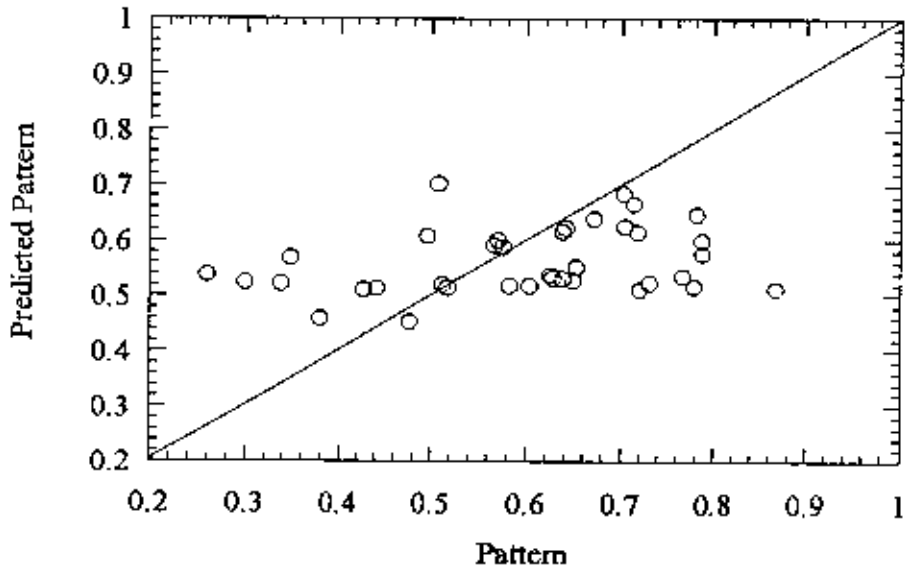

Figure 7.11: Permeability prediction versus pattern for Well 3 (Training set: Wells 1, 2, and 4).

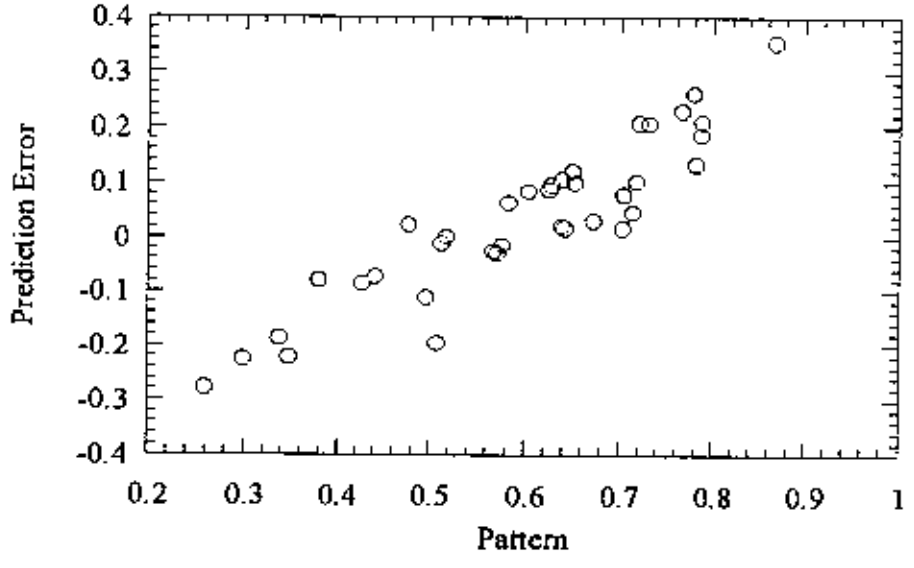

Figure 7.12: Permeability prediction error versus pattern for Well 3 (Training set: Wells 1, 2, and 4).

7.4 Well 4 Response Validation

Figures 7.13 and 7.14 show Well 4 permeability prediction and error plots when Well 4 data are included in training patterns. The predicted pattern and training pattern means are 0.522 and 0.591. Absolute prediction error varies from 0 to 0.17 with a mean value 0.065. Figures 7.15 and 7.16 are the prediction and error plots when Well 4 data is not used. The prediction and training pattern means are 0.670 and 0.552. Absolute mean error is 0.125 and it ranges from 0.01 to 0.27.

7.5 Prediction Space Characterization

In order to characterize the prediction space, we perform prediction exercises with different random number seeds and examine the scatter in the prediction. The scatter in the predicted values implicitly characterizes the uncertainty in prediction. Pattern-wise prediction scatters for Wells 1, 2, 3 and 4 are shown in Figures 7.17 to 7.24. The use of prior information is also examined in these plots. The uncertainty space is evidently larger when prior information is not used. It is intuitive.

7.6 Conclusion

- Model responses for Wells 1 and 2 are more reliable than those for Wells 3 and 4. The reason could be attributed to the amount of information available. However, there is always a possibility of not capturing information of spatial distribution with only 4 wells.

- Additional information may not lead to greater reduction in training RMSE. The reason could be existence of conflicting information in the data.

- Parameterization of the input space for the problem is not sufficient. Additional information regarding grain size and distribution, angularity, connectivity, geomechanical properties, diagenesis effect and so forth may improve the model responses.

- Uncertainty space in the prediction varies for one point to another within the reservoir.

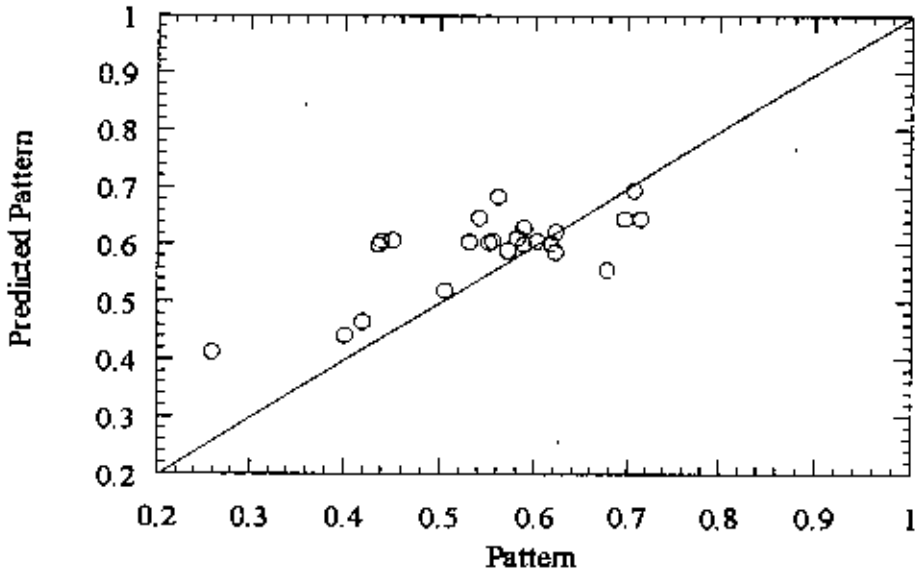

Figure 7.13: Permeability prediction versus pattern for Well 4 (Training set: Wells 1, 2, 3, and 4).

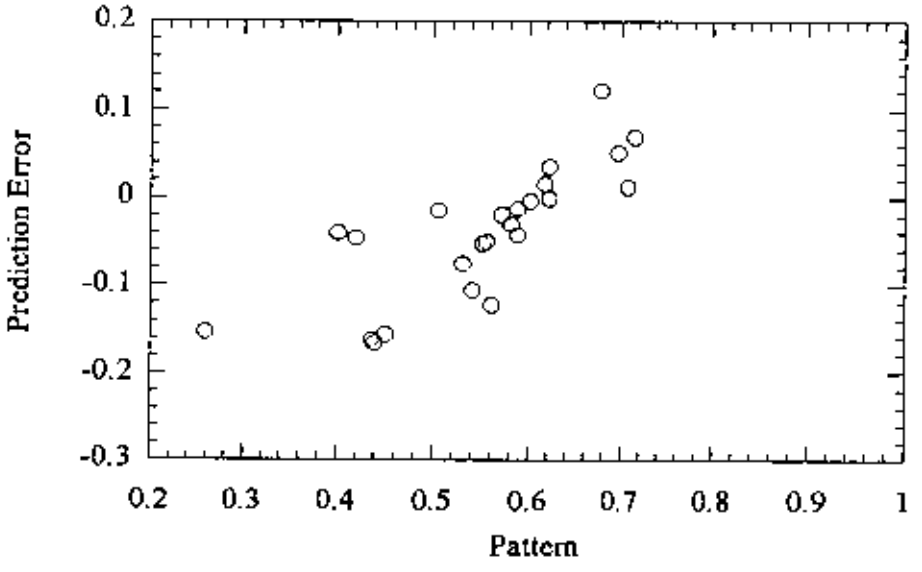

Figure 7.14: Permeability prediction error versus pattern for Well 4 (Training set: Wells 1, 2, 3, and 4).

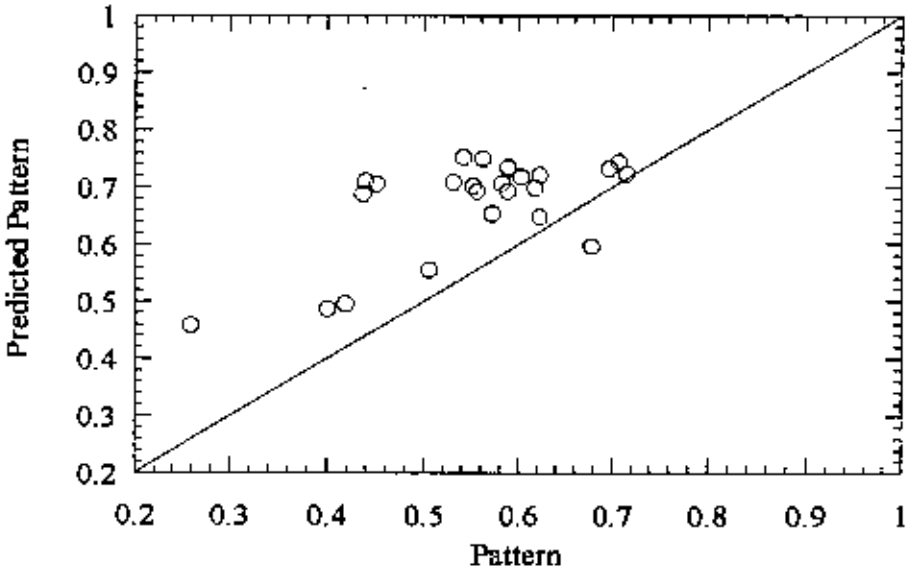

Figure 7.15: Permeability prediction versus pattern for Well 4 (Training set: Wells 1, 2, and 3).

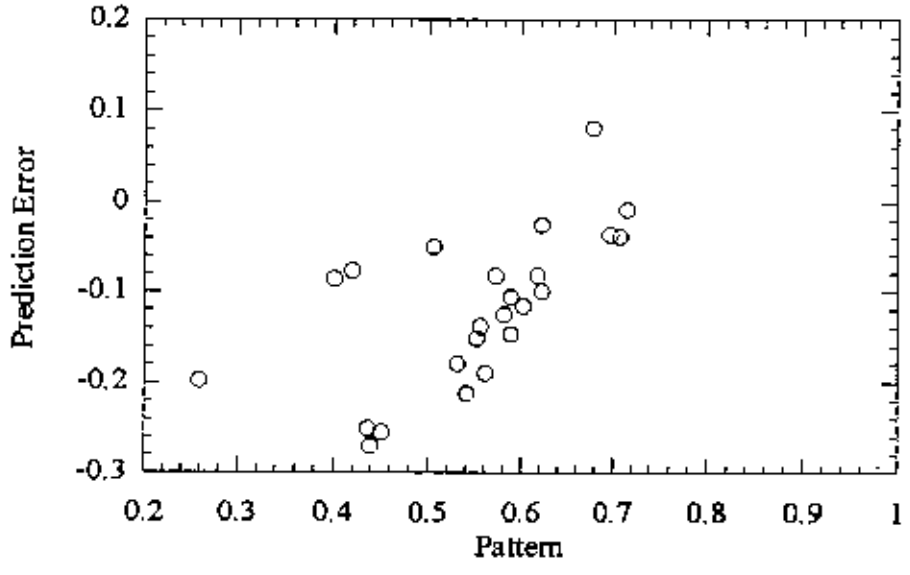

Figure 7.16: Permeability prediction error versus pattern for Well 4 (Training set: Wells 1, 2, and 3).

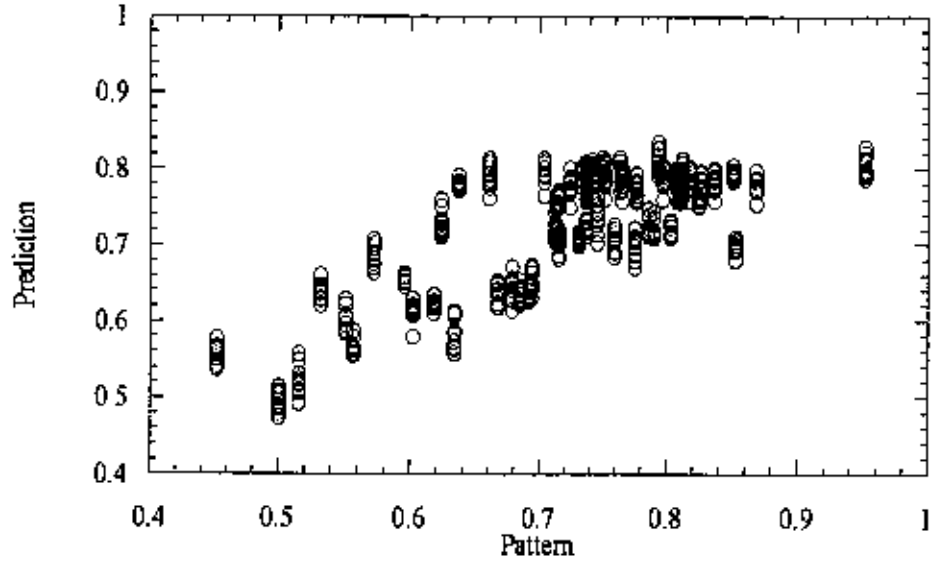

Figure 7.17: Permeability prediction scatters for Well 1 (Training set: Wells 1, 2, 3,. and 4).

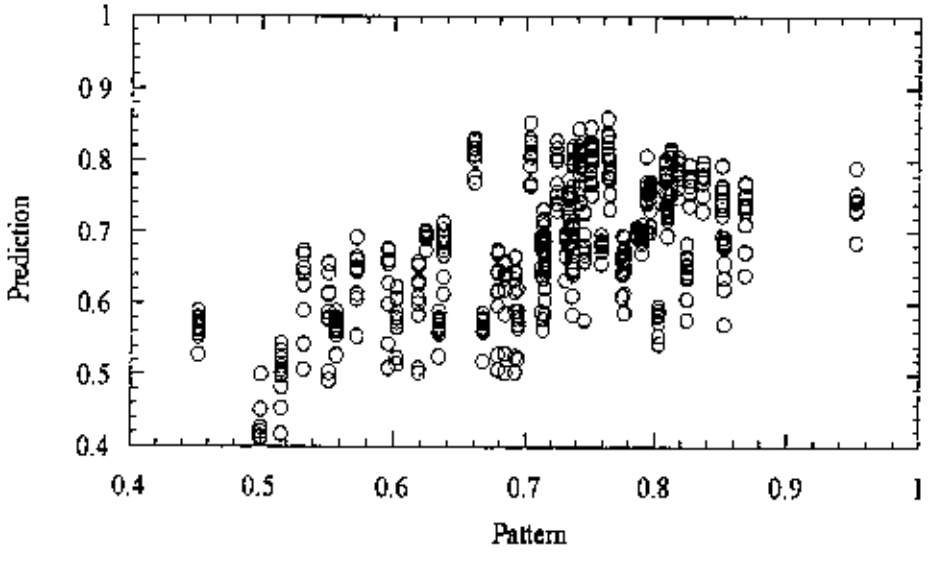

Figure 7.18: Permeability prediction scatters for Well 1 (Training set: Wells 2, 3, and 4).

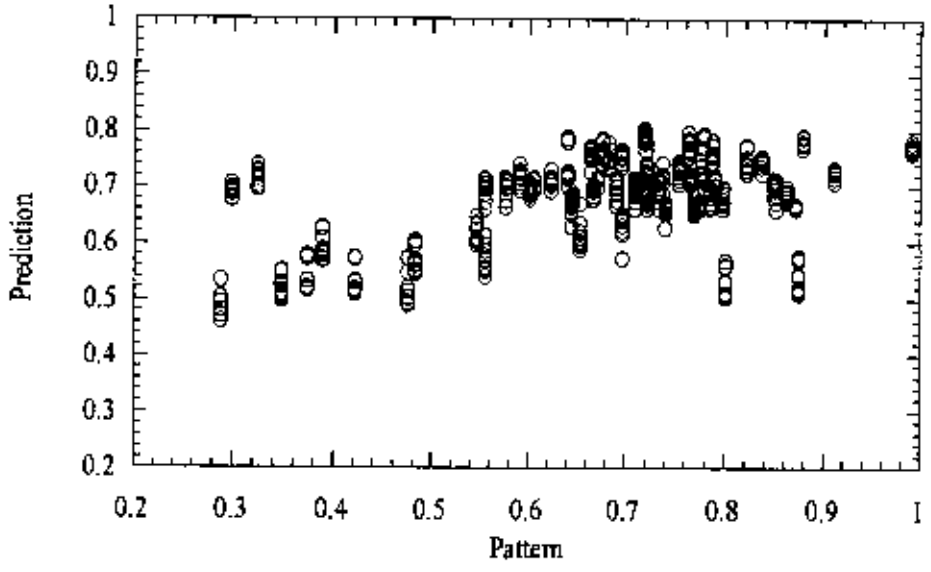

Figure 7.19: Permeability prediction scatters for Well 2 (Training set: Wells 1, 2, 3, and 4).

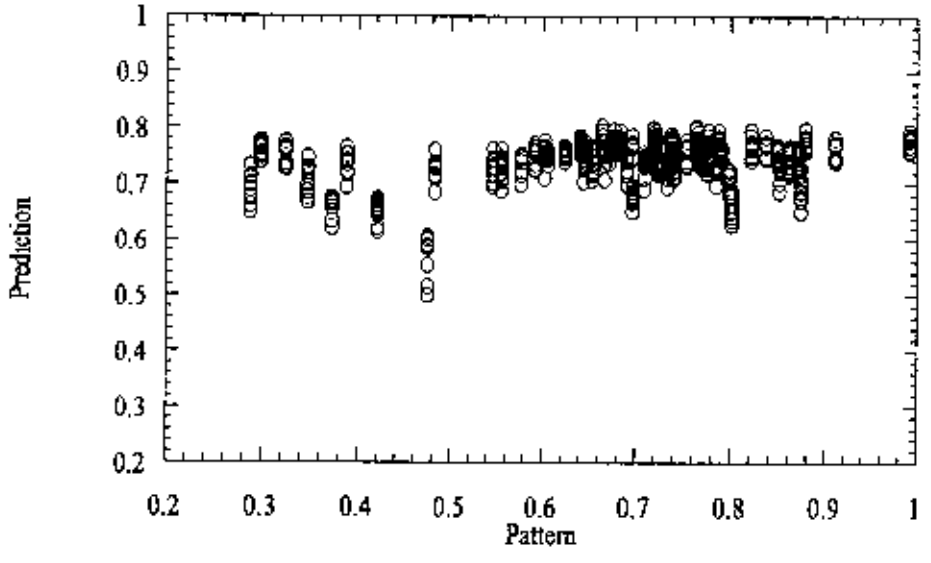

Figure 7.20: Permeability prediction scatters for Well 2 (Training set: Wells 1, 3, and 4).

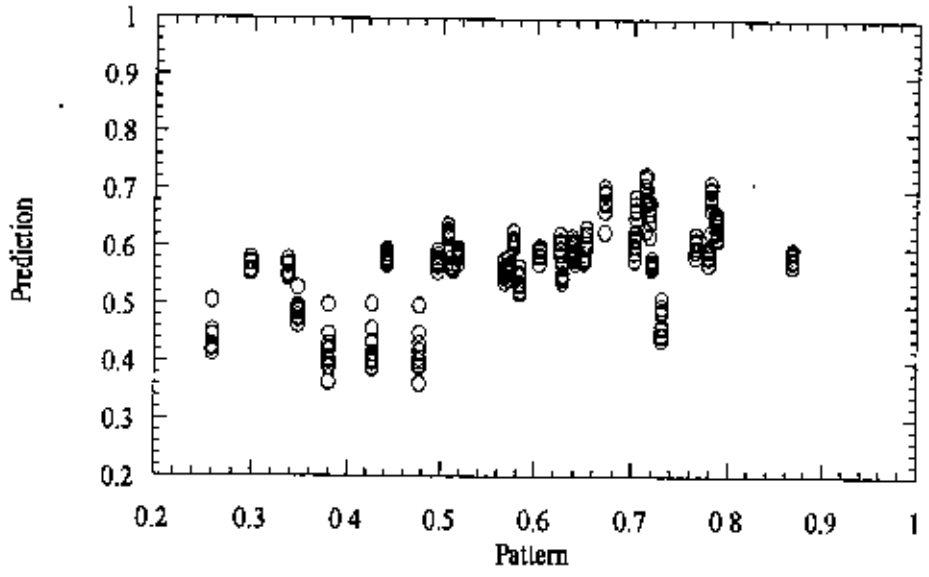

Figure 7.21: Permeability prediction scatters for Well 3 (Training set: Wells 1, 2, 3, and 4).

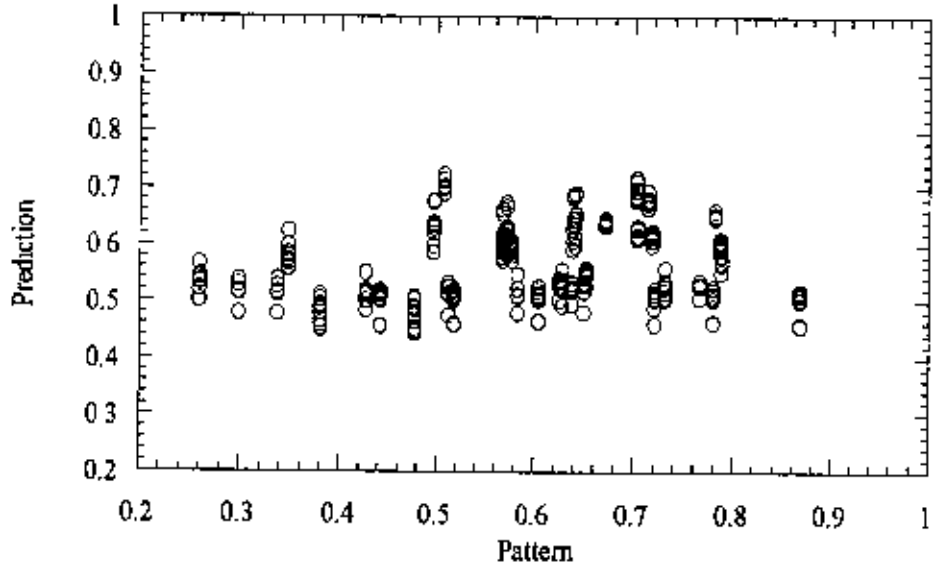

Figure 7.22: Permeability prediction scatters for Well 3 (Training set: Wells 1, 2, and 4).

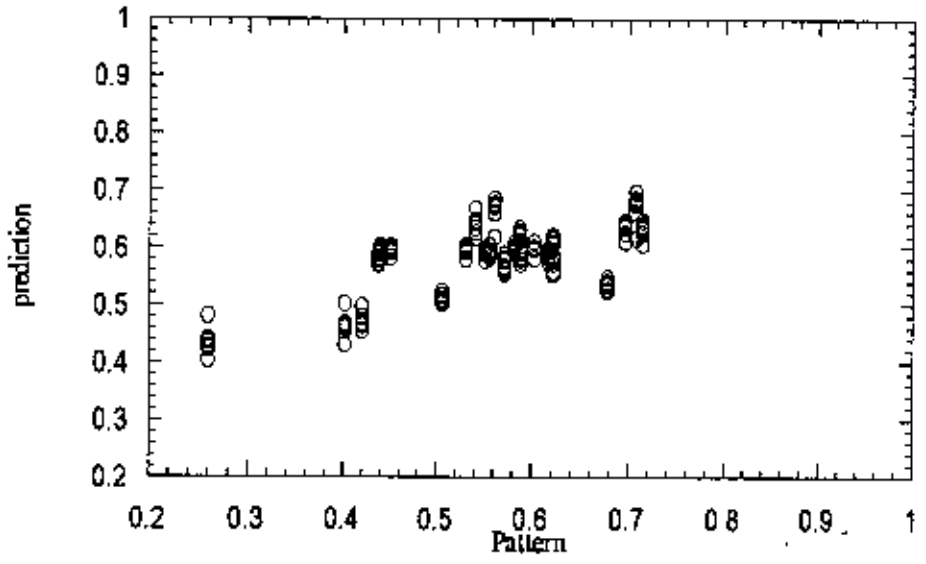

Figure 7.23: Permeability prediction scatters for Well 4 (Training set: Wells 1, 2, 3, and 4).

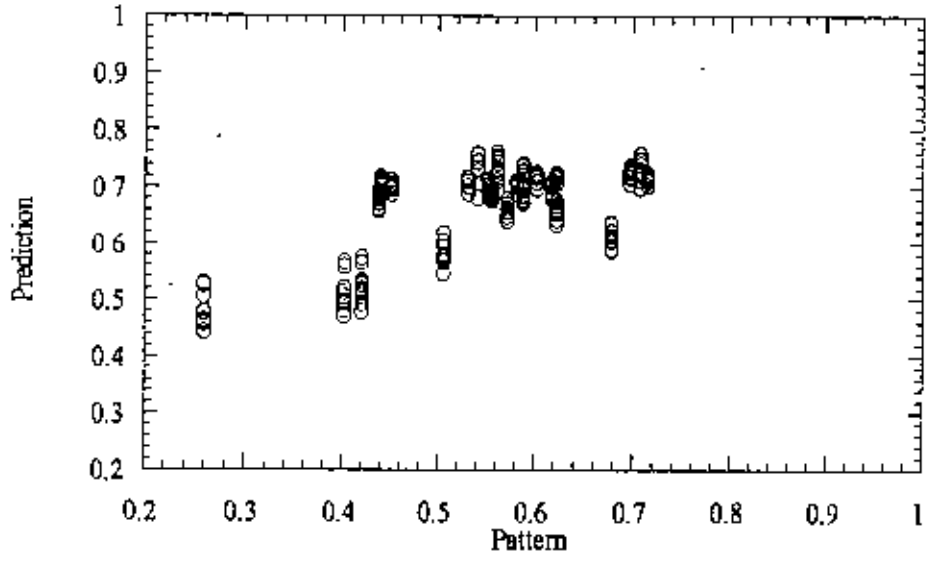

Figure 7.24: Permeability prediction scatters for Well 4 (Training set: Wells 1, 2, and 3).

CHAPTER 8

ERROR ANALYSIS

We analyze the prediction error in the neural network responses in this chapter. The objective is to investigate the source of errors and minimize the error. It is found that most of predictions tend towards the central values. Patterns with low and high (extreme) values have higher prediction errors. It was conjectured that affine scaling of the prediction histogram could improve the prediction. Affine scaling of the histogram is a deterministic method of introducing greater variability in the values keeping the shape of the histogram almost unchanged. Affine scaling has been used to lessen prediction error.

This chapter comprises of the following sections. Section 8.1 discusses neural network predictions. In Section 8.2, affine scaling is used to minimize error. Section 8.3 explores the sources of prediction error. Section 8.4 is the conclusion of error analysis.

8.1 Neural Network Predictions

We witness some interesting trend in the histograms of the patterns and predictions Figures 8.1 and 8.2 show histograms of the patterns and corresponding prediction for a typical neural network run. Pattern mean is 0.658. It ranges from 0.26 to 0 99 Prediction mean is 0.66 and the range is 0.40 to 0.81. Reproduction of mean is evident. In other words, there is very little bias in the neural network prediction. However, the prediction range is significantly lower than the pattern range.

We attempted to correct this error through the application of an affine scaling of the prediction histogram. In the following section, affine scaling is applied to minimize the neural network prediction error.

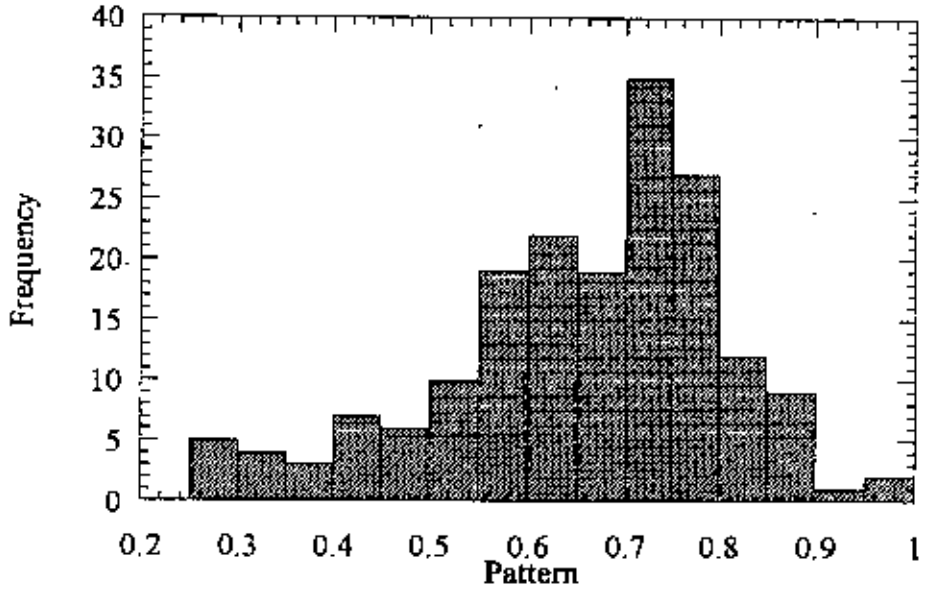

Figure 8.1 Histogram of training patterns.

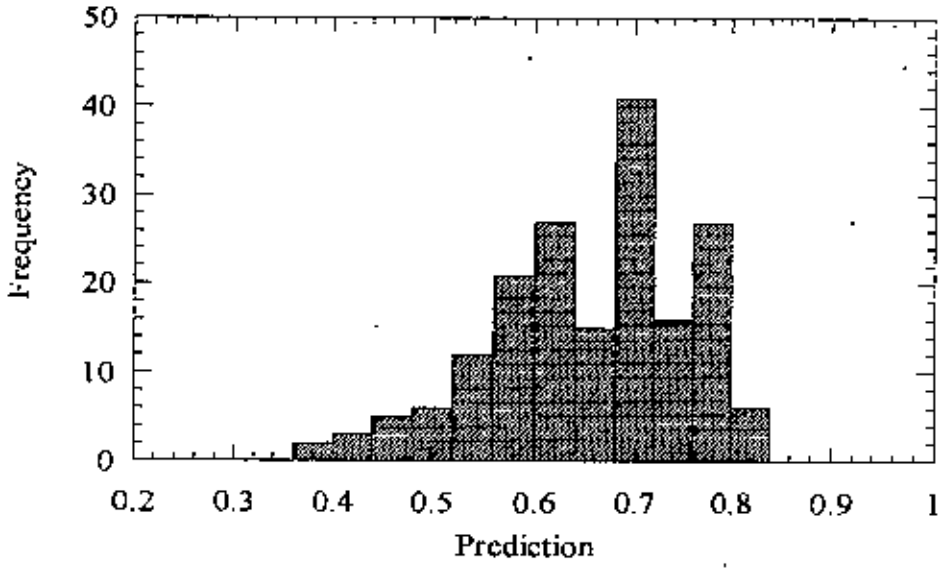

Figure 8.2 Histogram of neural network predictions.

8.2 Affine Scaling of Prediction Histogram

Affine scaling of the histogram is a deterministic method of introducing greater variability in the values keeping the shape of the histogram almost unchanged. A Fortran code for affine scaling of histogram, AFFINE, (Deutsch and Journel, 1998) was used.

Figure 8.3 shows neural network predictions and affine scaling corrected neural network prediction. The first column plots are prediction versus pattern and the second column are error versus pattern. Plots in the first row of the figure are neural network prediction and pattern-wise prediction error. The prediction mean is 0.659, and the pattern mean 0.658. The prediction ranges from 0.398 to 0.811 Correlation coefficient between pattern and prediction is 0.687. After affine scaling, prediction mean is 0.660 while the range is from 0.312 to 0.811. Correlation coefficient between pattern and prediction remains at 0.687. Although the range improves, there is hardly any improvement of the correlation. Table 8.1 presents neural network and affine scaling corrected neural network prediction results for 5 runs. Table 8.1 indicates similar trend in all the runs.

In order to see the effectiveness of affine scaling, we considered the patterns with neural network predictions error less than a threshold value. We arbitrarily fix the threshold at 0.15. The pattern and the prediction histograms and scatter plots are shown in Figures 8.4, 8.5 respectively In Figure 8.4, the first column plots are pattern histograms, and second column plots are prediction histograms. Plots in the first row are for neural network response. The pattern mean is 0.67 and the range is 0.26 to 0.95. The corresponding neural network prediction mean is 0 67 and the range is 0.40 to 0.81. Second row plots are those after affine scaling. The affine corrected prediction mean is 0.67. The prediction ranges from 0.31 to 0.86. Figure 8.5 shows prediction versus pattern in the first column and error versus pattern scatter plots in the second column for neural network prediction and for neural network prediction after affine scaling. The correlation between prediction and pattern is 0.816. After affine scaling histogram correction, the correlation between

prediction and pattern becomes 0.852, which is higher than the original neural network prediction.

This is evident in other four runs shown in Table 8.2 and 8.3. Affine scaling is effective to improve the neural network prediction performance for patterns with less than the threshold.

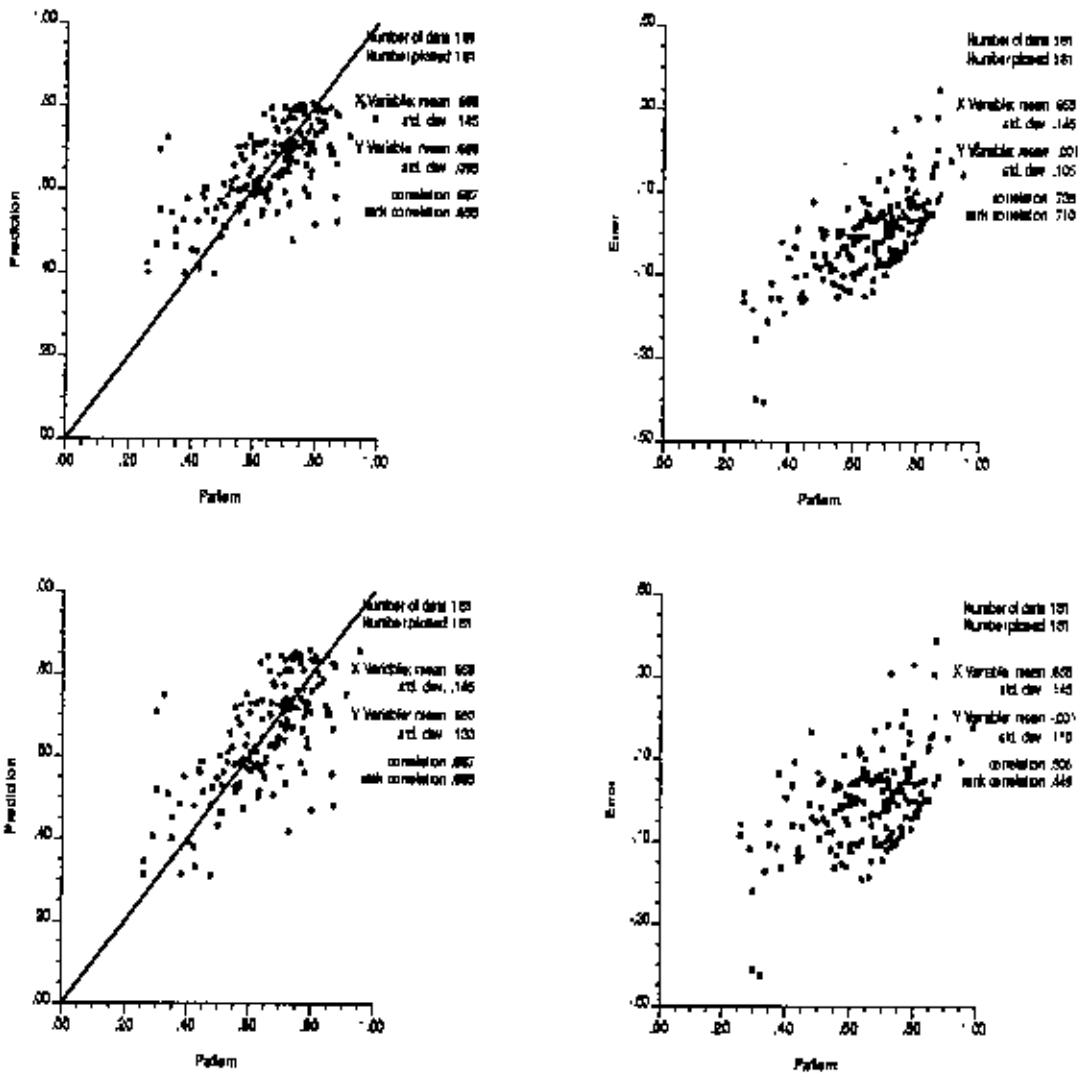

Figure 8.3: Artificial neural network prediction and affine scaling corrected neural network prediction.

Table 8.1: Neural network and affine scaling corrected prediction statistics.

Run	Mean			Range			Correlation Coefficient	
	Pattern	Neural Network Prediction	Affine Scaling Prediction	Pattern	Neural Network Prediction	Affine Scaling Prediction	Neural Network Prediction	Affine Scaling Prediction
1	0.658	0.659	0.660	.26 - .99	.416 – .780	.315 – .801	0.64	0 64
2	0 658	0.658	0.657	.26 - 99	.397 – .827	.318 – .878	0.683	0.683
3	0.658	0.660	0.660	.26 - .99	.402 – .825	.322 – .878	0.67	0 67
4	0.658	0.659	0 659	.26 - .99	.455 – .802	.364 – .867	0 648	0 648
5	0 658	0.659	0.660	.26 - 99	.398 – .811	312 – 811	0.687	0 687

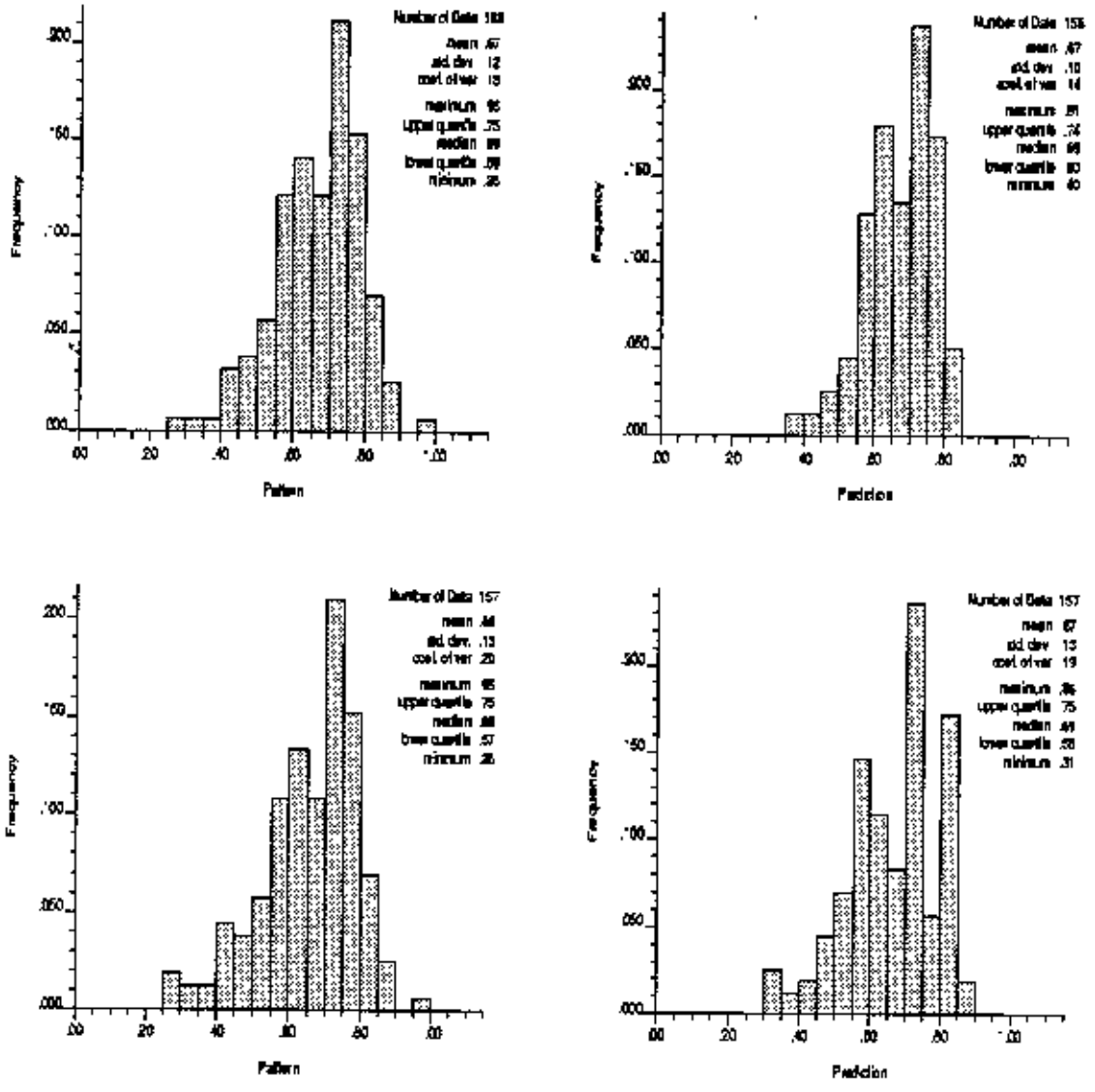

Figure 8.4: Neural network prediction and affine scaling corrected neural network prediction histograms.

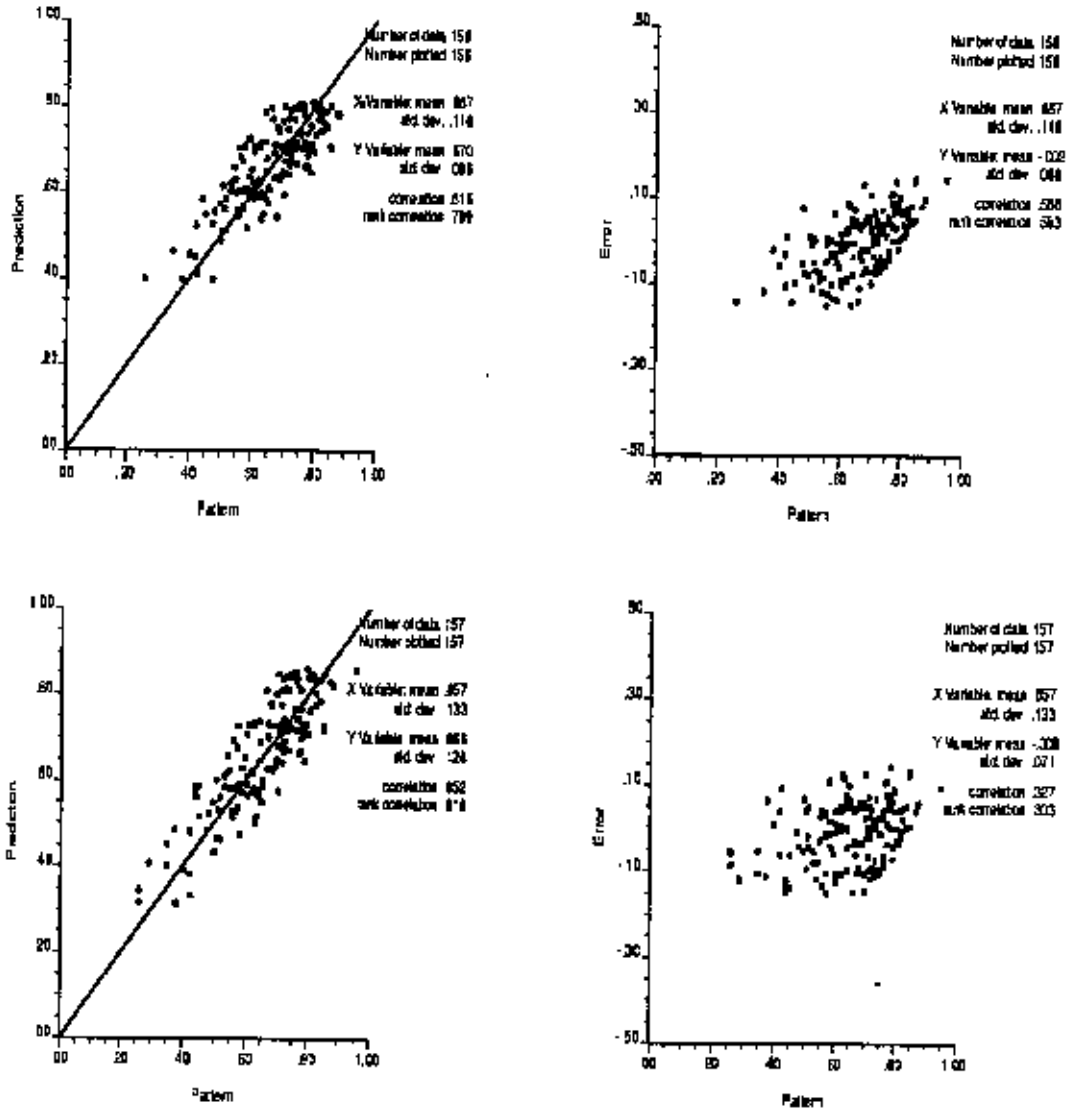

Figure 8.5: Neural network prediction and affine corrected neural network prediction for patterns with less than threshold.

Table 8.2: Neural network predictions with error less than the threshold.

Run	Mean		Range		Correlation Coefficient
	Pattern	Neural Network Prediction	Pattern	Neural Network Prediction	Neural Network Prediction
1	0 67	0.67	0.35 - 0.88	0.42 – 0.78	0.793
2	0.67	0.67	0.35 - 0.88	0.397 – 0.827	0.793
3	0.67	0.67	0.35 - 0.95	0 402 – 0.825	0.799
4	0.66	0.67	0 35 - 0.95	0.455 – 0.802	0.790
5	0.67	0.67	0.26 - 0.95	0.398 – 0.811	0.816

Table 8.3: Neural network predictions after affine scaling with error less than threshold.

Run	Mean		Range		Correlation Cofficient
	Pattern	Affine Scaling Prediction	Pattern	Affine Scaling Prediction	Affine Scaling Prediction
1	0.66	0.67	0.25 - 0.88	0.32 – 0.83	0.822
2	0.66	0.67	0.25 - 0.88	0.32 – 0.83	0.827
3	0.66	0.67	0.25 - 0.95	0.32 – 0.88	0.833
4	0.66	0.66	0.25 - 0.95	0.36 – 0.87	0.84
5	0.67	0.67	0.26 - 0.95	0.31 – 0.86	0.852

In the following section, we investigate the source of errors particularly of the patterns with prediction error greater than the threshold value.

8.3 Sources of High Prediction Errors

We explore the sources of high prediction errors. The objective was to determine the presence of some trend in the parameters that yield high errors. The parameters examined are the maximum permeability, porosity and facies. Figures 8.6, 8 7 and 8.8 show maximum permeability, porosity and facies histograms respectively for patterns with greater than the threshold value of prediction error. The threshold value is fixed at 0.15. The mean and standard deviation of maximum permeability is 119.32 and 259.2 and the range is 0.04 to 1320. Figure 8 6 shows that extreme values of permeability are responsible for higher prediction error. The porosity (see Figure 8.7) mean and standard deviation are 0.06 and 0.03 respectively and the range is 0.02 to 0.13. Figure 8.8 shows that Facies 1 is the major facies with prediction errors greater than the threshold. In fact, amount of information available of Facies 1 is greater those for other facies.

8.4 Conclusion

Some of the conclusions derived from the error analysis exercise are given below.

- Affine scaling of neural network prediction histogram appears to be an effective correction scheme for patterns with less than some threshold value.
- Prediction error is high for extreme values of permeability.
- Porosity and facies of the high prediction error patterns do not reveal any obvious trend.

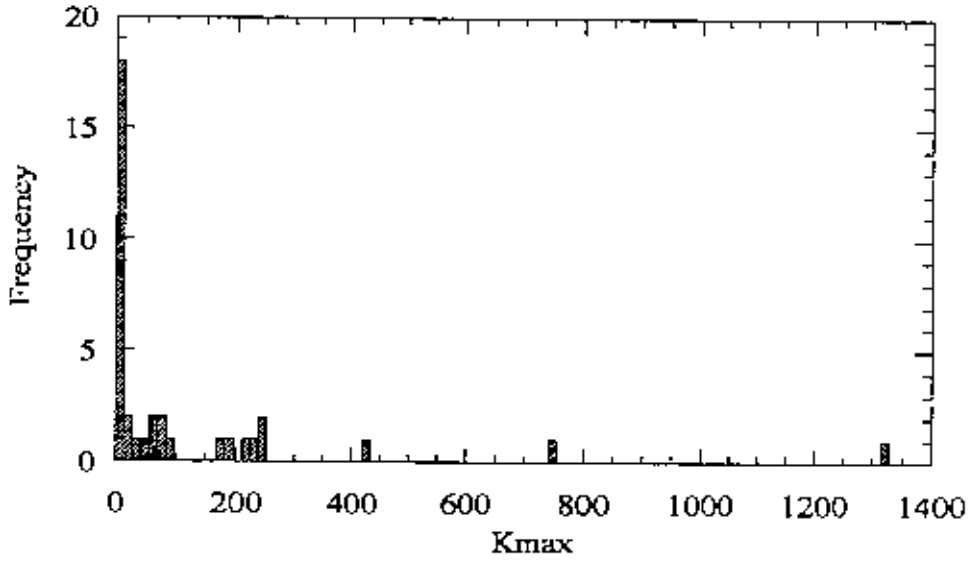

Figure 8.6: Maximum permeability histogram for patterns with prediction error greater than threshold.

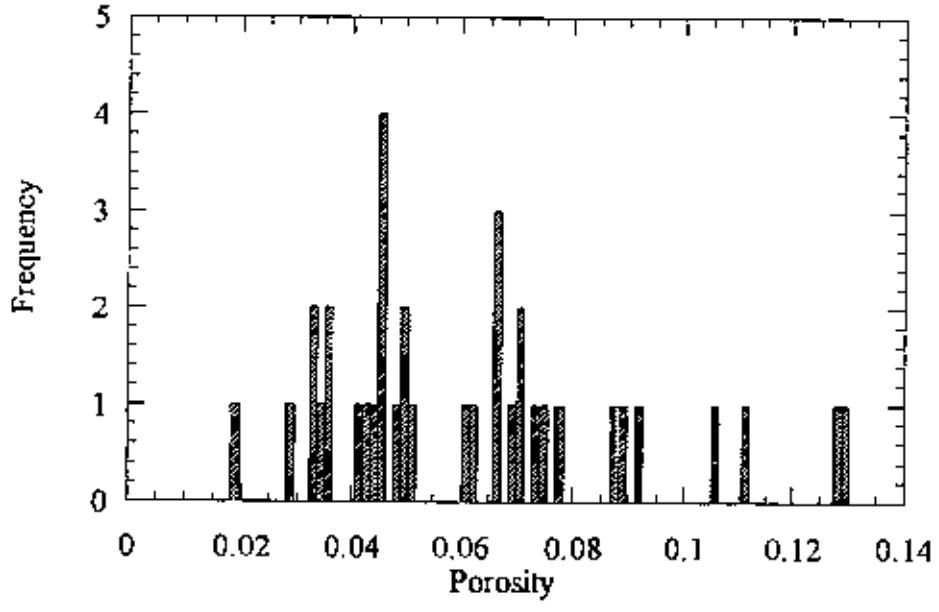

Figure 8.7: Porosity histogram for patterns with prediction error greater than threshold.

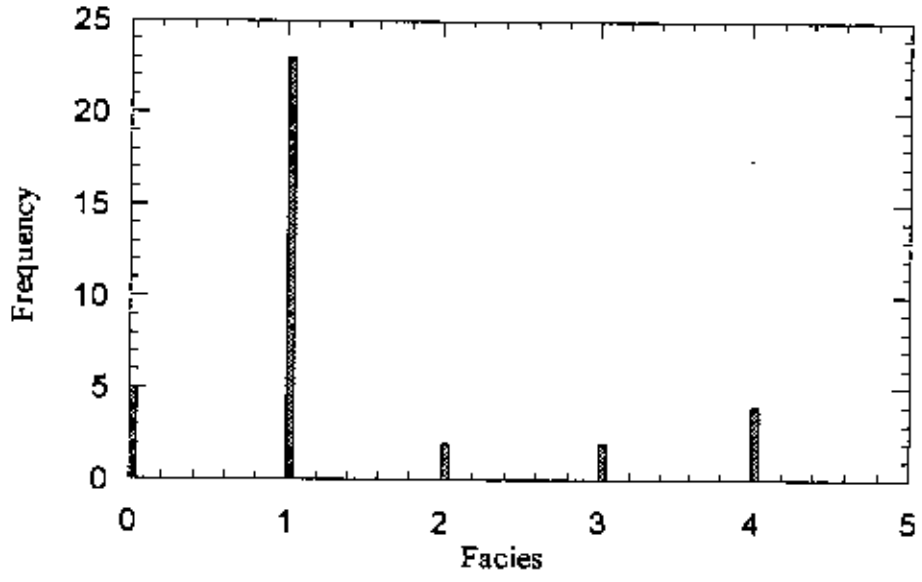

Figure 8.8 Facies histogram for patterns with prediction error greater than threshold.

CHAPTER 9

COMPARATIVE STUDY FOR PERMEABILITY DETERMINATION

This chapter compares the performance between developed neural network model and conventional methods used for permeability estimation. The conventional methods such as Carman-Kozeny equation and multiple linear regression are generally employed to estimate permeability. The performances of these methods have been compared with the neural network model responses to investigate their effectiveness in permeability determination.

This chapter comprises the following sections. Section 9.1 deals with Carman-Kozeny equation and prediction of permeability. In section 9.2, multiple linear regression model has been developed to predict permeability. Section 9.3 illustrates the comparison between these conventional and neural network models.

9.1 Permeability Determination Using Carman-Kozeny Equation

9.1.1 Basics of Carman-Kozeny Equatoin

The equation relating to measurable rock properties with permeability was first proposed by Kozeny. It was modified by Carman. Carman-Kozeny equation is used to provide framework for permeability estimation. Carman-Kozeny equation relates porosity, specific surface area, hydraulic radius and tortuosity According to Poiseuille's equation, flow through a conduit or capillary can be written as

$$Q = \frac{\pi r^4 \Delta P}{8 \mu L} \tag{9.1}$$

where r is the radius of the conduit, Q is the volumatric flow rate. From Darcy's equation, one can write

$$Q = \frac{k \pi r^2 \Delta P}{\mu L} \tag{9.2}$$

From Equation (9.1) and (9.2) one can write

$$k = \frac{r^2}{8} \tag{9.3}$$

Equation (9.3) should be corrected for porosity and tortuosity to give the interstitial velocity. So, the corrected permeability is

$$k = \frac{\phi r^2}{8\tau} \tag{9.4}$$

In porous rock medium, grain size is not constant. To overcome this problem, r is substituted by hydraulic radius. Hydraulic radius is defined by

$$r_h = \frac{\text{volume open to flow}}{\text{wetted surface area}}$$

For cylindrical capillary,

$$r_h = \frac{\pi r^2}{2\pi r} = \frac{r}{2}$$

$$r = 2r_h \tag{9.5}$$

For granular or porous medium specific area can be defined as

$$a_v = \frac{\text{wetted surface area}}{\text{solid volume}}$$

Specific surface area is an intrinsic property for a porous medium. It can be estimated using picnometer, porosity estimates, and grain size analysis. From the definitions of hydraulic radius and specific surface area, hydraulic radius can be derived as

$$r_h = \frac{\phi}{a_v(1-\phi)}$$

Using Equation (9.5)

$$r = 2r_h = \frac{2\phi}{a_v(1-\phi)} \tag{9.6}$$

From Equation (9.2) and (9.6)

$$k = \frac{\phi^3}{2\tau(1-\phi)^2 a_v^2} \tag{9.7}$$

Equation (9.7) is referred as Carman-Kozeny equation.

For porous medium with uniform sphere

$$a_v = \frac{\text{area of spheres}}{\text{volume of spheres}} = \frac{\pi D_p^2}{\frac{\pi}{6} D_p^3} = \frac{6}{D_p}$$

So, for a uniform porous medium, permeability can be estimated using Equation (9.7) as

$$k = \frac{D_p^2 \phi^3}{72\tau(1-\phi)^2} \tag{9.8}$$

From Equation (9.8), it can be inferred that tortuosity has weak influence on permeability due to its low power and it does not vary much. On the other hand, permeability increases dramatically with particle size and porosity. In practical cases, particle size is more uncertain than porosity. So, particle size is the main controlling factor in permeability estimation, not porosity.

Particle size and tortuosity data are not found in the field data, collected to perform the present research work. To overcome this problem the Kozeny constant has been considered in lumped form and it is

$$k = \frac{D_p^2 \phi^3}{72\tau(1-\phi)^2}$$

$$k = C \frac{\phi^3}{(1-\phi)^2} \tag{9.9}$$

where C is the Kozeny constant in lumped form. Equation (9.9) has been used to determine the Kozeny constant in lumped form (in field units) using field data for permeability and porosity.

9.1.2 Selection of Kozeny Constant

Figure 9.1 shows histogram for Kozeny constant (calculated using Equation (9.9) from field data, in field unit) in lumped form. Mean and standard deviation are found 200823.00 and 601084.80 respectively while the range is found to be 127.00 to 4930000.00. This histogram has been plotted to select Kozeny constants at different peaks. The Kozeny onstants are 1500, 10000, 30000, 100000, 400000, 200000. Using these constants, maximum permeability (K_{max}) has been determined.

9.1.3 Variation of Porosity and Permeability with Kozeny Constant

Figure 9 2 shows Kozeny constant (in lumped form) versus porosity scatter plot. Correlation is found to be –0.30. So, the Kozeny constant decreaes with porosity. Figure 9.3 shows maximum permeability (calculated) versus porosity scatter plots for the selected Kozeny constants. These plot shows that maximum permeability (K_{max}) increases slowly for low value of porosity, but permeability increases dramatically for higher values of porosity. Figure 9.3 shows that rate of permeability change is high for high values of the Kozeny constant.

In neural network model, the pattern (measured maximum permeability) and prediction (model output i.e., calculated maximum permeability) are in normalized form. To compare the performance of empirical model and neural network model, the calculated K_{max} from Carman-Kozeny equation and measured K_{max} are also tmasformed into normalized condition. To consider the entire population, the maximum and minimum values of log K_{max} were selected 4.50 and –3.00 respectively.

Figure 9.4 shows maximum permeability (measured) histogram. Mean and standard deviation are 0.54 and 0.12 respectively. The vlaue ranges from 0.21 to 0.82.

9.1.4 Determination of Permeability Using Carman-Kozeny Equaiton

9.1.4.1 Permeability Determination for Kozeny Constant 1500

Figure 9.5 shows histogram of calculated maximum permeability (K_{max}) Mean and standard deviation are 0.36 and 0.10. The difference for mean between calculated K_{max} and measured K_{max} is remarkable. So, the results from Carman-Kozeny equation for Kozeny constant 1500 are highly biased. The range of claculated K_{max} is from 0.11 to 0.56 which is narrow with respect to the measured K_{max} range. Figure 9.6 is calculated K_{max} versus measured K_{max} scatter plot with correlation 0.48 and RMSE 0.21.

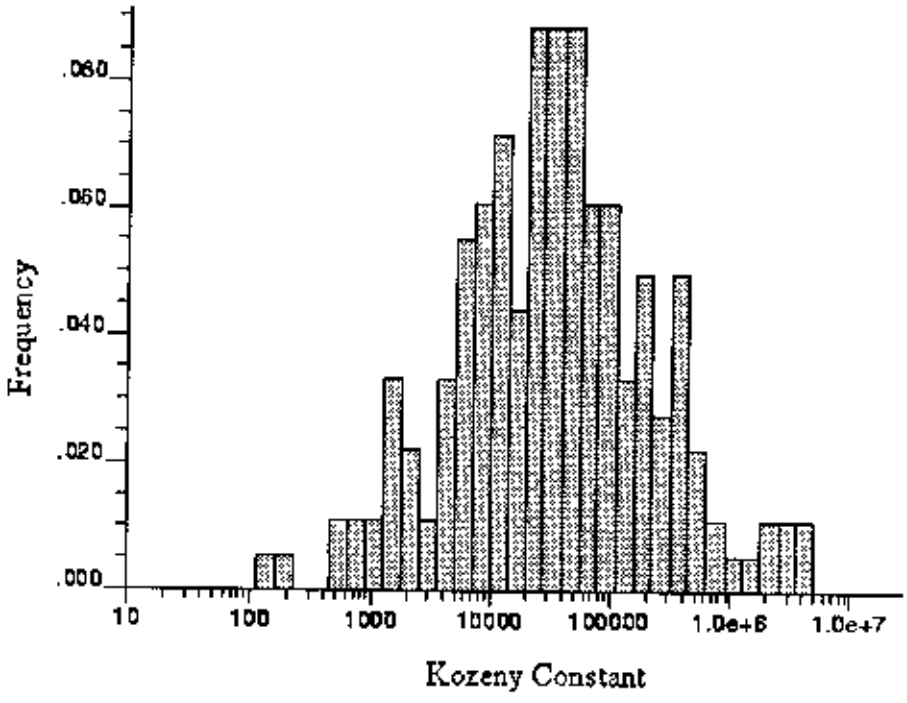

Figure9.1: Kozeny constant histogram.

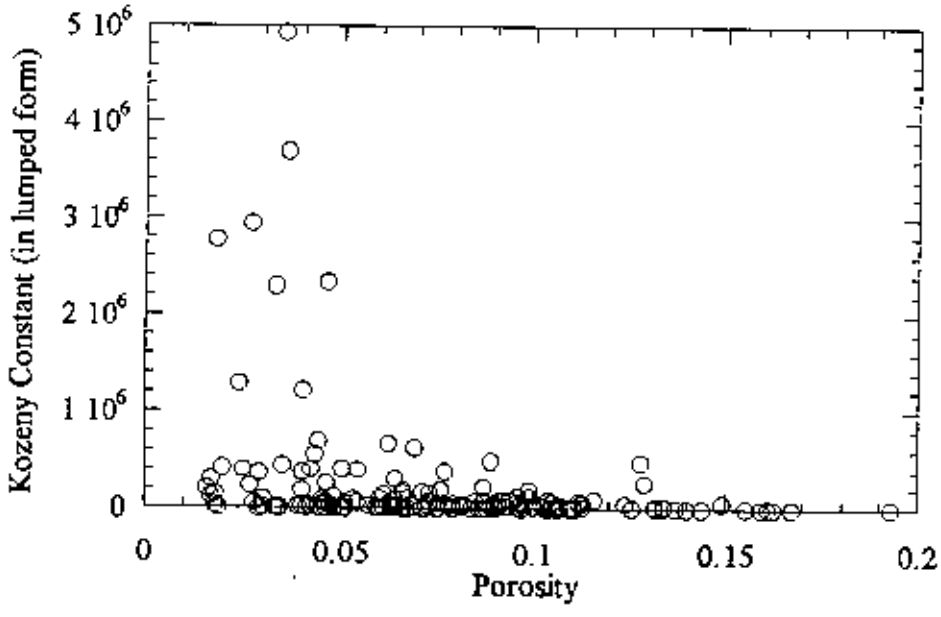

Figure 9.2: Kozeny constant versus porosity scattr plot.

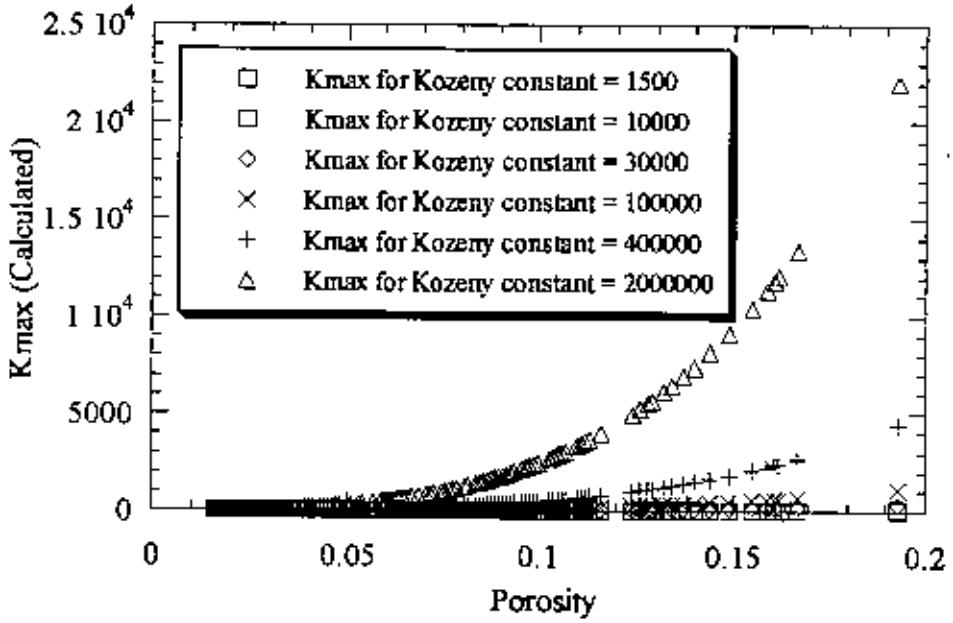

Figure 9.3: Maximum permeability versus porosity.

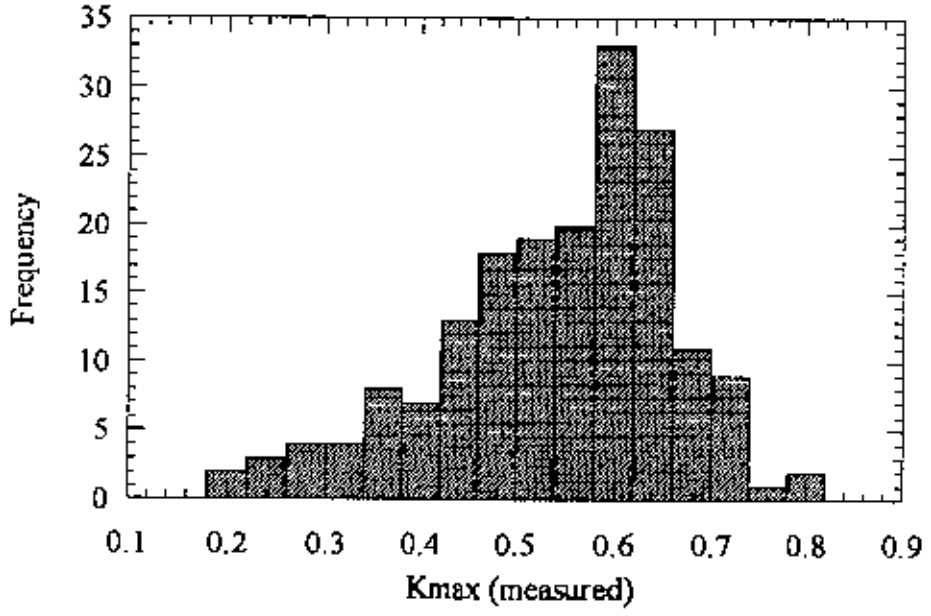

Figure 9.4: Maximum permeablity histogram.

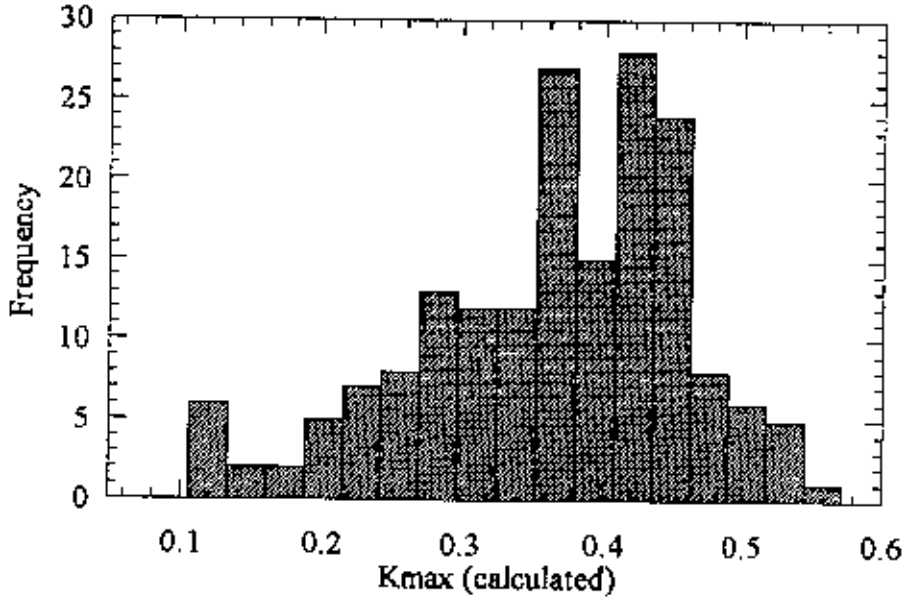

Figure 9.5: Maximum permeability histogram for Kozeny constant 1500.

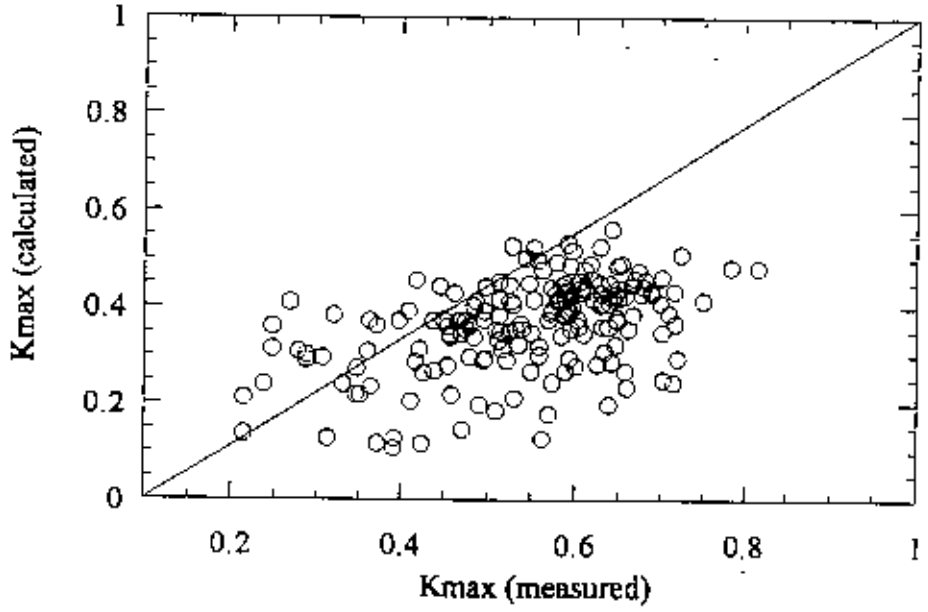

Figure 9.6: Calculated maximum permeability (for Kozeny constant 1500) versus measured maximum permeability.

9.1.4.2 Permeability Determination for Kozeny Constant 10000

Figure 9.7 shows histogram of calculated K_{max}. The mean and standard deviation of calculated K_{max} are 0.474 and 0.10. The calculated K_{max} ranges from 0.28 to 0.67 The calculated results are still biased and narrow ranged. Figure 9.8 is calculated K_{max} versus measured K_{max} scatter plot. Correlation and RMSE are 0.478 and 0 13. The reduction of RMSE is clearly evident.

9.1.4.3 Permeability Determination for Kozeny Constant 30000

Figure 9.9 is the histogram of calculated K_{max}. The mean and standard deviation are 0.54 and 0.10 respectively. The calculated K_{max} ranges from 0.28 to 0 74. So, the calculated results are relatively less biased and the range is moderately good. Figure 9.10 shows calculated K_{max} versus measured K_{max} scatter plot and the correlation is 0.48. RMSE is 0.11, farther reduction of RMSE.

9.1.4.4 Permeability Determination for Kozeny 100000

Maximum permeability histogram for Kozeny constant 100000 has been shown in Figure 9.11. It has a mean of 0.61 and a standard deviation of 0.10.The value ranges from 0 30 to 0.81. So, the calculated results show bias but the range is good. Figure 9 12 is the calculated K_{max} versus measured K_{max} scatter plot. The correlation is 0.48 and RMSE is 0.13. RMSE is more than the RMSE for Kozeny constant 30000.

9.1.4.5 Permeability Determination for Kozeny Constant 400000

Figure 9 13 is the histogram for calculated K_{max} for Kozeny constant 400,000 with mean and standard deviation 0.69 and 0.10 respectively. The minimum and maximum values are 0.43 to 0.89 respectively. Again the calculated result shows bias. The range has exceeded the measured permeability value. Figure 9.14 is the scatter plot between calculated K_{max} and measured K_{max} with correlation 0.48 and RMSE 0.18.

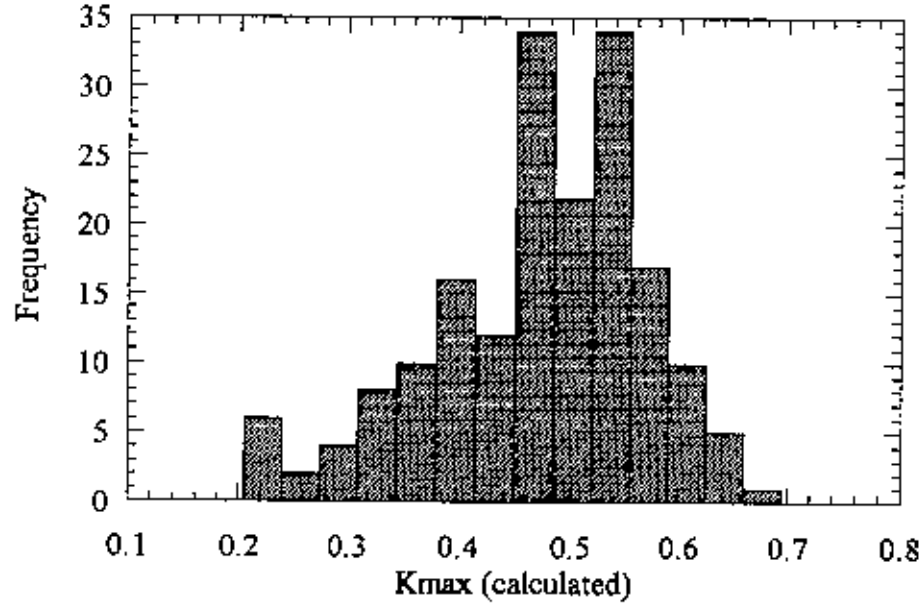

Figure 9.7: Maximum permeability hsitgram for Kozeny constant 10000

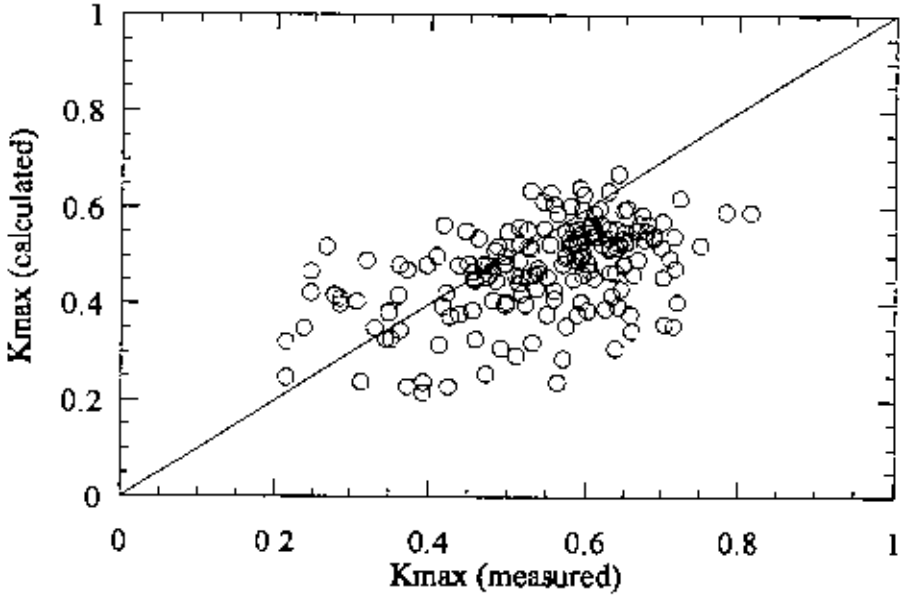

Figure 9.8: Calculated maximum permeability (for Kozeny constant 10000) versus measured maximum permeability.

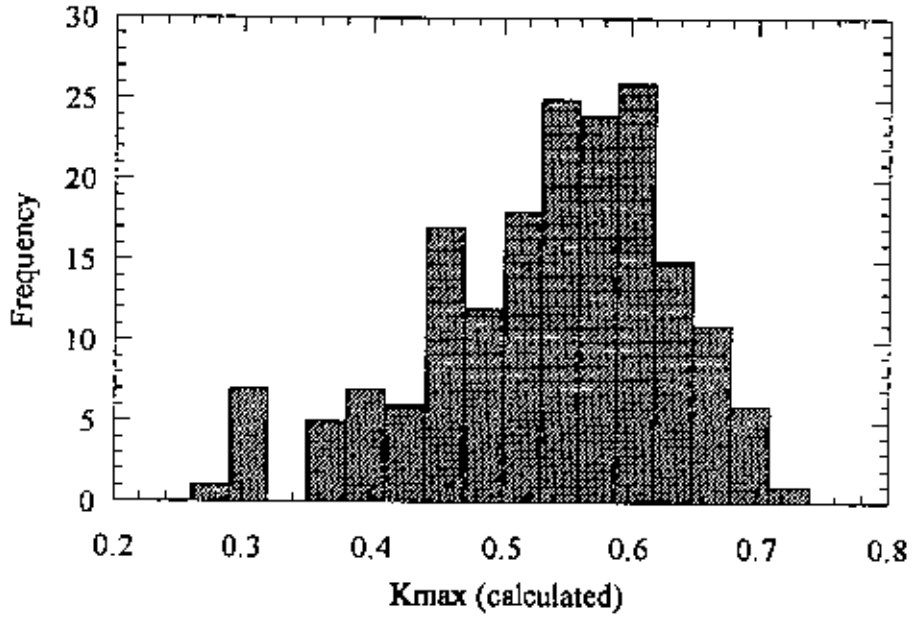

Figure 9.9: Maximum permeability histogram for Kozeny constant 30000.

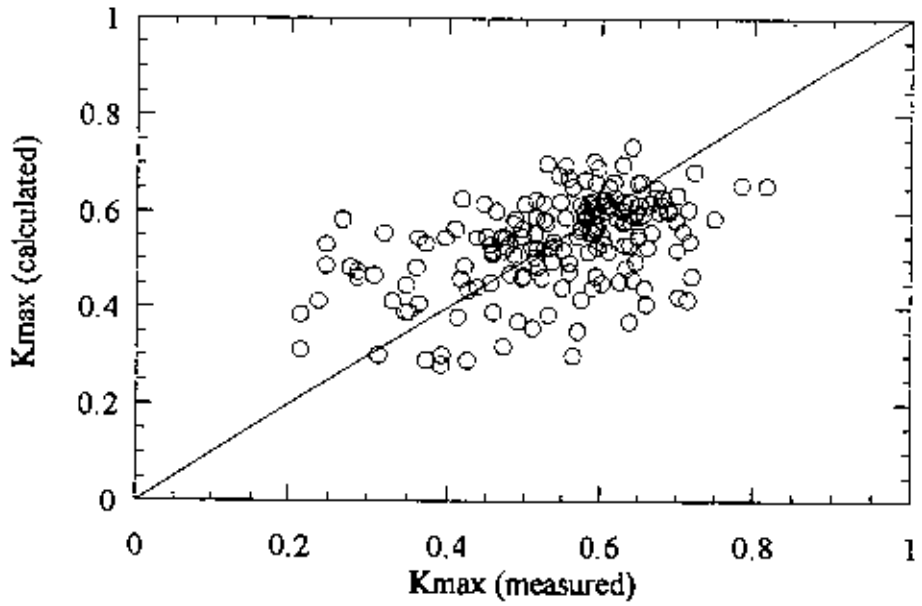

Figure 9.10: Calculated maximum permeability (for Kozeny constant 30000) versus measured maximum permeability.

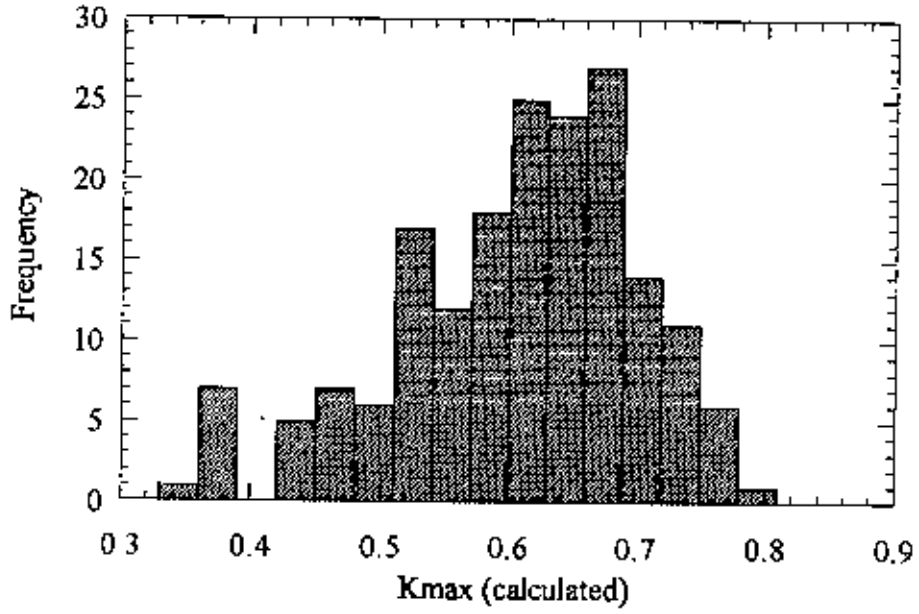

Figure 9.11: Maximum permeability hisotgram for Kozeny constant 100000

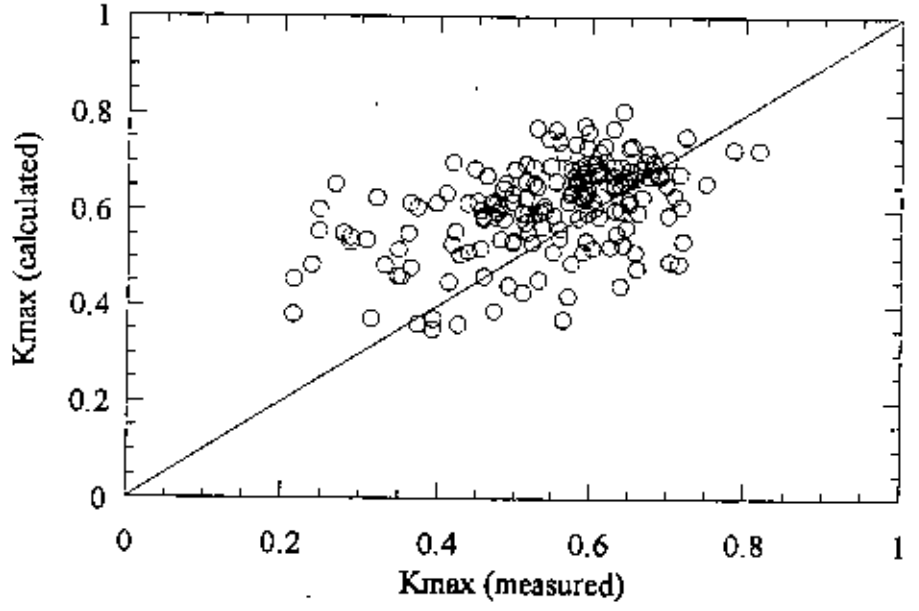

Figure 9.12: Calculated maximum permeability (for Kozeny constant 100000) versus measured maximum permeability.

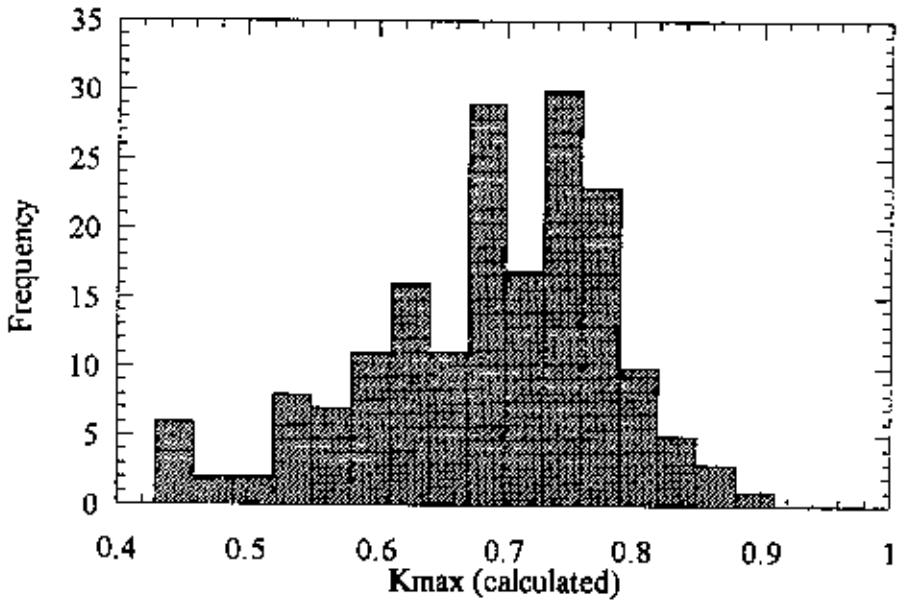

Figure 9.13: Maximum permeability histogram for Kozeny constant 400000.

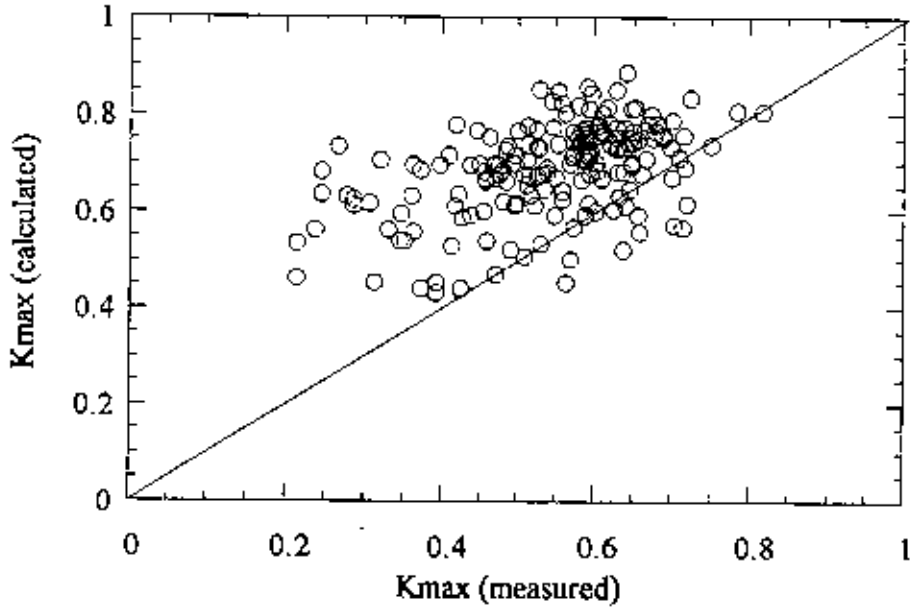

Figure 9.14: Calculated maximum permeability (for Kozeny constant 400000) versus measured maximum permeability.

9.1.4.6 Permeability Determination for 2000000

Figure 9.15 shows maximum permeability (K_{max}) histogram. It has mean and standard deviation 0.78 and 0.10 while the range is found to be 0.52 to 0.98. The calculated result is unable to generate mean i.e. biased and the range is very high with respect to the measured permeability range. Figure 9.16 indicates calculated K_{max} versus measured K_{max} scatter plot. Correlation is 0.48. RMSE is 0 26.

From the above results, it can be inferred that Carman-Kozeny equation performs better for Kozeny constant 30000. So, for comparison with other methods, the results for Kozeny constant 30000 has been used in Section 9.3.

The following Section 9.2 gives a statistical model, multiple linear regression for maximum permeability prediction.

9.2 Multiple Linear Regression Model

A useful extension of least square method is the multiple linear regression method where dependent variable, Y is linear function of two or more independent variables $(X_1, X_2, X_3, \ldots)$. The general form of multiple linear regression equation is

$$Y = B_0 + B_1 X_1 + B_2 X_2 + B_3 X_3 + e \qquad (9.10)$$

Where B_0, B_1, B_2, B_3 are the coefficients and e is the error term. The best values of coefficients are selected in such a manner so that the error term reduces to zero Under this study multiple linear regression model for permeability prediction has been developed based on least square method using well locations, relative stratigraphic depth, facies, porosity and thickness as independent variables. Normalization of the data has been done according to Table 6.1 The developed linear regression model is

$$K_{max} = 0.343172 - 0.32656X + 0.202984Y - 0.12247Z_{strata} - 0.09837(Facies) + 0.411706\varphi + 0.489087(Thickness) \qquad (9.11)$$

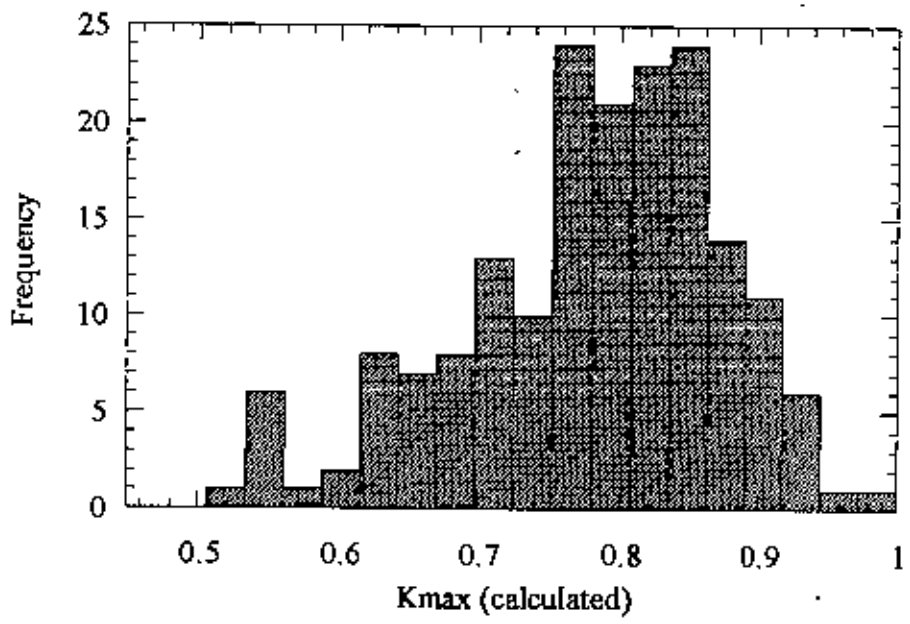

Figure 9.15: Maximum permeability histogram for Kozeny constant 2000000

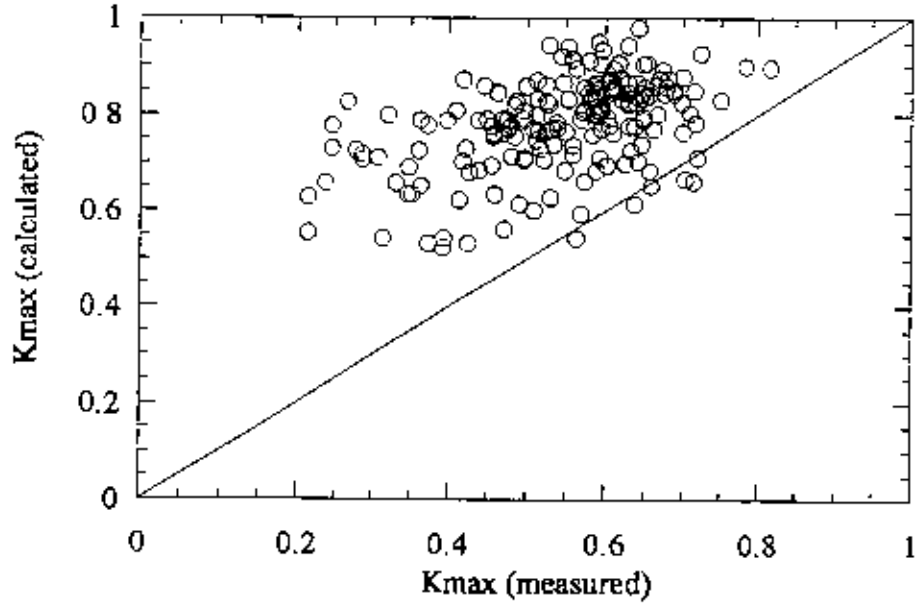

Figure 9.16: Calculated maximum permeability (for Kozeny constant 2000000) versus measured maximum permeability.

9.2.1 Validation of the Regression Model

Figure 9.17 is the histogram for maximum permeability (K_{max}) (measured). The measured K_{max} value ranges from 0.26 to 0.99 while the mean and standard deviation are 0.658 and 0.145 respectively. Figure 9.18 shows predicted maximum permeability histogram The predicted K_{max} value ranges from 0.504 to 0.845. The mean and standard deviation are 0.658 and 0.082. The developed regression model is capable to reproduce the mean. So, the developed regression model is more or less bias free but prediction range is very narrow.

Figure 9.19 shows predicted maximum permeability versus measured permeability scatter plot. The correlation and RMSE are 0.57 and 0.12 respectively.

9.3 Results and Discussions

The prediction statistics of neural network model, empirical Carman-Kozeny equation and multiple linear regression model have been shown in Table 9.1, 9 2 and 9.3 respectively. It is remarkable that neural network model and regression model are quite capable to reproduce the mean. So, their predictons are bias free. On the otherhand, the Carman-Kozeny equation is slightly biased. Prediction ranges for neural network and regession model are narrow. Comparatively, regression model prediction range is more narrow than that of neural network model. Prediction range for Carman-Kozeny equation is relatively better than the other two methods. The RMSE for neural network model (0.105) is less than that of Carman-Kozeny equation (0.11) and multiple linear regression model (0.12). On the other hand, neural network shows higher correlation (0 640 to 0.687) between predicted and measured maximum permeability than Carman-Kozeny Equatoin (0.48) and multiple linear regression model (0.57).

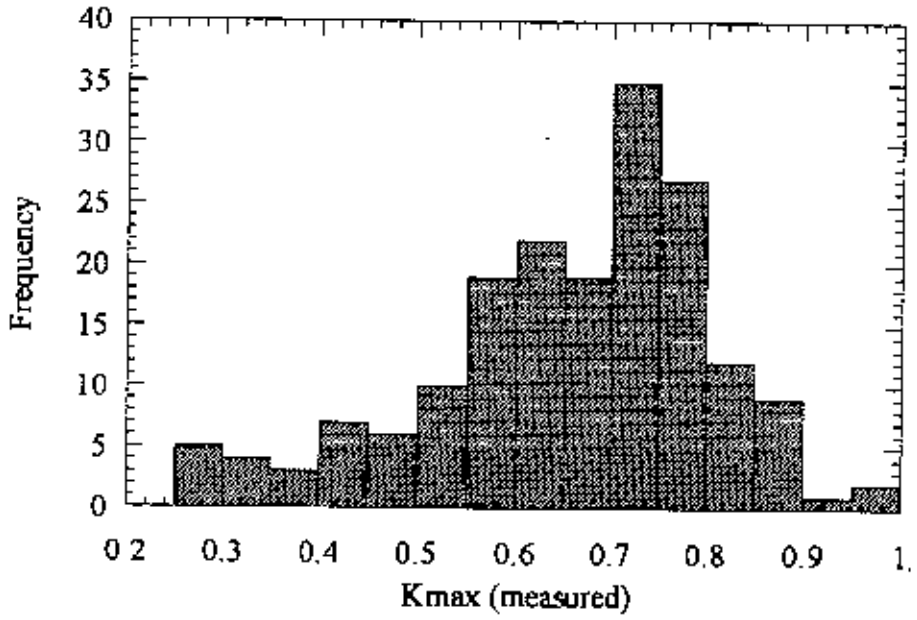

Figure 9.17: Maximum permeability histogram.

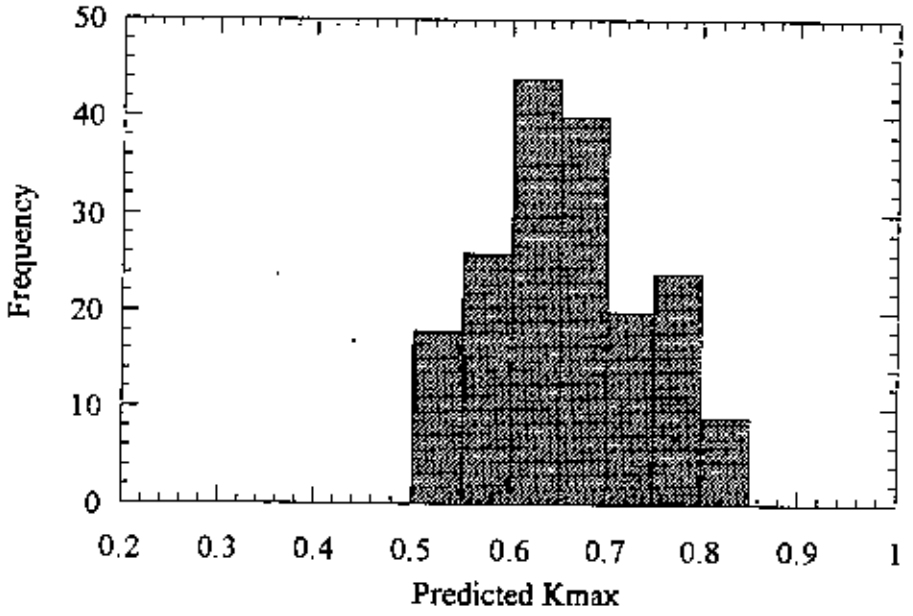

Figure 9.18: Multiple linear regression model predicted maximum permeability histogram.

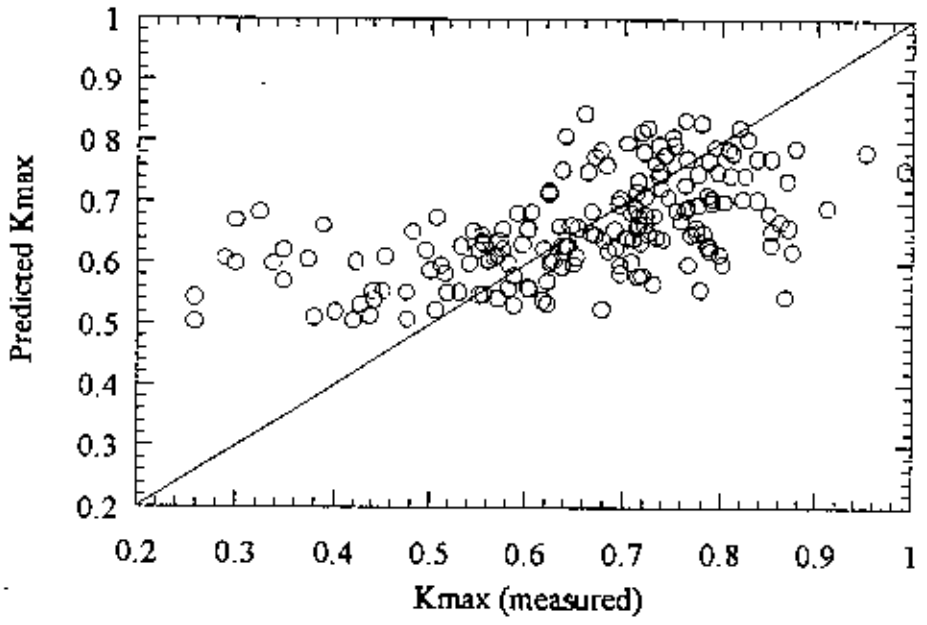

Figure 9.19: Multiple linear regression model predicted maximum permeability versus measured maximum permeability.

Table 9.1: Neural network prediction statistics.

Run	Mean		Range		RMSE	Correlation
	Mean K_{max} (measured)	Neural Network Prediction	K_{max} (measured)	Neural Network Prediction		
1	0.658	0.659	0.26 – 0.99	0.416 – 0.780	0.105	0.640
2	0.658	0.658	0.26 -0.99	0.397 – 0.827	0.105	0.683
3	0.658	0.660	0.26 – 0.99	0.402 – 0.825	0.105	0.670
4	0.658	0.659	0.26 – 0.99	0.455 – 0.802	0.105	0.648
5	0.658	0.659	0.26 -0.99	0.398 – 0.811	0.105	0.687

Table 9.2: Statistics of results from Carman-Kozeny equation for Kozeny constant 30000.

Mean		Range		RMSE	Correlation
Mean K_{max} (measured)	Calculated K_{max}	K_{max} (measured)	Calculated K_{max}		
0.543	0.538	0.213 - 0.816	0.28 - 0.736	0.11	0.48

Table 9.3: Multiple linear regression model prediction statistics.

Mean		Range		RMSE	Correlation
Mean K_{max} (measured)	Predicted K_{max}	K_{max} (measured)	Predicted K_{max}		
0.658	0.658	0.26 - 0.99	0 50 - 0.85	0.12	0.57

9.4 Conclusion

Neural network model predicts that porosity and relative strata are the most important and thickness is the least important input parameters for permeability prediction. This prediction is relevant with the exploratory data analysis and practical experiences. In contrast, multiple linear regression model shows that the thickness is the most important input parameter for permeability prediction. So, linear regressive model is unable to predict the important input parameters.

The conventional methods offer no prediction space. Neural network model can produce prediction space (Figure 7.17 to 7.24). So, it can reduce the artifact effect in the prediction space.

Neural network model is more efficient to capture the prediction space than the conventional methods Carman-Kozeny equation and multiple linear regression model. The RMSE of neural network model can be reduced as well as the correlation can be improved more by increasing the number of hidden layers.

CHAPTER 10

CONCLUSIONS AND RECOMMENDATIONS

In this study, a neural network model is developed for permeability prediction at a "uncored" well. The inputs to the neural network are porosity values, lithofacies identifiers, spatial coordinates, and thickness of the samples. A number of analytical studies are performed in this work. Some of the salient conclusions derived from the study are enumerated below. More detailed analyses are discussed in the previous chapters.

10.1 Conclusions

- Exploratory data analysis reveals the correlations between variables of interest are very poor. Under usual circumstances, this leads to poor models based on first principle or simple regressional methods. The use of highly nonlinear sophisticated tool such as neural networks was deemed to develop a model.

- Both sigmoidal function and hyperbolic tangent functions had been used for the activation function in the neural network model. It was found that sigmoidal function performs better of the two.

- Optimal parameters and their corresponding values were selected The optimum parameters were input to hidden node bias (0.50), hidden to output node bias (0 05), input to hidden node learning rate (2.0), hidden to output node learning rate (0.0005), number of hidden nodes (4).

- The relative influence of the input parameters was investigated. It is found that porosity and relative strata are the significant parameters amongst the others input parameters available for permeability prediction. On the other hand, contribution of multiple linear regression model about the importance of input parameter is misleading.

- Parameterization of the input space for the problem is not sufficient With available data, developed model shows high prediction error and less reliability. Of course, a major reason for high prediction error is the poor correlation between parameters.

- A number of uncertainty analyses were performed for permeability prediction using this study. This kind of studies is recommended as routine affair for any engineering investigation. Uncertainty characterization leads to better understanding of the problem being investigated.

- Prediction error is high for extreme values of permeability. Affine scaling has been applied to minimize prediction error of the developed model. Affine scaling is effective to minimize prediction error for patterns with values lower than some threshold

- One important thing over the conventional methods is that neural network model is error tolerant and it can minimize the artifacts in the prediction space.

- It is possible to predict permeability with more efficiently than conventional methods using neural network model.

10.2 Recommendations

In this study, a back propagation neural network model has been developed to predict permeability of an "uncored" well using mainly core data. Some of the recommendations realized from the work could be the following.

- The computer program developed in this work has a much wider applicability. The same code can be applied to some suitable petroleum engineering problems (mentioned in Chapter 3) with little modification.

- Number of hidden layer can be increased for better training as well as to get better responses of the developed neural network model.

- Additional information regarding grain size and distribution, angularity, connectivity, geomechanical properties, diagenesis effect etc. would have enhanced the model.

- For nonstationary systems, the developed neural network model should be modified to incorporate approaches similar to temporal Back Propagation Neural Network Model.

- More effective methods should be applied for error correction. Further investigation is required in this regard.

- This is the first documented application of neural network in petroleum engineering studies in Bangladesh. Techniques like neural networks discussed here and more sophisticated ones can be and should be applied extensively for the development of petroleum engineering in Bangladesh.

NOMENCLATURE

ALR = Learning rate between input and hidden layer

BLR = Learning rate between hidden and output layer

K_{max} = Maximum permeability

K_{vert} = Vertical permeability

RMSE = Root mean square error

Z_{strata} = Relative strata

a_v = Specific surface area in Equation (9.6)

b_j^n = Bias applied to neuron j

C = Kozeny constant in lumped form in Equation (9.8)

d^n = Desired output vector for time step n in Equation (4.10)

d_j^n = Desired response (output) for neuron j at time step n in Equation (4.15)

d_k^n = Desired response (output) for neuron k at time step n in Equation (4.29)

D_p = Particle diameter in Equation (9.8)

e_k^n = Refers to the error signal at the outpnt of neuron k for time step n time step in Equation (4.3)

e_j^n = Refers to the error signal at the output of neuron j for time step n in Equation (4.12)

f_i = Nonlinear function in Equation (4.1)

N = Total number of pattern in the training set in Equation (4.14)

n = Time step in Equation (4.3)

O_i = Output produced by neuron i in Equation (4.1)

ΔP = Pressure difference in Equation (9.1)

Q = Volumetric flow rate in Equation (9.1)

r = Radius of the conduit in Equation (9.1)

r_h = Hydraulic radius in Equation (9.5)

v_j^n = The induced local field (i.e , weighted sum of all synaptic inputs plus bias) of neuron j at iteration n in Equation (4.21)

w^* = Optimum weight vector in Equation (4.4)

w_{j0}^n = Synaptic weight for bias b_j^n at iteration n in Equation (4.15)

w_{ji}^n = Synaptic weight connecting the output of neuron i to the input of neuron j at iteration n in Equation (4.21)

Δw_{ji}^n = The correction applied to this weight at iteration n in Equation (4.23)

x^n = Input vector at iteration n Equation (4.10)

x_j^n = The j th element of the input vector (pattern) at iteration n in Equation (4.3)

x_{ij}^n = The input signal from neuron i to neuron j in Equation (4.1)

y_j^n = Function signal appearing at the output of neuron j at iteration n in Equation (4.12)

Greek Symbols

$\varphi_j(.)$ = The activation function associated with neuron j in Equation (4.16)

φ = Porosity in Equation (9.4)

η = The learning-rate parameter in Equation (4.3)

ε^n = Instantaneous value of the sum of squared errors in Equation (4.13)

$\varepsilon(w)$ = Cost function in Equation (4.4)

g = Gradient of cost function, $\varepsilon(w)$ in Equation (4.7)

ε_{av}^{n} = Average error energy in Equation (4.14)

Θ_i = Threshold for neuron i in Equation (4.2)

δ_j^n = Local gradient of neuron of neuron j at time n in Equation (4.24)

∂ = Partial

∇ = Differential operator

τ = Tortuosity in Equation (9.4)

Subscript

av = Average

i = Neuron i

j = Neuron j

k = Neuron k

max = Maximum

vert = Vertical

REFERENCES

Balan, B., Mohaghegh, S., Ameri, S. (1995): "State-Of-Art in Permeability Determination From Well Log Data: Parl-A Comparative Study, Model Development", SPE 30978, SPE Eastern Regional Conference and Exhibition, West Virginia, U.S.A , 17-21 September, 1995.

Bilgesu, H.I., Altmis, U. (1998): " A New Approach to Predict Bit Life Based on Tooth or Bearing Failures", SPE 51082, SPE Eastern Regional Meeting, Pittsburgh, PA,9-11 November 1998.

Bilgesu, H.I., Tetrick, L.T. (1997): " A New Approach for the Prediction of Rate of Penetration (ROP) Values", SPE39231, SPE Eastern Regional Meeting, Lexington, KY, 22-24 October, 1997.

Haykin, S. (2001): "Neural Networks", 2nd Edition, Pearson Education, Inc.

Huanng, Z , Shimeld, J. (1994). "Permeability Prediction Using a Computer Neural Network Application in the Venture Gas Field, Offshore Nova Scotia", Geological Survey of Canada, Atlantic, Nova Scotia, 1994.

Kartalopoulous S.V. (2002): Understanding Neural Networks and Fuzzy Logics, Eastern Economy Edition, 2nd Print, Prentice-Hall of India Private Limited, New Delhi, 2002.

McVey, Mohaghegh (1994): "Identification of Parameters Influencing the Response of Gas Storage Wells to Hydraulic Fracturing with the Aid of a Neural Network", SPE 29159, SPE Eastern Regional Conference & Exhibition, Charleston, WV, U.S.A., 8-10 November, 1994.

Mohaghegh, S. (1999): " Virtual Intelligence and Its Applications in Petroleum Engineering, 1. Artificial Neural Networks", West Virginia University, Distinguished Author Series, 1999.

Mohaghegh, S., Arefi (1994): "A Methodological approach For Reservoir Heterogeneity Characterization using Artificial Neural networks", SPE 28394, SPE Annual Technical Conference & Exhibition, New Orleans, LA, U.S.A., 25-28 September, 1994.

Mohaghegh, S., Arefi (1995): "Design and Development of an artificial Neural Network for Estimation of Formation Permeability", SPE 28237, The SPE Petroleum Computer Conference, Dallas, U.S.A., 31July-3 August, 1995.

Mohaghegh, S., Balan, B., Ameri, S. (1995b): " State-Of-The in Permeability Determination from Well Log Data: Part2-Verifiable, Accurate Permeability Predictions, the Touch-Stone of All Models", SPE 30979, SPE Eastern Regional Conference and Exhibition, West Virginia, U.S.A , 17-21 September, 1995.

Mohaghegh, S., Balan (1996): "A Hybrid, Neuro-Genetic Approach to Hybrid Fracture Treatment Design and Optimization", SPE 36602, SPE Annual Technical Conference & Exhibition, Denver, U.S.A, 6-9 October, 1996.

Mohaghegh, S.D., Gaskari (2001): " Identifying Best Practices in Hydraulic Fracturing Using Virtual Intelligence Techniques", SPE 72385, SPE Eastern Regional Meeting, Canton, Ohio, 17-19 October, 2001.

Mohaghegh, Gaskari (2002): " Identification of Successful Practices in Hydraulic Fracturing Using Intelligent Data Mining Tools; Application to the Codell Formation in the DJ-Basin", SPE 77597, SPE Annual Technical Conference and Exhibition, San Antonio, 29 September-2 October, 2002.

Mohaghegh, S.D., Goddard (2000): " Reservoir Characterization Through Synthetic Logs", SPE 65675, SPE Eastern Regional Meeting, Morgatown, West Virginia, 17-19 October, 2000.

Mohaghegh, S , Hefner, M.H. (1996): "Fracture Optimization eXpert (FOX)- How Computational Intelligence Helps the Bottom-Line in Gas Storage; a Case Study", SPE 37341, SPE Eastern Regional Conference, 23-8 October, 1996.

Mohaghegh, S.D., Hutchins, L.A. (2002): "Prudhoe Bay Oil Production Optimization Using Virtual intelligence Techniques, Stage One: Neural Model Building", SPE 77659, SPE Annual Technical Conference and Exhibition, San Antonio, Texas, 29 September-2 October, 2002.

Mohaghegh, S., McVey, D. (1995): "Predicting Well Stimulation Results in a Gas Field in the Absence of Reservoir Data, Using Neural Networks", SPE 31159, presented at SPE on a nonsolicited bases, 1995

Mohaghegh, S., Mohammad, K. (1999): " Performance Drivers in Resitmulation of Gas Storage Wells", SPE 57453, SPE Eastern Regional Meeting, Charleston, WV, 21-22 October, 1999.

Mohaghegh, S., Platon, V. (1998): " Candidate Selection for Stimulation of Gas Storage Wells Using Available Data With Neural Networks and Genetic Algorithms", SPE 51080, SPE Eastern Regional Meeting, Pittsburgh, PA, 9-11 November, 1998.

Mohaghegh, S., Popa, A. (1999a): " Reducing the Cost of Field-Scale Log Analysis Using Virtual Intelligence Techniques", SPE 57454, SPE Eastern Regional Meeting, Charleston, WV, 21-22 October,1999.

Mohaghegh, S., Popa, A. (1999b): " Intelligent System Can Design Optimum Fracturing Jobs", SPE 57433, SPE Eastern Regional Meeting, Charleston, WV, 21-22 October,1999.

Mohaghegh, S , Reeves, S. (2000): " Development of an Intelligent Systems Approach for Restimulation Candidate Selection", SPE 59767, SPE Gas Technology Symposium, Calgary, AB, 3-5 April, 2000.

Mohaghegh, S., Richardson, M. (1998): " Virtual Magnetic Imaging Logs: Generation of Synthetic MRI Logs from Conventional Well Logs", SPE 51075, SPE Eastern Regional Meeting, Pittsburgh, PA, 9-11 November, 1998.

Nikravesh, M., Kovsech, A.R. (1996): "Prediction of Formation Damage During Fluid Injection into Fractured, Low Permeability Reservoir via neural Network", SPE31103, SPE Formation Damage Symposium, Lafayete LA, 14-15 February, 1996.

Reeves, S.R., Bastian, P.A., Flumerfelt, R.W. (2000). "Benchmarking of Restimulation Candidate Selection Techniques in Layered, Tight Gas Sand Formations Using Reservoir Stimulation", SPE 63096, SPE Annual Technical Conference and Exhibition, Dallas, Texas, 1-4 October, 2000.

Reeves, S. R., Mohaghegh S.D. (2002): " Feasibility Assessment of a New Approach for Integrating Multiscale Data for High-Resolution Reservoir Characterization", SPE 77759, SPE Annual Technical Conference and Exhibition, San Antonio, Texas, 29 September-2 October, 2002.

Siripitayanannon, P., Hui-Chuan (2001): "A New Technique for Lithofacies Prediction: Back-Propagation Neural Network", Association for Computing Machinery, 2001.

Stundner, M., AL-Thuwani, J.S. (2001): " How Data- Driven Modeling Methods like Neural Networks can help to integrate different Types of Data into Reservoir Management", SPE 68163, SPE Middle East Oil Show, Bahrain, 17-20 March, 2001.

Subramanian, C., Manry, M.T. (1998): "Reservoir Inflow Forecasting Using Neural Networks", University of Texas at Arlington, 1998.

Ternyik, J., Bilgesu, H.I. (1995a): "Virtual Measurement in Pipes, Par1: Flowing Bottom Hole Pressure Under Multi-Phase Flow and Inclined Wellbore Conditions", SPE 30975, SPE Eastern Regional Conference & Exhibition, Morgaton, West Virginia, U.S.A., 18-20 September,1995.

Ternyik, J., Bilgesu, H.I. (1995b): "Virtual Measurement in Pipes, Part2: Liquid Holdup and Flow Pattern Correlations", SPE 30976, SPE Eastern Regional Conference & Exhibition, Morgaton, West Virginia, U.S.A., 18-20 September,1995.

Walls, J.D., Taner, M.T. (1999): "North Sea reservoir characterization using rock physics, seismic attributes, and neural networks; a case history", Houston, TX, SEG 1999 Expanded Abstracts.

White, A.C., Molnar, D. (1995): " The Application of ANN for Zone Identification in a Complex Reservoir", SPE 30977, SPE Eastern Regional Conference & Exhibition, Morgaton, West Virginia, U.S.A., 19-23 September,1995.

Zurada, J.M. (1996). Introduction to Artificial Neural Systems, 2nd Jaico Impression, Jaico Publishing House, India.