2017

Storage Capacity Estimation of Commercial Scale Injection and Storage of CO2 in the Jacksonburg-Stringtown Oil Field, West Virginia

Zhi Zhong

STORAGE CAPACITY ESTIMATION OF COMMERCIAL SCALE INJECTION AND STORAGE OF CO_2 IN THE JACKSONBURG-STRINGTOWN OIL FIELD, WEST VIRGINIA

Zhi Zhong

Dissertation submitted to the
Eberly College of Arts and Sciences
At West Virginia University
In partial fulfillment of the requirements
For the degree of

Doctor of Philosophy
In
Geology

Timothy R. Carr, Ph.D., Chair
Dengliang Gao, Ph.D.
Amy Weislogel, Ph.D.
Shikha Sharma, Ph.D.
Shahab Mohaghegh, Ph.D.

Department of Geology & Geography

Morgantown, West Virginia
2017

Keywords: 3D Geological Model, Artificial Neural Network, Support Vector Machine, Porosity and Permeability, Jacksonburg-Stringtown oil field

ABSTRACT

Geological capture, utilization and storage (CCUS) of carbon dioxide (CO_2) in depleted oil and gas reservoirs is one method to reduce greenhouse gas emissions with enhanced oil recovery (EOR) and extending the life of the field. Therefore CCUS coupled with EOR is considered to be an economic approach to demonstration of commercial-scale injection and storage of anthropogenic CO_2. Several critical issues should be taken into account prior to injecting large volumes of CO_2, such as storage capacity, project duration and long-term containment. Reservoir characterization and 3D geological modeling are the best way to estimate the theoretical CO_2 storage capacity in mature oil fields.

The Jacksonburg-Stringtown field, located in northwestern West Virginia, has produced over 22 million barrels of oil (MMBO) since 1895. The sandstone of the Late Devonian Gordon Stray is the primary reservoir.

The Upper Devonian fluvial sandstone reservoirs in Jacksonburg-Stringtown oil field, which has produced over 22 million barrels of oil since 1895, are an ideal candidate for CO_2 sequestration coupled with EOR. Supercritical depth (>2500 ft.), minimum miscible pressure (941 psi), favorable API gravity (46.5°) and good water flood response are indicators that facilitate CO_2-EOR operations. Moreover, Jacksonburg-Stringtown oil field is adjacent to a large concentration of CO_2 sources located along the Ohio River that could potentially supply enough CO_2 for sequestration and EOR without constructing new pipeline facilities.

Permeability evaluation is a critical parameter to understand the subsurface fluid flow and reservoir management for primary and enhanced hydrocarbon recovery and efficient carbon storage. In this study, a rapid, robust and cost-effective artificial neural network (ANN) model is constructed to predict permeability using the model's strong ability to recognize the possible interrelationships between input and output variables. Two commonly available conventional well logs, gamma ray and bulk density, and three logs derived variables, the slope of GR, the slope of bulk density and V_{sh} were selected as input parameters and permeability was selected as desired output parameter to train and test an artificial neural network. The results indicate that the ANN model can be applied effectively in permeability prediction.

Porosity is another fundamental property that characterizes the storage capability of fluid and gas bearing formations in a reservoir. In this study, a support vector machine (SVM) with mixed kernels function (MKF) is utilized to construct the relationship between limited conventional well log suites and sparse core data. The input parameters for SVM model consist of core porosity values and the same log suite as ANN's input parameters, and porosity is the desired output. Compared with results from the SVM model with a single kernel function, mixed kernel function based SVM model provide more accurate porosity prediction values.

Base on the well log analysis, four reservoir subunits within a marine-dominated estuarine depositional system are defined: barrier sand, central bay shale, tidal channels and fluvial channel subunits. A 3-D geological model, which is used to estimate theoretical CO_2 sequestration capacity, is constructed with the integration of core data, wireline log data and geological background knowledge. Depending on the proposed 3-D geological model, the best regions for coupled CCUS-EOR are located in southern portions of the field, and the estimated CO_2 theoretical storage capacity for Jacksonburg-Stringtown oil field vary between 24 to 383 million metric tons. The estimation results of CO_2 sequestration and EOR potential indicate that the Jacksonburg-Stringtown oilfield has significant potential for CO_2 storage and value-added EOR.

ACKNOWLEDGEMENTS

I would like to acknowledge the following people and organizations that helped me in various ways throughout the duration of this project.

First, I give my greatest appreciation to Professor Timothy Carr, my graduate advisor, and committee chairman. His skillful guidance, great support, and encouragement throughout my doctoral studies helped me to accomplish this challenging research project. Thanks to him again for his guiding and helping.

I will extend my special thanks to all my committee members; Dr. Dengliang Gao, Dr. Amy Weislogel, Dr. Shikha Sharma and Dr. Shahab Mohaghegh for their tremendous help and assistance on my dissertation. I also thank Dr. Kathy Brunner for her helpful comments in the early version of the manuscript. Special thanks give to Hope Stewart, Randy Crowe and Donna Titus for her assistance in my studies.

I would like thanks fellow students and people I met during my research at WWU, including Fei Shang, Ruiqian Cheng, Liaosha Song, Shuvajit Bhattacharya, Haibin Di, Shohreh Amini, Payam Kavousi, Thomas Paronish, Yaqian He, Fang Fang, Yixuang Zhu, Pragya Srivastava, and many others not listed for their great help.

I also thank my parents and my grandmother for their selfless love, endless supports, and encouragement in all the phases of my life. I also extend my thanks to my parents-in-law for their patience to take care of my cute son, Lucas.

Lastly and most importantly, I write of my wife, Yuyin Niu. We met 11 years ago and her tireless support and personal sacrifice are my strongest back to accomplish my research. She has earned my endless respect and appreciation for her unselfish support. Really thanks her, the one I love. I thank my cute son, Lucas. It is him that gave me more energy to go further.

Zhi Zhong

February 25, 2017

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

CHAPTER 3

LIST OF TABLES

CHAPTER 1

Application of Improved Back-Propagation Neural Network in Permeability Prediction: A Case Study in Jacksonburg-Stringtown Oil Field, West Virginia, USA

Zhi Zhong and Timothy R Carr[1]

[1] Department of Geology and Geography

West Virginia University, Morgantown

WV-26505, USA

(304)-282-9243

Email: zizhong1990@gmail.com; tim.carr@mail.wvu.edu

Abstract

Permeability evaluation is a critical parameter to understanding subsurface fluid flow, and reservoir management for primary and enhanced hydrocarbon recovery and efficient carbon storage. Accurate permeability values are measured in the laboratory from subsurface core samples or estimated from well test data. However these measurements are expensive and time-consuming and usually limited to a few wells and limited samples in a hydrocarbon field or carbon storage site. In order to create a rapid, robust and cost-effective model to predict permeability, intelligent techniques are applied that can recognize possible interrelationships between input and output variables. In this research, new back-propagation artificial neural networks (BPNN) are optimized using two evolutionary algorithms: genetic algorithms (GA's) and particle swarm optimization (PSO), to estimate permeability in potential carbon storage and enhanced oil recovery (EOR) operations in the Jacksonburg-Stringtown oil field, West Virginia, USA. The two evolutionary algorithms were applied to determine suitable initial connection weights and biases of a back-propagation neural network. Two commonly available conventional well logs, gamma ray and bulk density; and three logs derived variables, the slope of GR, the slope of bulk density and V_{sh} were selected as BPNN input parameters to better predict permeability. The predicted results of BPNN with particle swarm optimization model (PSO-BPNN) is compared with predicted results from BPNN with genetic algorithm (GA-BPNN). The results indicate that the PSO-BPNN model can be applied more effectively in permeability prediction with highest correlation coefficient (r of 0.9595), highest coefficient of determination (R^2 of 0.9208), lowest standard deviation (SD of 26.6584) and root mean square error (RMSE of 137.5647), mean error value (ME of 19.4389).

1.1 Introduction

Permeability (k) is a measure of how easily hydrocarbons, CO_2 and other fluids and gasses with a given viscosity flow through a subsurface unit's pore spaces in the presence of one or multiple fluid phases (e.g., oil and formation water). It is a critical parameter in reservoir characterization and determination of flow patterns. Reliable estimation of permeability is critical for seal evaluation and fluid-migration analysis in a field or reservoir (Helle et al., 2001). There are no geophysical well logs that provides a direct permeability measurement, and predicting permeability is one of selecting a model expressing *k* in terms of other measurable rock properties. A theoretical tube-like model of rock pore space known as the Kozeny-Carman relationship is applied to calculate the permeability. However, the result is unreliable due to numerous parameters, such as Kozeny constant and porosity value (Mauran et al., 2001). Pressure transient formation well testing including; pressure build up, drill stem testing (DST) and repeat formation testing (RFT) (Clark and Golf-Racht, 1985; Xu et al., 2008); and laboratory core measurements, provide relatively accurate measure of permeability, However, core and pressure transient tests are expensive, time-consuming and given limited data in a few wells may not be representative of the reservoir. Production history matching can also be used to estimate permeability, but this permeability value reflects the average permeability of the whole reservoir and ignores the reservoir's heterogeneity and complexity (Helle et al., 2001).

Recently, artificial neural networks (ANN) have been applied in the petroleum industry because of their strong ability of generalization and nonlinear approximation (Huang and Williamson, 1996; Mohaghegh and Ameri, 1995; Mohaghegh et al., 1996; Van der Baan and Jutten, 2000). This technology has been used extensively to improve prediction for: porosity and permeability (Al Moqbel and Wang, 2011; Helle et al., 2001; Huang et al., 1996; Huang and Williamson, 1997; Mohaghegh and Ameri (1995); Mohebbi et al., 2012); water saturation determination (Helle and Bhatt, 2002); lithofacies classification (Al Moqbel and Wang, 2011; Bhatt and Helle, 2002; Wang and Carr, 2012a; Wang et al., 2013); hydraulic fracture optimization (Mohaghegh and Ameri, 1995); reservoir pressure estimation (Chen et al., 2014; Sayyad et al., 2014); PVT property prediction (Gharbi and Elsharkawy, 1999); and other subsurface applications. Evaluation of ANN performance in predicting reservoir permeability using full suites of modern

geophysical well log data indicates good estimates even in heterogeneous reservoirs (Aminian and Ameri, 2005).

Artificial neural networks can be applied across almost any field or reservoir to provide a nonlinear mapping between inputs and outputs (Rosenblatt, 1961). The characteristics of ANN including learning feature from data, fast development, strong generalization and universal approximation ability, accurate nonlinear data fitting and regression capabilities, are major reasons for the rapid growth in number and diversity of applications (Kordon, 2009). ANN has been applied widely, but the process of minimizing convergence rate to zero during network training can cause overtraining, also known as overfitting, because of memorization of the training dataset (Geman et al., 1992; Moody, 1994; Scales and Snieder, 1998; Tu, 1996). Principal component analysis (PCA) and cross-validation methodologies have been applied to optimize ANN structure and reduce the overfitting effect (Jin et al., 2005; Zhang et al., 1999). However, these approaches significantly reduce the number of samples in the training dataset, and in the case of permeability with limited direct laboratory or well measurements can result in insufficient data for a robust training process. In order to deal with severely limited modern geophysical log data from the greater than 100 year-old Jacksonburg-Stringtown oil field in West Virginia, USA, we apply two evolutionary algorithms, namely genetic algorithms (GA's) and particle swarm optimization algorithm (PSO), to optimize the initial connection weights and biases of artificial neural networks (ANN), and develop hybrid GA-ANN and PSO-ANN regression models to predict the permeability. Five variables are evaluated: gamma ray (GR), density, the slope of GR, the slope of density and shale content (V_{sh}). The result of GA-BPNN is compared with the result of PSO-BPNN in order to determine the performance of evolutionary algorithms. A general BPNN is constructed and the result of BPNN is compared with the result of GA-BPNN and PSO-BPNN in order to illustrate the advantages of evolutionary algorithms' optimization ability.

The paper introduces the theoretical background of artificial neural network, evolutionary algorithms, illustrates the process of developing hybrid GA-BPNN and PSO-BPNN, and describes the process of how to evaluate the different BPNNs' performance. The geological background and the challenges of data acquisition of a case study are discussed and used to

evaluate the results of evolutionary algorithms in the super-mature reservoir that has been developed over more than a century.

1.2 Methodology

1.2.1 Principle of Artificial Neural Networks

Mathematical perceptron is the prototype of a neural network, which mimics biological neuron behavior (McCulloch and Pitts, 1943) (Figure 1-1a). Hérault and Jutten describe the process of biological neuron transiting signals from one neuron to others by the mathematical method (Hérault and Jutten, 1994). The mathematical neuron simplified the transiting process of biological neuron's signal (Figure 1-1b). The result of summing node is the weighted sum of input signals (Figure 1-1c). The final output signal is rescaled by various types of activation functions (Figure 1-1d). An artificial neural network is the combination of a series of mathematical neurons. There are many different types of ANNs, some of which are more popular than others (Agatonovic-Kustrin and Beresford, 2000). The most frequently used ANN is a fully connected, supervised network with a backpropagation learning rule, which generally is labeled as a back-propagation neural network (BPNN). BPNN normally consists of three layer types of neurons. The first layer is a single input layer and the last layer is a single output layer. The number of input neurons and output neurons is problem dependent. One or more hidden layers are located between input and output layers. The number of hidden layers and hidden layer's neurons vary, depending on the complexity of problem and training dataset' quality and size. A small number of neurons in hidden layer may lead the network to fall into a local minimum; conversely, many neurons will result in overfitting the network. These challenges can make trained networks lose their generalization ability (Jeirani and Mohebbi, 2006). Neurons of input layer are connected to the hidden layer by weights and biases in the same fashion as neurons between hidden layer and output layer. Individual neurons are connected by weights and biases, and selection of suitable weights and biases can avoid or reduce the overfitting effect. Normally, ANN structure is optimized by changing the number hidden layers and the number of hidden layers' neurons.

Learning processes are the main component of the ANN training process, in which weights and biases are adjusted continually until an expected output is produced or anticipated criteria are

achieved. The testing process is used to assess the generalization ability of the trained neural network (Saemi et al., 2007). Multiple BPNN models are constructed until training process and testing process achieves the required accuracy and generalization ability.

The error of ANN is a high dimensional surface with an extremely complex shape. The vertical dimension of each point on this surface corresponds to one error value and each point represents a vector of weights and biases. The well-trained BPNN model can be used to undertake prediction of unknown values with a defined error. A simplified 3D figure displaying the relationship between error, weight, and biases can illustrate how easy it is to fall into local minimum far removed from a global minimum (Figure 1-2). A suboptimum result can easily occur if there are numerous local minimums or if the initial weights and biases are far away from expected weights and biases. The convergent rate to a global optimum minimum can be extremely slow if the learning efficiency is low, and the neural network may forget the feature of the old sample when it is trained by the new sample. In order to deal with these challenges, several supervised learning algorithms are introduced to improve convergence and learning efficiency. Two evolutionary algorithms (GA and PSO) can be applied to optimize the initial weights and biases for each single neuron in well-constructed BPNN in terms of minimizing the effect of overfitting.

In order to evaluate the performance of training and testing process, mean squared error (MSE) is chosen as network performance function, which is defined as follows:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2 \quad , \qquad \text{Eq. (1)}$$

where $\hat{y}_i$ is the predicted value; y_i is the true value and n is the sample size.

1.2.2 Supervised Learning Algorithm

The process of training a neural network is a process of tuning the values of the weights and biases of the network to minimize the error between target and computed output in the neural network. Several algorithms have been developed; including Levenberg-Marquardt (LM), gradient descent (GD), gradient descent with momentum (GDM), scaled conjugate gradient method (SCG) and Quasi-Newton method (BFGS) (Wang and Carr, 2012b). To avoid the

network falling into local-minimums, two stochastic intelligence algorithms are used, genetic algorithm (GA) and particle swarm optimization (PSO). These algorithms optimize the weights and biases of the neural network by generating new offspring for GA or varying direction of movement and velocity for PSO.

1.2.3 Methods and Steps Combining GA with BPNN

1.2.3.1 The Description of Genetic Algorithm

Genetic algorithms (GA's) were first proposed in the 1970s (Holland, 1975). This global heuristic search algorithm is based on Darwinian evolutionary theory and is inspired by biological evolution and natural genetics. Over the last several decades, GA's have attracted attention, because of their potential as optimization techniques for complex functions (Dehghani *et al.*, 2008; Goldberg and Holland, 1988; Irani and Nasimi, 2011; Ravandi *et al.*, 2014; Velez-Langs, 2005; Whitley *et al.*, 1990).

As originally proposed, the formulation of a GA for a specific problem usually is composed of three main issues: the designation of chromosomes implying a potential solution to a given problem; the reproduction and breeding structures used to generate initial genomes; and the genetic operators including selection, cross-over and mutation used to generate new genomes (Velez-Langs, 2005). A population is composed of numerous individuals that are represented by different chromosomes representing a set of potential solutions to a given problem. The size of the population is problem dependent but normally ranges from 50 to 100 individuals.

GA's begin with an initial population. The fitness of each individual is calculated using a fitness function (usually MSE) in a decoded form in the current population (Saemi *et al.*, 2007). Quantitative fitness values in a population are the basis of probability allocation for the purpose of selecting the fittest individuals in a probabilistic manner (Ravandi *et al.*, 2014). Individuals' genetic information are recombined by a cross-over operation and modified through a mutation operation to give birth to offspring. The offspring is a new population. The fitness of new population is calculated using the same criterion, and this process is repeated until achieving termination conditions, such as detected convergence, fixed number of generations, allocated

budget, or the highest ranking solution's fitness via manual inspection (Kumar *et al.*, 2010; Polushina and Sofronov, 2011; Ravandi *et al.*, 2014; Velez-Langs, 2005).

1.2.3.2 Development of hybrid GA-BPNN model

A genetic algorithm is applied to searching the suitable initial weights and biases of BPNN to improve prediction performance and generalization. The flowchart for the hybrid GA-BPNN construction process is made up five steps based on the basic principle and working mechanisms of BPNN and GA (Figure 1-3). The first step is representing the connection weights and biases as coding chromosomes (Figure 1-4).

The second step is establishing the original populations for the corresponding neural network. For the initial random population, each gene (weight or bias) takes a random value which is subjected to a normal Gaussian distribution. Generally, the population size is 100, which means the population is composed of 100 chromosomes (individuals).

The third step involves the fitness of each chromosome based on the fitness function calculated (Eq. 1). The fitness function is used to evaluate the performance of each chromosome (potential solution) and represents the misfit between the predicted and the true permeability value.

The fourth step is applying the genetic operators such as selection, crossover, and mutation operators to generate a new population (Irani and Nasimi, 2011). The genetic operators determine the process of selection, which is the first step of the genetic evolution process. These values help GA to reserve high-grade individuals and eliminate bad individuals in each population. In the crossover, the genes of old individuals are exchanged to a gene in the purpose of generating new individuals, which have stronger search ability than any of previous individual. A roulette wheel selection is a common way to select two ancestor individuals from the population to generate two evolving individuals via crossover operators. The genetic operator, mutation is the last step of the evolutionary process. For a real number chromosomes, numerous mutation operators could be designed. In such a case, mutation operator can be randomly changed, which is subjected to a uniform probability distribution (Gholami et al., 2014; Velez-Langs, 2005).

Finally, decoding chromosomes convert the real-coded chromosomes to connection weights and biases. The process is repeated until specific criteria are satisfied. Typical parameters and values used in GA-BPNN are shown in Table 1-1.

1.2.4 Methods and Steps Combing PSO with BPNN

1.2.4.1 The Description of Particle Swarm Optimization Algorithm (PSO)

Based on stochastic search and optimization processes, the particle swarm optimization algorithm (PSO) was developed (Eberhart and Kennedy, 1995). PSO is an evolutionary algorithm that imitates human (or insects) social behavior. Individuals interact with one another while learning from their own experience, and gradually the population member moves into better regions of the problem space (Eberhart and Kennedy, 1995). Particles, indicating the potential solution, randomly locate in the architecture space (such as birds or fish randomly distributed in a specific open environment), and are utilized to calculate the global optima of the fitness function. Assuming D-dimensional architecture space, the population $x = \{x_1, x_2, x_3, ..., x_n\}$ is composed of n particles (x_i). Each $x_i = \{x_{i1}, x_{i2}, x_{i3}, ..., x_{iD}\}$ indicates its position and is also represents a potential solution. A global best value is $p_g = \{g_1, g_2, g_3, ..., g_D\}$ and personal best value is $p_i = \{p_{i1}, p_{i2}, p_{i3}, ... p_{iD}\}$. The velocity for particle x_i, representing the rate of position change, is written as $v_i = \{v_{i1}, v_{i2}, v_{i3}, ..., v_{iD}\}$ (Poli *et al.*, 2007).

During the optimizing process, each particle updates its position by velocity adjustment and fluctuates between the individual p_i and global p_g best values. When the particle swarm algorithm is running without restraining velocities, it rapidly increases to unacceptable levels within a few iterations. Some form of damping of the dynamics of particles (e.g., $v_{\max}$) is necessary (Cabrerizo *et al.*, 2013). To better allow an elegant and well-explained method for limiting the searching range, ensuring convergence, reducing and eliminating the importance of $v_{\max}$, an inertial weight algorithm was introduced (Shi and Eberhart, 1998) (Figure 1-5). The velocity updating formulae are defined as following:

$$\vec{v}_i^{\,new} = \omega_t \vec{v}_i^{\,old} + \phi_1 \otimes (\vec{p}_i - \vec{x}_i) + \phi_2 \otimes (\vec{p}_g - \vec{x}_i) \qquad \text{Eq. (2),}$$

$$\vec{x}_i^{new} = \vec{x}_i^{old} + \vec{v}_i^{new} \quad \text{Eq. (3),}$$

$$\omega_t = \omega_{\max} - (\omega_{\max} - \omega_{\min}) * t_{ite} / t_{\max} \quad \text{Eq. (4),}$$

where ω_{max} is initial inertia weight, ω_{min} is the final inertia weight, φ_1 is cognitive coefficient and φ_2 is social coefficient, t_{ite} is current iteration number, and t_{max} is maximum iteration number.

1.2.4.2 Development of Hybrid PSO-BPNN Model

In order to ameliorate training process and accelerate convergence rate, the PSO algorithm is combined with BPNN to optimize the initial weights and biases of BPNN. The number of weights and biases equals the dimension of each particle. The combination of a series of connection weights and biases is a particle and indicates the position of a particle (Sayyad et al., 2014). A flowchart of proposed the PSO-BPNN model for permeability prediction was developed (Figure 1-6).

The hybrid PSO-BPNN model searching mechanism and hybrid network are made up three stages: first, construction of BPNN, initializing the control parameters of PSO and randomly initializing all particles with a suitable size of population; second, training the BPNN by using each particle (weights and biases), and calculation of the fitness of each particle in the population based on the fitness function (Sedki et al., 2009). Unless termination conditions are achieved the velocity and position of each particle are updated based on a new variant of PSO to construct new generation (Shi and Eberhart, 1998). The fitness value is calculated again for renewed generation, then velocity and position for new particles are repeatability updated. This process will sustain until the stopping criteria are satisfied (Poli *et al.*, 2007). Finally, the best particle (weights and biases) are applied to BPNN, and the well-trained BPNN model is used to make predictions.

1.2.5 Performance Evaluation

The well-trained BPNN model is applied to estimate the permeability value in the Jacksonburg-Stringtown reservoir. The correlation coefficient (r) is an important criterion for evaluating the performance of regression, however, it is not sufficient to fully characterize a complex regression

problem (Zhong and Carr, 2016). Therefore, in order to verify the reliability and accuracy of a model, four additional statistical parameters were introduced as evaluation indices of the model performance (Jin et al., 2005; Jin et al., 2001). These performance evaluation parameters include; the coefficient of determination (R^2), root mean error (RMSE), average absolute error (AAE), and maximum absolute error (MAE) (Table 1-2). An R^2 (Coefficient of Determination) of 1 indicates a perfect regression model, while an R^2 of 0 indicates a completely random model (Oyerokum, 2002). Moreover, RMSE is used to evaluate overall performance, while AAE and MAE are used to determine the error range of the predicted results. A model with high r and R^2 values, and low RMSE, MAE and AAE values is considered to have good performance.

1.3 Application Case

1.3.1 Geological Setting

Jacksonburg-Stringtown field is situated along the axis of the Burchfield syncline in northwestern West Virginia (Figure 1-7b). The primary or secondary producing reservoir unit is Late Devonian Gordon Stray interval, which is contained within the middle to late Catskill deltaic complex (Catskill delta). During Early to Middle Devonian, crustal uplift in Acadian orogeny lead to further subsidence within the Appalachian foreland basin to the west and resulted in the deepening of the central Appalachian basin (Faill, 1997a, b). Deposition is interpreted to coincide with the heavy rainfall produced by the tropical climate, and sedimentary deposition accelerated during the Middle and Late Devonian (Blakey, 2008; McBride, 2004; Piotrowski and Harper, 1979). In the Late Devonian, five major delta systems prograded westwards and dominated the foreland ramp (Figure 1-7a). In the latest Devonian, maximum progradation of the Catskill delta complex was achieved west of the Acadian highlands (Boswell and Donaldson, 1988). During this period, Acadian tectonism ceased, and relative sea-level changes within the basin were controlled primarily by estuary sea-level fluctuations and variations in sedimentation (Coughlin, 2009). Non-marine red shale and most of the low-energy alluvial deposits are concentrated in the eastern portions of the Appalachian basin. Non-marine sediments increasingly advanced westward to cover marine beds, and near-shore deposits continued to prograde into the central Appalachian basin. In the area of Jacksonburg-Stringtown field, Gordon Stray/Gordon intervals are interpreted to be shoreline and shore-face sandstone

that occupied a broad structural trend at the time of maximum regression of the Acadian clastic wedge (Hohn, 2004).

The estimated original oil in place (OOIP) in Jacksonburg-Stringtown oil field is 88 million barrels of oil (MMBO). Since the discovery in 1895, primary production was estimated at 22 MMBO (Ameri et al., 2002), and the estimated oil recovery factor is 25%. A gas re-injection program took place in the field beginning in the mid-1930's and ended in the 1950's. A pilot waterflood program with a 35 acre dual 5-spot well pattern was conducted in 1981. After 1990, a full-scale waterflood was installed in a large portion of the field (Bergerud, 2011). Unlike modern oil and gas fields, which have abundant, high quality data, including full-suites of conventional and advanced well logs, and seismic data to construct 3D geological model, Jacksonburg-Stringtown oil field has predominately low quality well log data (i.e. gamma ray, bulk density), and highly limited high quality data (core measured porosity and permeability data) for reservoir modeling. The sparse modern subsurface data characteristic of super-mature fields such as Jacksonburg-Stringtown can inhibit the development of a robust geological model and effective evaluation of CO_2 storage capacity.

1.3.2 Data Acquisition

A model's stability and accuracy are largely dependent on the training dataset's reliability and comprehensiveness. In this research, data is extremely sparse with only 93 samples from 6 wells in the Jacksonburg-Stringtown field with both core data and conventional well logs (Figure 1-7b). To construct a reliable BPNN model, input data was divided into two parts: training data and testing data. Data belonging to the first five wells were used to train the BPNN model, while the remaining data of the last well was used to evaluate the stability and accuracy of the well-trained network. Based on the materials on hand, to construct the model two conventional well logs and three log derived variables were selected as BPNN input variables along with digitized permeability values. The input variables include two well logs, gamma-ray log (GR), density log, and three derived parameters; the slope of GR, the slope of density, and shale content (V_{sh}). In order to demonstrate the heterogeneity of this formation and to display the chaotic status of the information that existed, cross-plots of permeability versus each measured parameter were constructed. A simple linear relationship between the six measured parameters and permeability

is not apparent (Figure 1-8). The dataset was normalized to improve prediction results and improve the calculation and training speed. In this study, absolute scale is used for all input parameters. For permeability, a logarithmic scale instead of the absolute scale is used. The normalized formula is chosen as following:

$$X_i^{new} = [\frac{X - X_{\min}}{X_{\max} - X_{\min}}] \quad \text{Eq. (5)}$$

where X_i^{new} is the normalized input vector, $X_{\min}$ and $X_{\max}$ are the minimum and maximum and value respectively, X are the original input vector. The normalized input vector ranges from 0 to 1. When training and testing process are completed, the predicted permeability values by well-trained BPNN model also ranges from 0 to 1. In order to re-project the predicted permeability value into original order, data renormalization is required. The renormalization formula is following:

$$X^{renormalized} = 10^{(X_{\max} - X_{\min})X^{predicted} + X_{\min}} \quad \text{Eq. (6)}$$

where $X^{renormalized}$ is the renormalized output, $X^{predicted}$ is predicted output by well-trained BPNN model.

1.4. Results and Discussion

1.4.1 ANN Topology and Supervised Learning Algorithms

The designation of network architecture is a subjective task and problem-dependent. A priori selection of the best architecture of BPNN in a specific problem (e.g., porosity, permeability, and minimum miscible pressure) is challenging. Therefore, in order to obtain a reasonable BPNN architecture, several possibilities are considered in this study. The number of nodes in the input layer corresponds to the five basic input parameters, whereas the output layer node corresponds to the permeability value. Normally, two hidden layers can approximate most non-linear or linear regression problems (Kumar et al., 2002). The number of nodes in the hidden layer was varied from 20 to 50 for one hidden layer architecture. However, for the two hidden layer architecture

fifteen, twenty, twenty-five and thirty nodes were considered in first hidden layer, and ten, fifteen, twenty and twenty-five nodes were considered in second hidden layer (Figure 1-9). All five supervised learning algorithms and two evolutionary intelligent algorithms were applied to train all the BPNNs with varying topology.

BPNN with different architecture show various performance (Figure 1-9). The BPNN model trained by Levenberg-Marquardt algorithm performs best in various BPNN models trained by different supervised training algorithms (Figure 1-9, Table 1-3). This BPNN has relative high R^2 and r values, and low RMSE, MAE and AAE values. The BPNN that has one hidden layer with 25 nodes performed best with the highest R^2 (0.9131) and r (0.9555), and lowest RMSE (162.8002), MAE (66.1634) and AAE (of 0.7048).

After multiple runs of BPNN with different hidden layers and a various number of neurons, the best performance was achieved by 5-25-1 as final network architecture. Different supervised learning algorithms and two evolutionary intelligent algorithms were used to train the selected BPNN. Mean square error was used as the cost function to appraise the network's performance. BPNN trained by Levenberg-Marquardt algorithm and two evolutionary algorithm performs best (Figure 1-10), so a more detailed comparison was carried out.

1.4.2 Comparison of Permeability Prediction Base on Various Regression Model

In order to further evaluate the performance of the BPNN, GA-BPNN, and PSO-BPNN model, the predicted results obtained from well-trained models were compared with available core measured permeability datasets. To check and confirm the generalization capability and predicting precision in the estimation of permeability for different optimized or non-optimized BPNNs, 22 data points from well B-19, which were not used to constructing the BPNN model were estimated with the model. Generally, the error distribution fits the normal distribution. Therefore, the mean error values and the standard deviations presented here are those of a Gaussian model (Helle et al., 2001).

Histograms and cross-plots display the difference between permeability measured from the core and predicted permeability predicted by the input petrophysical parameters using different BP neural networks (Figure 1-11). In order to compare the result more concisely, permeability is

plotted on a traditional logarithmic scale. The comparison result between cores measured permeability and predicted permeability by non-optimized BPNN shows a mean error is approximately 28.5246 md with a standard deviation of 72.4197 md (Figure 1-11a and b). The coefficient of determination (R^2) and correlation coefficient (r) is 0.5177 and 0.7195 respectively. The comparison result between core measured permeability and predicted permeability by PSO-BP neural network has a smaller mean error than that from non-optimized BP neural network. Mean error for the PSO-BP is approximately 19.438 md with a standard deviation of 26.6584 md (Figure 1-11c and d). The cross-plot of core measured permeability and predicted permeability by PSO-BPNN regression model has higher R^2 (0.9208) and r (0.9596) compared with the standard BP neural network (Figure 1-10 d; Table 1-4). The BPNN optimized by PSO performs better than general BPNN. Also, GA-BPNN has a smaller mean error (26.3259), standard deviation (60.825), and higher R^2 (0.8328) and r (0.9126) compared with non-optimized BPNN (Figure 1-10e and f). Evolutionary algorithm optimized BP neural network optimized by GA improves performance compared to non-optimized BPNN.

The PSO-BPNN regression model provides a strong ability to predict permeability with high correlation, and the highest coefficient of determination among all regression models and lower mean square error (Figure 1-11 b, d, and f). The GA-BPNN model has the same R^2 value as PSO-BPNN model, but the mean error value and standard deviation in permeability of PSO-BPNN model are lower and provide better prediction of known values (Figure 1-11c and e). The predicted permeability by PSO-BPNN and GA-BPNN compared to core derived values plotted in depth show good agreement (Figure 1-12).

Though permeability values cover a wide range, the PSO-BPNN model is able to follow and recreate the core permeability values and trend very closely. The input petrophysical parameters have a relation to permeability. The gamma-ray log response provides evidence of clay that has an impact on permeability. The bulk density is inverse functions of porosity and shale content. The slope of gamma ray and bulk density represent the change rate of clay content. V_{sh} directly represents variations in the rock-clay contents.

1.5. Conclusions

In this paper, two evolutionary algorithms and five supervised learning algorithms were applied to optimize the BPNN to estimate permeability, a critical parameter for hydrocarbon and CO_2 storage reservoir characterization. In the BPNN model, using a highly constrained suite of logs, five derived parameters: GR, density, the slope of GR, the slope of density, and V_{sh} were selected as input parameters. BPNN architecture is optimized by adjusting hidden layers and number of neurons. By comparing different BPNNs, optimized BPNN performance can be determined as evolutionary algorithms search the best solution by different direction and converge toward a global minimum value. Through the foregoing analysis and discussion of predicted result, several conclusions are drawn:

(1) The PSO optimized BPNN regression model was successfully applied to predict the permeability values. Based on comparisons between two optimized BPNN, PSO–BPNN regression model provides more accurate results than GA-BPNN regression model.

(2) Different architecture of BPNN affects significantly the final performance of BPNN. The best ANN topology for permeability prediction is 5-25-1. Among five supervised learning algorithms, the Levenberg-Marquardt algorithm performed well in the training process.

(3) Based on the comparisons between non-optimized BPNN and evolutionary algorithm optimized BPNN, optimized BPNN performed well, both in training and testing process.

The main drawback of those methods is:

1) The construction of BPNN architecture is largely based on experience. Generally, the more inputs and the fewer hidden neurons, the better the prediction performance. Too few inputs or too many hidden neurons can lead the network to memorize, which means that it works well during the training process, but tests poorly and fails to generalize;

2) The selection of several key parameters in evolutionary algorithms, such as population size, generations, mutation possibility is critical. Normally, large population and generation will definitely increase the computing time for training and testing. On the other hand, once the network is established, the application requires a minimum of computing time. Higher mutation

possibility and crossover possibility can vary the solution, but also can lead to unstable solutions and failure to converge.

Acknowledgements

This work was funded in part by the US-China Clean Energy Research Center, Advanced Coal Technology Consortium, under grant DE-PI0000017 from the National Energy Technology Laboratory of the US Department of Energy.

Nomenclature

$\bar{U}(0,\phi_i)$	A vector of random numbers uniformly distributed in [0, ϕ_i]
$\otimes$	Component-wise multiplication
ω	Inertia weight
χ	Constriction coefficients
K_i	The number of neighbors for particle i
nbr_n	i's nth neighbor
$\omega_{\min}$	Initial values of the inertia weight
$\omega_{\max}$	Final value of the inertia weight
t_{ite}	Current iteration number
$t_{\max}$	Maximum number of iteration
m	Number of output nodes
$Y_j(k)$	Expected output from neural network
$T_j(k)$	Actual output from neural network

References

Agatonovic-Kustrin, S., Beresford, R., 2000. Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research. Journal Of Pharmaceutical And Biomedical Analysis 22, 717-727.

Al Moqbel, A., Wang, Y., 2011. Carbonate reservoir characterization with lithofacies clustering and porosity prediction. Journal of Geophysics and Engineering 8, 592-598.

Ameri, S., Aminian, K., Avary, K.L., Bilgesu, H.I., Hohn, M.E., McDowell, R.R., Patchen, D.L., 2002. Reservoir characterization of Upper Devonian Gordon sandstone, Jacksonburg, Stringtown Oil Field, Northwestern West Virginia. West Virginia University.

Aminian, K., Ameri, S., 2005. Application of artificial neural networks for reservoir characterization with limited data. Journal of Petroleum Science and Engineering 49, 212-222.

Bergerud, B., 2011. Reservoir model of the Jacksonburg-Stringtown Oil field, Northwestern West Virginia: potential for miscible CO_2 Enhanced Oil Recovery, Department of geology and geography. West Virginia University.

Bhatt, A., Helle, B.H., 2002. Determination of facies from well logs using modular neural networks. Petroleum Geoscience 8, 217-228.

Blakey, R., 2008. Paleogeography and geologic evolution of North America, Northern Arizona University.

Boswell, R.M., Donaldson, A.C., 1988. Depositional architecture of the Upper Devonian Catskill Delta complex: central Appalachian basin, U.S.A. Memoir Canadian Society of Petroleum Geologists 2, 19.

Cabrerizo, F.J., Herrera-Viedma, E., Pedrycz, W., 2013. A method based on PSO and granular computing of linguistic information to solve group decision making problems defined in heterogeneous contexts. European Journal Of Operational Research 230, 624-633.

Chen, G., Fu, K., Liang, Z., Sema, T., Li, C., Tontiwachwuthikul, P., Idem, R., 2014. The genetic algorithm based back propagation neural network for MMP prediction in CO_2-EOR process. Fuel 126, 202-212.

Clark, D., Golf-Racht, V., 1985. Pressure-derivative approach to transient test analysis: a high-permeability North Sea reservoir example (includes associated papers 15270 and 15320). Journal of Petroleum Technology 37, 2,023-022,039.

Coughlin, M.F., 2009. Subsurface mapping and reservoir analysis of the Upper Devonian Venango and Bradford groups in Westmoreland County, Pennsylvania, Department of Geology and Geography. West Virginia University, p. 114.

Dehghani, S.A., Vafaie Sefti, M., Ameri, A., Shojai Kaveh, N., 2008. Minimum miscibility pressure prediction based on a hybrid neural genetic algorithm. Chemical Engineering Research and Design 86, 173-185.

Eberhart, R.C., Kennedy, J., 1995. A new optimizer using particle swarm theory, Sixth international symposium on micro machine and human science. New York, NY, pp. 39-43.

Faill, R.T., 1997a. A geologic history of the north-central Appalachians, Part 1: Orogenesis from the Mesoproterozoic through the Taconic Orogeny. American Journal of Science 297, 551-619.

Faill, R.T., 1997b. A geologic history of the north-central Appalachians, Part 2: The Appalachian basin from the Silurian through the Carboniferous. American Journal of Science 297, 729-761.

Geman, S., Bienenstock, E., Doursat, R., 1992. Neural networks and the bias/variance dilemma. Neural computation 4, 1-58.

Gharbi, R.B.C., Elsharkawy, A.M., 1999. Neural network model for estimating the PVT properties of Middle East crude oils. Spe Reservoir Evaluation & Engineering 2, 255-265.

Gholami, R., Moradzadeh, A., Maleki, S., Amiri, S., Hanachi, J., 2014. Applications of artificial intelligence methods in prediction of permeability in hydrocarbon reservoirs. Journal of Petroleum Science and Engineering 122.

Goldberg, D.E., Holland, J.H., 1988. Genetic algorithms and machine learning. Machine Learning 3, 95-99.

Helle, H.B., Bhatt, A., 2002. Fluid saturation from well logs using committee neural networks. Petroleum Geoscience 8, 109-118.

Helle, H.B., Bhatt, A., Ursin, B., 2001. Porosity and permeability prediction from wireline logs using artificial neural networks: a North Sea case study. Geophysical Prospecting 49, 431-444.

Hérault, J., Jutten, C., 1994. Réseaux neuronaux et traitement de signal: Hermes édition. Traitement du signal.

Holland, J.H., 1975. Adaptation in natural and artificial systems: An introductory analysis with applications to biology, control, and artificial intelligence. University Michigan Press.

Huang, Z., Shimeld, J., Williamson, M., Katsube, J., 1996. Permeability prediction with artificial neural network modeling in the Venture gas field, offshore eastern Canada. Geophysics 61, 422-436.

Huang, Z., Williamson, M.A., 1996. Artificial neural network modeling as an aid to source rock characterization. Marine and Petroleum Geology 13, 277-290.

Huang, Z., Williamson, M.A., 1997. Determination of porosity and permeability in reservoir intervals by artificial neural network modeling, offshore eastern Canada. Petroleum Geoscience 3, 245-258.

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

Jeirani, Z., Mohebbi, A., 2006. Estimating the initial pressure, permeability and skin factor of oil reservoirs using artificial neural networks. Journal of Petroleum Science and Engineering 50, 11-20.

Jin, L., Kuang, X., Huang, H., Qin, Z., Wang, Y., 2005. Study on the overfitting of the artificial neural network forecasting model. Acta Meteorological Sinica 19, 216.

Jin, R., Chen, W., Simpson, T.W., 2001. Comparative studies of metamodelling techniques under multiple modeling criteria. Structural and Multidisciplinary Optimization 23, 1-13.

Kordon, A., 2009. Applying computational intelligence: how to create value. Springer Science & Business Media.

Kumar, M., Husian, M., Upreti, N., Gupta, D., 2010. Genetic algorithm: review and application. International Journal of Information Technology and Knowledge Management 2, 451-454.

Kumar, M., Raghuwanshi, N., Singh, R., Wallender, W., Pruitt, W., 2002. Estimating evapotranspiration using artificial neural network. Journal of Irrigation and Drainage Engineering 128, 224-233.

Mauran, S., Rigaud, L., Coudevylle, O., 2001. Application of the Carman–Kozeny correlation to a high-porosity and anisotropic consolidated medium: The compressed expanded natural graphite. Transport in Porous Media 43, 355-376.

McBride, P.S., 2004. Facies analysis of the Devonian Gordon Stray sandstone in West Virginia, Department of Geology and Geography. West Virginia University, p. 145.

McCulloch, W.S., Pitts, W., 1943. A logical calculus of the ideas immanent in nervous activity. The bulletin of mathematical biophysics 5, 115-133.

Mohaghegh, S., Ameri, S., 1995. Artificial neural network as a valuable tool for petroleum engineers, SPE Petroleum Computer Conference, Texas, SPE. Citeseer.

Mohaghegh, S., Arefi, R., Ameri, S., Aminiand, K., Nutter, R., 1996. Petroleum reservoir characterization with the aid of artificial neural networks. Journal of Petroleum Science and Engineering 16, 263-274.

Mohebbi, A., Kamalpour, R., Keyvanloo, K., Sarrafi, A., 2012. The prediction of permeability from well logging data based on reservoir zoning, using artificial neural networks in one of an Iranian heterogeneous oil reservoir. Petroleum Science and Technology 30, 1998-2007.

Moody, J., 1994. Prediction risk and architecture selection for neural networks, From Statistics to Neural Networks. Springer, pp. 147-165.

Oyerokum, A.A., 2002. A new approach for training and testing artificial neural networks for permeability prediction, College of engineering and mineral resources. West Virginia University, p. 107.

Piotrowski, R.G., Harper, J.A., 1979. Black shale and sandstone facies of the Devonian Catskill clastic wedge in the subsurface of western Pennsylavnia, EGSP. Pennsylvania Department of Environmental Resources Bureau of Topographic and Geologic Survey, p. 42.

Poli, R., Kennedy, J., Blackwell, T., 2007. Particle swarm optimization. Swarm intelligence 1, 33-57.

Polushina, T., Sofronov, G., 2011. Change-point detection in biological sequences via genetic algorithm, Evolutionary Computation (CEC), 2011 IEEE Congress on. IEEE, pp. 1966-1971.

Ravandi, E.G., Nezamabadi-Pour, H., Monfared, A.E.F., Jaafarpour, A.M., 2014. Reservoir characterization by a combination of fuzzy logic and genetic algorithm. Petroleum Science and Technology 32, 840-847.

Rosenblatt, F., 1961. Principles of neurodynamics. perceptrons and the theory of brain mechanisms. DTIC Document, Washington, D.C.

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

Sayyad, H., Manshad, A.K., Rostami, H., 2014. Application of hybrid neural particle swarm optimization algorithm for prediction of MMP. Fuel 116, 625-633.

Scales, J.A., Snieder, R., 1998. What is noise? Geophysics 63, 1122-1124.

Sedki, A., Ouazar, D., El Mazoudi, E., 2009. Evolving neural network using real coded genetic algorithm for daily rainfall–runoff forecasting. Expert Systems with Applications 36, 4523-4527.

Shi, Y., Eberhart, R., 1998. A modified particle swarm optimizer, Evolutionary Computation Proceedings, 1998. IEEE World Congress on Computational Intelligence., The 1998 IEEE International Conference on. IEEE, pp. 69-73.

Tu, J.V., 1996. Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes. Journal of clinical epidemiology 49, 1225-1231.

Van der Baan, M., Jutten, C., 2000. Neural networks in geophysical applications. Geophysics 65, 1032-1047.

Velez-Langs, O., 2005. Genetic algorithms in oil industry: An overview. Journal of Petroleum Science and Engineering 47, 15-22.

Wang, G., Carr, T.R., 2012a. Methodology of organic-rich shale lithofacies identification and prediction: A case study from Marcellus Shale in the Appalachian basin. Computers & Geosciences 49, 151-163.

Wang, G.C., Carr, T.R., 2012b. Marcellus Shale Lithofacies Prediction by Multiclass Neural Network Classification in the Appalachian Basin. Mathematical Geosciences 44, 975-1004.

Wang, G.C., Cheng, G.J., Carr, T.R., 2013. The application of improved NeuroEvolution of Augmenting Topologies neural network in Marcellus Shale lithofacies prediction. Computers & Geosciences 54, 50-65.

Whitley, D., Starkweather, T., Bogart, C., 1990. Genetic algorithms and neural networks: Optimizing connections and connectivity. Parallel computing 14, 347-361.

Xu, L., Cai, J., Guo, S., Yi, P., Xiao, D., Dai, Y., Yang, S., Khong, C., Fujisawa, G., Dong, C., 2008. Real Time Carbon Dioxide Quantification Using Wireline Formation Tester to Optimize Completion and Drill Stem Testing Decisions. SPE IPTC 12081.

Zhang, G., Hu, M.Y., Patuwo, B.E., Indro, D.C., 1999. Artificial neural networks in bankruptcy prediction: General framework and cross-validation analysis. European Journal Of Operational Research 116, 16-32.

Zhong, Z., Carr, T.R., 2016. Application of mixed kernels function (MKF) based support vector regression model (SVR) for CO_2 – Reservoir oil minimum miscibility pressure prediction. Fuel 184, 590-603.

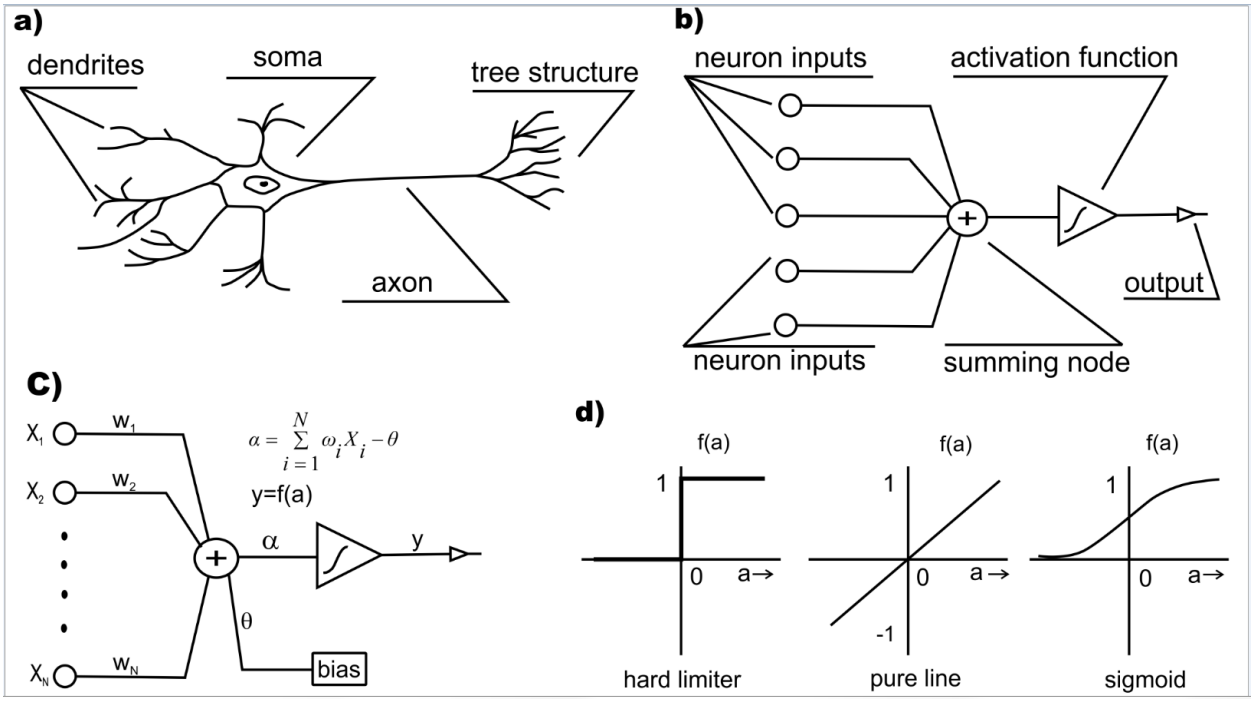


Figure 1-1.The biological and the mathematical neuron. The mathematical neuron (b) mimics the behavior of the biological neuron (a). The weighted sum of the inputs is rescaled by an activation function (c), of which several examples are shown in (d) (After Hérault and Jutten, 1994).

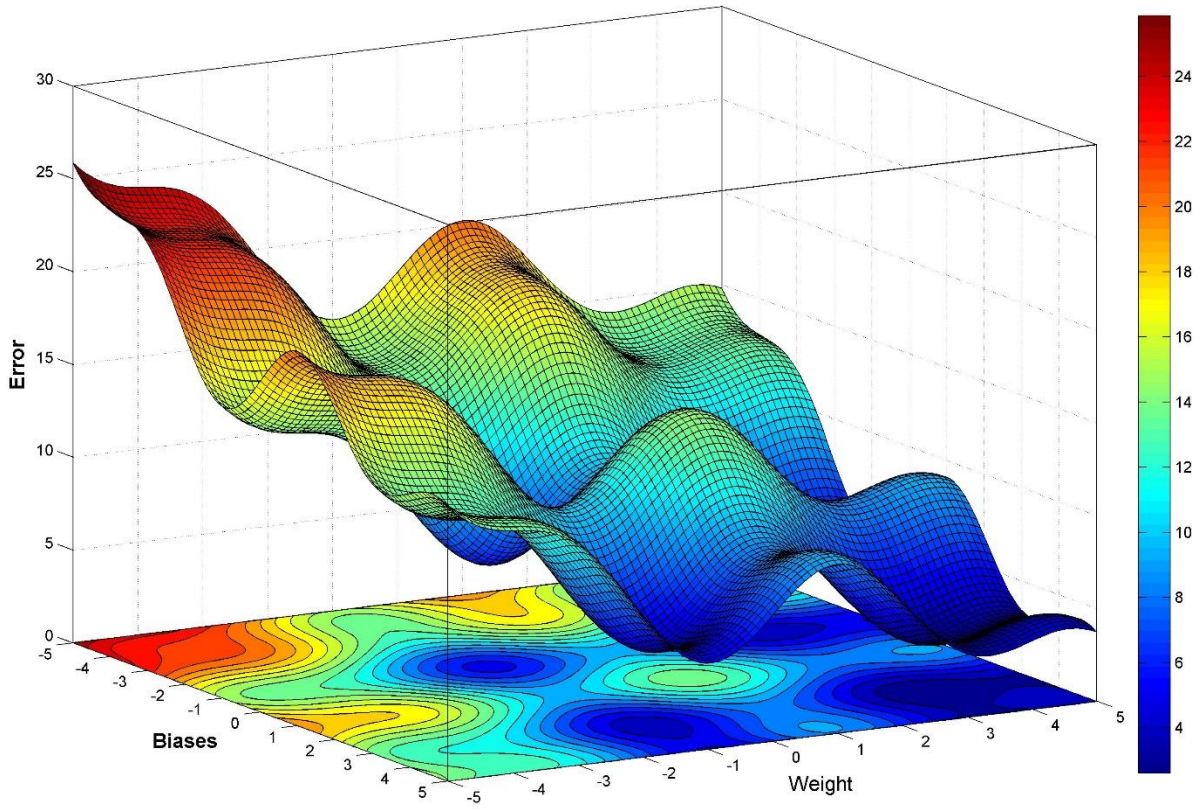


Figure 1-2. The simplified error surface illustrating the relationship between error and weight/biases of artificial neural network and the existence of areas of local suboptimum minimums.

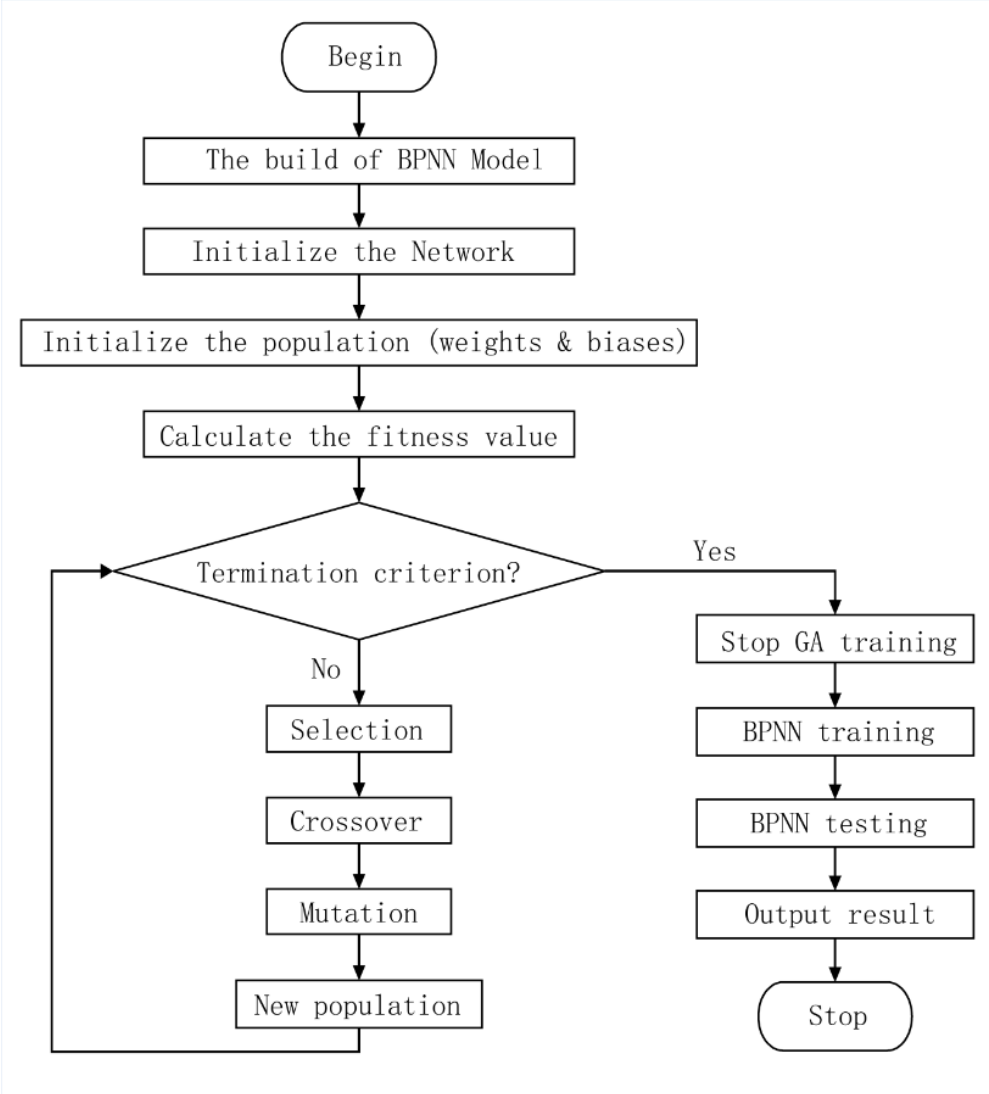


Figure 1-3.Flow chart of the hybrid GA-BPNN model used for permeability prediction.

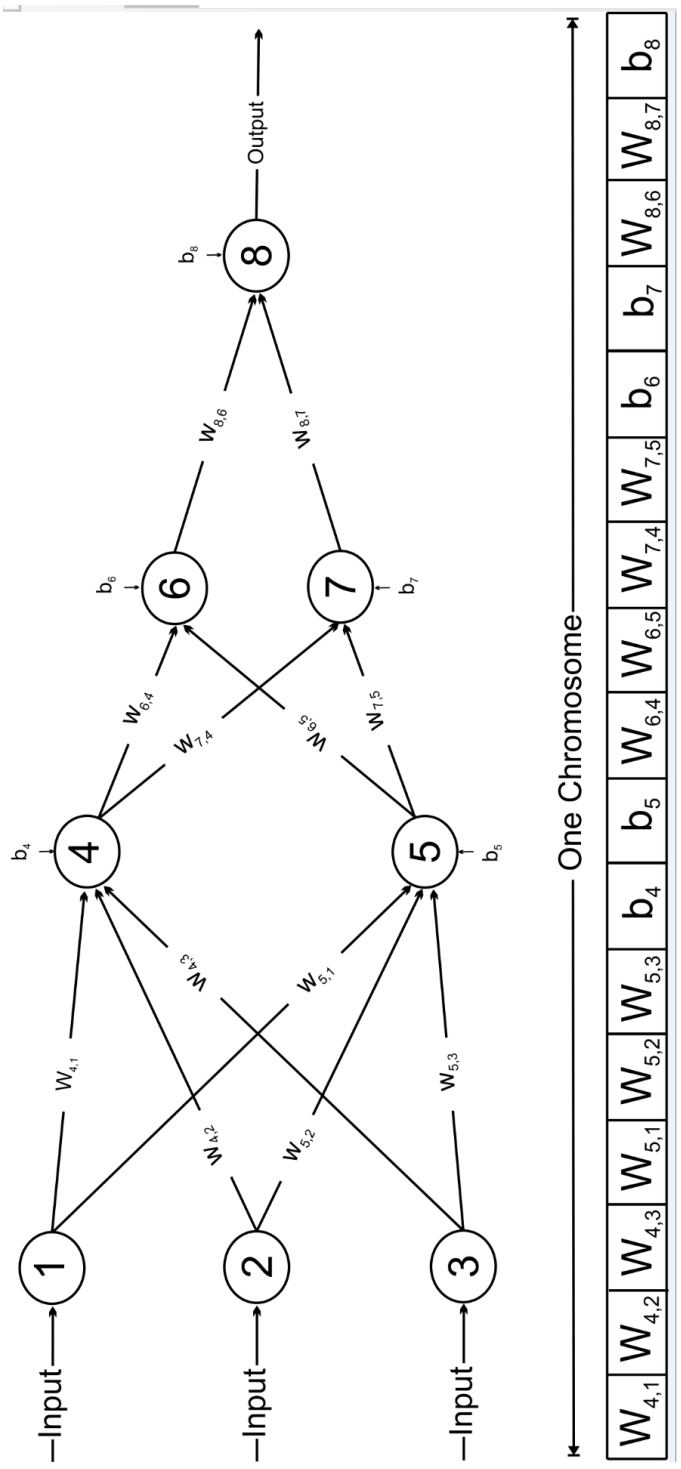


Figure 1-4. An example of an artificial neural network, with the connection weights ($w_{i,j}$) and the bias (b_k) of each neuron node (a); genetic chromosome representation of (a), one chromosome represents one individual (b)(after Dehghani et al., 2008).

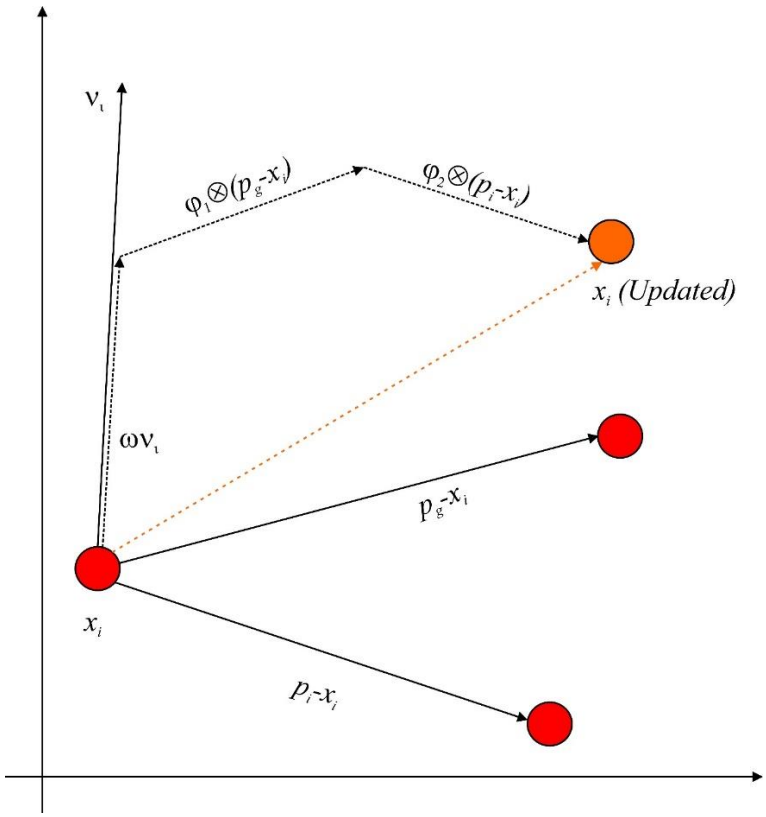


Figure 1-5. Graphical illustration of the particle swarm optimization algorithm PSO velocity and particle position update for particle xi in a two-dimensional search space.

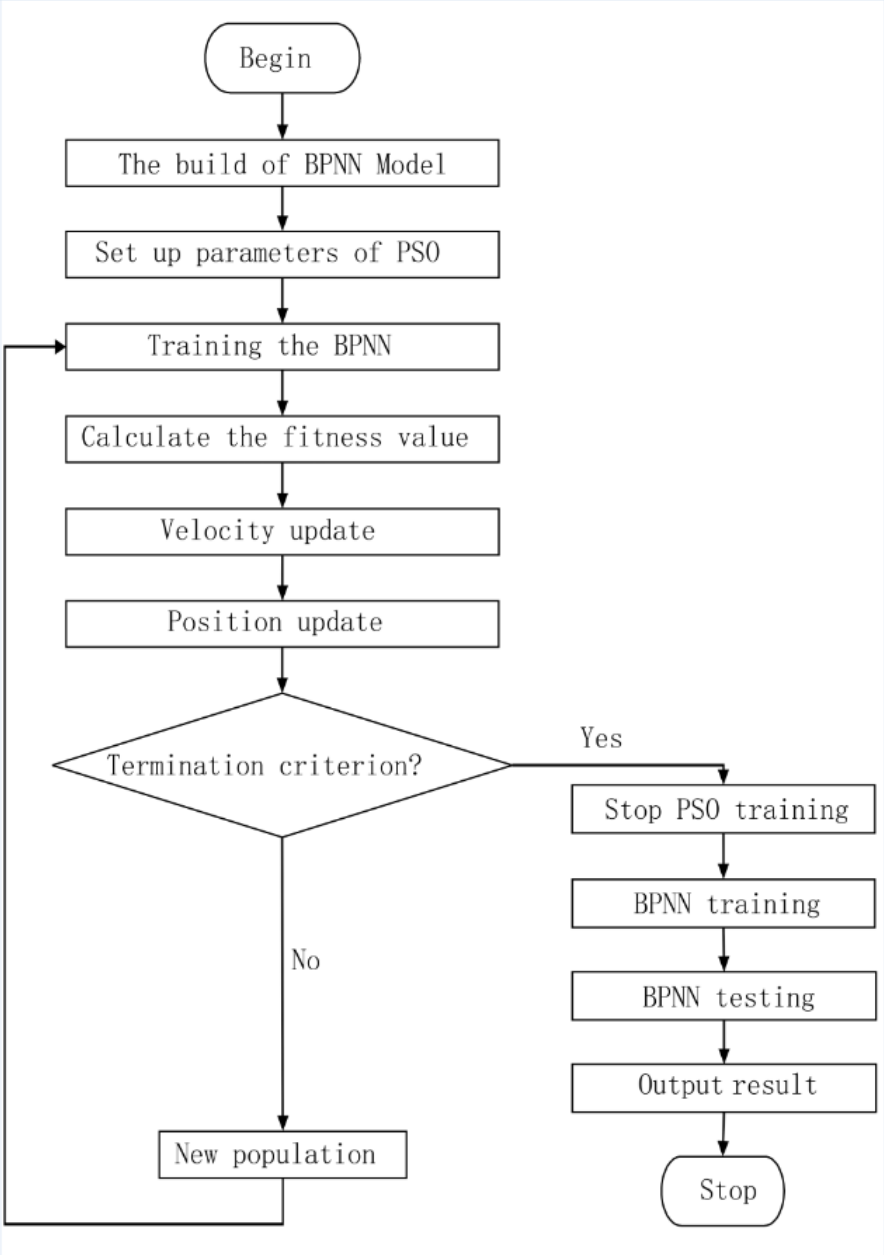


Figure 1-6. Flowchart of BPNN optimized with particle swarm optimization (PSO-BPNN) for permeability prediction.

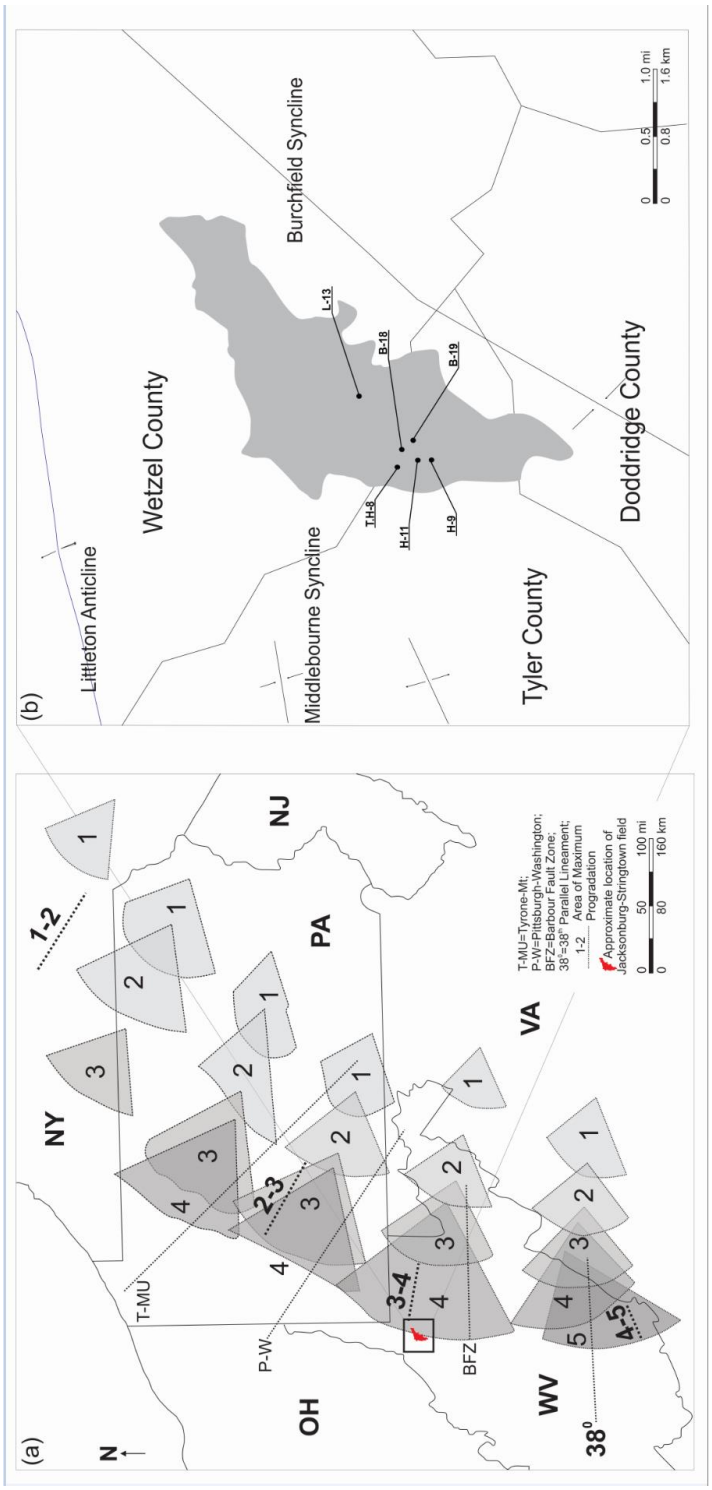


Figure 1-7. Late Devonian paleogeography of study area and five major delta systems in Appalachian foreland basin, black line indicates geographical state boundary (a); location of Jacksonburg-Stringtown field in northwestern West Virginia, black dots mark the location of cored wells (b) (Boswell and Donaldson, 1988).

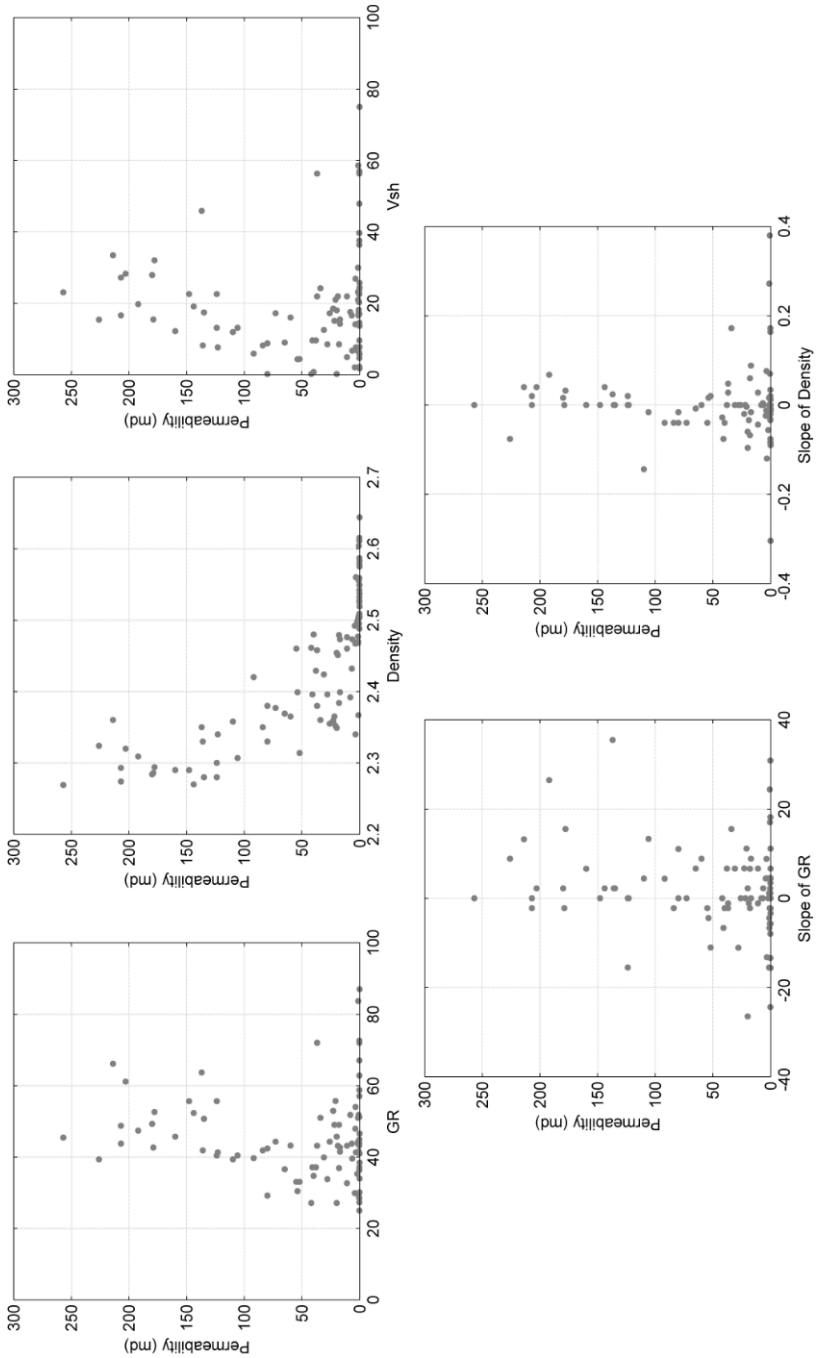


Figure 1-8. Permeability vs. different input variables which are used as input parameters in this study. It appears that there are not clear linear relationships between permeability and each input variables.

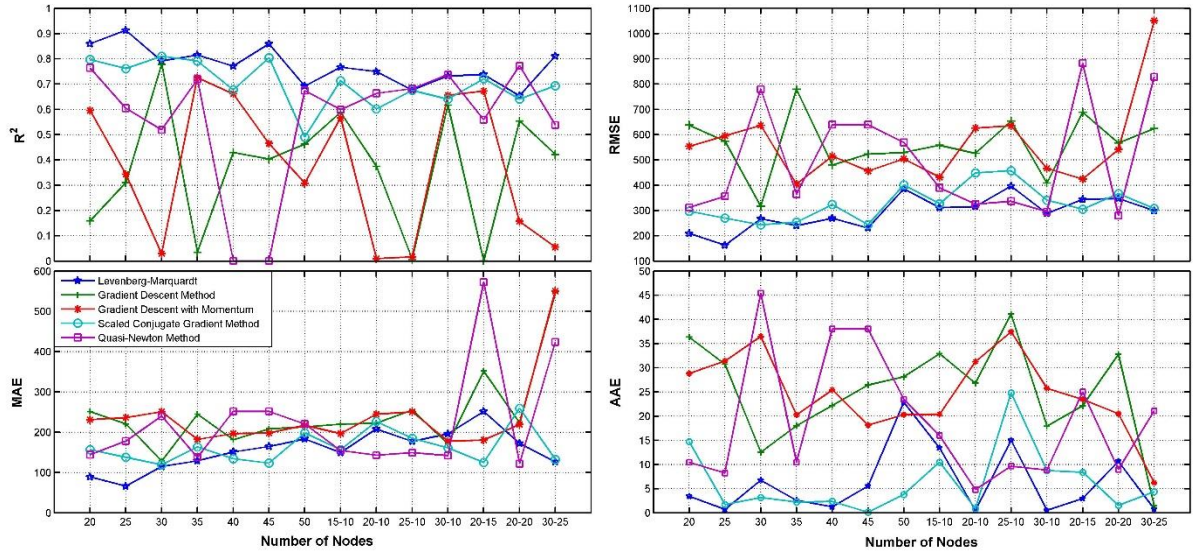


Figure 1-9. BPNN performance with various hidden layer(s) and different numbers of neurons. For one hidden layer, neurons changed from 20 to 50; while for two hidden layers, neurons in first layer changed from 15 to 30 and neuron in second layer changed from 10 t

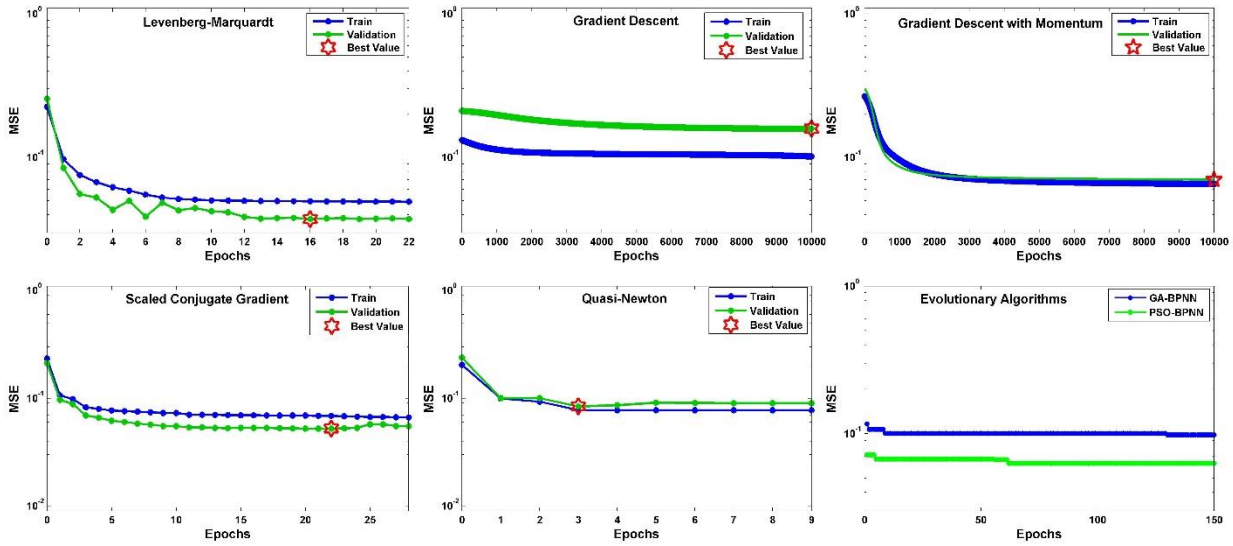


Figure 1-10. MSE decline curves of BPNN with 5-25-1 architecture which are trained by five supervised learning algorithms and two evolutionary intelligent algorithms.

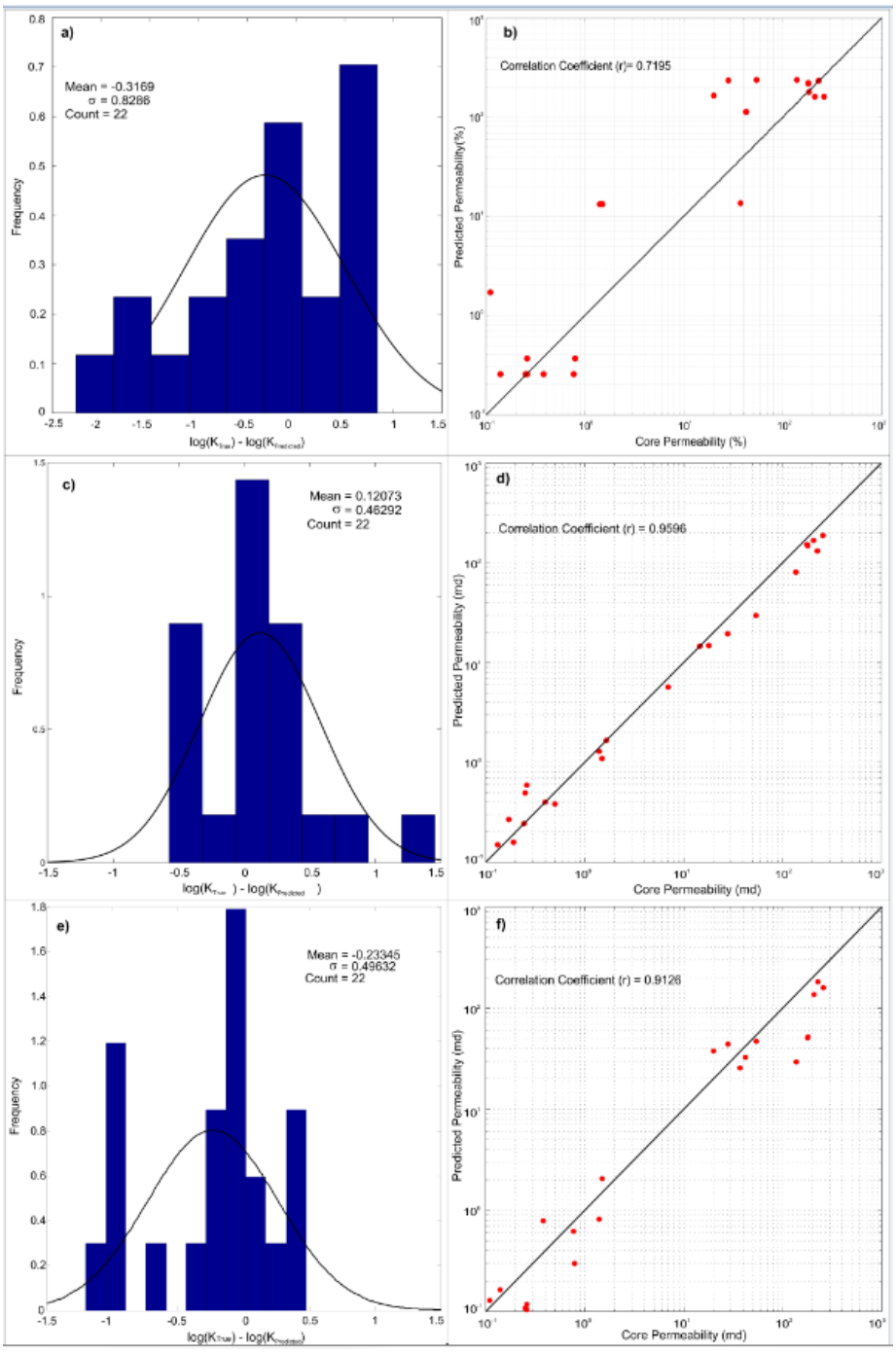


Figure 1-11. Histogram and cross-plot displaying the difference between the permeability values obtained from core measurement and back-propagation neural network. Figures (a) and (b) show the difference between the core measured permeability and predicted permeability by standard BPNN; (c) and (d) show the difference between the core measured permeability and predicted permeability by PSO-BPNN; (e) and (f) show the difference between the core measured permeability and predicted permeability by GA-BPNN.

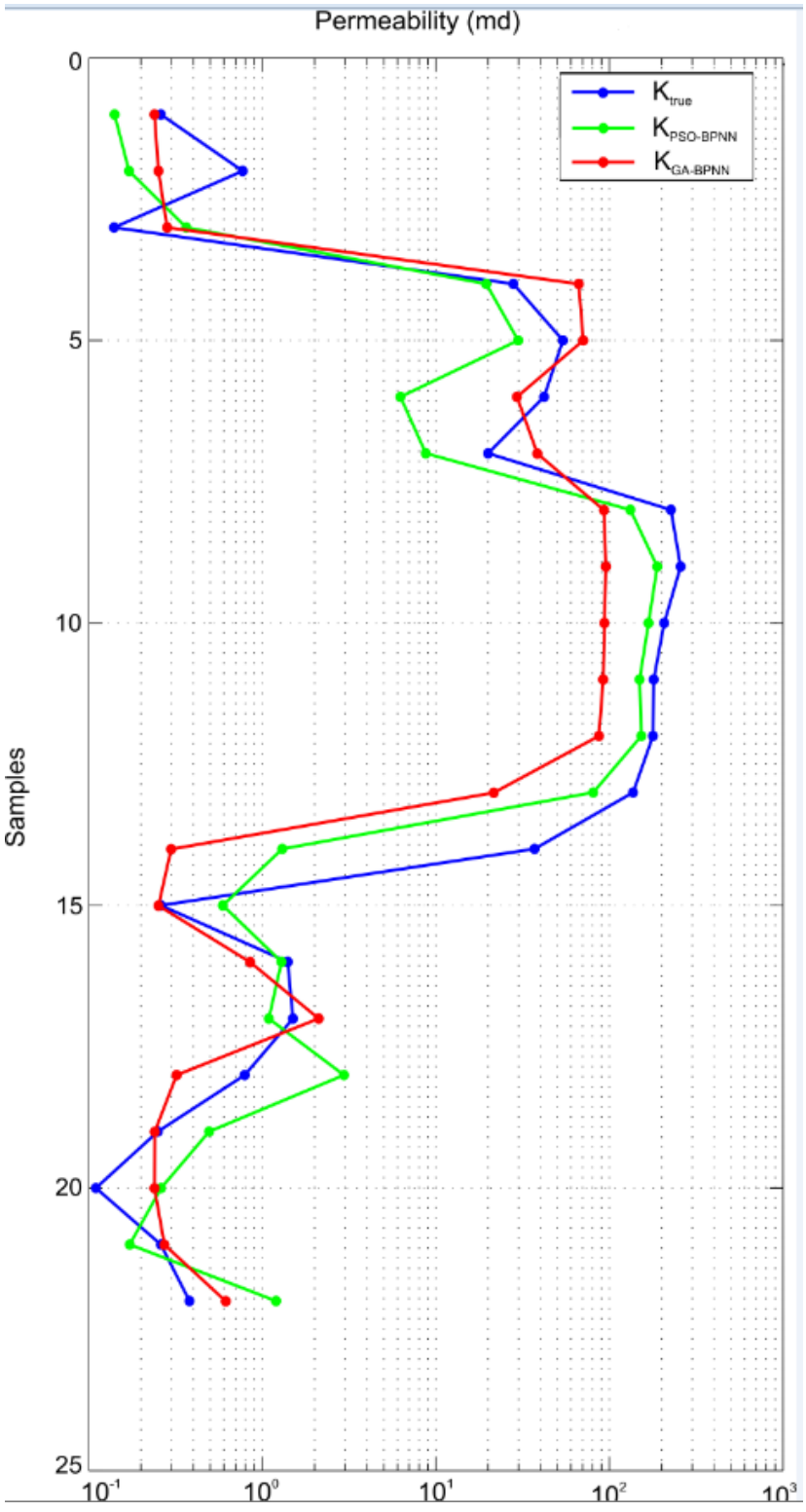


Figure 1-12. Predicted permeability by the artificial neural network with particle swarm optimization and genetic algorithm compared to permeability values from core data.

Table 1-1. Typical parameters used for training BPNN regression model with different learning algorithms including genetic algorithm (GA-BPNN) and particle swarm optimization (PSO-BPNN).

Parameter	Value	Parameter	Value
Input layers nodes	5	Hidden layers nodes	4
Output layers nodes	1	Termination criterion (e)	0.001
Special parameters for genetic algorithm			
Maximum Generation	150	Population size	100
Selection probability	Rand value (0,1)	Crossover probability	0.4
Mutation probability	0.1	Chromosome length	176
Special parameters for particle swarm optimization			
Maximum Generation	150	Population size	50
Cognitive efficient (ω_1)	1.5	Social efficient (ω_2)	1.5
Initial inertia weights W_{start}	0.9	Final inertia weight W_{end}	0.4

Table 1-2. Error measures for accuracy assessment to evaluate model performance.

Accuracy measure	Mathematical expression
Coefficient of Determination, R^2	$R^2 = 1 - \frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)}{\sum_{i=1}^{N}(y_i - averg(y_i))}$
Correlation coefficient, r	$r = \frac{\sum_{i=1}^{N}(y_i - \bar{y}_i)(\hat{y}_i - \bar{\hat{y}}_i)}{\sum_{i=1}^{N}(y_i - \bar{y}_i)^2 \sum_{i=1}^{N}(\hat{y}_i - \bar{\hat{y}}_i)^2}$
Root mean square error, RMSE	$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}$
Average error, AE	$AE = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)$
Maximum absolute error, MAE	$MAE = \max\lvert y_i - \hat{y}_i \rvert, i = 1,2,...,N$

Table 1-3. Permeability prediction results by BPNN with various hidden layer(s) and different numbers of neurons. LM: Levenberg–Marquardt; GD—Gradient descent; GDM—gradient descent with momentum; SCG—scaled conjugate gradient; BFGS-- Quasi-Newton method r (correlation coefficient) is used to evaluate the performance of BPNN with different architecture. The bold color indicates the highest correlation coefficient.

Architecture / Algorithms	Back-Propagation Topology Architecture													
	One hidden layer							Two hidden layers						
	20	25	30	35	40	45	50	15-10	20-10	25-10	30-10	20-15	20-20	30-25
LM	0.927069	**0.955522**	0.88925	0.902969	0.878167	0.926796	0.832022	0.875672	0.865662	0.822324	0.855567	0.859015	0.808497	0.900868
GD	0.398776	0.55715	0.882161	0.185025	0.654926	0.634988	0.679796	0.767001	0.612204	0.083334	0.78537	0.007926	0.744505	0.649279
GDM	0.771949	0.585894	0.172796	0.851726	0.813062	0.681737	0.554836	0.752989	0.097994	0.128824	0.809781	0.820165	0.395699	0.235026
SCG	0.892977	0.872495	0.900268	0.889572	0.823224	0.896662	0.70027	0.844629	0.775671	0.821716	0.800213	0.849063	0.800546	0.832869
BFGS	0.873924	0.77756	0.72043	0.848095	0.003734	0.029147	0.821091	0.774067	0.814255	0.825858	0.8586	0.747789	0.87878	0.733683

Table 1-4.Statistical parameters of the developed Artificial Neural Networks to determine the permeability for well B-18.

	ME (Mean Error)	STD (Standard Deviation Error)	RMSE (Root of Mean Square Error)	R^2 (Coefficient of Determination)
BPNN	28.5246	72.4197	166.2519	0.5177
PSO-BPNN	19.4389	26.6584	137.5647	0.9208
GA-BPNN	26.3529	60.825	158.9218	0.8328

CHAPTER 2

Application of A New Hybrid PSO-Mixed Kernels Function (MKF) Based Support Vector Machine (SVM) Model for Reservoir Porosity Prediction: A Case Study in Jacksonburg-Stringtown Oil Filed, West Virginia, USA

Zhi Zhong and Timothy R Carr[1]

[1] Department of Geology and Geography

West Virginia University, Morgantown

WV-26505, USA

(304)-282-9243

Email: zizhong1990@gmail.com; tim.carr@mail.wvu.edu

Abstract

Porosity is a fundamental property that characterizes the storage capability of fluid and gas bearing formations in a reservoir. An accurate porosity value can be measured from core sample in the laboratory, however, core analysis is expensive and time-consuming. Usually, the available core is limited to incomplete vertical sampling in only a few wells in a field. Well logs can be used to calculate porosity, but in mature fields availability of log suites are often limited in types making control a challenge for the theoretical relationships resulting from natural lithologic heterogeneity and to a lesser degree fluid content in the pores. Therefore, robust porosity prediction requires integration of core-measured porosity with modern well log suites to control for changes in lithology and fluid content. In this study, a support vector machine (SVM) model is used to improve estimates of porosity by constructing the relationship between limited conventional well log suites and sparse core data. The kernel function is the key technology in SVM, different kernel functions are applied to construct a reasonable SVR model. A new mixed kernel function is introduced that is a convex combination of the radial basis function kernel and the polynomial kernel function. This mixed kernel function not only preserves a strong extrapolating ability extended from radial basis function kernel but also possesses good interpolating capacity inherited from polynomial function kernel. Porosity, the desired output, uses two conventional well log responses (gamma ray and bulk density) and three well log derived parameters (slope of gamma ray, slope of density, and V_{sh}) as the input training and testing parameters. A grid searching method was applied to find the best control parameter (gamma and C) for each normal kernel function, which determines the performance of SVM. However, the extended computation time to find appropriate values of five control parameters in SVR with mixed kernels function restricts applications. Therefore, a global stochastic searching algorithm, particle swarm optimization, was applied to improve the efficiency of locating the global optimum. The results of SVM with different kernel functions were compared, and the SVM model with a mixed kernel function provided an improvement over the SVR with a single kernel. To confirm the advantage of the hybrid PSO-MKF-SVM model, results from three models: 1) radial basis function based least square support vector machine (RBF-LS-SVM), 2) multilayer perceptron artificial neural network (MLP-NN), and 3) radial basis function artificial neural network (RBF-NN), are compared with the result of the hybrid PSO-MKF-SVM model

and traditional SVM model with radial basis kernel function and polynomial kernel function. The results indicate that the hybrid PSO-MKF-SVM model improves porosity prediction with highest correlation coefficient (r of 0.9560), highest coefficient of determination (R^2 of 0.9140), lowest root mean square error (RMSE of1.6505), average absolute error value (AAE of 1.4050) and maximum absolute error (MAE of 2.717).

Key Words: support vector regression; mixed kernel function; radial basis function neural network; multilayer perceptron neural network; least squares support vector machine.

2.1. Introduction

A petroleum reservoir is a heterogeneous geological system with large intrinsic complexity that can be considered as a nonlinear regression problem (Al-Anazi and Gates, 2010c; Saljooghi and Hezarkhani, 2014). Porosity is a key parameter for characterizing the storage capacity of liquid and gas bearing formations. Accurate porosity values can be measured in the laboratory from reservoir core samples, however, acquisition and analysis are expensive and time-consuming. The number and density of core measurements in a reservoir or field are normally very limited and does not provide an adequate sampling of the reservoir. Therefore, a cheaper and faster method to estimate porosity is necessary. Porosity is estimated from well logs, such as bulk density, neutron porosity, and sonic, but many of these logs are not widely available in wells of mature reservoirs. In addition, wells without a modern suite of well logs can provide erroneous porosity because the parameters of theoretical physical models or empirical equations are not controlled for natural heterogeneity of lithology and fluid content, and nonlinearity of reservoir (Helle et al., 2001; Huang and Williamson, 1997). Without modern log suites, the spatial relationships of different reservoir properties affecting porosity determination are difficult to quantify (Verma et al., 2012).

Methods applied to construct reasonable porosity models with high accuracy and strong generalization ability vary from ‘hard computation’, such as empirical prediction (“rules of thumb”), and multilinear regression (Bloch, 1991; Byrnes and Wilson, 1991; Scherer, 1987; Wendt et al., 1986), to computer-based intelligence ‘soft computation’, which include neural network and machine learning (Al-Anazi and Gates, 2010b, c; Lim, 2005; Lim and Kim, 2004; Ravandi et al., 2014; Verma et al., 2012; Zargari et al., 2013). One drawback of empirical

prediction is that it is strongly localized to a region or formation (i.e., the empirical equation works well only in a specific region or formation). Multilinear regression constructs a linear relationship between dependent variables and independent variable, which facilitate the analysis of the cause-and-effect relationships. However one should detect the exact relationship between each dependent variable and independent variable first, which is a challenge and experience dependence. Moreover, multilinear regression ignores the cross relationship (covariance) between each independent variables. Artificial neural network analysis (ANN), including radial basis function neural network, multilayer perceptron neural network, is computer-based intelligence method that has attracted attention for porosity prediction and shows strong generalization ability (Gardner and Dorling, 1998; Jeirani and Mohebbi, 2006). The most serious disadvantage of ANN is overfitting due to memorization of the training set (Geman et al., 1992; Scales and Snieder, 1998; Zhang et al., 1999). In addition, the training process minimizes empirical risk and leads the process to fall into a local minimum and induce large prediction errors. Though some global searching algorithms, such as genetic algorithm (Saemi et al., 2007) and particle swarm optimization algorithms (Zhong and Carr, 2016) are applied to optimize the searching process, the optimization algorithms remain prone to capture by a local minimum.

Support vector machine (SVM) is based on statistical-learning theory, and has been applied to pattern recognitions and function approximation in the petroleum industry, including lithofacies classification (Al-Anazi and Gates, 2010a; Al-Anazi et al., 2011; Wang et al., 2014); reservoir permeability and porosity prediction (Al-Anazi and Gates, 2010b, c), minimum miscible pressure prediction (Zhong and Carr, 2016). SVM was developed at AT&T Bell laboratories by Vapnik and co-workers (Boser et al., 1992; Burges, 1998; Cortes and Vapnik, 1995; Drucker et al., 1997; Smola, 1996; Vapnik et al., 1997). This algorithm was introduced to solve pattern recognitions problems by projecting the original nonlinear data into higher dimensional feature space by kernel functions to locate an optimal hyperplane that separates the data in the feature space (Vapnik et al., 1997). SVM was extended to regression problems to find an optimal hyperplane on which projected targets can be located within ε deviation in feature space (Fu and Cheng, 2011). To avoid large-scale quadratic programming problems the complexity of optimization process was reduced by introducing least-squares SVM (LS-SVM) (Hemmati-Sarapardeh et al., 2014; Wang and Hu, 2005). LS-SVM avoided the complexity of optimization process, but the kernels function's generalization ability was still not optimized.

Separating hyperplane, maximum-margin hyperplane, soft margin and kernel function are four basic concepts for understanding SVM (Noble, 2006). Different kernel types will determine the various performance of SVM. There are two kernel functions: one is local kernel function, such as radial basis function kernel (RBF), which has a strong ability to extrapolate; the other is global kernel function, such as the polynomial kernel function, which is good at interpolating (Huang et al., 2012; Lian et al., 2013). A mixture of polynomial and RBF kernels will have better performance than either single one (Smits and Jordaan, 2002). In this study, two conventional well log (gamma ray, bulk density) and three derived-logs (slope of gamma, slope of bulk density, V_{sh}) are used as input data, and porosity as output data to train and test SVM.

Support vector machine (SVM) with mixed kernels function, a variant of SVM, is constructed for reservoir porosity prediction from conventional well logs. Particle swarm optimization (PSO), one of the evolutionary algorithms is introduced to optimize the SVM's structure parameters (i.e., mixing coefficient, penalty(C), gamma, epsilon and polynomial degree). More specifically, SVM with mixed kernels function is compared to SVM with regular kernel function, such as linear, polynomial, radial basis function, and sigmoid kernel function.

2.2 Methodology

2.2.1 Artificial Neural Network (ANN)

Artificial neural networks are a branch of artificial intelligence. Multilayer perceptron neural network (MLP-NN) and radial basis function neural network (RBF-NN) are two most wildly used artificial neural network. They are both the feed-forward neural network and can be used in the similar application with different performing structures.

2.2.1.1 Multilayer Perceptron Network (MLP-NN)

Multilayer perceptron neural network (MLP-NN), one type of feed-forward neural network, consists of three different kinds of layers. The first layer is a single input layer and the last layer is a single output layer. The number of input neurons and output neurons is problem dependent. One or more hidden layers are located between input and output layers. The number of hidden layers and number of hidden layer's neurons vary, depending on the complexity of problem and

training dataset' quality and size (Majumdar et al., 2008). A small number of neurons in hidden layers may lead the network to fall into a local minimum, in which case the network does not have sufficient time to learn the dataset's feature. Conversely, a large number of neurons will result in overfitting the network, in which case network does learn but memorizes. These challenges can result in trained networks losing their generalization ability (Gardner and Dorling, 1998; Jeirani and Mohebbi, 2006). A fully connected multilayer perceptron with two hidden layers is shown in Figure 2-1a, which represents a nonlinear mapping between an input vector and output vector (Gardner and Dorling, 1998).

Learning processes are the main component of the MLP-NN training process, in which weights and biases are adjusted continually until expected output is produced or anticipated criteria are achieved. The back-propagation training algorithm is the pivotal algorithm for the training process. Gradient descent algorithm assisted in finding the best weights and biases at which the minimum error between the desired and actual output is achieved. The testing process is used to assess the generalization ability of the trained neural network (Saemi et al., 2007). The designation of neural network architecture is a subjective task and problem-dependent. Since the number input layer (five input parameters) and output layer (porosity) is constant, the hidden layer neurons contribute a significant part to the performance of the MLP-NN as they behave as feature detectors (Tatar et al., 2016).

2.2.1.2 Radial Basis Neural Network (RBF-NN)

Radial basis function neural network (RBF-NN) is a special type of feed-forward neural network and is based on localized basis function and iterative function approximation (Tatar et al., 2013). RBF-NN consists of only three layers: input, hidden and output layer (Aljarah et al., 2016; Gardner and Dorling, 1998). The input layer is composed of an input vector. With only one hidden layer between input and output layer, the number of hidden neurons are strongly determined based on the specific problem. The activation function for each hidden neuron is RBF function, which calculates the similarity between the input and a stored prototype in that neuron. In order to increase the accuracy of the model in training and testing process, more prototypes should be used in the hidden layer. This process not only increases the complexity of decision boundary but also increases the computation time to evaluate the network (Aljarah et

al., 2016). The output layer is a linear combination of previous results from hidden layer. Compared with MLP-NN, RBF-NN has a simpler structure, and improved generalization, higher tolerance of input noises and the ability of online learning (Singh and Rao, 2005). Figure 2-1b displays the structure of RBF-NN, in which X_i is the input vector, b_i is biases, and ϕ_n is the activation function. The activation function is formed as following:

$$\phi_n(\|x - c_i\|) = \exp(-\frac{\|x - c_i\|^2}{2\sigma_i^2}) \quad , \qquad \text{Eq. (1),}$$

where the c_i is prototype of center of the i^{th} hidden neuron, σ is the bandwidth of i^{th} kernel node and $\|x - c_i\|$ denotes the Euclidean norm.

2.2.2 Support Vector Machine (SVM)

Support vector machine, based on the statistic-learning theory, was first proposed in Russia during the 1960s (Vapnik, 1963; Vapnik and Chervonenkis, 1964; Vapnik and Kotz, 1982). It is constructed to deal with pattern recognitions problems, where it uses adaptive margin-based loss functions, projects the learning data (non-linearly) into higher dimensional feature space, and locates a decision rule with good generalization ability (Zhong and Carr, 2016). The SVM decision rule actually consists of classification functions that are expanded on a subset of support vectors (Al-Anazi and Gates, 2010c; Boser et al., 1992; Cortes and Vapnik, 1995; Scholkopf et al., 1997; Vapnik et al., 1997). The projecting functions are called kernel functions – $K(x_i, x_j)$, which are satisfied with the Mercer's condition. Support vector regression (SVR) is another variant of support vector machines, which involves nonlinear regression and time series prediction (Drucker et al., 1997; Müller et al., 1997; Smola, 1996; Smola et al., 1998; Vapnik et al., 1997).

Assume we are given training data $\{(x^1, y^1), (x^2, y^2), ..., (x^m, y^m)\} \subset \chi \times \Re$, where χ denotes the space of the input patterns (e.g. $\chi = \Re^d$). For the case of a nonlinear function f, taking the form:

$$f(x) = \langle \omega, \varphi(x) \rangle + b, \qquad \omega \in \chi, b \in \Re \qquad \text{Eq. (2),}$$

where $\langle .,. \rangle$ represent dot product; $\varphi(x)$ represent the nonlinear function that applies linear regression; b and ω are bias terms and weight vector, respectively.

The SVR basic concept projects the original nonlinear data into higher n-dimensional feature space, and then a linear model- $f(x)$, is established in this feature space. The linear regression hyperplane in high dimensional feature space is in reality a nonlinear regression hypersurface in original input space (Asoodeh and Bagheripour, 2013; Na'imi et al., 2012). As first described, the objective of ε - SVM regression model is to find a function f(x) by which the deviations between estimated values of output and actual training output data equal to or less than ε (Vapnik et al., 1997). The complexity of the regression functions is essentially controlled by ε. In other words, the smaller value it is, the larger portion of the training data will be penalized, which will generate a tighter SV regression model; while the larger ε is, the smaller portion of training data that will be penalized, which will produce a looser SV regression model. It is like a tube, into which errors are accepted while the points will be penalized if any deviation is larger than ε and falls outside the tube (Smola and Schölkopf, 2004). Slack variables ξ_i and ξ_i^* have been introduced as asymmetric bound to satisfy constraints on the function instead of the 'hard margin' lose function (Bennett and Mangasarian, 1992). The SVM for regression using a kernel function and the ε -insensitive loss function is formulated as:

$$\min_{\omega} \quad \frac{1}{2}\|\omega\|^2 + C\sum_{i=1}^{m}(\xi_i + \xi_i^*) \qquad \text{Eq. (3),}$$

$$s.t. \quad \begin{cases} y_i - \langle \omega, \varphi(x_i) \rangle - b \le \varepsilon + \xi_i, i = 1,2,.....m \\ \langle \omega, \varphi(x_i) \rangle + b - y_i \le \varepsilon + \xi_i^*, i = 1,2,.....m \\ \xi_i, \xi_i^* \ge 0, i = 1,2,....m \end{cases} \qquad \text{Eq. (3a),}$$

The first term of Eq. (3) is the Vapnik-Chervonenkis (VC) confidence interval, whereas the second one is the empirical risk (Al-Anazi and Gates, 2010c; Maleki et al., 2014). The regularization constant C in Eq. (3) not only decides the complexity of the SVM model but also makes a compromise between the confidence degree and the empirical risk minimization. An improper C value will weaken the generalization capability of an SVM (Yuan and Chu, 2007). Vapnik's ε-insensitivity loss function defines a tube in high dimensional feature space. The

points outside the ε-tube contribute to the loss, which is equal to the gap between the point values and the radius of the tube. However, the points inside the ε-tube contribute nothing to the cost, which means the loss (error or cost) is zero (Zhong and Carr, 2016) (Figure 2-2, left).

In order to solve the mathematical optimization problem given by Eq. (3), it is necessary to construct Lagrange function from the primal function and corresponding constraints (Bazaraa et al., 2013; Luenberger and Ye, 2008; Smola, 1996). A dual formulation transformed from it primal function is generated by introducing a dual set of variables. The saddle point in this function represents the solution of primal and dual variables. The new dual objective function is defined as following:

$$\max_{\alpha,\alpha^*}\min_{\omega} L(w,b,\xi,\xi^*,\alpha,\alpha^*,\beta,\beta^*) = \frac{1}{2}\|\omega\|^2 + C\sum_{i=1}^{m}(\xi_i+\xi_i^*) - \sum_{i=1}^{m}(\beta_i\xi_i+\beta_i^*\xi_i^*) - \sum_{i=1}^{m}\alpha_i\,(\varepsilon+\xi_i-y+<\omega,\varphi(x_i)>+b) - \sum_{i=1}^{m}\alpha_i^*(\varepsilon+\xi_i^*-y+<\omega,\varphi(x_i)>+b) \quad \text{Eq. (4),}$$

where is L the Lagrange function and α*, α, β* and β are positive Lagrange multipliers, which represent the virtual forces resulting from the constraints of primal problem. Based on Karush-Kuhn-Tucker (KKT) theorem, only active constraints may result in Lagrange multipliers not equal to zero, which means only the data points with non-vanishing Lagrange multipliers have to be taken into account (Smola, 1996), meanwhile the partial derivatives of L with respect to the primal variable (ω,b,ξ,ξ^*) vanish at the optimum (actually a saddle point).

$$\begin{cases} \dfrac{\partial L}{\partial b} = 0 \rightarrow \sum_{i=1}^{m}(\alpha_i^*-\alpha_i) = 0 \\ \dfrac{\partial L}{\partial \omega} = 0 \rightarrow \omega - \sum_{i=1}^{m}(\alpha_i^*-\alpha_i)\varphi(x_i) = 0 \\ \dfrac{\partial L}{\partial \xi_i^{(*)}} = 0 \rightarrow C-\alpha_i^{(*)}-\beta_i^{(*)} = 0 \end{cases} \quad \text{Eq. (5),}$$

Substituting equation (4) into equation (3) generates the following dual optimization problem.

$$\max_{0<\alpha,\alpha^*<C} \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{m}(\alpha_i-\alpha_i^*)(\alpha_j-\alpha_j^*)\langle\varphi(x_i),\varphi(x_j)\rangle - \varepsilon\sum_{i=1}^{m}(\alpha_i+\alpha_i^*) + \sum_{i=1}^{m}y_i(\alpha_i-\alpha_i) \quad \text{Eq. (6),}$$

$$s.t. \qquad \sum_{i=1}^{m}(\alpha_i^* - \alpha_i) = 0 \qquad \text{Eq. (6a),}$$

After elimination of ω, the regression model is defined as following:

$$f(x) = \sum_{i=1}^{m}(\alpha_i - \alpha_i^*)\langle\varphi(x_i), \varphi(x_j)\rangle + b \qquad \text{Eq. (7),}$$

Based on Mercer's condition, the inner product $\langle\varphi(x_i), \varphi(x_j)\rangle$ can be defined through a kernel $K(x_i, x_j)$. So substituting $K(x_i, x_j)$ into Equation 7 the support vector expansion of regression estimation model becomes

$$f(x) = \sum_{i=1}^{m}(\alpha_i - \alpha_i^*)K(x_i, x_j) + b \qquad \text{Eq. (8),}$$

2.2.2.1 Least Squares Support Vector Machine (LS-SVM)

As discussed above, resolving a large-scale quadratic programming problem is inevitable. To deal with this challenge, Least-Squares SVM (LS-SVM) is introduced, which is modified from the traditional SVM (Suykens and Vandewalle, 1999, 2000). Avoiding large-scale quadratic programming problems, this technique solves linear equations and reduces the complexity of optimization process (Hemmati-Sarapardeh et al., 2014; Wang and Hu, 2005). In this LS-SVM technique, Suykens and Vandewalle (Suykens and Vandewalle, 1999, 2000) reformulated the SVM as follows:

$$\min_{\omega} \qquad \frac{1}{2}\|\omega\|^2 + \gamma\sum_{i=1}^{m}e_i^2 \qquad \text{Eq. (9),}$$

$$s.t. \qquad y_i = \langle\omega, \varphi(x_i)\rangle + b + e_i, i = 1,2,.....m \qquad \text{Eq. (9a),}$$

Where $\gamma > 0$ is a regularization constant; and e_i are error variables for i^{th} output. After the primal function is transformed into its dual formulation, the Lagrange for this problem is as follows:

$$\max_{\alpha}\min_{\omega} L(w,b,e,\alpha)=\frac{1}{2}\|\omega\|^2+\frac{1}{2}\gamma\sum_{i=1}^{m}e_i^2-\sum_{i=1}^{m}\alpha_i\,(e_i+<\omega,\varphi(x_i)>+b-y_i) \quad \text{Eq. (10)},$$

where L is the Lagrange function, and α is positive Lagrange multipliers? Based on Karush-Kuhn-Tucker (KKT) theorem, only active constraints may result in Lagrange multipliers not equal to zero. Thus the partial derivatives of L with respect to the primal variable (ω,b,α_i,e_i) are determined as following:

$$\begin{cases}\dfrac{\partial L_{LS\_SVM}}{\partial b}=0\rightarrow\sum_{i=1}^{m}\alpha_i=0\\ \dfrac{\partial L_{LS\_SVM}}{\partial \omega}=0\rightarrow\omega-\sum_{i=1}^{m}\alpha_i\varphi(x_i)=0\\ \dfrac{\partial L_{LS\_SVM}}{\partial e_i}=0\rightarrow\alpha_i-\gamma e_i=0\\ \dfrac{\partial L_{LS\_SVM}}{\partial \alpha_i}=0\rightarrow e_i+<\omega,\varphi(x_i)>+b-y_i=0\end{cases} \quad \text{Eq. (11)},$$

These conditions are similar to traditional SVM optimality conditions in Eq. (5), except for the condition $\alpha_i=\gamma e_i$, for which the sparseness property has been lost in LS-SVM (Figure 2-2 right). After elimination of ω and e, the inner product $\langle\varphi(x_i),\varphi(x_j)\rangle$ can be defined through a kernel $K(x_i,x_j)$. So the function of LS-SVM model becomes:

$$f(x)=\sum_{i=1}^{m}\alpha_i K(x_i,x_j)+b \quad \text{Eq. (12)},$$

2.2.2.2 Mixed Kernel Function Based Support Vector Machine (MKF-SVM)

The projection function in SVM model consists of kernel functions- $K(\vec{x},\vec{x}^*)$. Kernel function maps the original linearly or non-linearly learning data from original space into high dimensional feature space, in which all of the data can be presented linearly (Al-Anazi et al., 2011). A kernel function must meet Mercer's condition (Boser et al., 1992; Burges, 1998; Cortes and Vapnik, 1995; Smola and Schölkopf, 1998; Smola and Schölkopf, 2004), as following:

Suppose $K(\vec{x},\vec{x}^*) \in L_2(\Re^n) \otimes L_2(\Re^n)$

$\exists \forall g(x) \in L_2(\Re^n)$,

$$\iint_{L_2 \otimes L_2} K(\vec{x},\vec{x}^*) g(\vec{x}) g(\vec{x}^*) d\vec{x} d\vec{x}^* \geq 0 \qquad \text{Eq. (13),}$$

Based on this theorem, Table 2-1 listed four kernels, which determined the various characteristics of SVM model. According Smola and Schölkopf (1998), there are two types of kernels including local kernels and global kernels (Smola and Schölkopf, 1998). As shown in Figure 2-3a, data points that are far away from each other have a significant influence on the kernel values in a global kernel. One typical example of a global kernel is polynomial kernel in Eq. (14) (Smits and Jordaan, 2002; Zheng *et al.*, 2004).

$$K_{ploy}(x_i, x_j) = (\langle x_i, x_j \rangle + 1)^d \qquad \text{Eq. (14),}$$

Kernel's interpolation capacity is positively correlated to the degree of polynomial kernel, while its extrapolation capacity is negatively correlated to the degree of polynomial kernel. As Figure 2-3b displays, a local kernel allows data points that are close to each other to have a significant influence on the kernel values. The radial basis function kernel (RBF) is the typical local kernel, as defined in Eq. (15).

$$K_{rbf}(x_i, x_j) = \exp(-\gamma \left\| x_i - x_j \right\|^2) \qquad \text{Eq. (15),}$$

Local kernel's interpolation ability is positively correlated to the gamma (γ) value. The smaller value of γ is, the worse its interpolation ability will be and vice versa. In another word, no single value of kernel parameter polynomial degree or γ will provide a model with both strong interpolation and extrapolation properties (Smits and Jordaan, 2002).

One corollary can be generated based on the Mercer's conditions. Assuming $K_1(\vec{x},\vec{x}^*)$, $K_2(\vec{x},\vec{x}^*)$ are admissible support vector kernels and $C_1, C_2 \geq 0$,

$$K(\vec{x},\vec{x}^*) = C_1 K_1(\vec{x},\vec{x}^*) + C_2 K_2(\vec{x},\vec{x}^*) \qquad \text{Eq. (16),}$$

$K(\vec{x},\vec{x}^*)$ is an admissible kernel [55]. Numerous investigators have proposed the mixed kernels function $K(\vec{x},\vec{x}^*)$, which combines the good characteristics of both $K_1(\vec{x},\vec{x}^*)$ and $K_2(\vec{x},\vec{x}^*)$ (Huang et al., 2012; Lian et al., 2013; Smits and Jordaan, 2002; Smola and Schölkopf, 2004; Yang et al., 2013; Zhu et al., 2005).

The mixtures of the RBF and polynomial kernels can be defined as Eq. (17, 18)

$$K_{mix} = mK_{poly} + (1-m)K_{rbf} \qquad 0 \le m \le 1 \qquad \text{Eq. (17),}$$

$$K_{mix}(x_i, x_j) = m(\gamma_1 \langle x_i, x_j \rangle + 1)^d + (1-m)\exp(-\gamma_2 \left\| x_i - x_j \right\|^2) \quad 0 \le m \le 1 \quad \text{Eq. (18),}$$

where K_{poly} is a polynomial kernel and K_{rbf} is a radial basis function kernel, m is the mixing coefficient. If m=1, then the mixed kernels function will become the polynomial kernel, which means $K_{mix} = K_{ploy}$. If m=0, then the mixed kernels function will equal the RBF kernel, which means $K_{mix} = K_{rbf}$. The effect of the mixing a polynomial kernel with an RBF kernel is shown in Figure 2-3c, where $x_i = 0.2, \gamma = 0.1, d = 1$. The mixed kernels function has characteristics of both the polynomial kernel and the RBF kernel and improves fitting and generalizing ability.

2.2.3 The Proposed Hybrid PSO-MKF-SVM Model

Based on stochastic search and optimization processes, Eberhart and Kennedy (1995) first introduced the concept of particle swarm optimization algorithm (PSO) (Eberhart and Kennedy, 1995). PSO imitated human (or insect) social behavior as a typical evolutionary algorithm. Individuals interact with one another while learning from their own experience, and gradually the population members' move into better regions of the problem space (Eberhart and Kennedy, 1995; Zhong and Carr, 2016). Particles indicate the potential solution and randomly locate in the architecture space (like birds randomly distributed in a specific open environment). Global optima of fitness function for each generation can be calculated based on the position of particles. Because the selection of kernel function parameters determine the performance of

SVR, both particle swarm optimization and n-fold cross-validation re-sampling method are employed to optimize the parameters, which include mixing coefficient (ρ), penalizing coefficient (C), RBF kernel parameter (ϒ), polynomial degree (d) and (ε) in ε-insensitivity function. This algorithm starts at the random position, by updating velocity and position repeatedly, the particles will search the global extremum in the searching space. In this paper, the positions of particles, P (ρ, ϒ, d, C, ε), represent the values of parameters which will be optimized. In the actual optimization process, the SVM parameter ρ is limited in [0, 1], ϒ in [2^{-4}, 2^{4}], d in [1, 3], C in [2^{-5}, 2^{5}] and ε in [0.001, 0.01], and the PSO searches the optimal values of the SVM parameters in these areas. Figure 2-4 illustrate the detailed process of PSO-MKF-SVM model for porosity prediction. Fitness function for PSO algorithm is MSE, which is formulated as:

$$F_{Fitness} = \frac{1}{m}\sum_{i=1}^{m}(\hat{y}_i - y_i)^2 \qquad \text{Eq. (19)},$$

where $\hat{y}_i$ is predicted value, y_i is true value and m is number of samples. Figure 2-4 shows the workflow for the optimization of the parameter in mixed kernel function. For more details about PSO and workflow, referring (Eberhart and Kennedy, 1995; Poli et al., 2007; Sayyad et al., 2014).

2.2.4 Predictive Performance Evaluation Index

In order to test that the newly hybrid PSO-mixed kernel function based support vector machine (MKF-SVM) performs better than other models, five evaluation indices including correlation coefficient (r), coefficient of determination (R^2), average absolute error (AAE), root mean square error (RMSE), and maximum absolute error (MAE) were chosen to evaluate the performance of various regression models (Jiao et al., 2016; Zhong and Carr, 2016) (Table 2-2). An R^2 (Coefficient of Determination) of 0 indicates a completely inaccurate model, while an R^2 of 1 indicates a perfect regression model. Moreover, RMSE is used to evaluate overall performance, while AAE and MAE are used to determine the error range of the predicted results. When the values of R^2 is higher and the values of AAE, MAE and RMSE are smaller, the forecast performance is better.

2.3 Case Study

2.3.1 Geological Background

Jacksonburg-Stringtown field is situated along the axis of the Burchfield syncline in northwestern West Virginia (Figure 2-5b). The estimated original oil in place (OOIP) in Jacksonburg-Stringtown oil field is 88 million barrels of oil (MMBO). Since discovered in 1895, estimated about 22 MMBO oil has been produced, and the estimated oil recovery factor is 25%. The primary and/or secondary producing reservoir unit is Late Devonian Gordon Stray interval, which is contained within the Middle to Late Catskill deltaic complex (Catskill delta). During Early to Middle Devonian, crustal uplift in Acadian orogeny lead to further subsidence within the Appalachian foreland basin to the west and resulted in the deepening of the central Appalachian basin (Faill, 1997a, b). Deposition coincided with the heavy rainfall produced by the tropical climate, sedimentary deposition accelerated during the Middle and Late Devonian (Blakey, 2008; McBride, 2004; Piotrowski and Harper, 1979). In the Late Devonian, five major delta systems prograded westwards and dominated the foreland ramp (Figure 2-5a). In the latest Devonian, maximum progradation of the Catskill delta complex was achieved west of the Acadian highlands (Boswell and Donaldson, 1988). During this period, Acadian tectonism ceased, and relative sea-level changes within the basin were controlled primarily by sea-level fluctuations and variations in sedimentation in an estuary (Coughlin, 2009). Non-marine red shale and most of the low-energy alluvial deposits are concentrated in the eastern portions of the Appalachian basin. Non-marine sediments increasingly advanced westward to cover marine beds, and near-shore deposits continued to prograde into the central Appalachian basin. In the area of Jacksonburg-Stringtown field, Gordon stray/Gordon interval are interpreted to be shoreline/shore-face sandstone that occupied a broad structural trend at the time of maximum regression of the Acadian clastic wedge (Hohn, 2004).

Gas re-injection program took place in the field beginning in the mid-1930's and ended in the 1950's. First pilot waterflood program with a 35 acre dual 5-spot well pattern was conducted in 1981. After 1990, a full-scale waterflood was installed over a large portion of the field (Bergerud, 2011). Unlike modern oil and/or gas fields, which have abundant, high-quality data, including conventional and advanced well logs, seismic data to construct 3D geological model,

Jacksonburg-Stringtown oil field has plentiful low quality well log data (i.e. gamma-ray, bulk density), and limited high quality core data (core measured porosity and permeability data) for reservoir modeling. The sparse modern subsurface data can constrain the development of a robust geological reservoir volumetric model and effective evaluation of potential of enhanced oil recovery and CO_2 geologic storage capacity.

2.3.2 Pre-processing of the Dataset

A total of 94 samples with both core data and conventional logs were collected from wells T.H8, H9, B19, B18, LM13 and H11 in the Jacksonburg-Stringtown field. The relative locations of each well are shown in Figure 2-5b. To establish a reliable PSO-MKF-SVM regression model, the dataset is divided into two parts: training and testing data. The procedure used to train and test support vector regression technique is summarized as follows:

(1) Training and testing data generation: there are a total 6 cases as shown in Table 2-3. In each case, data combined from five wells were used to train the PSO-MKF-SVM regression model, while remaining data of last well were applied to evaluate the accuracy and stability of the trained regression model. Based on the materials on hand, two conventional well logs and three log-derived variables were chosen as PSO-MKF-SVM input parameters and core-based porosity as a scalar output. Two well logs are gamma-ray log (GR), bulk density log, and three derived variables include the slope of GR, the slope of density, and V_{sh}. Figure 2-6 shows the heterogeneity of Gordon Stray, there are no clear linear relationships between input parameters and output porosity value.

(2) Data normalization: Dataset normalization is an important process before training and testing proposed model in order to improve prediction accuracy and training speed. Absolute scale is used for all input parameters which are normalized into range of [0, 1]. Following normalized formula were chosen:

$$X_i^{new} = [\frac{X_i^{old} - X_{\min}}{X_{\max} - X_{\min}}] \qquad \text{Eq. (20),}$$

where X_i^{new} is the normalized input vector, $X_{\min}$ and $X_{\max}$ are the minimum and maximum and value respectively, X_i^{old} are the original input vector.

2.4 Results

In this study, improved ε-SVM with normal kernel function and mixed kernel function, listed in Table 2-1, are applied for constructing a regression model to predict reservoir porosity (φ) based on conventional well log data, including gamma ray, bulk density well log data and three log-derived data, such as slope of gamma ray, slope of density, and shale volume (V_{sh}). The LIBsvm toolbox (Chang and Lin, 2011) was used to assist in completing this study by Matlab software (MATLAB Release 2012b). Grid searching algorithm was used to search best control parameters of normal kernel functions as listed in Table 2-1. Table 2-4 displays the typical parameters range for training SVR. Based on experiments, 4-fold cross validation is more appropriate.

For SVR with traditional kernel function, such as linear, polynomial, RBF, sigmoid kernel function, finding the global optimum of gamma and C value is really a challenge. Figure 2-7 shows the process of optimizing penalty C and gamma value by grid searching algorithm for well case 1 as described in Table 2-2. The color of the contour lines in the figure indicates the mean square error (MSE), the blue color is smaller MSE, and the red color is larger MSE. The smaller MSE indicates a more accurate model.

However, it takes a long time for grid searching algorithm to determine a global optimum of all parameters in the mixed kernel function. The particle swarm optimization algorithm is applied to optimize the control parameters in mixed kernel function. Figure 2-8 shows optimization process of mixing coefficient (ρ), penalizing coefficient (C), RBF kernel parameter (ϒ), polynomial degree (d) and ε in ε-insensitivity function in mixed kernel function for well case 1.

As Figure 2-7 shows, the minimum MSE of SVR models with normal kernel function are greater than 4, however, the minimum MSE of SVR model with mixed kernel function is 2.1362 (Figure 2-8). The SVR model with mixed kernel function performance is improved over the SVR with normal kernel function. All ϒ and C values are exponentially 2 in SRV with linear, polynomial, sigmoid and RBF kernel. This phenomenon is caused by its optimizing algorithm. Grid searching algorithm can just increase searching steps by exponential 2, and cannot search the value between those values. But particle swarm optimization is based on stochastic searching, which means the parameters can take any value.

Figure 2-9 indicates the correlation coefficient of SVR for training and testing stage with five kernel functions for six well cases, respectively. High value of correlation coefficient demonstrates that the SVR method performed very well as a regression technique in both training and testing stage. Comparison results using coefficient of determination (R^2), correlation coefficient (r), root mean square error (RMSE), average absolute error (AAE) and maximum absolute error (MAE) are shown in Table 2-5. Table 2-5 list the RMSE, AAE and MAE for SVR with different kernel function for SVR training process for well case 1 through 6. Compared with the performance of other kernel functions, the mixed kernel function works better with lower errors. Table 2-6 list all the well-trained control parameters, which can be used for further testing and future prediction. As in Table 2-6 shown, a little change will influence the final result dramatically. Table 2-7 list the RMSE, AAE and MAE for SVR with different kernel function with well-trained parameters for well case 1 through 6. Though the errors of testing runs are larger than training runs, the performance of mixed kernel is an improvement over other kernel function.

2.5 Comparison and Discussion

To confirm the model's forecast ability, the comparison between the proposed model with the predicted results calculated from the newly hybrid PSO-MKF-SVM model is necessary. Those proposed models include MLP-NN, RBF-NN, RBF-SVM, POLY-SVM and RBF-LS-SVM models. Multilayer perceptron neural network is widely used in petroleum industry, especially in reservoir characterization problem (e.g. porosity and permeability prediction) because of its strong feature learning and generalization ability. The radial basis function neural network (RBF-NN) is the special improvement of MLP-NN, which just has one hidden layer and the active function of hidden neurons are a Gaussian radial basis function. The RBF-NN has strong approximation ability and high convergence speed. RBF-SVM and POLY-SVM are both standard support vector machine, but using a different kernel function. The RBF kernel function uses local kernels that have stronger interpolation ability, while polynomial kernel function belongs to a group of global kernels that have stronger extrapolation ability (Zhong and Carr, 2016). Different kernel function performance varies, so both kernel functions are selected to compare with newly developed hybrid PSO-MKF-SVM model. RBF-LS-SVM is further improved SVM with RBF kernel function. Because this technique applies last squared technique

to reduce complex optimization process, LS-SVM has fast learning speed and good generalization ability.

To evaluate the accuracy and stability of PSO-MKF-SVM model, three evaluation parameters (e.g. RMSE, MAE, AAE) serve as the criteria to assess the stability of PSO-MKF-SVM model and one parameter (R^2) were applied to evaluate the accuracy of this proposed hybrid model. Their effects are displayed in Figure 2-10, in this situation case 5 was applied. In training process, PSO-MKF-SVM model has highest R^2 and smallest RMSE, AAE, and MAE. Those indexes indicate that hybrid PSO-MKF-SVM model performs best. In contrast, MLP-NN performs worst (Figure 2-10a). In testing process, PSO-MKF-SVM model still has highest R^2 and smallest RMSE, AAE, and MAE (Figure 2-10b). The hybrid PSO-MKF-SVM is the best mode in those models. Other models are not accurate or stable neither because of low R^2 or large RMSE, AAE or MAE. The training process learns the feature from the training dataset, and testing process verifies the process, which reflects the generalization ability of a model. After detailed comparison of the all above-listed regression models, the hybrid PSO-MKF-SVM regression model is a reasonable, accurate and stable regression model because all indexes (RMSE, MAE, AAE and R^2) are lower. A comparison between core measured porosity and the predicted porosity by the six proposed models are displayed in Figure 2-11. Figure 2-11a displays a good agreement between core measured porosity and PSO-MKF-SVM predicted porosity with highest R^2 (0.9140) and r (0.9560), and lowest RMSE (1.6505), AAE (1.4050) and MAE (2.717).

2.6 Conclusions

For reservoir characterization and evaluation of applications such as enhanced oil recovery and CO_2 geologic storage in mature oil fields, accurate calculation of fluid and gas volumes is dependent on accurate porosity estimation from often very limited log and core data. It is necessary to develop an effective method to improve prediction with limited subsurface data of reservoir porosity across wells in a reservoir. The hybrid PSO-MKF-SVM regression model is introduced and is applied to predict the porosity in the Jacksonburg-Stringtown depleted oil field. The result demonstrates that this hybrid model is a potentially accurate and robust methodology to improve porosity prediction with limited subsurface data. In a comparison, the newly hybrid

PSO-MKF-SVM model outperforms the LS-SVM model, SVM model, BRF-NN and MLP-NN model. Through the foregoing analysis and discussion the prediction result, several conclusions were developed -

(1) The hybrid PSO-MKF-SVM regression model provides a reliable way of estimating porosity from limited conventional well log data.
(2) Support vector regression model performance varies due to different kernels functions and control parameters. Polynomial and RBF kernels are typical global and local kernels. Mixed kernel function takes advantages of RBF kernel function and polynomial kernel function to increase the applicability of SVM.
(3) PSO, as an evolutionary algorithm, can improve the efficiency and accuracy of locating the optimal value in searching space. However, caution is required since premature convergence and suboptimal search results can occur prior to locating the global optimal (Figure 2-8). Further investigation is needed.
(4) The comparison between hybrid PSO-MKF-SVM and other proposed models show that this hybrid model has excellent performance and great generalization ability with higher R^2 (0.9140) and r (0.9560), and smaller RMSE (1.6505), AAE (1.4050) and MAE (2.717).

Acknowledgement

This work was funded by the US-China Clean Energy Research Center, Advanced Coal Technology Consortium, under grant DE-PI0000017 from the National Energy Technology Laboratory of the US Department of Energy.

References

Al-Anazi, A., Gates, I., 2010a. A support vector machine algorithm to classify lithofacies and model permeability in heterogeneous reservoirs. Engineering Geology 114, 267-277.

Al-Anazi, A.F., Gates, I.D., 2010b. Support-vector regression for permeability prediction in a heterogeneous reservoir: A comparative study. Spe Reservoir Evaluation & Engineering 13, 485-495.

Al-Anazi, A.F., Gates, I.D., 2010c. Support vector regression for porosity prediction in a heterogeneous reservoir: A comparative study. Computers & Geosciences 36, 1494-1503.

Al-Anazi, A.F., Gates, I.D., Azaiez, J., 2011. Support vector machines for petrophysical modelling and lithoclassification.

Aljarah, I., Faris, H., Mirjalili, S., Al-Madi, N., 2016. Training radial basis function networks using biogeography-based optimizer. Neural Computing and Applications, 1-25.

Asoodeh, M., Bagheripour, P., 2013. Fuzzy classifier based support vector regression framework for Poisson ratio determination. Journal of Applied Geophysics 96, 7-10.

Bazaraa, M.S., Sherali, H.D., Shetty, C.M., 2013. Nonlinear programming: theory and algorithms. John Wiley & Sons.

Bennett, K.P., Mangasarian, O.L., 1992. Robust linear programming discrimination of two linearly inseparable sets. Optimization methods and software 1, 23-34.

Bergerud, B., 2011. Reservoir model of the Jacksonburg-Stringtown Oil field,Northwesrern West Virginia: potential for miscible CO_2 Enhanced Oil Recovery, Department of geology and geography. West Virginia University.

Blakey, R., 2008. Paleogeography and geologic evolution of North America, Northern Arizona University.

Bloch, S., 1991. Empirical Prediction of Porosity and Permeability in Sandstones (1). AAPG Bulletin 75, 1145-1160.

Boser, B.E., Guyon, I.M., Vapnik, V.N., 1992. A training algorithm for optimal margin classifiers, Proceedings of the fifth annual workshop on Computational learning theory. ACM, pp. 144-152.

Boswell, R.M., Donaldson, A.C., 1988. Depositional architecture of the Upper Devonian Catskill Delta complex: central Appalachian basin, U.S.A. Memoir Canadian Society of Petroleum Geologists 2, 19.

Burges, C.J.C., 1998. A tutorial on support vector machines for pattern recognition. Data mining and knowledge discovery 2, 121-167.

Byrnes, A., Wilson, M., 1991. Aspects of porosity prediction using multivariate linear regression. AAPG Bulletin (American Association of Petroleum Geologists);(United States) 75.

Chang, C.-C., Lin, C.-J., 2011. LIBSVM: a library for support vector machines. ACM Transactions on Intelligent Systems and Technology (TIST) 2, 27.

Cortes, C., Vapnik, V., 1995. Support-vector networks. Machine Learning 20, 273-297.

Coughlin, M.F., 2009. Subsurface mapping and reservoir analysis of the Upper Devonian Venango and Bradford groups in Westmoreland County, Pennsylvania, Department of Geology and Geography. West Virginia University, p. 114.

Drucker, H., Burges, C.J.C., Kaufman, L., Smola, A., Vapnik, V., 1997. Support vector regression machines. Advances in neural information processing systems 9, 155-161.

Eberhart, R.C., Kennedy, J., 1995. A new optimizer using particle swarm theory, Sixth international symposium on micro machine and human science. New York, NY, pp. 39-43.

Faill, R.T., 1997a. A geologic history of the north-central Appalachians, Part 1: Orogenesis from the Mesoproterozoic through the Taconic Orogeny. American Journal of Science 297, 551-619.

Faill, R.T., 1997b. A geologic history of the north-central Appalachians, Part 2: The Appalachian basin from the Silurian through the Carboniferous. American Journal of Science 297, 729-761.

Fu, Y., Cheng, Y., 2011. Application of an integrated support vector regression method in prediction of financial returns. International Journal of Information Engineering and Electronic Business (IJIEEB) 3, 37.

Gardner, M.W., Dorling, S., 1998. Artificial neural networks (the multilayer perceptron)—a review of applications in the atmospheric sciences. Atmospheric Environment 32, 2627-2636.

Geman, S., Bienenstock, E., Doursat, R., 1992. Neural networks and the bias/variance dilemma. Neural computation 4, 1-58.

Helle, H.B., Bhatt, A., Ursin, B., 2001. Porosity and permeability prediction from wireline logs using artificial neural networks: a North Sea case study. Geophysical Prospecting 49, 431-444.

Hemmati-Sarapardeh, A., Shokrollahi, A., Tatar, A., Gharagheizi, F., Mohammadi, A.H., Naseri, A., 2014. Reservoir oil viscosity determination using a rigorous approach. Fuel 116, 39-48.

Hohn, M.E., 2004. Petroleum reology and resrvoir characterization of the Upper Devonian gordon sandstone, Jacksonburg-Stringtown Oil Field, Northwestern West Virginia. West Virginia Geological and Economic Survey, Morgantown.

Huang, H., Ding, S., Jin, F., Yu, J., Han, Y., 2012. A novel granular support vector machine based on mixed kernel function. International Journal of Digital Content Technology and its Applications 6, 484-492.

Huang, Z., Williamson, M.A., 1997. Determination of porosity and permeability in reservoir intervals by artificial neural network modeling, offshore eastern Canada. Petroleum Geoscience 3, 245-258.

Jeirani, Z., Mohebbi, A., 2006. Estimating the initial pressure, permeability and skin factor of oil reservoirs using artificial neural networks. Journal of Petroleum Science and Engineering 50, 11-20.

Jiao, G., Guo, T., Ding, Y., 2016. A new hybrid forecasting approach applied to hydrological data: a case study on precipitation in northwestern China. Water 8, 367.

Lian, C., Zeng, Z., Yao, W., Tang, H., 2013. Displacement prediction of landslide based on PSOGSA-ELM with mixed kernel, Advanced Computational Intelligence (ICACI), 2013 Sixth International Conference on. IEEE, pp. 52-57.

Lim, J.-S., 2005. Reservoir properties determination using fuzzy logic and neural networks from well data in offshore Korea. Journal of Petroleum Science and Engineering 49, 182-192.

Lim, J.-S., Kim, J., 2004. Reservoir porosity and permeability estimation from well logs using fuzzy logic and neural networks, SPE Asia Pacific Oil and Gas Conference and Exhibition. Society of Petroleum Engineers.

Luenberger, D.G., Ye, Y., 2008. Linear and nonlinear programming. Springer.

Majumdar, A., Ciocoiu, M., Blaga, M., 2008. Modelling of ring yarn unevenness by soft computing approach. Fibers and Polymers 9, 210-216.

Maleki, S., Ramazia, H.R., Moradi, S., 2014. Estimation of Iron concentration by using a support vector machineand an artificial neural network-the case study of the Choghart deposit southeast of Yazd, Yazd, Iran. Geopersia 4, 201-212.

McBride, P.S., 2004. Facies analysis of the Devonian Gordon Stray sandstone in West Virginia, Department of Geology and Geography. West Virginia University, p. 145.

Müller, K.-R., Smola, A.J., Rätsch, G., Schölkopf, B., Kohlmorgen, J., Vapnik, V., 1997. Predicting time series with support vector machines, Artificial Neural Networks. Springer, pp. 999-1004.

Na'imi, S.R., Shadizadeh, S.R., Riahi, M.A., Mirzakhanian, M., 2012. Support vector machine approach and petroleum engineering. Journal of American Science 8.

Noble, W.S., 2006. What is a support vector machine? Nature biotechnology 24, 1565-1567.

Piotrowski, R.G., Harper, J.A., 1979. Black shale and sandstone facies of the Devonian Catskill clastic wedge in the subsurface of western Pennsylavnia, EGSP. Pennsylvania Department of Environmental Resources Bureau of Topographic and Geologic Survey, p. 42.

Poli, R., Kennedy, J., Blackwell, T., 2007. Particle swarm optimization. Swarm intelligence 1, 33-57.

Ravandi, E.G., Nezamabadi-Pour, H., Monfared, A.F., Jaafarpour, A., 2014. Reservoir characterization by a combination of fuzzy logic and genetic algorithm. Petroleum Science and Technology 32, 840-847.

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

Saljooghi, B.S., Hezarkhani, A., 2014. Comparison of WAVENET and ANN for predicting the porosity obtained from well log data. Journal of Petroleum Science and Engineering 123, 172-182.

Sayyad, H., Manshad, A.K., Rostami, H., 2014. Application of hybrid neural particle swarm optimization algorithm for prediction of MMP. Fuel 116, 625-633.

Scales, J.A., Snieder, R., 1998. What is noise? Geophysics 63, 1122-1124.

Scherer, M., 1987. Parameters influencing porosity in sandstones: a model for sandstone porosity prediction. AAPG Bulletin 71, 485-491.

Scholkopf, B., Sung, K.-K., Burges, C.J., Girosi, F., Niyogi, P., Poggio, T., Vapnik, V., 1997. Comparing support vector machines with Gaussian kernels to radial basis function classifiers. IEEE Transactions on Signal Processing, 45, 2758-2765.

Singh, V., Rao, S.M., 2005. Application of image processing and radial basis neural network techniques for ore sorting and ore classification. Minerals Engineering 18, 1412-1420.

Smits, G.F., Jordaan, E.M., 2002. Improved SVM regression using mixtures of kernels, In Neural Networks, 2002. IJCNN'02. Proceedings of the 2002 International Joint Conference on. IEEE, pp. 2785-2790.

Smola, A.J., 1996. Regression estimation with support vector learning machines. Master's thesis, Technische Universit at Munchen.

Smola, A.J., Schölkopf, B., 1998. Learning with kernels. MIT.

Smola, A.J., Schölkopf, B., 2004. A tutorial on support vector regression. Statistics and Computing 14, 199-222.

Smola, A.J., Schölkopf, B., Müller, K.-R., 1998. General cost functions for support vector regression, In Proceedings of the 8th International Conference on Artificial Neural Networks. Citeseer.

Suykens, J.A., Vandewalle, J., 1999. Least squares support vector machine classifiers. Neural Processing Letters 9, 293-300.

Suykens, J.A., Vandewalle, J., 2000. Recurrent least squares support vector machines. IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications 47, 1109-1114.

Tatar, A., Halali, M., Mohammadi, A., 2016. On the estimation of the density of brine with an extensive range of different salts compositions and concentrations. J Thermodyn Catal 7, 2.

Tatar, A., Shokrollahi, A., Mesbah, M., Rashid, S., Arabloo, M., Bahadori, A., 2013. Implementing Radial Basis Function Networks for modeling CO_2-reservoir oil minimum miscibility pressure. Journal of Natural Gas Science and Engineering 15, 82-92.

Vapnik, V.N., 1963. Pattern recognition using generalized portrait method. Automation and remote control 24, 774-780.

Vapnik, V.N., Chervonenkis, A.J., 1964. On the one class of the algorithms of pattern recognition. Automation and remote control 25.

Vapnik, V.N., Golowich, S.E., Smola, A., 1997. Support vector method for function approximation, regression estimation, and signal processing. Advances in neural information processing systems, 281-287.

Vapnik, V.N., Kotz, S., 1982. Estimation of dependences based on empirical data. Springer-Verlag New York.

Verma, A.K., Cheadle, B.A., Routray, A., Mohanty, W.K., Mansinha, L., 2012. Porosity and Permeability Estimation using Neural Network Approach from Well Log Data, GeoConvention Vision Conference, Canada.

Wang, G.C., Carr, T.R., Ju, Y.W., Li, C.F., 2014. Identifying organic-rich Marcellus Shale lithofacies by support vector machine classifier in the Appalachian basin. Computers & Geosciences 64, 52-60.

Wang, H., Hu, D., 2005. Comparison of SVM and LS-SVM for regression, International Conference on Neural Networks and Brain. IEEE, pp. 279-283.

Wendt, W., Sakurai, S., Nelson, P., 1986. Permeability prediction from well logs using multiple regression. Academic Press, Inc., Orlando, Florida.

Yang, X., Peng, H., Shi, M., 2013. SVM with multiple kernels based on manifold learning for breast cancer diagnosis, Information and Automation (ICIA), 2013 IEEE International Conference on. IEEE, Yinchuang, China, pp. 396-399.

Yuan, S.-F., Chu, F.-L., 2007. Fault diagnostics based on particle swarm optimisation and support vector machines. Mechanical Systems And Signal Processing 21, 1787-1798.

Zargari, H., Poordad, S., Kharrat, R., 2013. Porosity and permeability prediction based on computational intelligences as artificial neural networks (ANNs) and adaptive neuro-fuzzy

inference systems (ANFIS) in southern carbonate reservoir of Iran. Petroleum Science and Technology 31, 1066-1077.

Zhang, G., Hu, M.Y., Patuwo, B.E., Indro, D.C., 1999. Artificial neural networks in bankruptcy prediction: General framework and cross-validation analysis. European Journal Of Operational Research 116, 16-32.

Zheng, S., Liu, J., Tian, J., 2004. An SVM-based small target segmentation and clustering approach, Machine Learning and Cybernetics, 2004. Proceedings of 2004 International Conference on. IEEE, pp. 3318-3323.

Zhong, Z., Carr, T.R., 2016. Application of mixed kernels function (MKF) based support vector regression model (SVR) for CO_2 – Reservoir oil minimum miscibility pressure prediction. Fuel 184, 590-603.

Zhu, Y., Tian, L., Mao, Z., Wei, L., 2005. Mixtures of kernels for SVM modeling, Advances in Natural Computation. Springer, pp. 601-607.

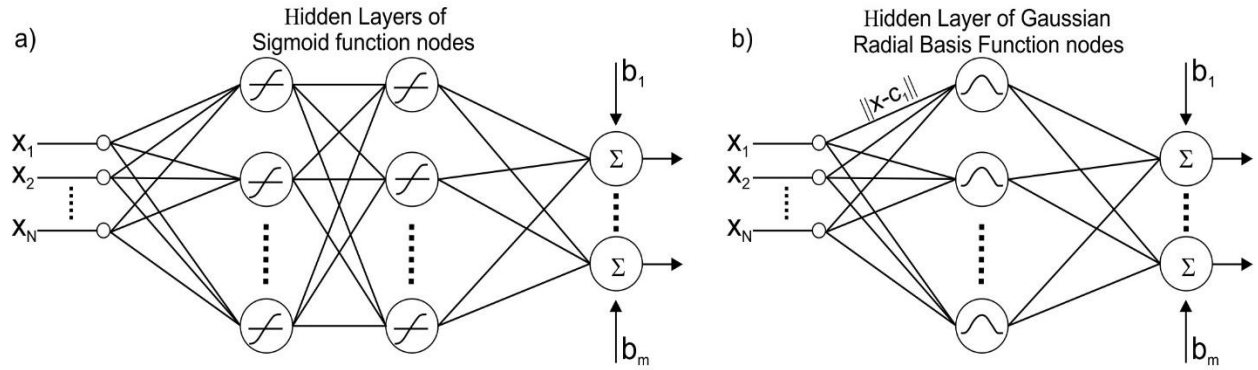


Figure 2-1. (a) The structure of fully connected multilayer perceptron with two hidden layers. The active function of hidden neurons is sigmoid function. (b) The structure of fully connected RBF neural network, the active function of hidden neurons is Gaussian radial basis function (modified from Aljarah et al. 2016).

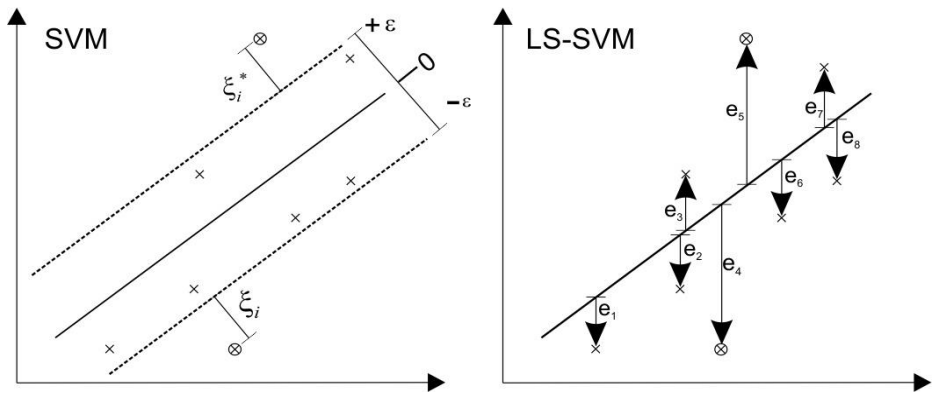


Figure 2-2. Comparison of SVM and LS-SVM for linear regression. Standard SVM (left) use ε-insensitive loss function, in which data points lying on or outside of ε-tube of decision function are support vectors, and the gap between data points and the radius of tube are slack variables ($\xi^{(*)}$). LS-SVM (right) involves equality constraints and uses least square loss function, in which the ε-tube and slack variables are replaced by error variable (e_i) (modified from Wang et al. 2005).

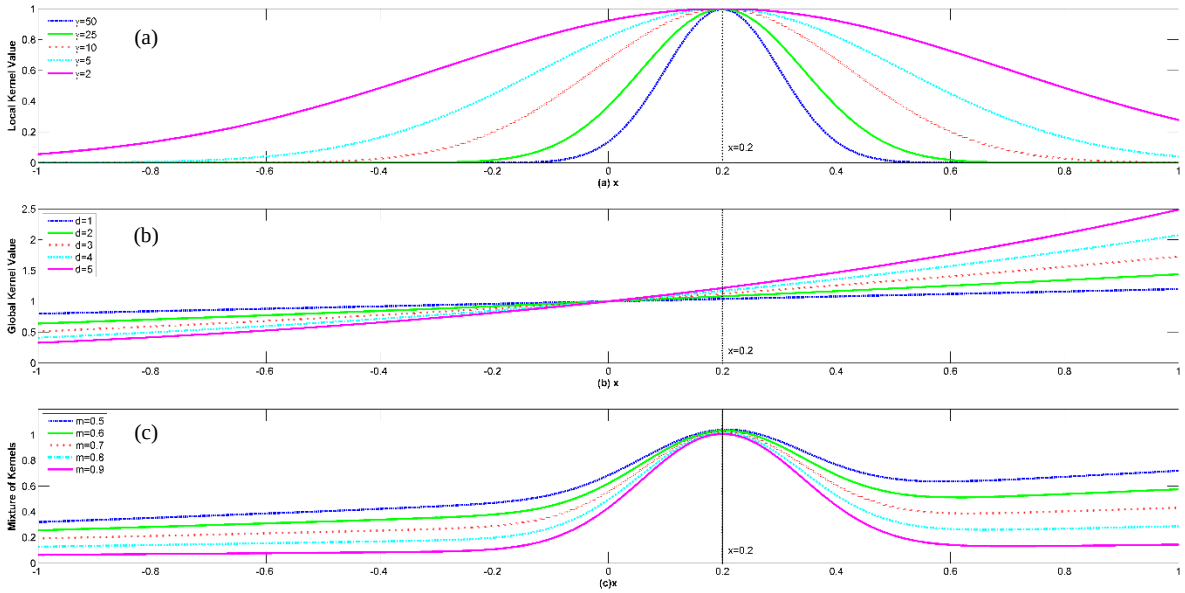


Figure 2-3. Mapping features of polynomial, radial basis function and mixed kernel function. (a) is radial basis function kernel, x = 0.2 is test point. Various values of ϒ was selected, the points adjacent to the test point have a great influence on the kernel values; (b) is polynomial kernel, d is the operation degree, x = 0.2 is test point. Various values of d were selected, only the points that are far enough from test point will have an effective influence on the kernel value; (c) is mixed kernel function. m is mixing coefficient, x = 0.2 is the test point, ϒ is 0.1 and d = 1. Various values of mixing coefficient (m) were selected, data points that are both far away from the test point and adjacent to the test point have a great influence on the kernel values.

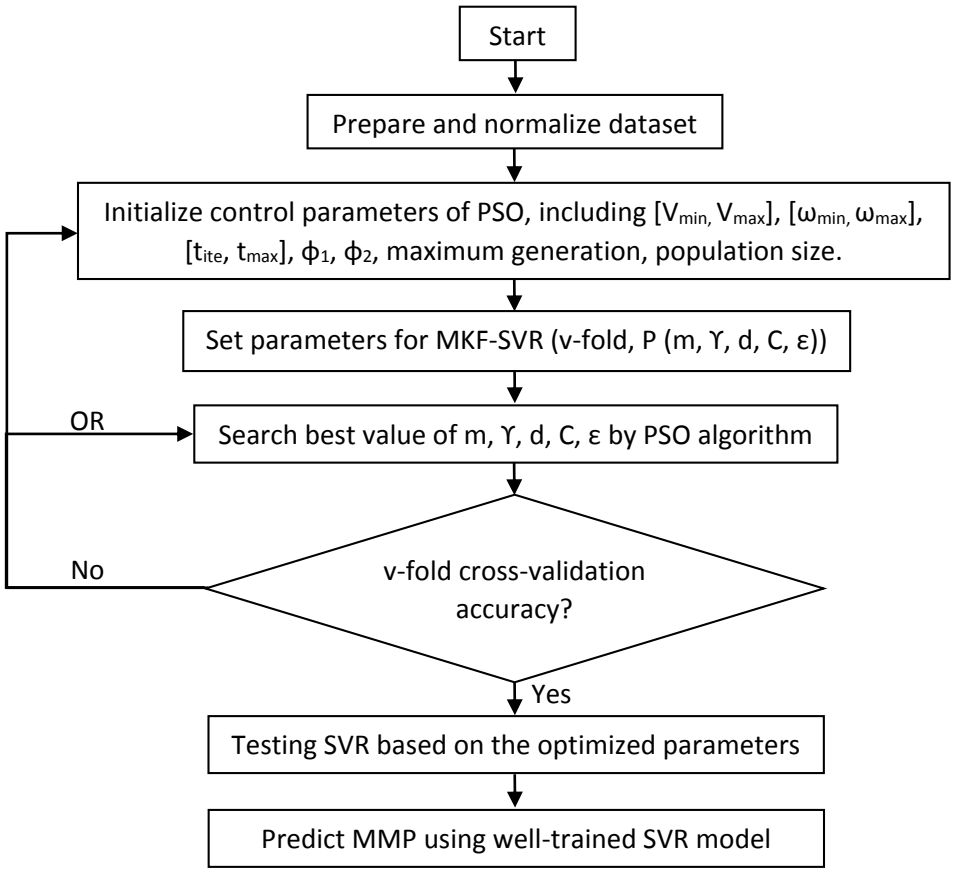


Figure 2-4. Workflow of PSO to optimize parameters of mixed kernels function.

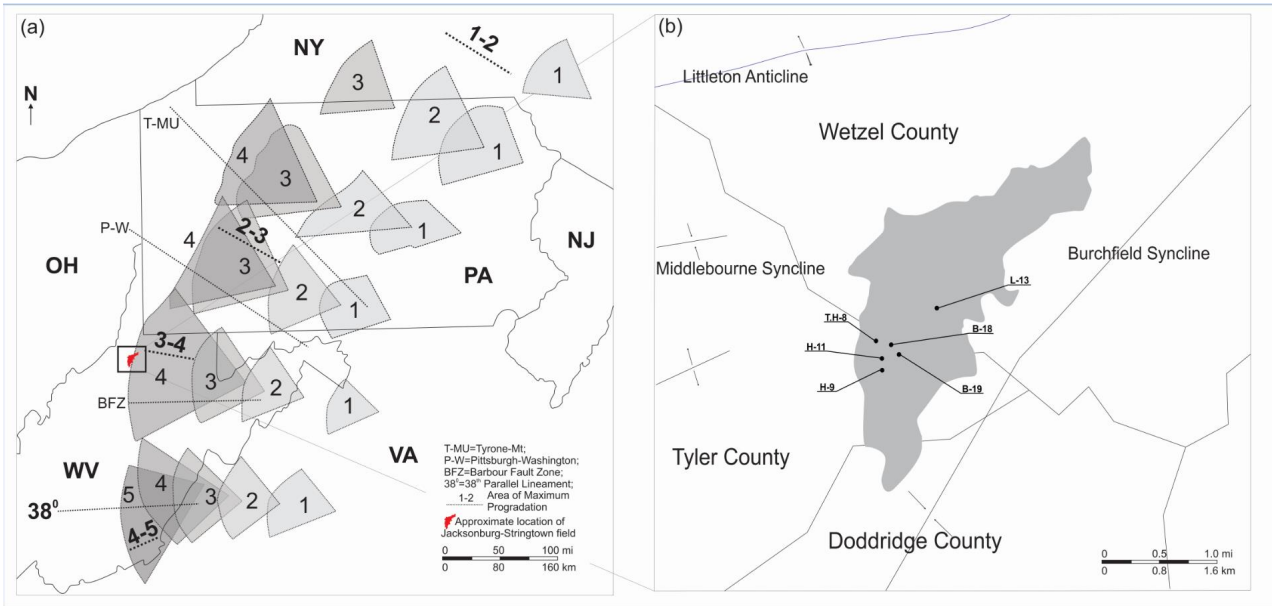


Figure 2-5. (a) Late Devonian paleogeography of study area and five major delta systems in Appalachian foreland basin, black lines indicates geographical state boundaries and approximate location of Jacksonburg-Stringtown field is highlighted in red. (b) Location of cored wells in the Jacksonburg-Stringtown field in northwestern West Virginia that were used in the study. Figure 2-5a modified from Bowell, 1988.

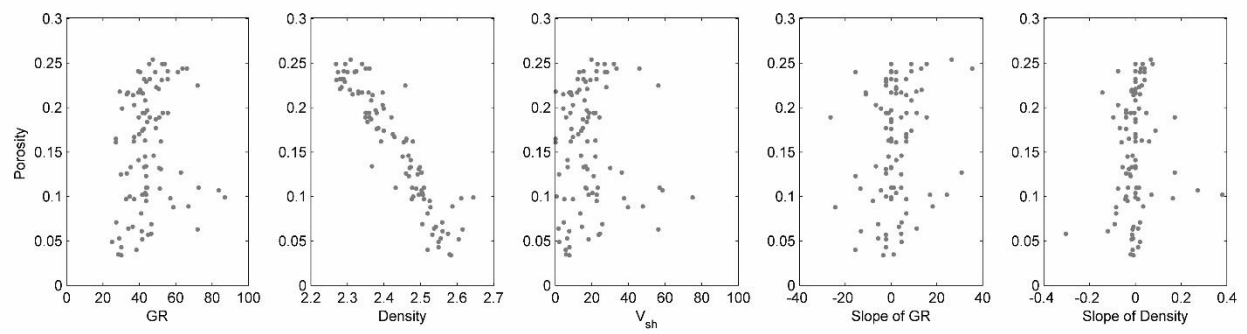


Figure 2-6. Porosity vs. different log derived variables which are used as input parameters in this study. It indicates that there are not strong linear relationships between porosity and each input variables.

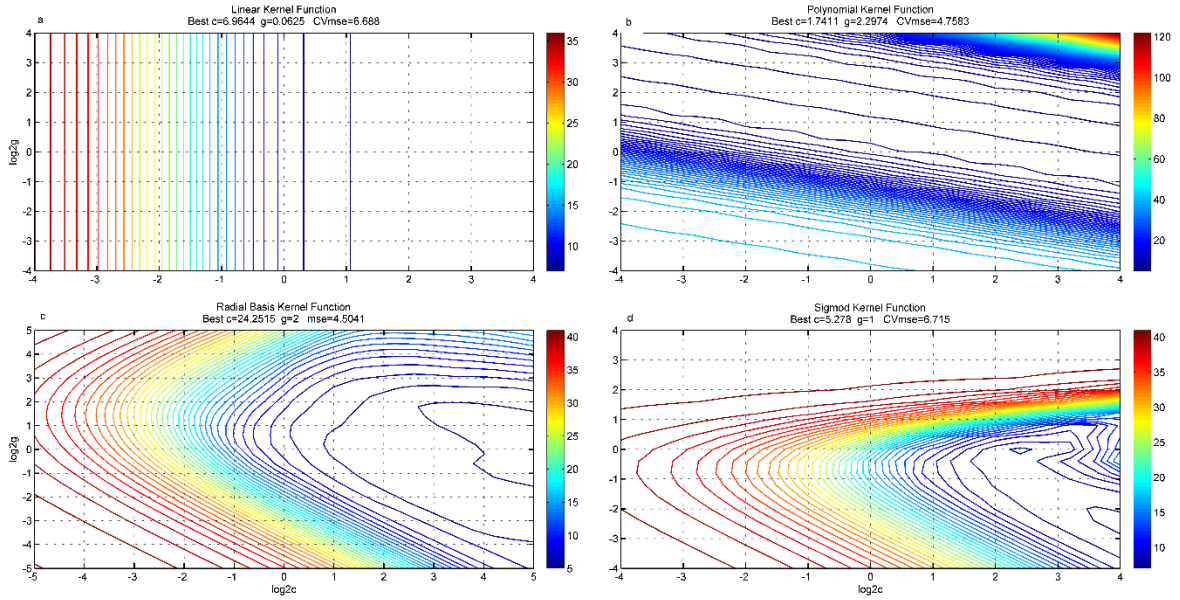


Figure 2-7.The process of searching for best gamma and cost parameters by grid searching algorithm for SVR with linear kernel (a), polynomial kernel (b), RBF kernel (c), sigmoid kernel (d) and n-fold of 4. The color of the contour lines in the figure indicated the associated cross-validation mean square error. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

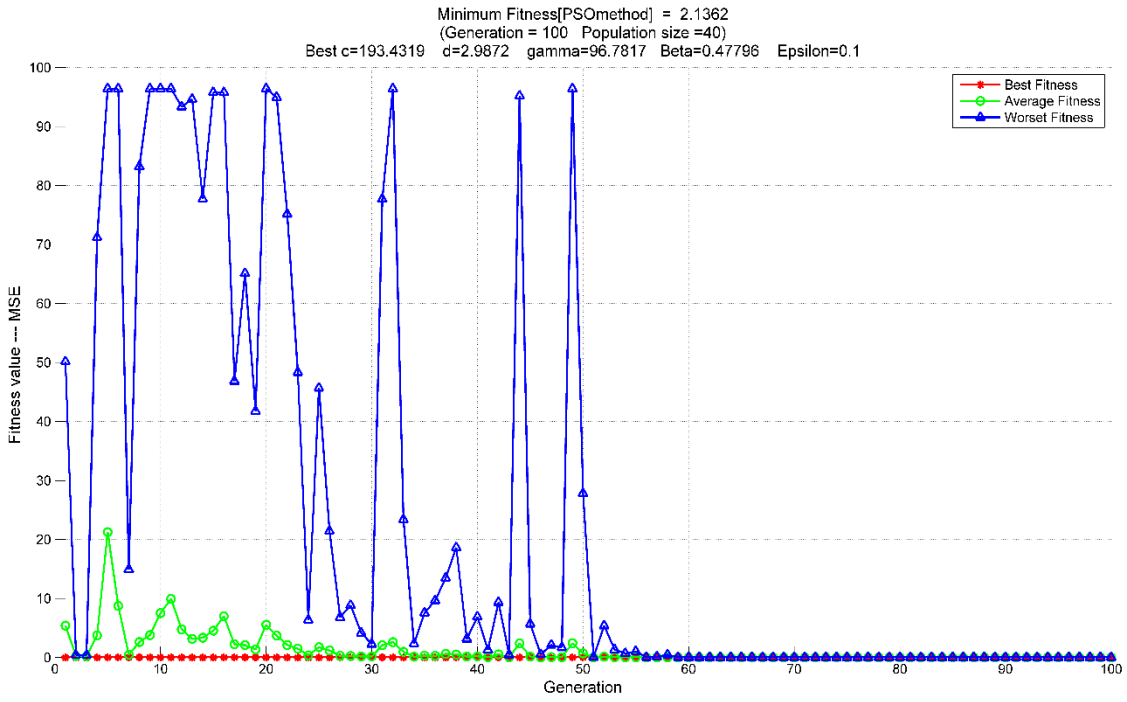


Figure 2-8.The process of searching for best gamma, penalty, mixing coefficient, degree and epsilon by particle swarm optimization (PSO) for SVR model with mixture of kernels and n-fold of 4.

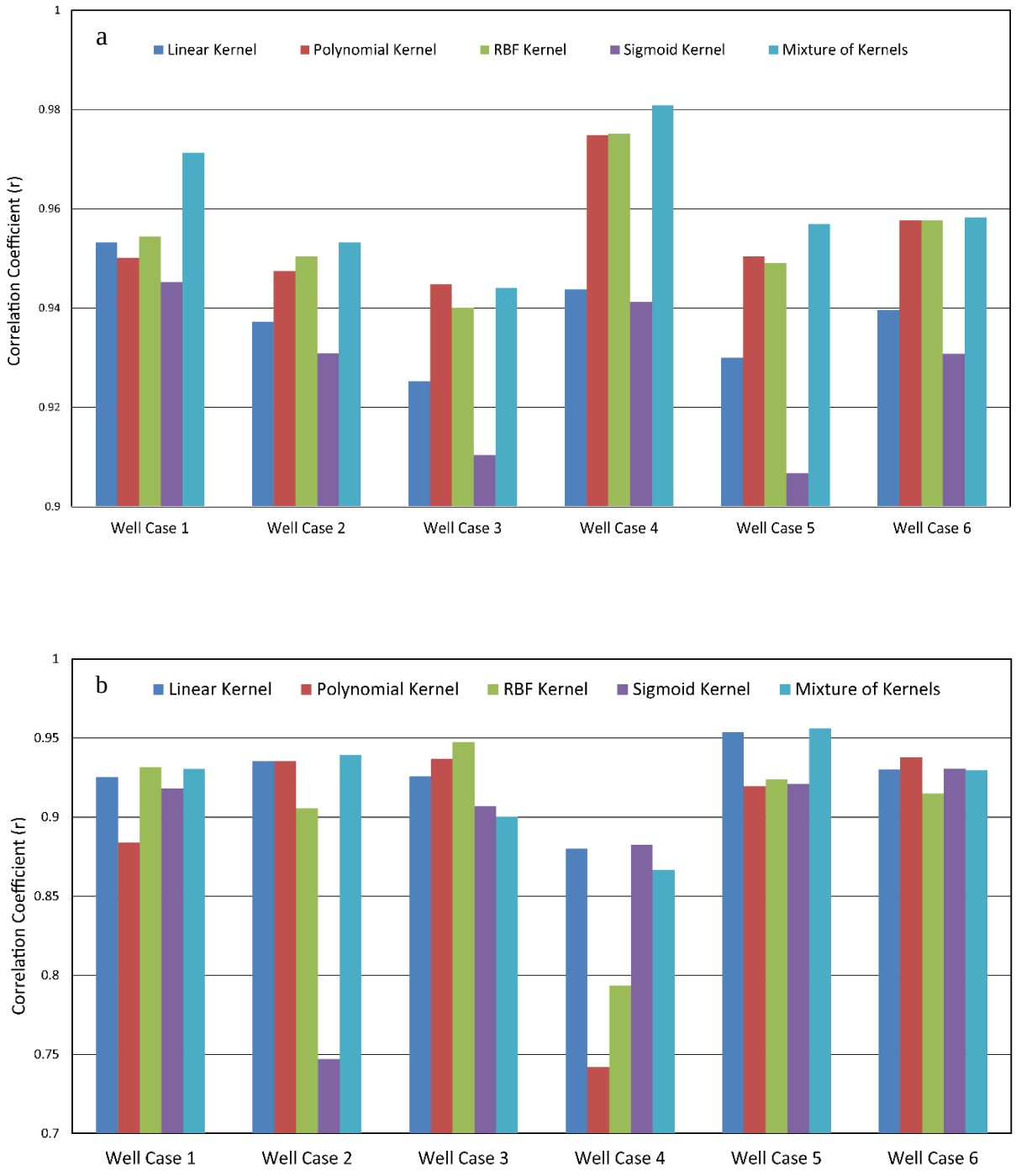


Figure 2-9. Correlation coefficient of training (a) and testing (b) for SVR with different kernel function.

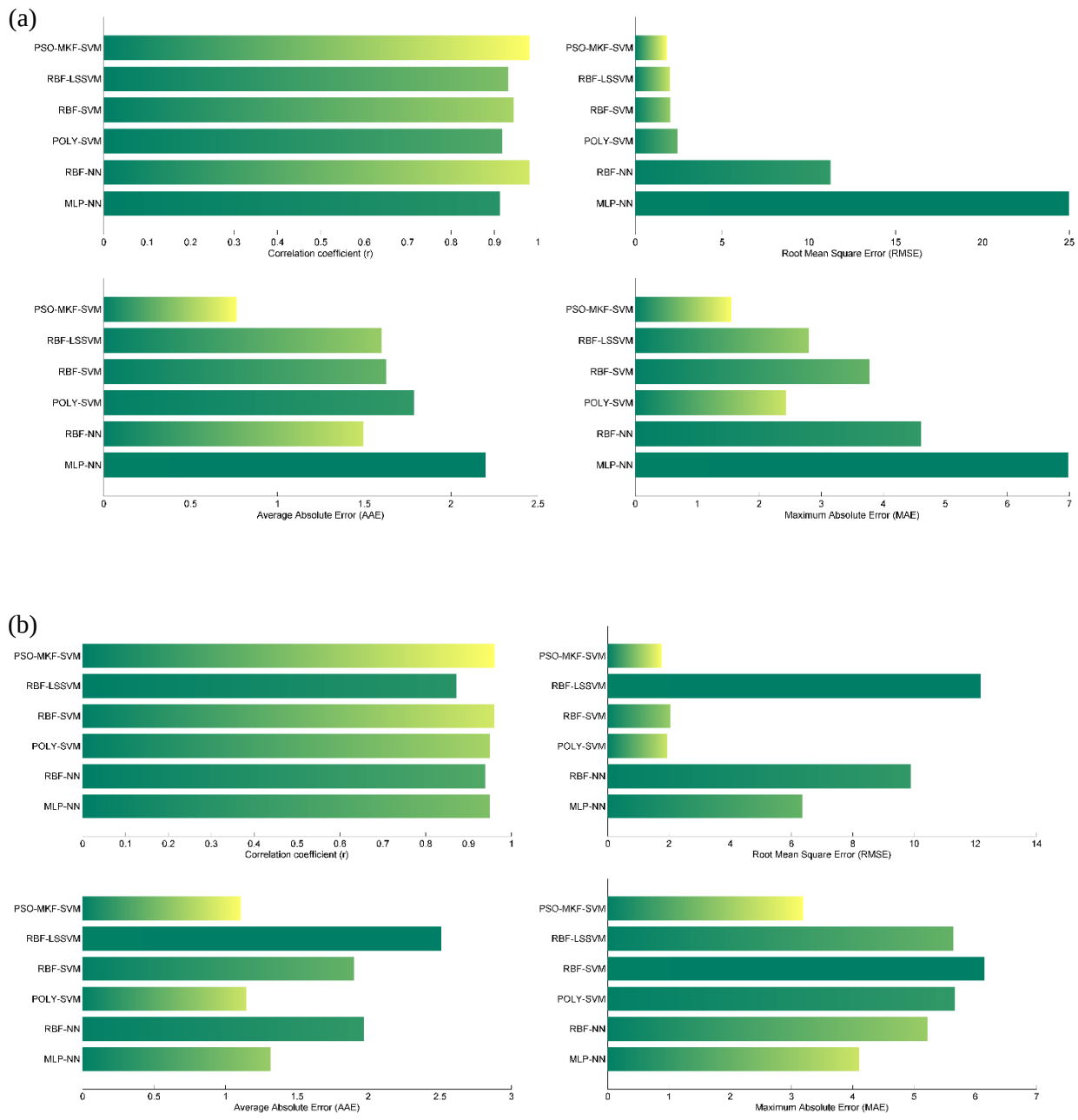


Figure 2-10. Evaluation index bar chart for (a) training process and (b) testing process for different kernel functions.

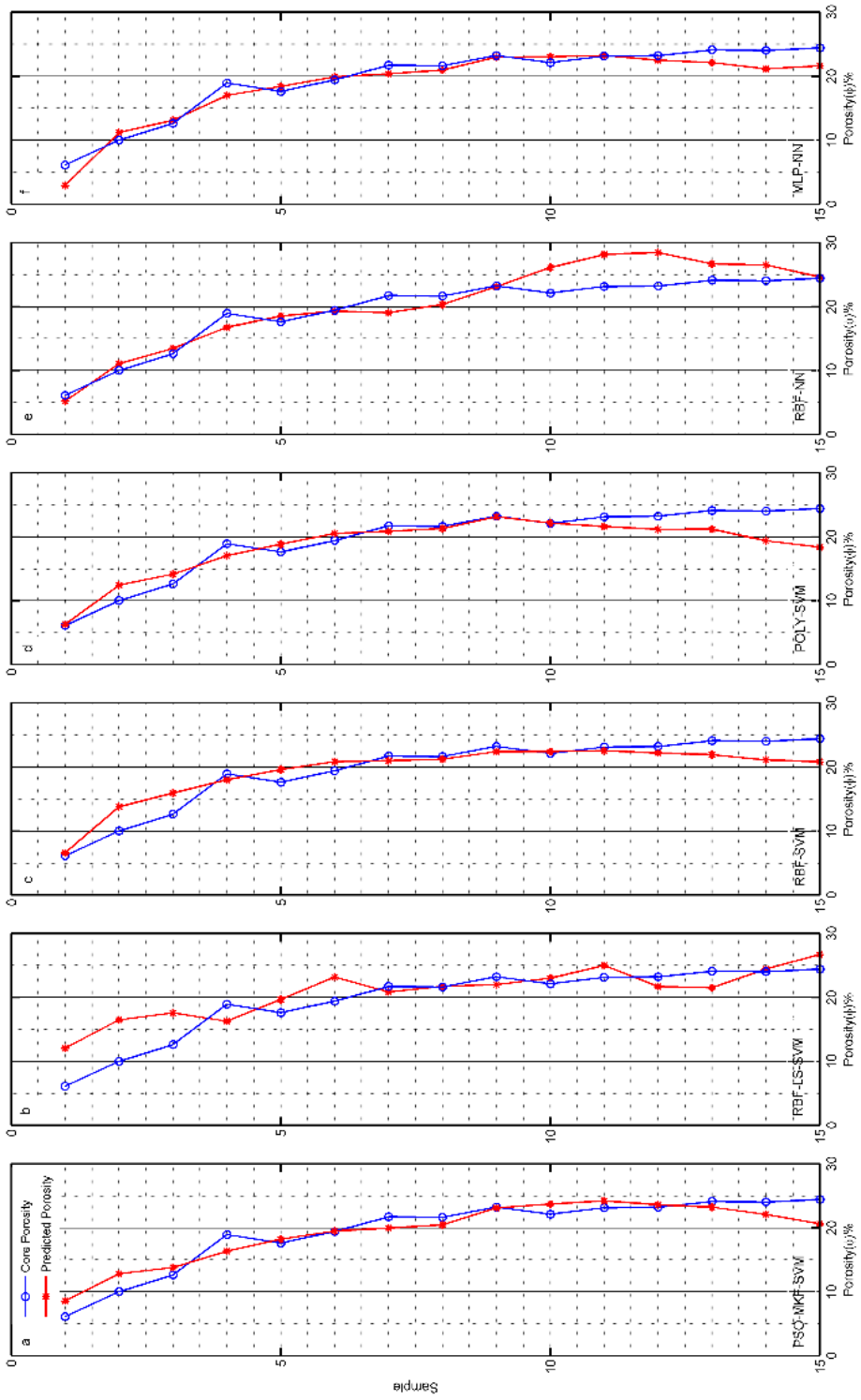


Figure 2-11. Single model forecast effect chart showing core porosity compared with predicted porosity for (a) PSO-MKF-SVM; (b) RBF-LS-SVM; (c) RBF-SVM; (d) POLY-SVM; (e) RBF-NN; and (f) MLP-NN models.

Table 2-1. Common kernel functions and corresponding mathematical expressions.

Linear Kernel	$K(x_i, x_j) = \langle x_i, x_j \rangle + c$
Polynomial Kernel	$K(x_i, x_j) = (\gamma \langle x_i, x_j \rangle + c)^d$
Radial Basis Function Kernel	$K(x_i, x_j) = \exp(-\gamma \left\| x_i - x_j \right\|^2)$
Sigmoid Kernel	$K(x_i, x_j) = \tanh(\gamma \langle x_i, x_j \rangle + c)$

Table 2-2. Various error measures used for accuracy assessment on this paper.

Accuracy measure	Mathematical expression
Coefficient of Determination, R^2	$R^2 = 1 - \frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)}{\sum_{i=1}^{N}(y_i - averg(y_i))}$
Correlation coefficient, r	$r = \frac{\sum_{i=1}^{N}(y_i - \bar{y}_i)(\hat{y}_i - \bar{\hat{y}}_i)}{\sum_{i=1}^{N}(y_i - \bar{y}_i)^2 \sum_{i=1}^{N}(\hat{y}_i - \bar{\hat{y}}_i)^2}$
Root mean square error, RMSE	$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}$
Average absolute error, AAE	$AAE = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)$
Maximum absolute error, MAE	$MAE = \max\left\|y_i - \hat{y}_i\right\|, i = 1,2,...,N$

Table 2-3. Training and testing wells from the Jacksonburg-Stringtown field used in this paper.

Well case index	Training well	Testing well
1	TH_8,H_9,B_19,LM_13,H_11	B_18
2	TH_8,H_9,B_19,LM_13,B_18	H_11
3	TH_8,H_9,B_19, H_11,B_18,	LM_13
4	TH_8,H_9,LM_13,H_11,B_18	B_19
5	TH_8,B1_9,LM_13,H_11,B_18	H_9
6	H_9,B_19,LM_13,H_11,B_18	TH_8

Table 2-4.Typical parameters used for training SVM regression model with different learning algorithms.

Common parameters for training SVR			
Sample size	94	Termination criterion (e)	0.005
Gamma bound	$2^{(-4)}$-$2^{(4)}$	Cost bound	$2^{(-4)}$-$2^{(6)}$
r in polynomial and sigmoid function	1	n-fold	4
Special parameters for grid searching algorithm			
Gamma searching step	$\Delta Log2(gamma) = 0.4$	Cost searching step	$\Delta Log2(\cos t) = 0.4$
Special Parameters for particle swarm optimization (PSO)			
Maximum generation	100	Population size	40
Cognitive efficient (c_1)	1.4	Social efficient (c_2)	1.7
Initial inertia weights W_{start}	0.9	Final inertia weight W_{end}	0.4

Table 2-5. Error measures, RMSE, AAE and MAE, resulting from SVM regression model training with different kernel function.

Well case index	Linear Kernel			Polynomial Kernel			RBF Kernel			Sigmoid Kernel			Mixture of Kernels		
	RMSE	AAE	MAE	RMSE	AAE	MAE	RMSE	AAE	MAE	RMSE	AAE	MAE	RMSE	AAE	MAE
1	1.9500	1.5689	5.0335	1.9952	1.4936	5.5082	1.9298	1.5388	5.1839	2.1115	1.7313	4.8846	1.5285	1.0287	4.7721
2	2.3594	1.8171	5.4489	2.1125	1.5191	4.9851	1.6245	0.9765	6.2904	2.4592	1.8277	6.3440	2.1074	1.5662	5.4573
3	2.3050	1.6346	5.0849	1.9486	1.1962	5.5820	1.5563	0.8826	5.9702	2.4827	1.7704	6.1409	2.0940	1.4381	6.9185
4	2.2237	1.6400	3.9323	1.4604	0.9056	3.7720	0.7668	0.4246	2.0947	2.3076	1.8031	5.1416	1.5860	0.9915	5.3350
5	2.3537	1.8076	4.9793	1.9471	1.1606	5.6995	1.9017	1.1960	5.6380	2.6604	1.9417	5.9941	1.8493	1.2238	5.3077
6	2.3102	1.8108	5.2645	1.9170	1.2368	6.2455	1.7230	1.0277	6.7035	2.4820	1.8991	5.2730	1.9362	1.4222	5.4454

Table 2-6. Basic control parameters for SVM regression model training with different kernel function.

Well case index	Linear Kernel	Polynomial Kernel			RBF Kernel			Sigmoid Kernel			Mixture of Kernels					
	R^2	R^2	γ	C	R^2	γ	C	R^2	γ	C	R^2	β	γ	d	C	ε
1	0.9085	0.9028	2.2974	0.1436	0.9109	0.0313	97.0059	0.8935	0.3299	16	0.9434	0.4792	96.3658	3	168.1266	0.0794
2	0.8785	0.8976	3.0314	0.1088	0.9033	2.6390	13.9288	0.8666	1.000	5.2780	0.9086	0.6594	14.0913	1.4169	137.8577	0.0102
3	0.8562	0.8926	2.2974	1	0.8837	3.4822	13.9289	0.8288	0.1895	16	0.8912	0.5107	54.4831	1.6856	560.1957	0.0543
4	0.8908	0.9503	9.1896	0.1436	0.9509	2.6399	388.0234	0.8859	0.1088	16	0.9620	0.1107	86.9035	3	762.511	0.0990
5	0.86500	0.9033	2.2974	1.7411	0.9007	0.0947	2702.3522	0.8221	1	5.2780	0.9158	0.6247	19.9453	2.6327	636.1171	0.0834
6	0.8830	0.9172	5.2780	0.1436	0.9172	2.6390	10.5561	0.8664	0.3299	16	0.9183	0.6067	40.1341	2.1118	828.9583	0.0457

Table 2-7. Error measures, RMSE, AAE and MAE, resulting from SVM regression model testing with different kernel function.

Well case index	Linear Kernel				Polynomial Kernel				RBF Kernel				Sigmoid Kernel				Mixture of Kernels			
	RMSE	AAE	MAE	R^2	RMSE	AAE	MAE	R^2	RMSE	AAE	MAE	R^2	RMSE	AAE	MAE	R^2	RMSE	AAE	MAE	R^2
1	3.3734	2.3961	2.2001	0.8561	4.2564	3.3518	5.8817	0.7814	3.1998	2.2878	1.9303	0.8679	3.4272	2.4752	2.5732	0.8432	3.2831	2.5629	2.7151	0.8657
2	2.6378	2.1776	0.4219	0.8747	2.2953	1.7724	0.9271	0.8749	2.5851	2.1644	1.7099	0.8199	3.4543	2.8783	3.9375	0.5582	2.6321	2.1053	0.3312	0.8825
3	2.6498	2.3094	4.2308	0.8573	2.6876	2.2972	4.3005	0.8778	1.1813	1.5580	2.95000	0.8976	3.1367	2.7777	4.8361	0.8223	4.0314	3.5074	60.524	0.8104
4	2.3882	1.9793	4.8206	0.7748	3.5974	2.20032	13.1476	0.5506	3.8257	2.6801	11.6690	0.6300	2.4114	2.0250	4.7866	0.7791	2.5655	2.1117	4.8584	0.7510
5	1.6880	1.2707	1.8986	0.9099	2.4026	1.7601	2.3399	0.8459	2.2904	1.7143	2.4709	.08539	2.2604	1.6859	2.6461	0.8487	1.6505	1.4050	2.7171	0.9140
6	1.9345	1.4770	5.2645	0.8656	2.0045	1.7393	6.2455	0.8793	2.2927	1.9052	6.7035	0.8372	1.9520	1.4707	5.2729	0.8665	2.0625	1.6196	5.4454	0.8642

CHAPTER 3

Geostatistical 3-D Geological Model Construction to Estimate the Capacity of Commercial Scale Injection and Storage of CO2 in Jacksonburg-Stringtown Oil Field, West Virginia, USA.

Zhi Zhong and Timothy R Carr[1]

[1] Department of Geology and Geography

West Virginia University, Morgantown

WV-26505, USA

(304)-282-9243

Email: zizhong1990@gmail.com; tim.carr@mail.wvu.edu

Abstract

As concerns around global warming increase, carbon capture, utilization and geological storage (CCUS) is a promising way to reduce the emissions of anthropogenic CO_2 into the atmosphere. Sequestering the CO_2 into depleted hydrocarbon reservoirs with associated enhanced oil recovery (EOR) is the most achievable approach under current economic constraints since it increases recovery of existing oil reserves, and bridges the gap between regional-scale CO_2 capture and geologic sequestration. The Upper Devonian fluvial sandstone reservoirs in the Jacksonburg-Stringtown oil field in West Virginia, which have produced over 22 million barrels of oil since 1895, are ideal candidates for CO_2 sequestration coupled with EOR. Reservoir storage capacity and oil recovery factors are keys for the evaluation of coupled CO_2 storage and CO_2-EOR process. In this research, a static 3D reservoir model, which integrates detailed geological knowledge and existing legacy geological data from Jacksonburg-Stringtown oil field, is constructed to estimate theoretical CO_2 storage capacity. Regression relationship between wireline logs and core measured data for porosity, permeability are constructed by artificial neural network and support vector machine in core-scale; then those regression models are extended from core-scale to well-scale, where wells do not have porosity and permeability wireline logs. Finally a 3D static geological model is generated based on the Random Gaussian Function simulation method and well-established variogram models generated by detailed data analysis. Depending on the proposed 3-D geological model, the best regions for coupled CCUS-EOR are located in southern portions of the field, and the estimated CO_2 theoretical storage capacity for Jacksonburg-Stringtown oil field ranges from 24 to 383 million metric tons. The estimated results of CO_2 sequestration indicate that the Jacksonburg-Stringtown oilfield has significant potential for CO_2 storage and value-added EOR.

3.1 Introduction

Svante Arrhenius proposed 100 years ago the basic model in which the variations of Earth's surface temperature are related to the concentration of carbon dioxide (CO_2) (Crawford, 1997; Rodhe et al., 1997). The greenhouse model was resuscitated in the 1970s due to the concern of global warming resulting from the increased greenhouse gasses emissions from industrial sources, fossil fuel combustion and land-use change (Falkowski et al., 2000; Quéré et al., 2014;

Quéré et al., 2013b). In order to maintain future global warming below 2℃ above pre-industrial levels, approximately 9.5 gigatons (Gt) carbon/year or 35 Gt CO_2/year will need to be captured and stored before entering the atmosphere (Peters et al., 2012; Quéré et al., 2013a).

The common approaches to mitigate potential global climate change induced by anthropogenic emissions of CO_2 and other greenhouse gases is to reduce CO_2 emissions or to increase storage in CO_2 sinks (Bachu, 2000). CO_2 emissions reduction while satisfying increasing global energy demands is not viable on current technological and engineering conditions. Carbon capture, utilization and storage (CCUS) captures the anthropogenic carbon dioxide (CO_2) from large stationary carbon sources, such as coal and natural gas power plants or resource extraction industries. Then the CO_2 is either piped or transported to specific sits, and utilized or stored into deep geological media, oceans, or settling through surface mineral carbonation, thus achieving a long-term isolation of CO_2 from the atmosphere (Bachu, 2000; Jiang et al., 2014).

One of the most effective sites to store CO_2 are depleted gas and oil reservoirs because the characteristics of hydrocarbon reservoirs are generally better known as a result of the extensive history of exploration and production (Zhao and Liao, 2012). In addition, CO_2 can also enhance oil recovery (EOR) and can cover part of the cost of sequestration. Theoretical storage capacity is sensitive to the porosity model, and effective storage capacity is sensitive to the porosity and permeability model, thus constructing a reasonable 3D geological reservoir model is critical for the evaluation of coupled CO_2 storage and CO_2-EOR process. Detailed sources to sink matching analysis, which is matching storage sites to large emission sources is also an important step for accurate estimation of CO_2 storage capacity.

Unlike modern oil and/or gas fields, which have abundant, high quality data, including conventional and advanced well logs, seismic data, production data and geological framework to construct 3D geological models, a super-mature oil field, such as Jacksonburg-Stringtown oil field, has plentiful low quality data, and limited high quality data. The sparse modern subsurface data can constrain the development of a robust 3D static geological model when evaluating the CO_2 storage capacity.

The focus of this research is to evaluate the capacity of CO_2 sequestration and the potential of CO_2-EOR in a highly mature and depleted oil reservoir, Jacksonburg-Stringtown oil field in West Virginia, USA. This research will ultimately aid in the determination of the technical and economic suitability of the field for long-term storage of supercritical CO_2. In addition, the proposed research will construct and validate a basic workflow for CO_2 storage capacity evaluation that can be applied to other mature gas and/or oil fields, which have limited high quality modern data. Once this detailed assessment has been achieved, it will be possible to express capacity at a regional level as a annual sustainable rate of injection, not just as a total volume (Bradshaw et al., 2007).

Specific objectives within the scope of the proposed research include:

- to predict permeability values with the aid of artificial neural network (ANN),
- to estimate porosity values by using support vector machine (SVM), and
- to construct a 3D geologic reservoir model and to calculate theoretical CO_2 storage capacity.

3.2 Geologic Background

The Jacksonburg-Stringtown field, located in northwestern Doddridge, southeastern Wetzel and eastern Tyler counties, West Virginia, sits along the western edge of the Burchfield syncline (Figure 3-1b). The primary and/or secondary producing reservoir units of Jacksonburg-Stringtown oil field are in the middle to late Catskill deltaic complex (Catskill Delta), which is a thick sediment wedge deposited during the Late Devonian (Ameri et al., 2002; Boswell, 1988; Bridege and Willis, 1994). This clastic wedge includes the Early Mississippian Price-Rockwell deltaic complex (Price Delta) composed the Acadian clastic wedge, which was deposited as various marine and continental beds associated with the Early and Middle Devonian Acadian orogeny (Bjerstedt and Kammer, 1987).

During Early to Middle Devonian, crustal uplift in Acadian orogeny lead to further subsidence within the Appalachian foreland basin to the west and resulted in the deepening of the central

Appalachian basin (Faill, 1997a, b). Deposition coincided with the heavy rainfall produced by the tropical climate, sedimentary deposition accelerated during the Middle and Late Devonian (Blakey, 2008; McBride, 2004; Piotrowski and Harper, 1979). A series of river systems aided transport of sediment westward into the Appalachian basin (McBride, 2004).

In the Late Devonian, five major delta systems prograded westwards and dominated the foreland ramp (Figure 3-1a). Delta systems are separated by the interdeltaic shale (Coughlin, 2009; Dennison, 1985). Two southernmost delta complexes deposited the Bradford and Venango groups (Boswell and Donaldson, 1988). In the late Devonian, maximum progradation of the Catskill delta complex was achieved west of the Acadian highlands (Boswell, 1988; Tassell, 1987). During this period, Acadian tectonism ceased, and relative sea-level changes within the basin were controlled primarily by eustatic sea-level fluctuations and variations in sedimentation (Coughlin, 2009). Non-marine red shale and most of the low-energy alluvial deposits are concentrated in the eastern portions of the Appalachian basin. Non-marine sediments increasingly advanced westward to cover marine beds, and near-shore deposits continued to prograde into the central Appalachian basin (Figure 3-2). Marine shale units are mainly deposited in western portions of Appalachian basin in Kentucky and Ohio (Donald L. Woodrow, 1983; Moore, 2009). The shore zone became the discernable feature of the Catskill delta system, which can be used to illustrate the progradation direction and evolutionary history of this delta's structure (Faill, 1997a).

The Gordon stray and/or Gordon interval, our target CO_2 storage formation, is thick part of Venango Group in the Catskill delta, as shown in Figure 3-2. In the area of Jacksonburg-Stringtown field, Gordon stray/Gordon intervals are interpreted to be shoreline/shoreface sandstone that occupied a broad structural trend at the time of maximum regression of the Acadian clastic wedge (Hohn, 2004). Like many of the oilfields in West Virginia, oil is trapped downdip of gas accumulations along the structural highs.

This field was discovered in 1895, the original oil in place (OOIP) is estimated at 88 million barrels of oil (MMBO), and primary production is estimated at 22 MMBO. The estimated oil recovery factor is 25%. Total Jacksonburg-Stringtown field covers 15,386 Acres, but the effective area is 4,388 acres (Ameri et al., 2002). Average well spacing is 13 acres per well. Gas

re-injection program took place in the field beginning in the mid 1930's and ended in the 1950's. The first pilot waterflood program with a 35 acre dual 5-spot well pattern was conducted in 1981. After 1990, full-scale waterfloods were installed in a large portion of the field (Bergerud, 2011).

The signature of the Gordon Stray intervals follows an idealized estuarine vertical succession (Figure 3-3). By examining the logs from the Jacksonburg-Stringtown field (Ameri et al., 2002; Boswell, 1988; Boswell and Donaldson, 1988), the Upper Devonian Gordon stray and the Gordon sands are inferred to be shoreline and deltaic deposits ranging between dip trending lobes of a fluvial dominated system to the strike-parallel sandstone bodies of a wave-dominated estuarine environment (Ameri et al., 2002; Boswell, 1988; Boswell, 1985; Boswell and Donaldson, 1988). A detailed comparison between idealized estuarine vertical successions with well logs from the Jacksonburg-Stringtown field also support this conclusion (Figure 3-3) (Buatois et al., 2002; Buatois et al., 1999; Dalrymple et al., 1992). As Figure 3-3 shows, the depositional sequence is consistent with incised valley estuarine system can be identified from well logs. The thinner upper sand is interpreted as the estuary mouth deposit, meanwhile the much thicker lower sand can be identified as tidal dominated deltaic deposits, and the inner interval is estuarine shale. The lowest subunit is interpreted as a fluvial deposit.

3.2.1 Carbon Source

At normal atmospheric conditions, CO_2 is a thermodynamically very stable gas with density of 1.98kg/m^3, which is heavier than air. For temperatures greater than 31.1°C and pressures greater than 7.38 MPa (critical point), CO_2 is in a supercritical state. At supercritical conditions, CO_2 has totally different properties than in either liquid or gaseous phase. When supercritical CO_2 behaves like a gas by filling all the available volume, but has a "liquid" density that increases (Holloway and Savage, 1993; Bachu, 2000, 2008). Since both temperature and pressure increase with depth in the subsurface, CO_2 can be stored underground either as a compressed gas, liquid or in supercritical phase, depending on reservoir temperature and pressure. Along the Ohio River, there are numerous power plants, which emit thousands of tons carbon dioxide per year (Figure 3-4a and b). More than 90% of stationary CO_2 emissions are from electric power plants and industrial manufacturing facilities (Figure 3-4b and c). Location of viable geologic storage

sites for CO_2 near the sources will provide economic efficiencies and reduce risks associated with transporting CO_2 long distances from source to sink.

3.3.2 Available Data

The database gathered for this project includes core, well log and production data and other well information. In many cases data from wells are incomplete or not available. Most of the data required for this study was obtained from the "Reservoir Characterization of Upper Devonian Gordon Sandstone" project (Gil, 2000; Alla, 2002; Ameri et al., 2002; Oyerokum, 2002; Thomas, 2002). The core data includes core samples and a petrophysics analysis (Figure 3-5). Ten wells have cores which were described in the study. Moreover, a total of 93 core samples collected from six well as listed in Table 3-1, had measured porosity and permeability. Relative permeability data was available from the PH-9 core well. In a number of wells, conventional well logs were available; including Gamma Ray (GR), Bulk Density (RHOB), Neutron Porosity (NPHI) and Induction logs. Of these, GR and RHOB were used in this study because of their availability in 179 wells. Well logs used in this study are all raster logs. Raster logs were digitized subsequently and auto-corrected to check for accuracy. The slopes of GR and RHOB logs were calculated from first derivatives of GR and RHOB with respect to depth using the three point method. Production data for oil, gas and water collected by WVGS, was used for history matching, well development planning and calculation of ultimate recovery. The data for oil API gravity and viscosity, the formation temperature, the water viscosity were collected.

3.3 Methodology

In this research, the workflow for the three dimensional geological model for CO_2 sequestration capacity calculation used in this research includes: (1) constructing models of porosity and permeability between conventional well logs and limited core-measured data; (2) extending porosity, permeability models to predict porosity and permeability at well-scale; (3) building a 3D petrophysical model via Random Gaussian Function Simulation method with the geologic constrains, and (4) calculating the theoretical CO_2 sequestration capacity based on specific formulation proposed by Bachu et al., (2007)(Figure 3-6).

It was necessary to conduct depth calibration between wireline logs and cored samples before establishing the relationships between wireline logs and the core-measured dataset. Two conventional well logs with gamma ray and bulk density plus three logs derived logs (slope of GR, slope of bulk density, and V_{sh}) were selected as input training and testing for the model; porosity and permeability were the desired output.

In this project, a new back-propagation artificial neural network (BPNN) optimized by two evolutionary algorithms: particle swarm optimization (PSO) and genetic algorithm (GA) were proposed to estimate permeability in Jacksonburg-Stringtown oil field, in West Virginia, USA. Those two evolutionary algorithms were applied to determine the best initial weights and biases for the back-propagation neural network (Zhong and Carr, 2017 a). A support vector machine (SVM) was proposed to estimate porosity by constructing the relationship between conventional well logs and limited core data. A new mixed kernels function (MKF), which is a convex combination of a radial basis function kernel and a polynomial function kernel, was introduced to construct SVM regression model. One global stochastic searching algorithm, particle swarm optimization, was applied to determine the global optimum of five control parameters of the newly developed model (Zhong and Carr, 2017b). The well-trained BPNN is used to predict the permeability and SVM model is used to predict the porosity at the well scale.

Corresponding well tops are determined from the wireline logs of 179 wells and used to construct structure maps of units in the Gordon Stray formation. A 3D structural model of target formation (Gordon Stray formation) is constructed to cover the Jacksonburg-Stringtown oil field (Figure3-1) to provide a framework for 3D facies model. A 3D facies model is established by integrating the 3D structure model and various lithofacies. The arithmetic average method is used to upscale porosity values and harmonic average method is applied to upscale permeability values. Variogram models for upscaled porosity and permeability in each facies are generated by geostatistical analysis, which can be used to illustrate the lateral and vertical porosity and permeability distribution pattern (Wang and Carr, 2013). 3D porosity and permeability models are generated based on a stochastic simulation method called Random Gaussian Function Simulation.

3.4 Fluid and Rock Properties

To demonstrate the heterogeneity of this formation and to display the chaotic status of the information that existed, cross-plots of permeability and porosity versus each measured parameter were constructed (Figure 3-7). The plots indicate that a simple linear relationship is not evident among the six measured parameters and permeability. It shows a complex geologic environment and nonlinear relationship between wireline logs and porosity, permeability.

The RHOmaa/Umaa crossplot derived from bulk density, neutron density and photoelectric wireline logs is common method to estimate matrix mineral compositions (Doveton 1994, Bergerud, 2012). Three wells that have bulk density, neutron density and photoelectric wireline logs, were utilized to analyze the mineral composition of Gordon Stray formation (Figure 3-5). Figure 3-8a shows the RHOmaa/Umaa ternary diagram for four Gordon Stray subunits including; barrier sand, central bay shale, tidal channel, and lower Gordon Stray (fluvial channel), as shown in Figure 3-3, with different colors. The lower Gordon Stray subunit has high quartz content consistent with a clean sandstone reservoir (red circle). The tidal channel subunit consists of a mixture of quartz-rich and illite-rich content, as well as, small concentration of calcite, which could be interpreted as tidally-influenced bay-head delta depositional environment. The central bay shale subunit shows low quartz and high illite content, which is consistent with model of a low-energy depositional environment. Finally, the barrier sand subunit has lower quartz and higher calcite content which indicates a marine influence. Figure 3-8b shows mineral composition of target formation using the bulk density, neutron density and photoelectric wireline logs for well 4710301547. As expected, Lower Gordon Stray (marked as red star) shows high quality of quartz and porosity volume, which indicates that this interval has high capacity of CO_2 storage.

The minimum miscibility pressure (MMP) of oil within the Jacksonburg-Stringtown oil field was determined by slim tube experiment. Figure 3-9a shows the MMP is 941 psi at a reservoir temperature of 80 °F, which means that CO_2 and oil will be miscible completely at a pressure over 941 psi. In Appalachian region, the general geothermal gradient is 20°C/km, and hydrostatic pressure gradient is 10.52 Mpa/km (Johnsson, 1986). This equates to an approximate minimum subsurface depth of 700-800m for supercritical conditions which means that CO_2 should be

injected at depths greater than 800m (Holloway and Savage, 1993) (Figure 3-9b). Based on the true measured depth map of lower Gordon Stray top structure, a pressure map (Figure 3-9c) is generated with the aid of an average pressure gradient of 10.52 Mpa/km (0.465 psi/ft.). The minimum pressure in Jacksonburg-Stringtown oil field is estimated to be 1230 psi, which when compared to the empirical MMP value of 941 psi indicates that CO_2 is miscible with the reservoir oil.

3.5 Petrophysical Properties Prediction by Support Vector Machine and Artificial Neural Network

Porosity is a fundamental and essential property to characterize the storage capability of hydrocarbon bearing formations in reservoirs. Permeability evaluation has a significant impact on injection rates, reservoir management, and flow patterns determination. Thus, accurate porosity and permeability estimates are the key features for reservoir characterization and geological modeling at the field scale.

As discussed, artificial neural network and support vector machine, which are excellent means of dealing with complex nonlinear problems in the petroleum industry, are utilized to construct the regression relationship between wireline logs and core data (Zhong and Carr, 2017 a). Hybrid particle swarm optimization (PSO) mixed kernels function (MKF) based support vector machine (SVM) is used to establish a porosity prediction model. The mixed kernels function has both strong extrapolation and interpolation ability, and compensate for the weakness of any single kernel function. Moreover, the particle swarm optimization algorithm is utilized to locate global optima of training and testing parameters in MKF-SVM. A permeability prediction model is established with the aid of artificial neural network, which is very flexible in the design of network architecture and selection of training and testing parameters. After comparing the performance of different ANN architecture and leaning algorithm, an ANN with one-hidden layer and 25 neurons is constructed and two evolutionary algorithms (genetic algorithm and particle swarm optimization algorithm) are applied to optimize the weights and biases of ANN. There were a variations between core measured data and predicted data set, but the trends of porosity and permeability are consistent with those observed in the formation. Those porosity

and permeability regression models were extended from core-scale (centimeter or inch) to well-scale (meter or foot) to the field-scale (kilometer or mile) (Figure 3-10).

3.6 Three Dimensional Geological Modeling

The borehole data from Gordon Stray formation reflect the petrophysical properties in one-dimension, and provide insufficient information to reflect horizontal heterogeneity and connectivity of Gordon Stray formation; the interpolated two-dimensional cross-sections show the large-scale lateral and vertical distribution of the petrophysical properties, but still cannot quantitatively reflect the distribution of heterogeneity in more detail. Therefore, a 3-D petrophysical properties model is needed to characterize the internal and external architecture of Gordon Stray formation. Deterministic and stochastic method are the two modeling algorithms for continuous property simulation. The deterministic method includes moving average and kriging algorithms. A deterministic model is a model where no randomness is involved in the modeling process, thus the same result will be repeated for every modeling run with the same input data. Moving average uses an interpolation technique, which finds an average of the input data and weighs according to the distance from the wells. A Kriging algorithm can produce repeatable results, which honors local data, and is the preferred approach in areas of abundant hard and soft data. However, very few subsurface areas are understood in such sufficient detail. Thus a stochastic approach is a more common approach (Gunnarsson, 2011; Norden and Frykman, 2013; Schlumberger, 2011). The stochastic method is pixel-based technique, which includes sequence Gaussian simulation (SGS) algorithm and Gaussian random function simulation (GRFS). Typically, sequence Gaussian simulation produces a realization of the targeted property, which honors the well data and also honors a target histogram for the property, usually derived from the well data. The degree of continuity is controlled by the variogram. A long range on the variogram will mean that points spatial far apart are related, while a short range will mean that points are less spatially related. The type of variogram governs the smoothness of the realization. A Gaussian random function simulation (GRFS) performance runs on the same principle as sequence Gaussian simulation does, but it is faster to run and has a fast on-the-fly tabulator to update the model as the correlation coefficient changes (Qi et al., 2007; Wang and Carr, 2013, Schlumberger, 2011).

3.6.1 Well Correlation and Map

Data sets of 179 wells covering the Gordon Stray interval have been correlated across the study area. The Gordon Stray is divided into 4 subunits based on log response. The depositional environments of the units are interpreted to include; barrier sandstone, central bay shale, tidal channels and fluvial sandstone. Numerous cross-sections were built to correlate units with the response of the commonly available logs (Figure 3-11). Preliminary formation top maps for all formations were created (Figure 3-12). Based on the formation tops, an estuarine environment is determined to originate from the northeast and cease in a southward direction close to the southern parts of the field. The interpretation spatially coincides with the large syncline that is evident in the structural map (Figure 3-1b), which can be explained as the earlier tectonic action and later enhanced incision of underlying strata due to rapidly regressing sea. An incised valley floor can serve as the depocenter for the transgressing estuarine sequence.

3.6.2 Stratigraphic Framework

A three-dimensional static model can improve our understanding of the spatial distribution and geometry of porosity and permeability within the Gordon Stray formation. A reasonable 3D static geologic model is constrained by a reasonable and robust stratigraphic framework. The model can be used to calculate the theoretical CO_2 storage capacity and for future reservoir fluid simulation to evaluate EOR potential. A 3D geologic model is a grid-based 3D volume, thus each grid will just have one value for each reservoir property. The size of the grid will reflect the resolution of 3D geologic model. The appropriate selection of the grid size is important to represent the vertical and lateral range of the study area. However, high resolution with small grid size will increase computation time and computer resources. Therefore keeping the balance between computation time and model's resolution is largely based on the suitable grid size.

As interpreted in figure 3-12, four original structural surfaces of Gordon Stray formation were used to construct the stratigraphic framework reflecting the main reservoir intervals (Figure 3-13). The grid size of x and y dimensions are equal at 150 ft. (50m), and the grid size of vertical direction is 1ft. Thus there are 110 grid blocks in x-direction, 74 grid blocks in y-direction and 64 layers in vertical direction. The cell layers are parallel to the underlying surface in each

reservoir interval, with cells truncating against upper bounding surfaces. The resulting total number of cells in the model is 539,904.

3.6.3 Geostatistical Analysis of Gordon Stray Formation

Extrapolation of petrophysical properties from wells to areas of limited well control is the essential step to construct 3-D gridded petrophysical property model for the study area (i.e., porosity model and permeability model). Vertical and lateral distribution of those properties was determined based on the knowledge of geological background, depositional environment, and geostatistical analysis from conventional well logs (e.g., Brett and Baird, 1996; Lash, 2008; Boyce and Carr, 2010; Lash and Engelder, 2011). The deposition trend of Gordon Stray formation is major north-northeast direction. As a result, we set the major direction of geostatistical ellipse to 36° for the 3-D geologic modeling the Gordon Stray formation. In the vertical direction, the distribution of porosity and permeability is really depended on the wave and river energy.

The semi-variogram is a simplified representation of porosity and permeability distribution pattern. I analyzed and created different variogram for each Gordon Stray subunit (Figure 3-13). The variogram developed for the vertical distribution of porosity and permeability was applied to control the spatial extrapolation of Gordon Stray formation in stochastic method

Based on the core measurement and conventional well log analysis, the porosity and permeability values ranges from 3.4% to 25.4% and 0.11 mD to 257 mD, respectively. For Gordon Stray formation, the data distribution was analyzed abd shown as histograms. Comparing core and combined core-log evaluated porosity and permeability for Gordon Stray formation, Figure 3-14 represent a generally good agreement between core measure data and predicted date for porosity and permeability. Based on the histograms (Figure 3-14a and b), predicted permeability and core measured permeability have nearly identical trends, and the statistical parameters are comparable. The QQ-plots for comparing the distribution of the core and log-based permeability have a slightly different shape, and data points almost locate on the 45^0 line (Figure 3-14e). Those figures show that core measured porosity distribution values are comparable to the log-predicted permeability distribution. In terms of porosity, core measured porosity and log-predicted porosity still have same trend (Figure 3-14b and d). As Figure 3-14f

shows, data points locate along the diagonal line, reflecting similar shapes of the porosity distribution.

For upscaling of the well-log data, an arithmetic average was applied to the well log derived porosity into the grid with 1 feet vertical scale. The effect is illustrated in Figure 3-15(a), and shows the original well-log derived porosity and the model-grid upscaled porosity. The process of upscaling has a minimal effect on the porosity distribution, which can be seen by comparing histograms of the two data sets.

To model the spatial distribution of permeability, a harmonic average method was utilized to upscale the permeability value. Figure 3-15(b) shows the distribution of permeability after and before upscaling. The shape of distribution for upscaled and raw well log is comparable, indicating that upscaling process has minimal negative impact on the final 3D petrophysical model.

3.6.4 Petrophysical Properties Modeling

The petrophysical properties model is composed of porosity and permeability model. Porosity is predicted based on support vector machine, and permeability is predicted based on the artificial neural network. There are four common simulation algorithms that are used to construct models of continuous variables including kriging, moving average, sequence Gaussian simulation (SGS), Gaussian random function simulation (GRFS). Kriging is the primary deterministic approach used for porosity and permeability modeling and works well with high-density data to avoid over-interpretation of available data. Kriging generated broad regional trends with sharp boundaries among areas of similar porosity and permeability. Kriging is a deterministic method and cannot generate local variation, thus the simulation result will identical. Moving average is another deterministic approach that can be used for porosity and permeability modeling. This algorithm is fast and can create values for individual cells; however it cannot generate values larger or smaller than the minimum and maximum values of the input data. SGS and GRFS are the common stochastic approaches for constructing models of continuous variables, require similar degree of geologic supervision and computational overhead. GRFS can produce local variation and generate numerous variations of output models depending on the number of runs.. The Gaussian random function simulation approach is beneficial for uncertainty analysis, if the

suite of models is reasonable. Another advantage GRFS is that GRFS is faster to run compared with SGS, because of parallelized computation. In this study, GRFS is selected as the simulation method and the up-scaled porosity, permeability and experimental variogram generated by geostatistical analysis are the primary input to generate 3D geological model.

Based on the stratigraphic framework, the distribution of petrophysics properties in Gordon Stray unit were constructed using Random Gaussian Function Simulation method based on well data and derived data. Three-dimensional results illustrate the reservoir architecture, porosity, permeability and convective of this formation in a high-resolution grid (Figures 4-17 and 4-18). In the 3D porosity model the 3D porosity model (Figure 4-17) shows a high porosity reservoir located in the southern part of Jacksonburg-Stringtown oil field. This is the area with the highest potential for CO_2 storage and EOR in the field. In addition, the 3D permeability model (Figure 4-18) shows higher permeability values, which should improve injectivity and assist the CO_2-EOR process.

3.6.5 Theoretical CO_2 Storage Capacity

Depending on the well-constructed 3-D static geological reservoir model, using CLSF calculation method, theoretical CO_2 storage capacity can be estimated (Table 3-2). The P_{90} confidence storage capacity is 24 million tons, and the most risked storage capacity P_{10} is 383 million tons.

3.7 Summary and Conclusions

Accurate calculation of theoretical and practical CO_2 storage capacity in mature oil fields depends on the accurate porosity and permeability estimation and geostatistical distribution of petrophysical properties. Reasonable and robust 3D static geological model help to improve the understanding of CO_2/oil flow patterns during CO_2 enhanced oil recovery process, and to determine optimal CO_2 storage locations. Artificial neural network and support vector machine, used detailed core and well log data to construct the relationship between log data and core measured porosity and permeability data. Geostatistical methods were used to generate the porosity and permeability distribution pattern across the field. Pixel-based simulation was used to assign porosity and permeability to subunits according to the spatial relationship built from

semi-variogram analysis. A 3D fluid model can be constructed by up-scaling the well-defined 3D static geologic model, and can be used to explore the fluid flow pattern and to calculate the final oil recovery factor and practical CO_2 storage capacity.

The following specific conclusion are drawn from this study:

(1) Porosity and permeability can be predicted based on core and conventional wireline logs including gamma ray, bulk density, slope of gamma ray, slope of density and V_{sh}.
(2) An artificial neural network was constructed to predict permeability value. Tow evolutionary algorithm including particle swarm optimization algorithm and genetic algorithm were applied to optimize the weights and biases in order to get best regression performance.
(3) A support vector machine regression model was trained and tested to estimate porosity value. A newly developed mixed kernels function was applied to improve the generalization ability of support vector machine. Particle swarm optimization algorithm was also used to locate the best training and testing parameters.
(4) A regional 3D structure and Lithofacies framework were constructed by integrating well logs and detailed core description, which can be used to constrain the future 3D petrophysical
(5) Gaussian random function simulation with a suitable variogram model of porosity and permeability was applied to generate 3D petrophysical model in Gordon Stray formation.
(6) The distribution of porosity and permeability shows that the best regions for CO_2 storage and CO_2-EOR are located in southern regions of Jacksonburg-Stringtown oil field due to the high porosity and high permeability.
(7) Based on the calculation, the most confidence theoretical storage capacity is 24 million tons, and most risk theoretical storage capacity is 383 Mt.

Acknowledgement

This work was funded by the US-China Clean Energy Research Center, Advanced Coal Technology Consortium, under grant DE-PI0000017 from the National Energy Technology Laboratory of the US Department of Energy.

Reference

Ameri, S., Aminian, K., Avary, K.L., Bilgesu, H.I., Hohn, M.E., McDowell, R.R., Patchen, D.L., 2002. Reservoir characterization of Upper Devonian Gordon sandstone, Jacksonburg-Stringtown Oil Field, Northwestern West Virginia. West Virginia University.

Bachu, S., 2000. Sequestration of CO_2 in geological media: criteria and approach for site selection in response to climate change. Energy Conversion and Management 41, 953-970.

Bachu, S., Bonijoly, D., Bradshaw, J., Burruss, R., Holloway, S., Christensen, N.P., Mathiassen, O.M., 2007. CO_2 storage capacity estimation: Methodology and gaps. International Journal of Greenhouse Gas Control 1, 430-443.

Bergerud, B., 2011. Reservoir model of the Jacksonburg-Stringtown Oil field; Northwestern West Virginia: potential for miscible CO_2 Enhanced Oil Recovery, Department of geology and geography. West Virginia University.

Bjerstedt, T.W., Kammer, T.W., 1987. Genetic stratigraphy and depositional systems of the Upper Devonian-lower Mississippian Price-Rockwell deltaic complex in the central Appalachians, U.S.A. Sedimentary Geology 54, 36.

Blakey, R., 2008. Paleogeography and geologic evolution of North America, Northern Arizona University.

Boswell, R., 1988. Basin analysis of the Acadian clastic wedge in northern West Virginia and adjacent areas, College of Arts and Sciences Vest Virginia University.

Boswell, R.M., 1985. Stratigraphy and sedimentation of the Acadian clastic wedge in northern West Virginia, College of Arts and Sciences West Virginia University.

Boswell, R.M., Donaldson, A.C., 1988. Depositional architecture of the Upper Devonian Catskill Delta complex: central Appalachian basin, U.S.A. Memoir Canadian Society of Petroleum Geologists 2, 19.

Bridege, J.S., Willis, B.J., 1994. Marine transgerssions and regressions recorded in middle Devonian shore-zone deposits of the Catskill clastic wedge. GSA Bulletin 106, 18.

Buatois, L.A., Mángano, G.M., Alissa, A., Carr, T.R., 2002. Sequence stratigraphic and sedimentologic significance of biogenic structures from a late Paleozoic marginal-to open-marine reservoir, Morrow Sandstone, subsurface of southwest Kansas, USA. Sedimentary Geology 152, 99-132.

Buatois, L.A., Mangano, M.G., Carr, T.R., 1999. Sedimentology and inchology of paloezoic estuarine and shoreface reservoirs, Morrow sandstone, Lower Pennsylvanian of southwest kansas, USA, Lawrence, KS. Kansas Geological Survey.

Coughlin, M.F., 2009. Subsurface mapping and reservoir analysis of the Upper Devonian Venango and Bradford groups in Westmoreland County, Pennsylvania, Department of Geology and Geography. West Virginia University, p. 114.

Crawford, E., 1997. Arrhenius' 1896 model of the greenhouse effect in context. AMBIO-STOCKHOLM 26, 6-11.

Dalrymple, R.W., Zaitlin, B.A., Boyd, R., 1992. Estuarine facies models: conceptual basis and stratigraphic implications. Journal of Sedimentary Research 62.

Dennison, J.M., 1985. Catskill Delta shallow marine strata. Geological Society of America Special Papers 201, 91-106.

Donald L. Woodrow, A.M.I., 1983. Facies, topography, and sedimentary processes in the Catskill Sea (Devonian),. Geological Society Of America Bulletin 94, 11.

Faill, R.T., 1997a. A geologic history of the north-central Appalachians, Part 1: Orogenesis from the Mesoproterozoic through the Taconic Orogeny. American Journal of Science 297, 551-619.

Faill, R.T., 1997b. A geologic history of the north-central Appalachians, Part 2: The Appalachian basin from the Silurian through the Carboniferous. American Journal of Science 297, 729-761.

Falkowski, P., Scholes, R., Boyle, E.e.a., Canadell, J., Canfield, D., Elser, J., Gruber, N., Hibbard, K., Högberg, P., Linder, S., 2000. The global carbon cycle: a test of our knowledge of earth as a system. Science 290, 291-296.

Gunnarsson, N., 2011. 3D modeling in Petrel of geological CO_2 storage site, Department of Earth Sciences. Uppsala University, p. 52.

Hohn, M.E., 2004. Petroleum reology and reservoir characterization of the Upper Devonian Gordon sandstone, Jacksonburg-Stringtown Oil Field, Northwestern West Virginia. West Virginia Geological and Economic Survey, Morgantown.

Jiang, P., Li, X., Xu, R., Wang, Y., Chen, M., Wang, H., Ruan, B., 2014. Thermal modeling of CO_2 in the injection well and reservoir at the Ordos CCS demonstration project, China. International Journal of Greenhouse Gas Control 23, 135-146.

McBride, P.S., 2004. Facies analysis of the Devonian Gordon Stray sandstone in West Virginia, Department of Geology and Geography. West Virginia University, p. 145.

Norden, B., Frykman, P., 2013. Geological modeling of the Triassic Stuttgart Formation at the Ketzin CO 2 storage site, Germany. International Journal of Greenhouse Gas Control 19, 756-774.

Peters, G.P., Andrew, R.M., Boden, T., Canadell, J.G., Ciais, P., Le Quéré, C., Marland, G., Raupach, M.R., Wilson, C., 2012. The challenge to keep global warming below 2 °C. Nature Climate Change 3, 4-6.

Piotrowski, R.G., Harper, J.A., 1979. Black shale and sandstone facies of the Devonian Catskill clastic wedge in the subsurface of western Pennsylvania, EGSP. Pennsylvania Department of Environmental Resources Bureau of Topographic and Geologic Survey, p. 42.

Qi, L., Carr, T.R., Goldstein, R.H., 2007. Geostatistical three-dimensional modeling of oolite shoals, St. Louis Limestone, southwest Kansas. AAPG Bulletin 91, 69-96.

Quéré, C.L., Andres, R.J., Boden, T., Conway, T., Houghton, R., House, J.I., Marland, G., Peters, G.P., van der Werf, G., Ahlström, A., 2013a. The global carbon budget 1959–2011. Earth System Science Data 5, 165-185.

Quéré, L.C., Moriarty, R., Andrew, R.M., Peters, G.P., Ciais, P., Friedlingstein, P., Jones, S.D., Sitch, S., Tans, P., Arneth, A., 2014. Global carbon budget 2014. Earth System Science Data Discussions 7, 521-610.

Quéré, L.C., Peters, G.P., Andres, R.J., Andrew, R.M., Boden, T., Ciais, P., Friedlingstein, P., Houghton, R.A., Marland, G., Moriarty, R., 2013b. Global carbon budget 2013. Earth Syst. Sci. Data Discuss 6, 689-760.

Rodhe, H., Charlson, R., Crawford, E., 1997. Svante Arrhenius and the greenhouse effect. Ambio, 2-5.Zhao, X., Liao, X., 2012. Evaluation method of CO_2 sequestration and enhanced oil recovery in an oilrReservoir, as applied to the Changqing Oilfields, China. Energy & Fuels 26, 5350-5354.

Wang, G.C., Carr, T.R., 2013. Organic-rich Marcellus Shale lithofacies modeling and distribution pattern analysis in the Appalachian Basin. AAPG Bulletin 97, 2173-2205.

Zhong, Z., Carr, T.R., 2017 a. Application of improved back-propagation neural network in permeability prediction: a case study in Jacksonburg-Stringtown oil field, West Virginia. Computer & Geoscience, under review.

Zhong, Z., Carr, T.R., 2017 b. Application of A New Hybrid PSO-Mixed Kernels Function (MKF) Based Support Vector Machine (SVM) Model for Reservoir Porosity Prediction: A Case Study in Jacksonburg-Stringtown Oil Filed, West Virginia, USA. under Review.

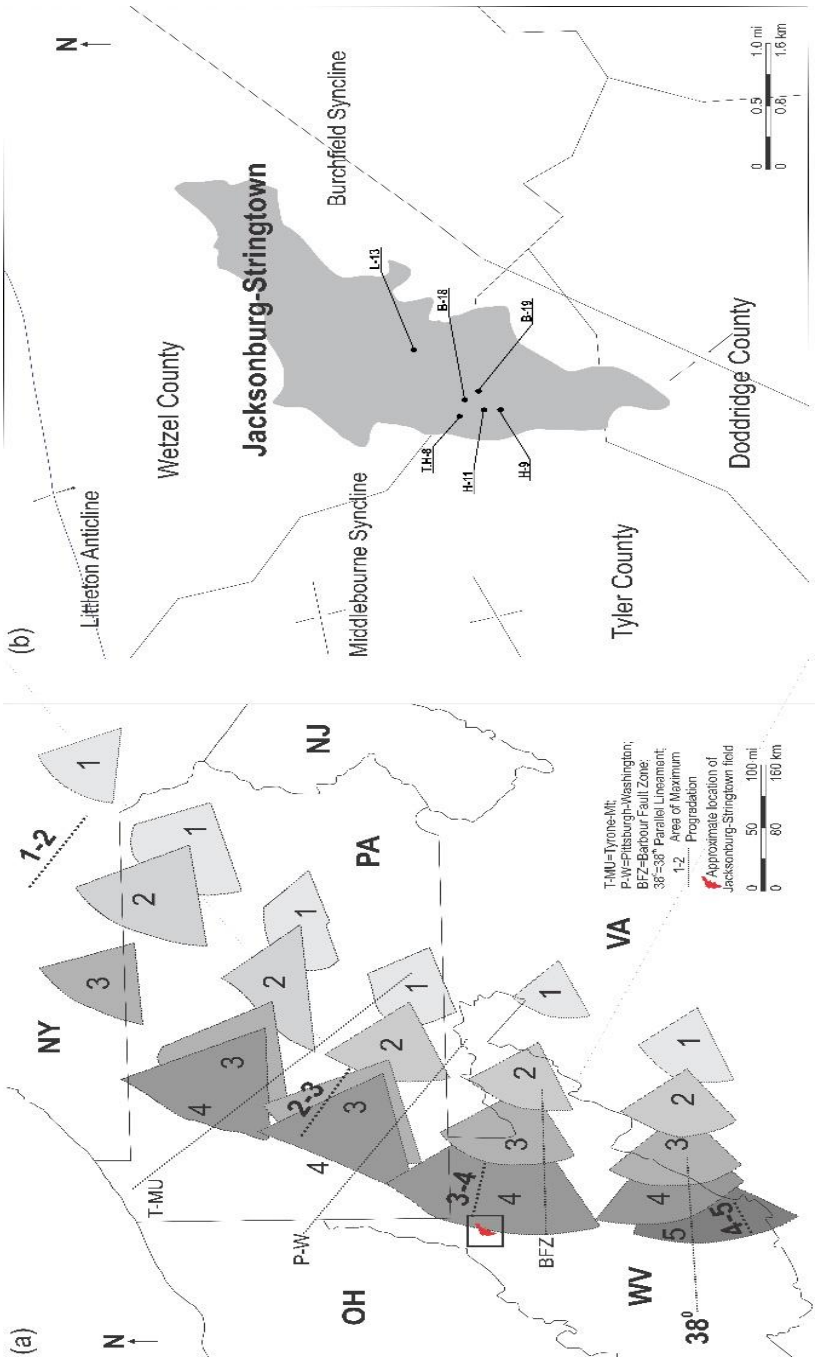


Figure 3-1. (a) Late Devonian paleography of study area in Appalachian basin, five major delta systems prograded westwards and dominated the foreland ramp. Delta systems are separated by the interdeltaic shale; (b) Location of Jacksonburg-Stringtown oil fields in northwestern West Virginia. Black dots indicate cored well with core measured porosity and permeability.

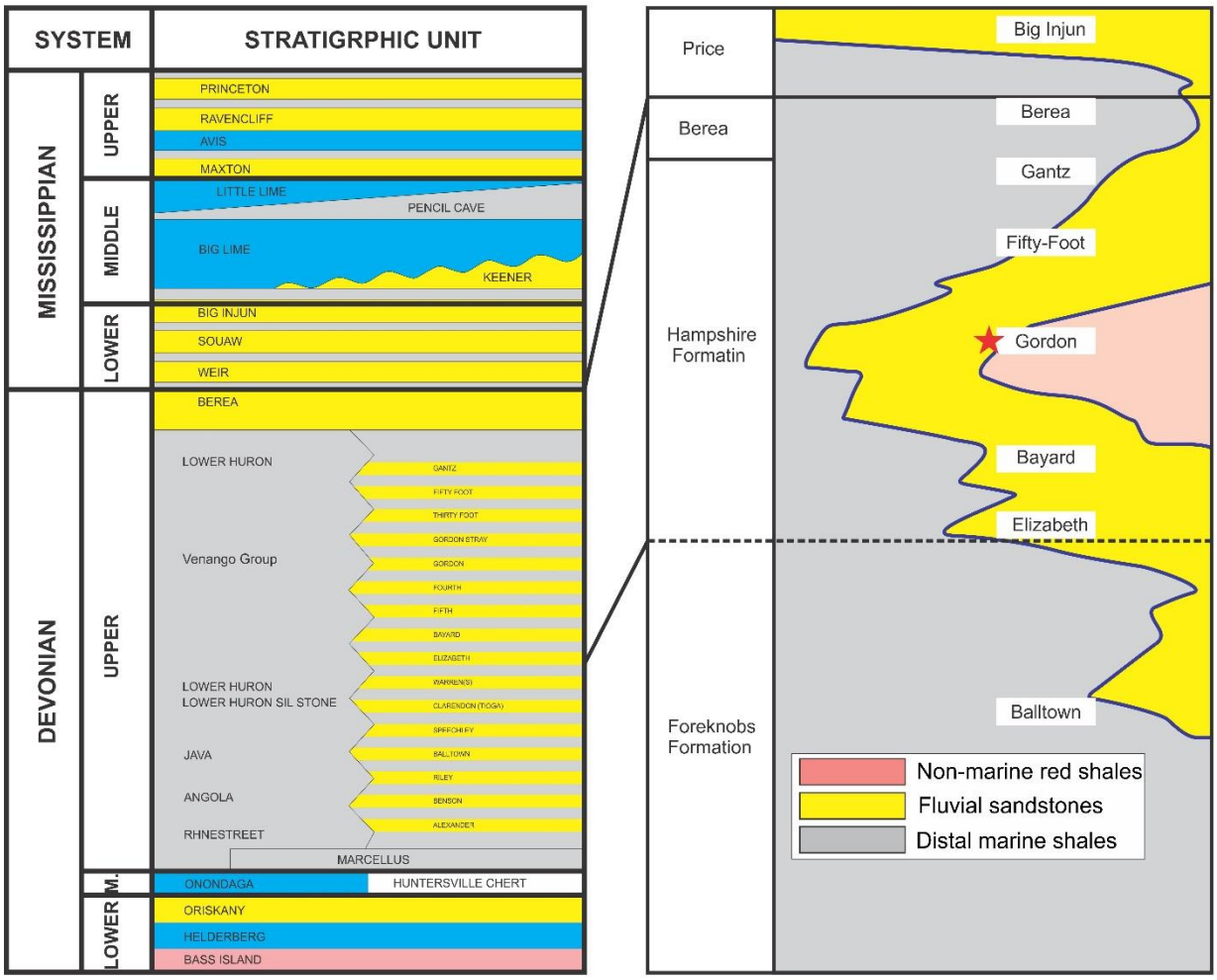


Figure 3-2. Generalized stratigraphic column for the Later Devonian of the Appalachian basin showing subsurface (right) and outcrop (left) terminology. From east to west, red color, yellow and blue color represents non-marine red shales, fluvial sandstones, and distal marine shales respectively (After Ameri et al., 2002).

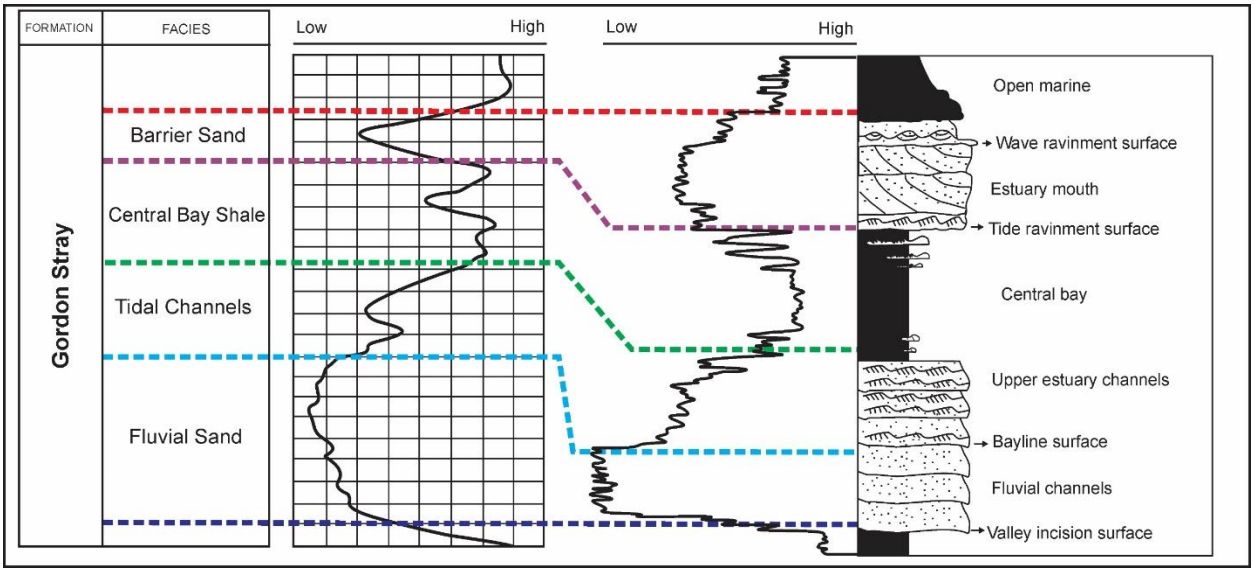


Figure 3-3. Comparison of example Gordon Stray interval gamma log signature from the Jacksonburg-Stringtown field with idealized model.

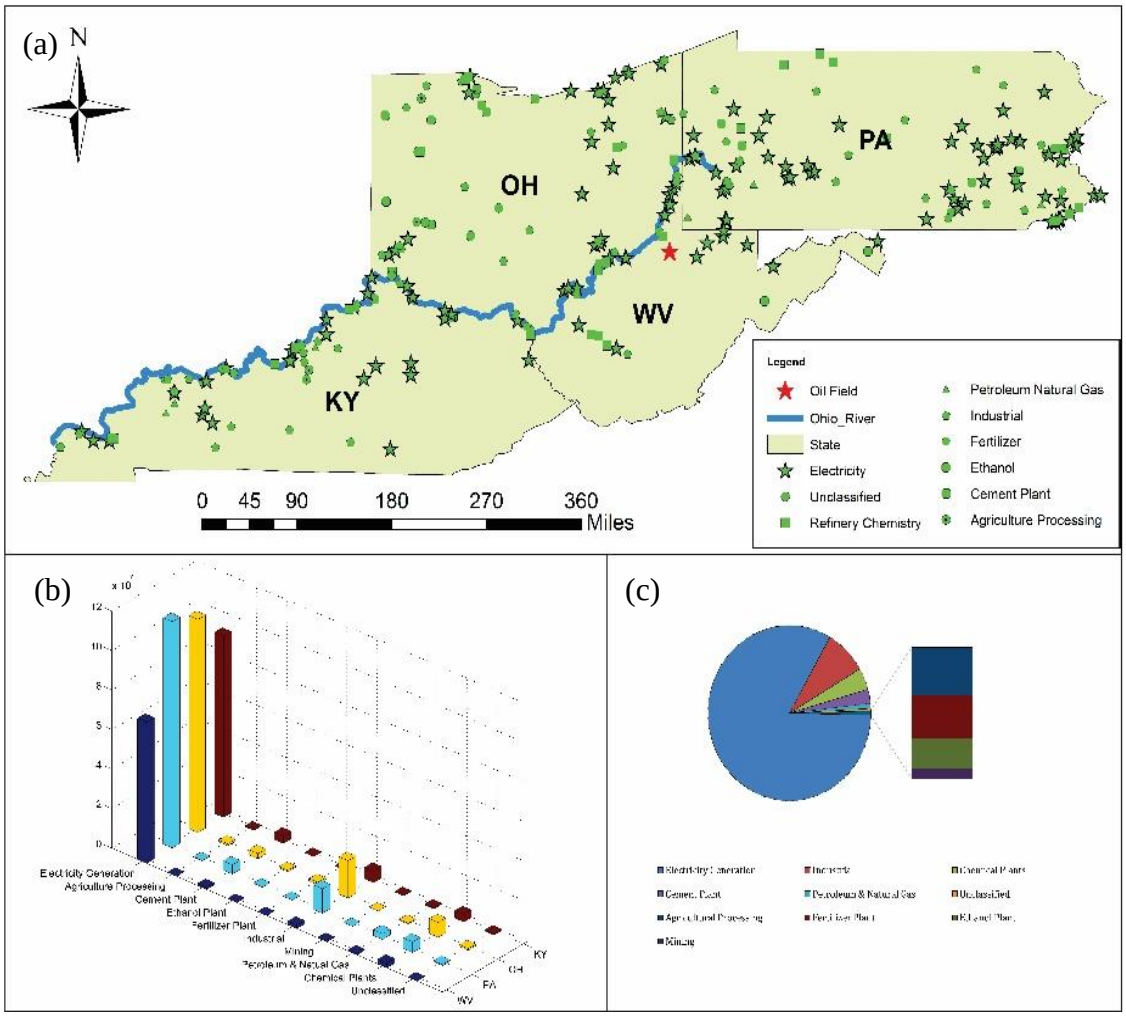


Figure 3-4. (a) Large CO2 stationary sources (metric tons) in the Appalachian basin. Location of the Jacksonburg-Stringtown field is highlighted. Data from the US DOE Carbon Storage and Utilization Atlas; (b) Along the Ohio River, there are numerous power plants, which emit thousands of tons carbon dioxide per year. (c) More than 90% of stationary CO_2 emissions are from electric power plants and industrial manufacturing facilities. Location of viable geologic storage sites for CO_2 in close proximity to the sources will provide economic efficiencies and reduce risks associated with transporting CO_2 long distances from source to sink.

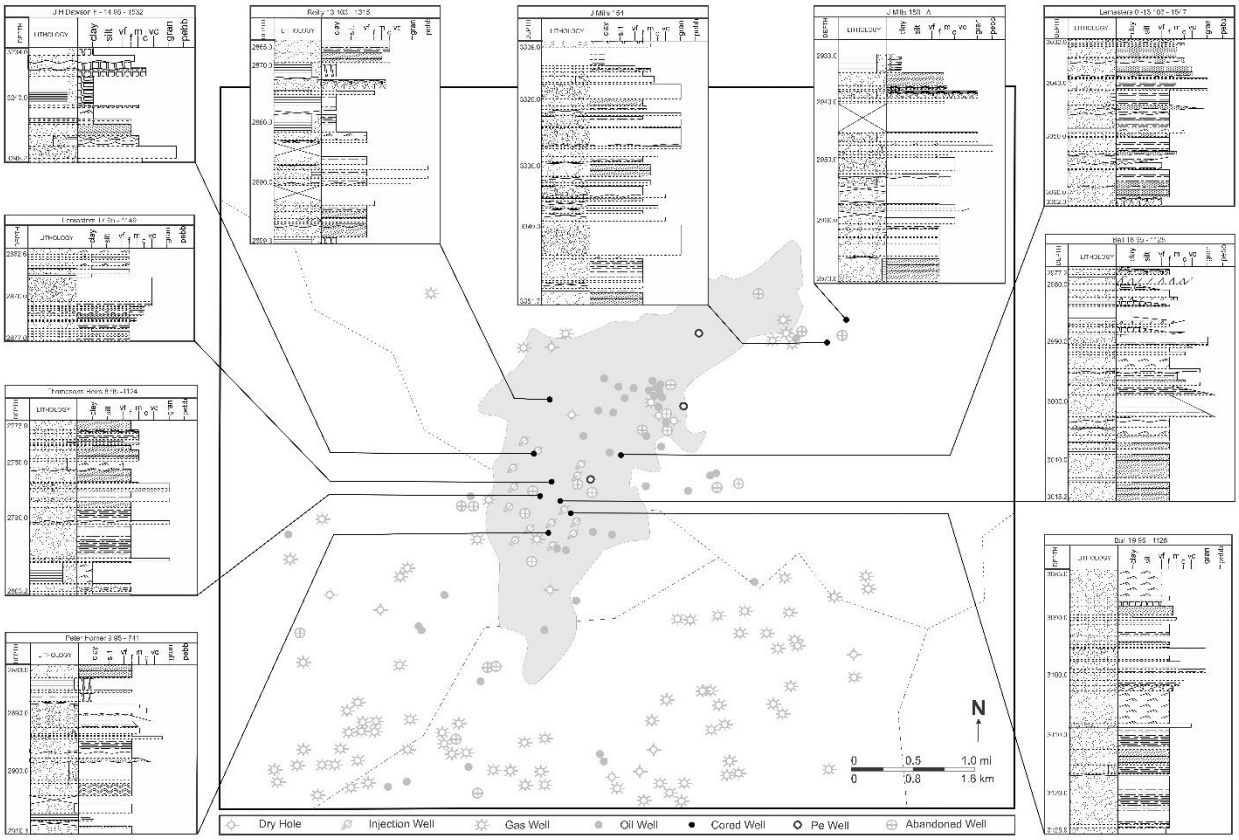


Figure 3- 5. Well locations in research area. Black dots represent cored well which have detailed core description; black circles represent the wells which have photoelectric wireline logs; other markers represent the wells which have common wireline logs, including gamma ray and bulk density wireline logs (modified from McBride et al., 2004).

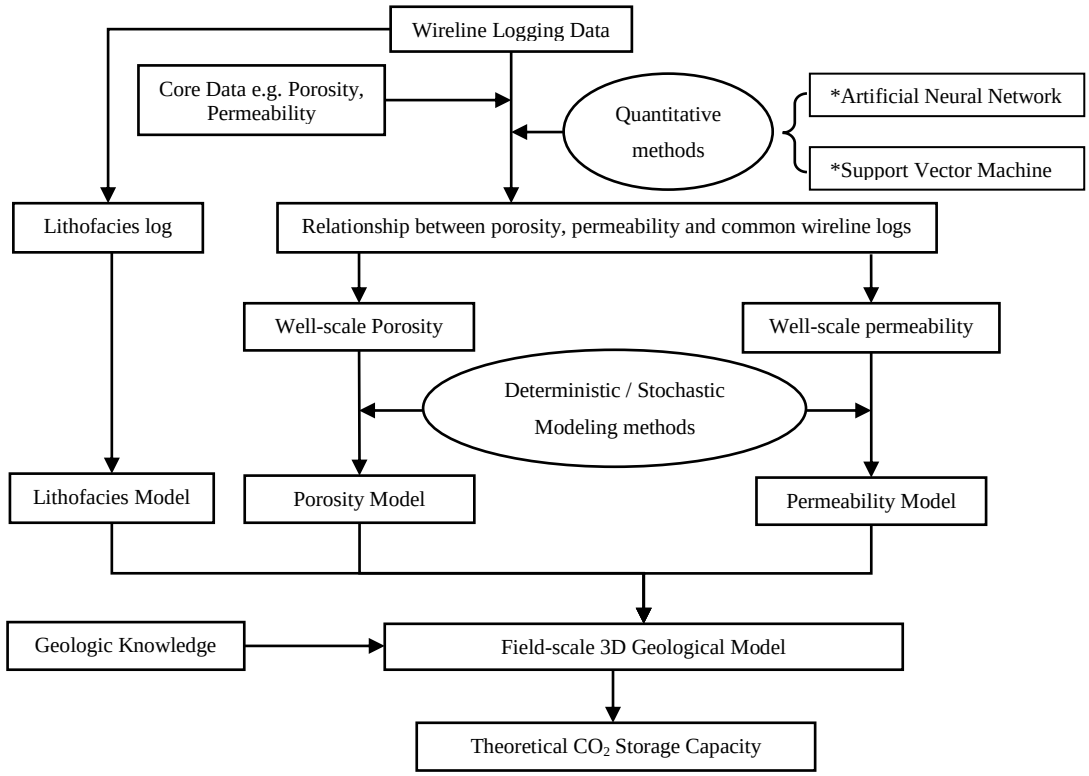


Figure 3-6. Whole workflow of 3D geological model to calculate the theoretic CO_2 storage capacity.

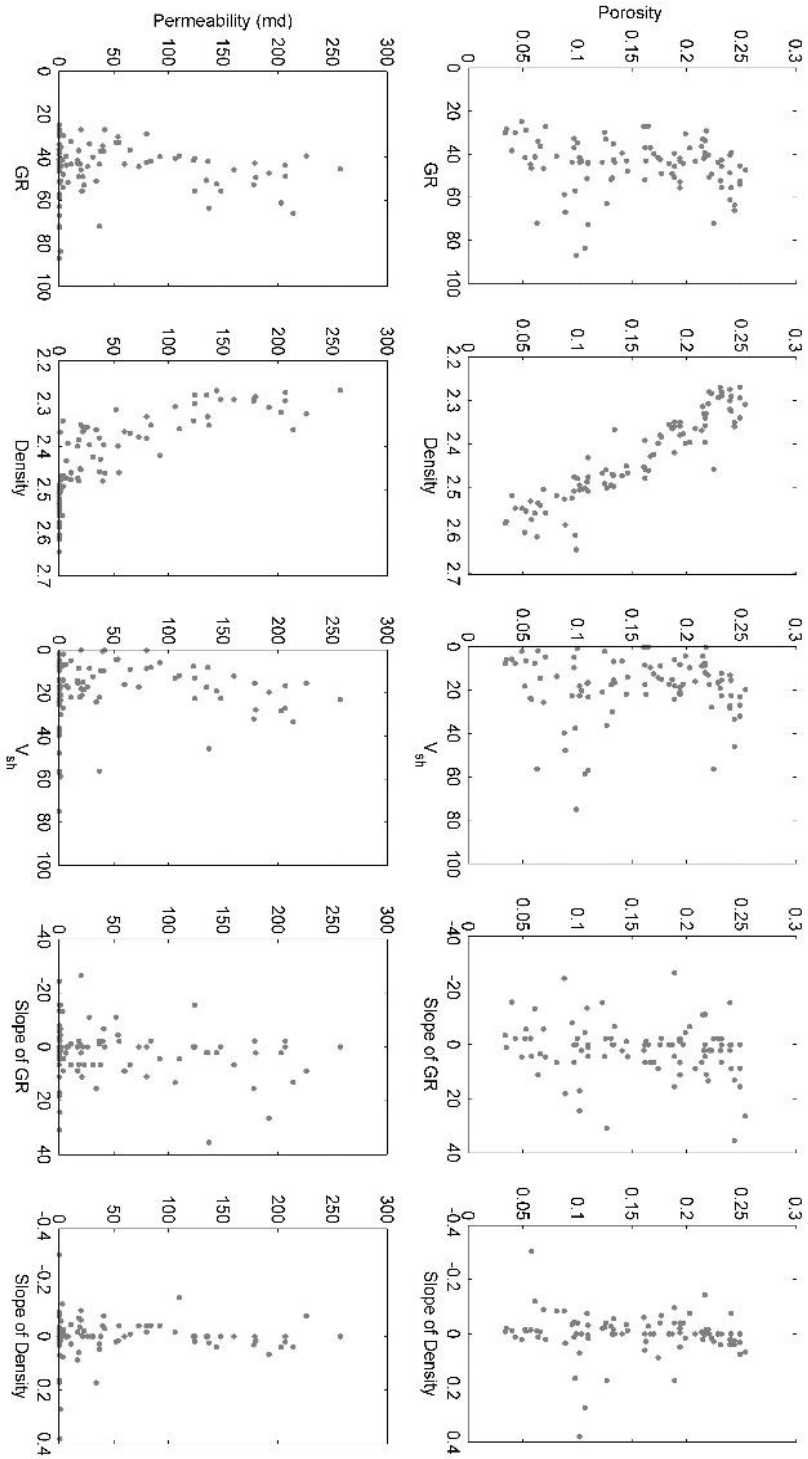


Figure 3-7. Cross-plot of permeability and porosity versus each measured parameters.

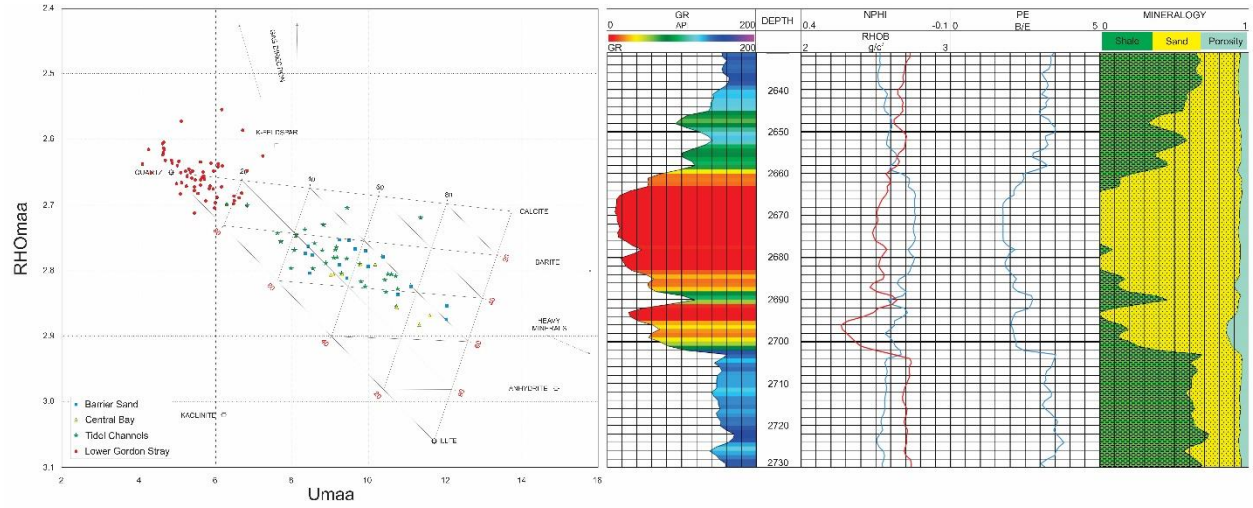


Figure 3-8. (Left)(a) RHOmaa/Umaa plots of Gordon Stray interval. RHOmaa values are measured in g/cm3, and Umaa values are in barnes/electron. a) Sample composition ranges from 0%-20% calcite, 80%-100% quartz, 0%-20% illite. High content of quartz indicate a clean sandstone. Points above the ternary plot indicates the gas in this fluvial sand subunit. (b) Sample composition ranges from 10%-70% calcite, 20%-80% quartz, and 5%-45% illite. The change of lithology is related to variable concentration of sand and shale in tidal channel subunit. (c) Sample composition ranges from 20%-45% calcite, 10%-45% quartz, and 20%-50% illite. High content of illite and calcite in central bay shale subunit indicate a low-energy marine depositional environment. (d) Sample composition ranges from 10%-50% calcite, 10%-55% quartz, and 20%-50% illite. High content of quartz indicate the influence of longshore current, and high calcite levels are a result of marine influence. (Right) Mineral composition of Gordon stray intervals was analyzed with the aid of bulk density, neutron density and photoelectric logs. As expected, lower Gordon Stray (marked as red star) shows high quality of quartz and porosity volume, which indicates that this intervals has high capacity of CO_2 storage.

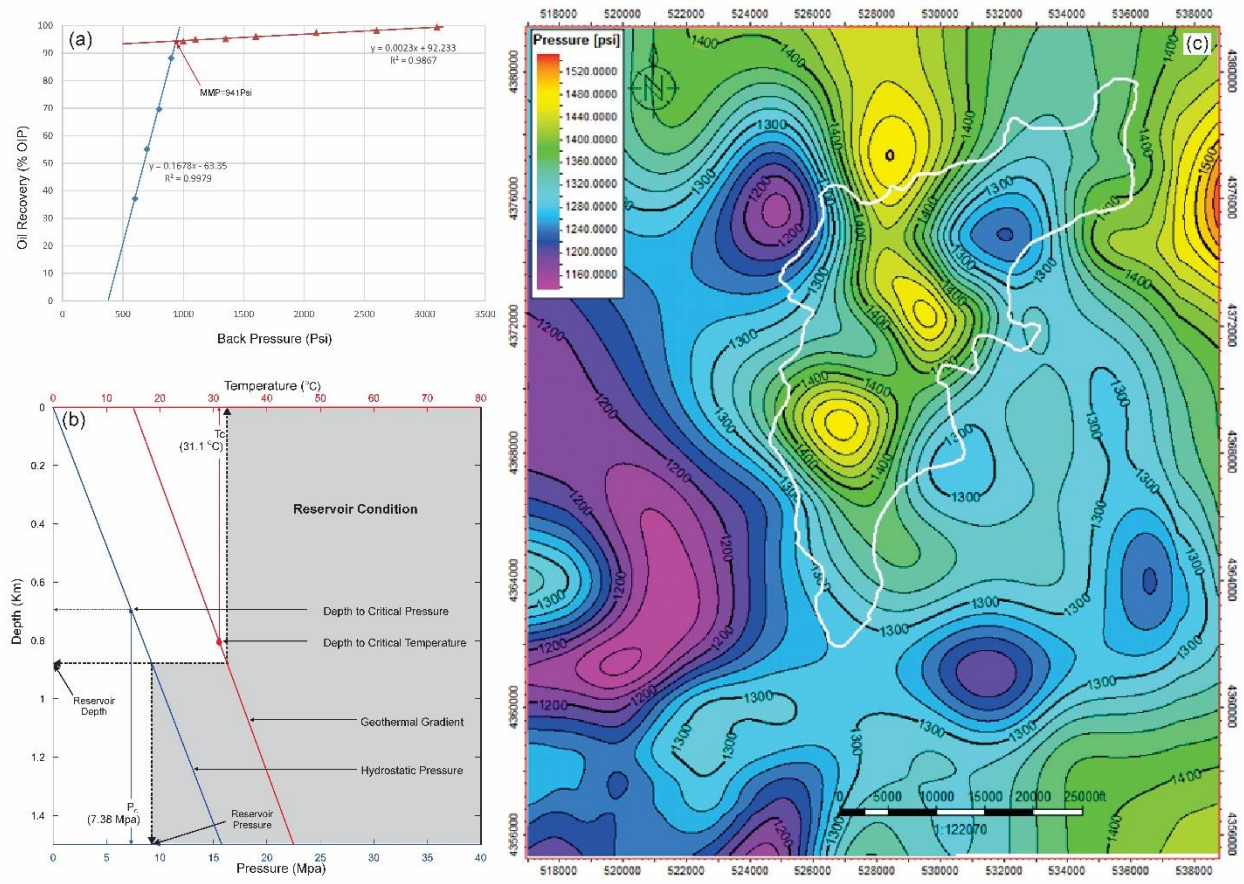


Figure 3-9. (a) Minimum miscibility pressure (MMP) chart with a sample of oil from the Jacksonburg-Stringtown oil field. Performed by Special Core Analysis Laboratories, Inc. (SCAL). (b) Pressure and temperature vs. depth on regional geothermal and pressure gradient. (c) Reservoir pressure map calculated based on statistical pressure gradient for Lower Gordon Stray subunit.

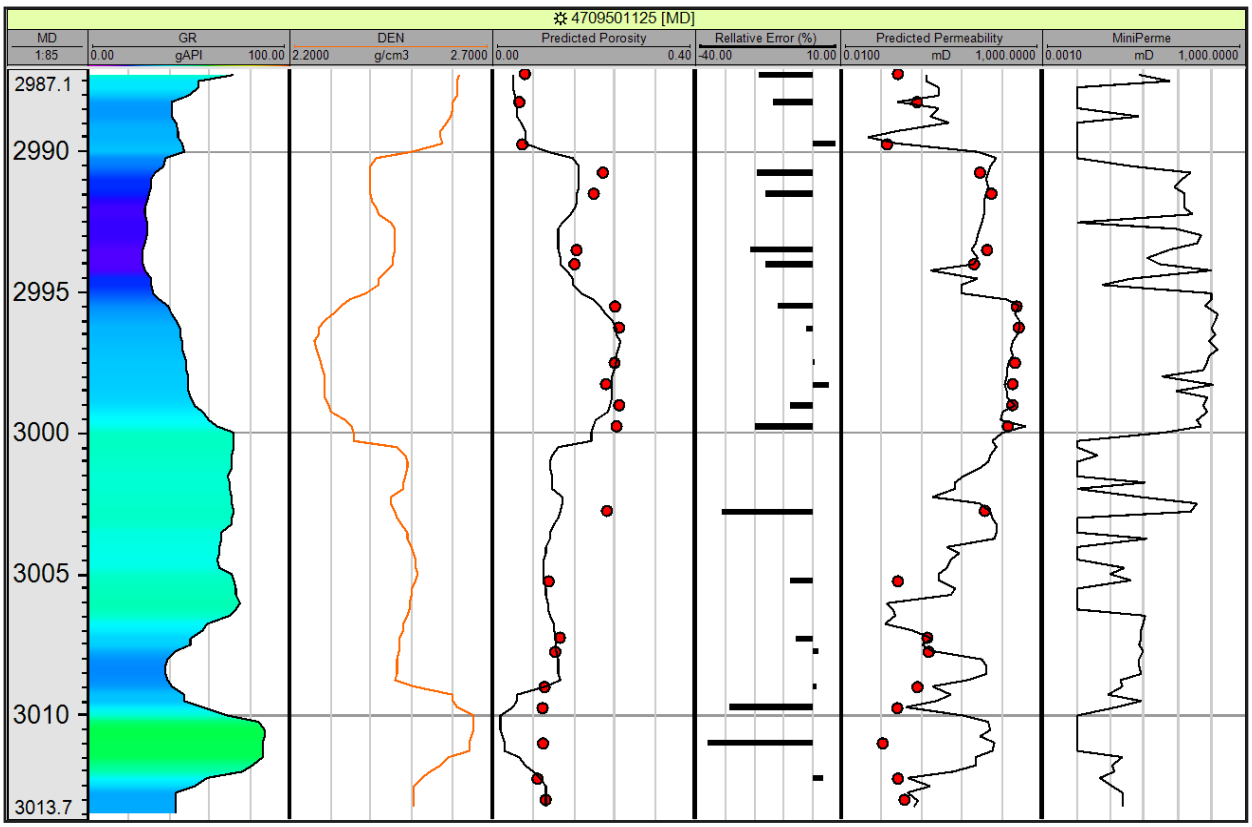


Figure 3-10. Example of well section from Gordon Stray Formation, which shows the predicted porosity and permeability curves by applying laboratory measured core data and conventional well logs as training and testing data set. The predicted porosity and permeability are closely to true measured core data.

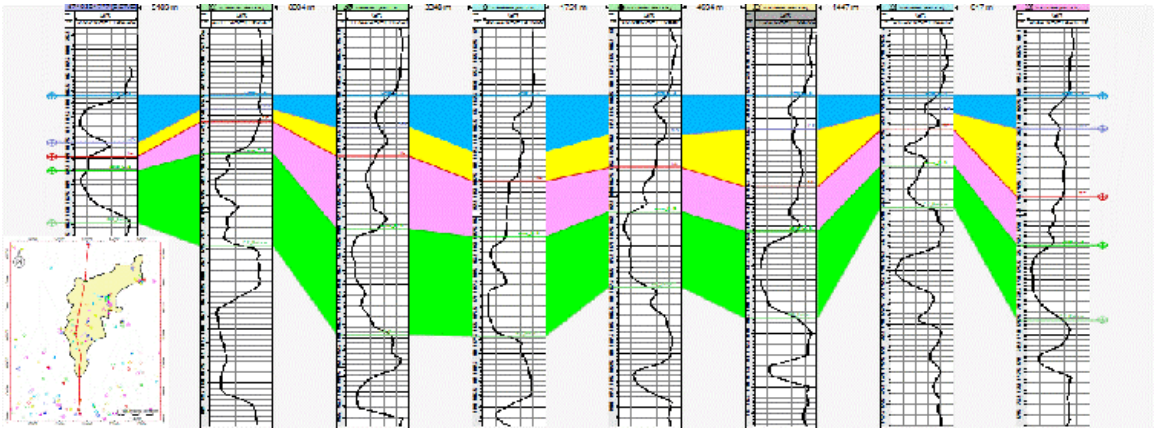


Figure 3-11. Example of cross-section of the Gordon Stray interval and the location of the cross-section (shown as red line).

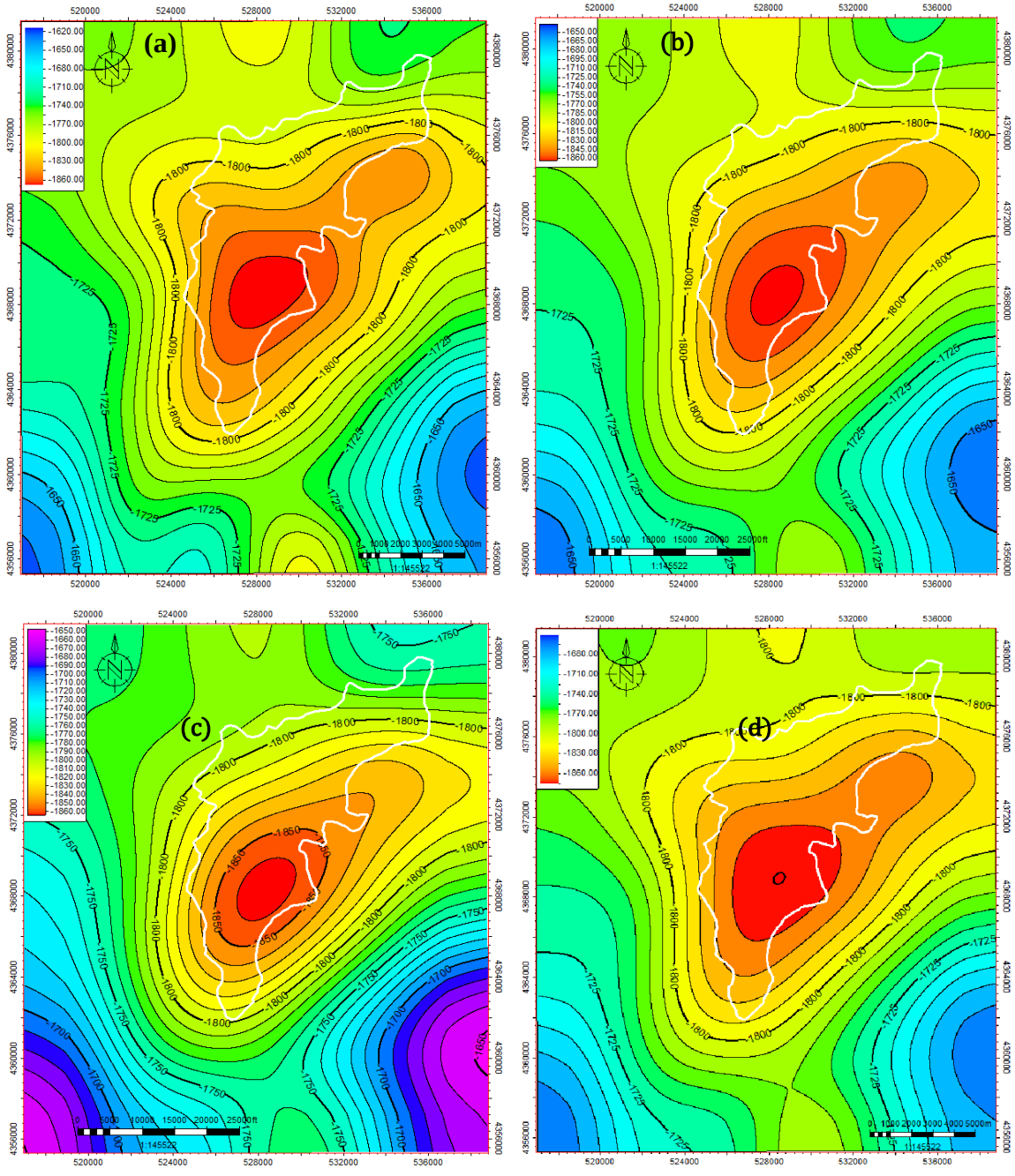


Figure 3-12. (a) Gordon Stray sequence top structure map (10 ft. contour intervals). The white color boundary represents the boundary of Jacksonburg-Stringtown oil field; central Bay shale top structure map; (c) tidal Channels top structure map and (d) lower Gordon Stray top structure map.

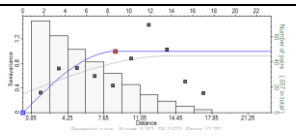
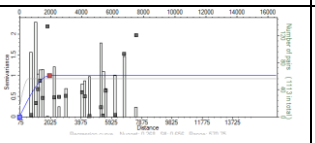
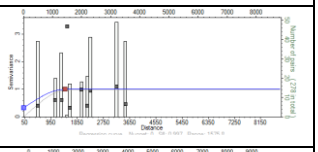
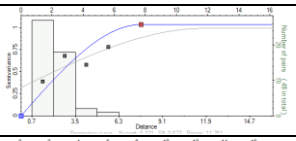
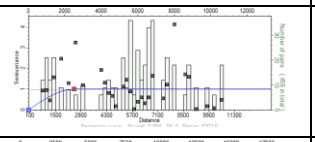
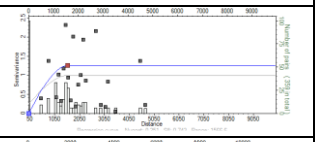
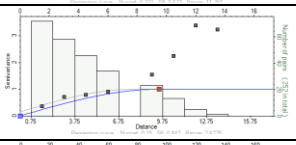
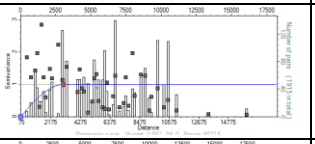
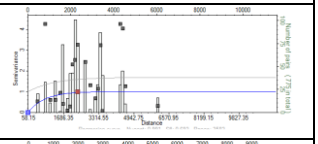
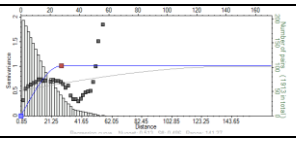
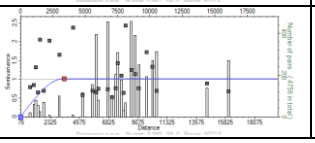
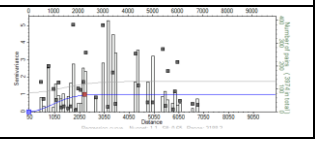
Direction / Subunit	Vertical	Major	Minor
Barrier Sand Subunit			
Central Bay Shale Subunit			
Tidal Channels Subunit			
Fluvial Sand Subunit			

Figure 3-13. Experimental variogram for different Gordon Stray Subunit in Jacksonburg-Stringtown oil field, which is developed based on the up-scaled porosity and permeability logs.

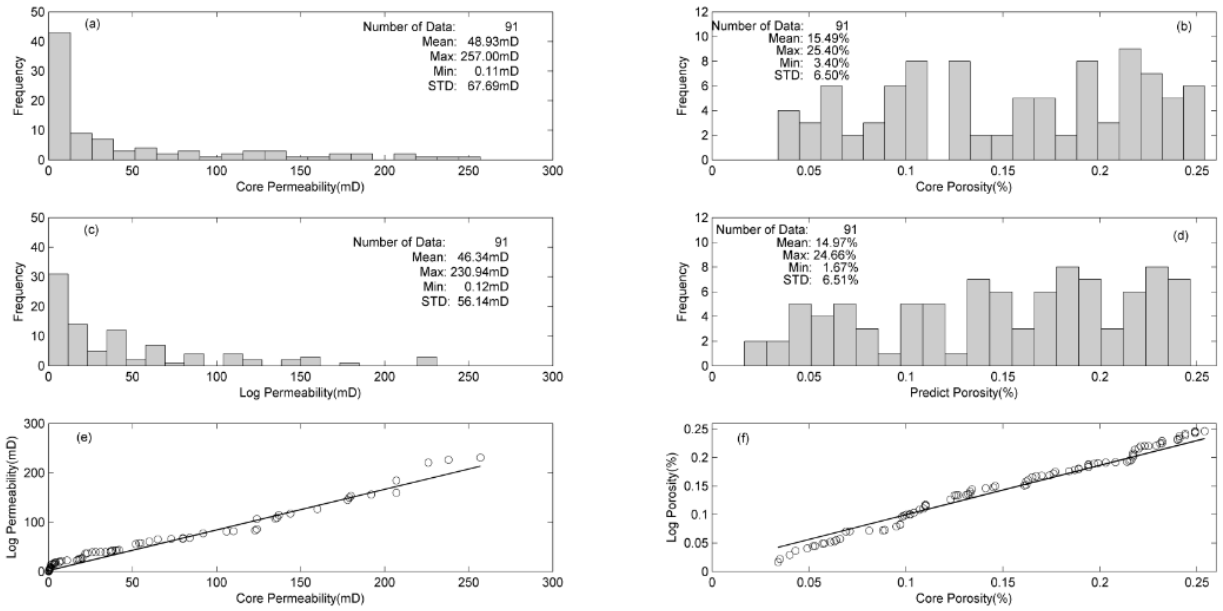


Figure 3-14. Histograms and quantile–quantile (QQ)-plots of total porosity (PHI) and permeability (PERM) for both core and log-evaluated data.

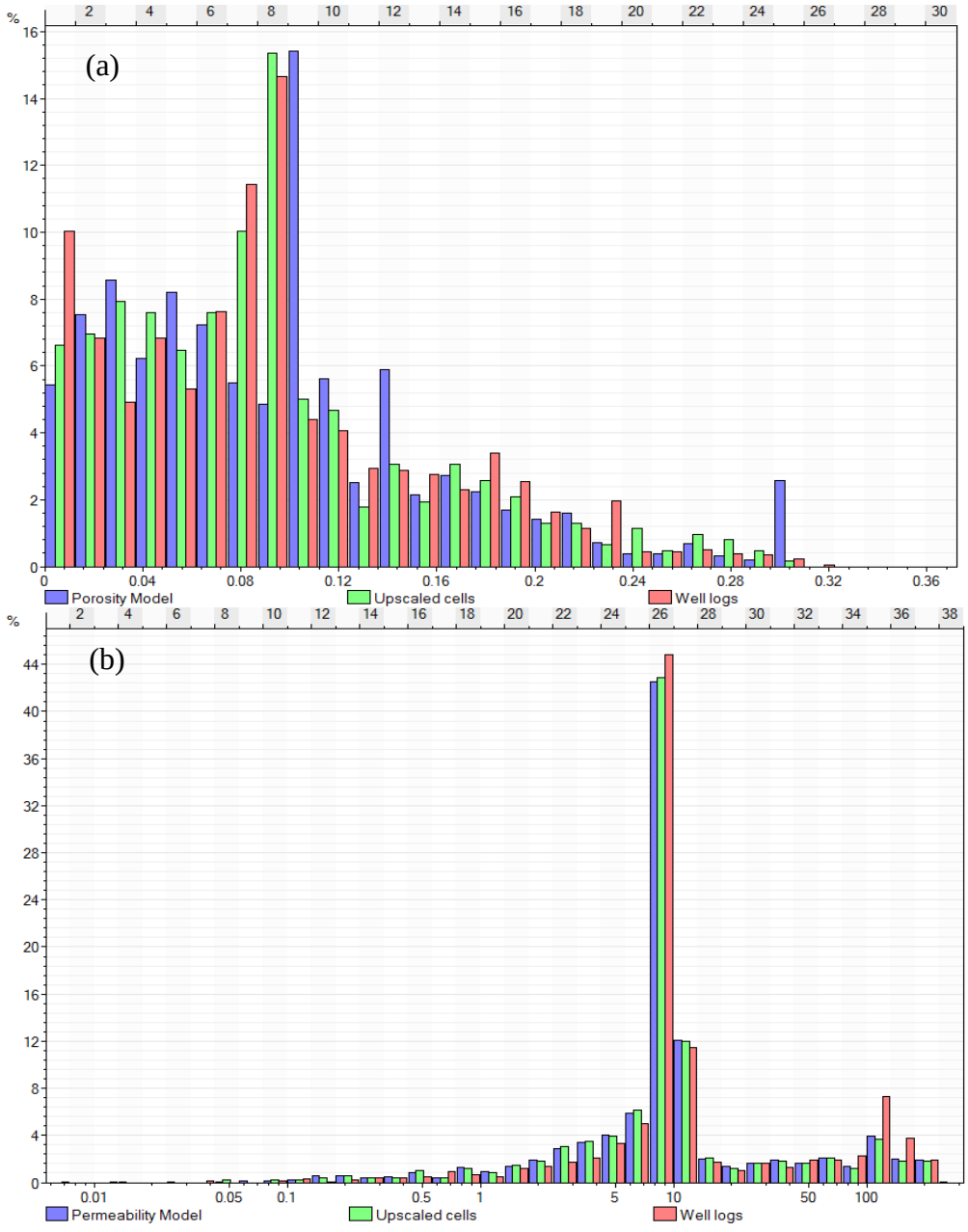


Figure 3-15. A histogram shows the distribution of porosity values (a) and permeability values (b) for raw logs (red color), upscaled cells (green color) and petrophysical model data (purple color).

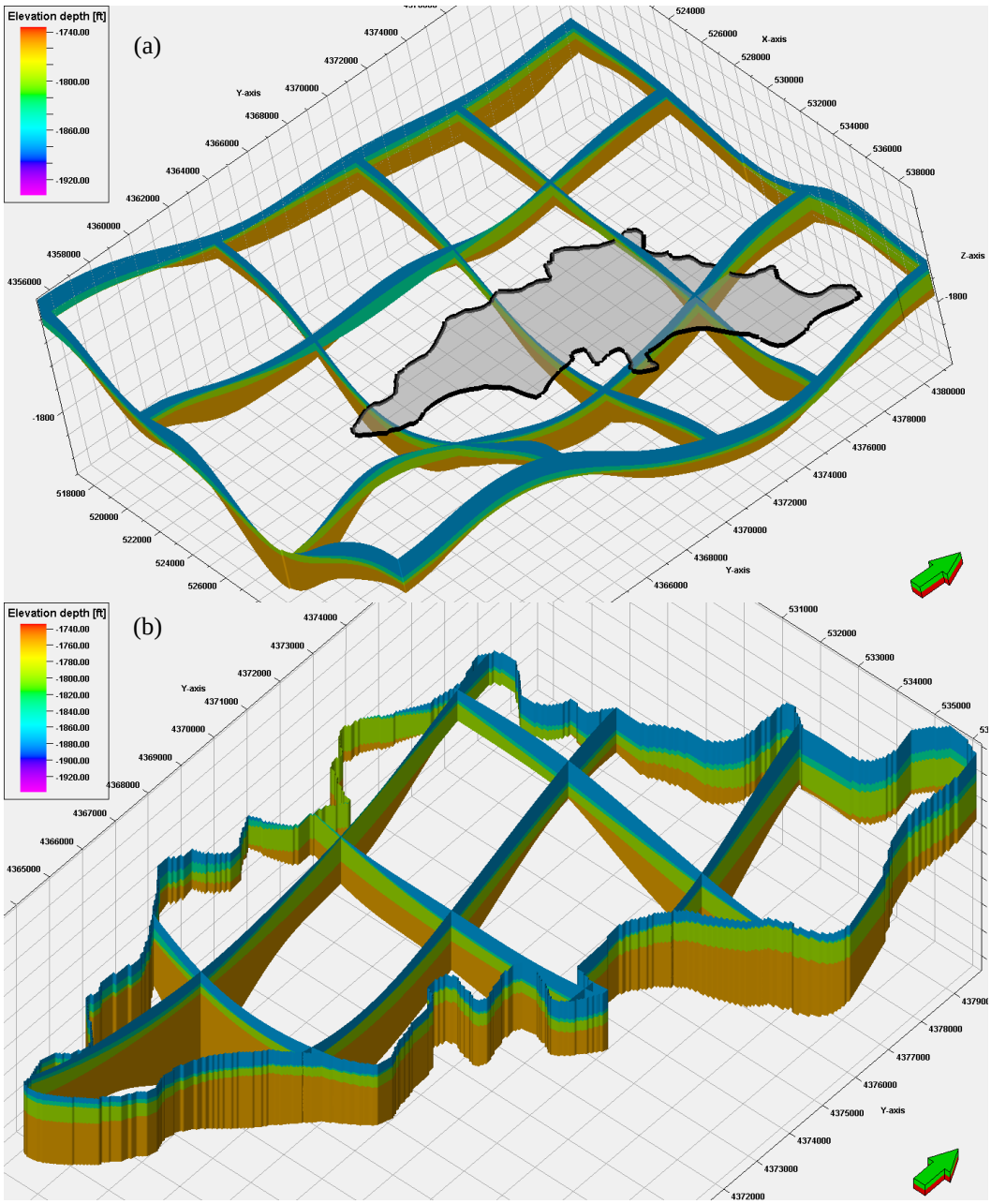


Figure 3-16. (a) Stratigraphic framework of study area and location of Jacksonburg-Stringtown oil field. Each color represents one subunit, which are fluvial subunit, tidal channel subunit, central bay shale subunit and barrier sand subunit. (b) detailed stratigraphic framework of Jacksonburg-Stringtown oil field.

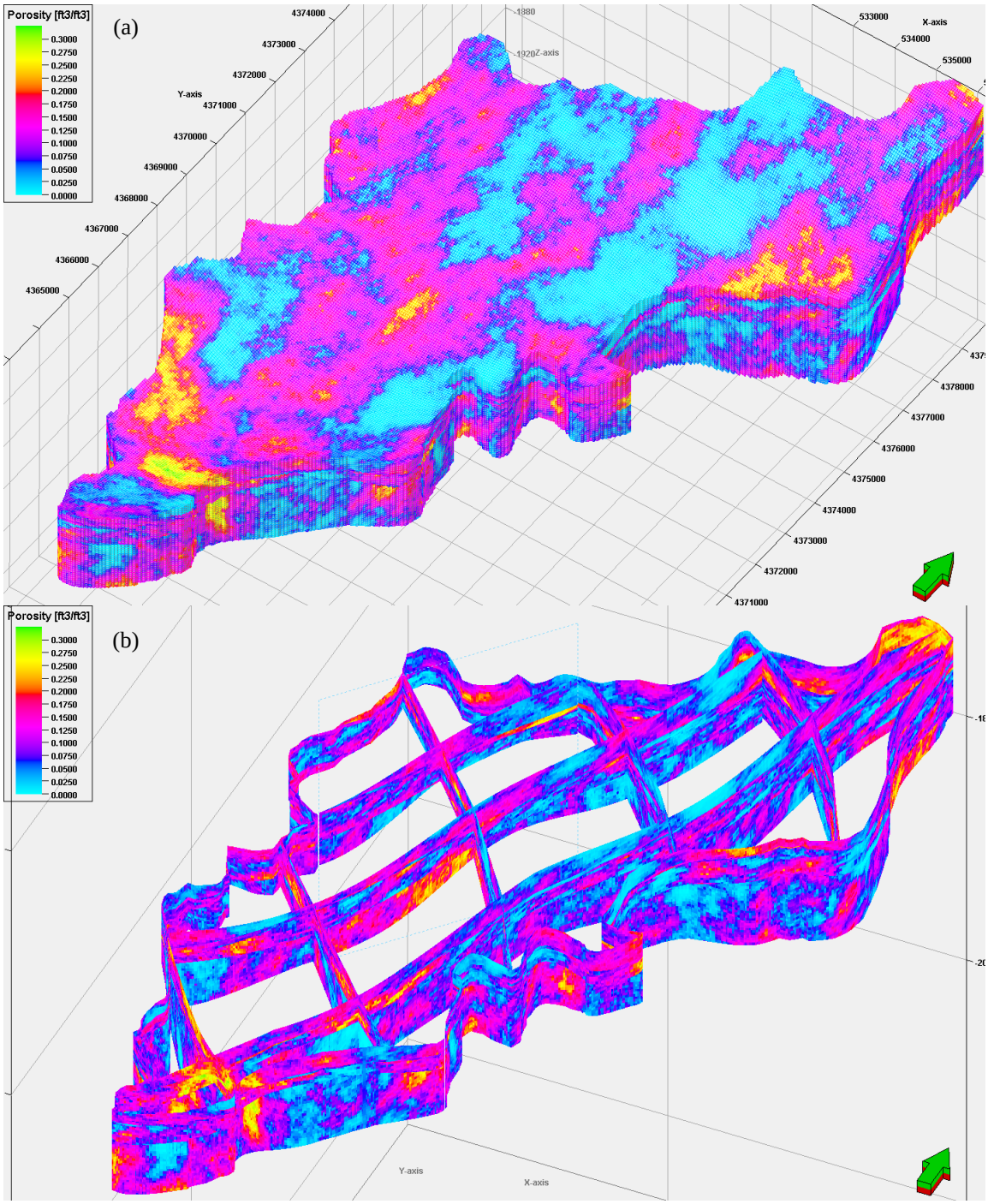


Figure 3-17. (a)Fence diagram shows the distribution of porosity in each subunit; (b) shows a 3-D porosity model with Gaussian Random Function Simulation. The horizontal scale are various, but the vertical scale is 50:1.

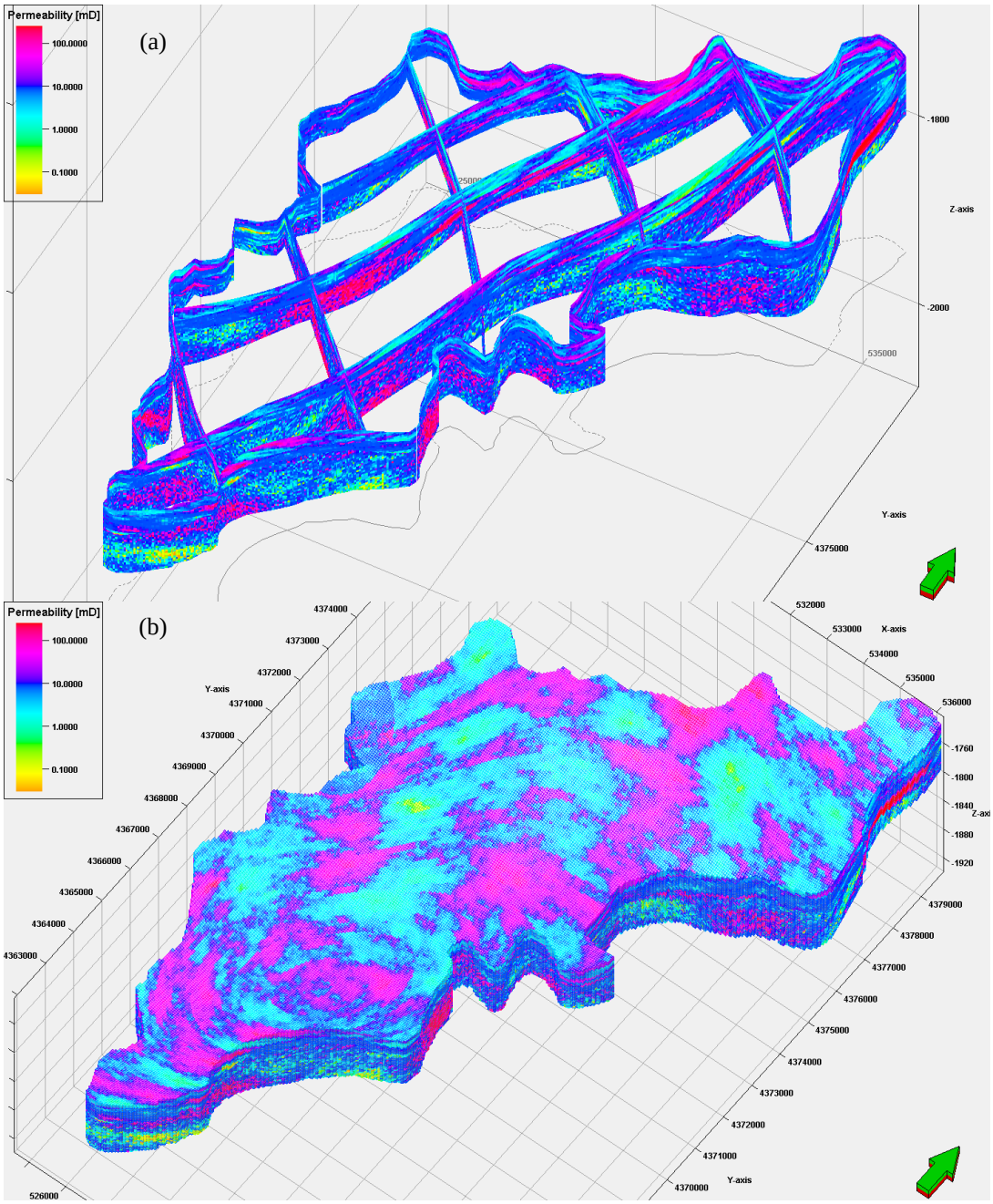


Figure 3-18. (a) Fence diagram shows the distribution of permeability in each subunit; b) shows a 3-D permeability model with Gaussian Random Function Simulation. The horizontal scale are various, but the vertical scale is 50:1.

Table 3-1. Summary of Jacksonburg-Stringtown research cores

Well	Cored Interval, ft.	Avg. Porosity %	Avg. Permeability
B-18	2988.5-3014	14.7	52
B-19	3086-3115	14.9	41
H-9	2980-2908	18.2	106
H-11	3083.4-3093.4	18.8	72
T-8	2781-2797	12.4	6.5
L-13	3032.4-3061.5	8.4	2.5

Table 3-2. Theoretical CO2 storage for P10, P50, and P90 in Jacksonburg-Stringtown oil field

Volume Statistics					
Parameters	Symbol	Unit	P_{10}	P_{50}	P_{90}
Total Pore Volume (ft^3)	V_{pv}	Tft^3	1.1	1.3	1.408
Water Saturation	S_{wi}	%	0.35	0.25	0.1
Formation Volume Factor	B	Bbl/STB	1.4	1.4	1.4
Average CO_2 density	P_{CO2}	Lbs/ft^3	48.0	48.0	48.0
Efficiency Factor	E	%	0.1	0.5	0.9
Reservoir CO_2 storage capacity	M_{CO2}	Mt	24.0	163.8	383.2

CHAPTER 4

Application of Mixed Kernels Function (MKF) Based Support Vector Regression Model (SVR) for CO_2 – Reservoir Oil Minimum Miscibility Pressure Prediction.

Zhi Zhong and Timothy R Carr[1]

[1] Department of Geology and Geography

West Virginia University, Morgantown

WV-26505, USA

(304)-282-9243

Email: zizhong1990@gmail.com; tim.carr@mail.wvu.edu

Abstract

CO_2 injection into oil reservoirs is considered a mature enhanced oil recovery (EOR) technique for conventional reservoirs. The local displacement efficiency of the CO_2-EOR process is highly dependent on the minimum miscibility pressure (MMP), estimating this parameter is critical to design of the CO_2 injection process. Traditional empirical methods to test the CO_2-oil MMP are time consuming and expensive; derived correlations are fast but not accurate. Therefore, an efficient and reliable method to determine MMP is beneficial. In this study, a mixed kernels function (MKF) based support vector regression (SVR) model was developed and used to predict the MMP for both pure and impure CO_2 injection cases. Four parameters were chosen as input parameters: 1) reservoir temperature; 2) average critical temperature; 3) molecular weight of C_{5+} fraction of crude oil, and; 4) the ratio of volatile components to intermediate components in crude oil. MMP was selected as the desired output parameter to train and test this newly developed model. The performance of basic kernels function based SVR model is compared with that of this newly developed MKF-SVR model. The well-trained MFK-SVR was compared with three well-established published correlations, demonstrated the highest correlation coefficient (R of 0.9381), lowest root mean square error (RMSE of 1.9151), smallest average absolute error (AAE of 1.1406) and maximum absolute error (MAE of 4.6291). We believe that the proposed MFK-SVM model is a more reliable and stable regression model to predict MMP. In addition, a sensitivity analysis was conducted to evaluate the physical correctness and indicates that the predicted results from the newly developed model are in excellent agreement with previous empirical work.

Keywords: CO_2-oil minimum miscibility pressure, CO_2 enhanced oil recovery, support vector regression, mixed kernels function, particle swarm optimization algorithm

4.1 Introduction

One of the first viable CO_2 commercial EOR applications was in 1972 in the Kelly-Snyder oil field (Norman, 2001). Today, CO_2 injection is an established enhanced oil recovery (EOR) technique for recovering residual oil in older and/or mature oil field. In 2014, 136 CO_2-EOR projects were underway in the U.S. and producing approximately 300,000 barrels of oil per day (BOPD) (Kuuskraa and Wallace, 2014). CO_2 injection can facilitate the reduction of greenhouse gas emissions by storing the CO_2 into geological media, such as depleted oil reservoirs. CO_2 - EOR is a preferred EOR method because CO_2 is miscible with crude oil increasing displacement efficiency, and less expensive than other similarly miscible fluids (NETL, 2010).

The displacement efficiency of CO_2 is highly dependent on the minimum miscible pressure (MMP) (Yuan and Johns, 2005; Yuan et al., 2004). MMP is the minimum pressure at which injected gas can develop multi-contact miscibility with the reservoir oil (Al-Wahaibi, 2010; Egwuenu, 2004; Hutchinson and Braun, 1961; Stalkup Jr, 1983a, b). Above this pressure, interfacial tension between these two phases is zero, and there is no difference between the densities of oil and injected gas (Al-Wahaibi, 2010; Fazlali et al., 2013). In contrast, at pressure lower than MMP, CO_2 will no longer be miscible with oil and the displacement efficiency decreases. Accurate CO_2–oil MMP estimation is critical in selecting suitable oil and gas reservoirs for CO_2–EOR processes.

Laboratory empirical methods are commonly used to estimate the MMP (Shokrollahi et al., 2013). The petroleum industry standard for estimating the MMP is slim-tube, because its results are most reliable and comparable with established industry data (Wang et al., 2015; Yellig and Metcalfe, 1980). In the early 1980s, Christiansen et al. (Christiansen and Haines, 1986) developed a new rapid and fast experimental method for MMP estimation using a rising-bubble apparatus (RBA) (Christiansen and Haines, 1987). Both slim-tube and rising-bubble methods are expensive, time consuming and depend on large amount of data (Huang et al., 2003; Metcalfe, 1982). Rao et al. (Rao, 1997) proposed a fast and simpler experimental method using experimental approach labeled vanishing interfacial tension (VIT) (Hemmati-Sarapardeh et al., 2013; Orr and Jessen, 2007; Rao, 1997; Rao and Lee, 2002). The accuracy of the measurement

depends on the composition of the mixture created fluid, and may give an higher MMP result than actual, because the composition cannot be selected in advance(Orr and Jessen, 2007).

Numerous empirical correlations have been developed to describe the phase behavior of reservoir oil-CO_2 and to minimize the difference between the transformed dependent variable (CO_2-oil MMP) and the sum of the transformed independent variables (Alston et al., 1985; Cronquist, 1978; Emera and Sarma, 2005; Glass, 1985; Lee, 1979; Orr Jr and Jensen, 1984; Yellig and Metcalfe, 1980; Yuan et al., 2005). However each correlation is restricted for specific oil reservoir conditions, such as temperature and oil composition. Where the geological conditions are very complex, empirical correlations cannot meet the various requirements for different oil reservoirs (Chen et al., 2014). Thus the application of those techniques is limited, and more adaptable, reliable correlation methods are required to overcome these limitations.

Support Vector Regression (SVR) is highly effective in representing a system's complexity; it also has a generalizing ability in function approximation (Al-Anazi and Gates, 2010b; Al-Anazi et al., 2011; Wang et al., 2014). Vapnik et al. (Boser et al., 1992a; Burges, 1998; Cortes and Vapnik, 1995; Drucker et al., 1997; Smola, 1996; Vapnik et al., 1997) first developed Support Vector Machine (SVM) at AT&T Bell laboratories based on statistical-learning theory to solve pattern recognitions problems (Vapnik et al., 1997). This algorithm was later extended to solve regression problems (Fu and Cheng, 2011). Al-Anazi et al. (Al-Anazi and Gates, 2010a, c) applied SVR to predict reservoir permeability and porosity. The principle of SVR is to find an optimal hyperplane in which all projected training data from the original data space is located within ε deviation in a high dimensional feature space (Vapnik et al., 1997).

The four basic elements of SVR are the separating hyperplane, maximum-margin hyperplane, soft margin and kernel function (Noble, 2006). The most important part of SVR is kernel function. There are two different groups of kernels: global kernels and local kernels. Global kernels have stronger extrapolation ability and local kernels have stronger interpolation ability. Standard kernels that can simultaneously extrapolate and interpolate are to some degree inaccurate. In this paper, a process is introduced that uses a mixed kernels function (MKF), which combine the local and global kernel function to overcome this drawback mentioned above (Smits and Jordaan, 2002; Huang et al., 2012; Lian et al., 2013). Mixing coefficient, gamma,

penalty(C), polynomial degree and epsilon are the important parameters in the MKF-SVR model. The determination of these parameters is the key step to achieve high performance from MKF-SVR. A grid searching algorithm (GSA) (Wang et al., 2014) and particle swarm optimization (PSO) (Eberhart and Kennedy, 1995) algorithm were applied to optimize these parameters for fast convergence in the training process.

The mixed kernels function (MKF) based support vector regression (SVR) is used to evaluate the pure and impure CO_2/oil MMP in crude oil. Four variable are evaluated: (a) molecular weight of C_{5+} fraction; (b) reservoir temperature; (c) volatile oil fraction (methane and nitrogen gas); and (d) intermediate oil fraction (C_2–C_4 and CO_2, H_2S) in crude oil. The results of MKF-SVR are compared to the results from SVR based on each basic kernel function. Moreover, the result predicted from the MKF-SVM model is compared to calculated values from traditional empirical correlations reported in the published literature. Sensitivity analysis is conducted to check out the most important input variables.

4.2 Basic Description of Support Vector Regression and Particle Swarm Optimization Algorithm

4.2.1 Support Vector Regression Principles

The support vector machine (SVM), based on the statistic-learning theory (SLT), and was first introduced in 1960s in Russia (Vapnik, 1963; Vapnik and Chervonenkis, 1964; Vapnik and Kotz, 1982). Initially it was modified to solve pattern recognitions problems. Later it was extended to the case of nonlinear regression and time series prediction problems (Drucker et al., 1997; Müller et al., 1997; Smola, 1996; Smola et al., 1998; Vapnik et al., 1997), which is support vector regression (SVR). Using adaptive margin-based loss functions and projecting the learning data (linearly or non-linearly) into higher dimensional feature space, SVR finds the best decision rule that has good generalization ability (Figure 4-1a). The projecting functions satisfied with the Mercer's condition are called Kernels – $K(x_i, x_j)$ (Zhu et al., 2005).

For example, assume training data $\{(x^1, y^1),(x^2, y^2),...,(x^m, y^m)\} \subset \chi \times \Re$, where χ denotes the space of the input parameters (e.g. $\chi = \Re^d$). For the case of linear function f takes the form:

$$f(x) = \langle \omega, x \rangle + b, \qquad \omega \in \chi, b \in \Re \qquad \text{Eq. (1),}$$

In the high dimensional feature space, the optimization problem for SVR with ε-insensitive loss function is:

$$\min_{\omega} \qquad \frac{1}{2}\|\omega\|^2 + C\sum_{i=1}^{m}(\xi_i + \xi_i^*) \qquad \text{Eq. (2),}$$

$$s.t. \qquad \begin{cases} y_i - \langle \omega, x_i \rangle - b \le \varepsilon + \xi_i, i = 1,2,.....m \\ \langle \omega, x_i \rangle + b - y_i \le \varepsilon + \xi_i^*, i = 1,2,.....m \\ \xi_i, \xi_i^* \ge 0, i = 1,2,....m \end{cases} \qquad \text{Eq. (3),}$$

The first part of Eq. (2) is the Vapnik-Chervonenkis (VC) confidence interval, whereas the second part is the empirical risk (Al-Anazi and Gates, 2010c). The regularization constant C in Eq. (2) is used to determine the complexity of the SVM model. The constant C also makes a compromise between the empirical risk minimization and the confidence degree. The epsilon (ε , essentially controls the complexity of the regression functions mentioned above). ξ_i and ξ_i^* are slack variables introduced by Bennett et al. (Bennett and Mangasarian, 1992), as opposed to symmetric boundary produced by the 'hard margin' loss function.

Vapnik's ε-insensitivity loss function as illustrated by Figure 4-1b and 1c defines a multi-dimensional tube. Only the points outside the ε-tube contribute to the loss, which is equal to the gap between the point values and the radius of the tube. The deviations are penalized in a linear fashion. However, the points inside the ε-tube contribute nothing to the cost, which means the loss (error or cost) is zero.

4.2.2 Global Kernels, Local Kernels and Mixed Kernels

Kernels $K(\vec{x}, \vec{x}^*)$ actually are a projection function. The function projects the original linearly or non-linearly learning data into high dimensional feature space, where all of the data can be presented linearly (Burges, 1998). A kernel function must meet Mercer's conditions (Boser et al., 1992b; Burges, 1998; Cortes and Vapnik, 1995; Smola and Schölkopf, 1998; Smola and Schölkopf, 2004), as following:

Suppose $K(\vec{x},\vec{x}^*) \in L_2(\Re^n) \otimes L_2(\Re^n)$

$\exists \forall g(x) \in L_2(\Re^n)$,

$$\iint_{L_2 \otimes L_2} K(\vec{x},\vec{x}^*) g(\vec{x}) g(\vec{x}^*) d\vec{x} d\vec{x}^* \geq 0 \qquad \text{Eq. (4),}$$

Based on this theorem, several different types of kernels, which determine the various characteristics of the model, are listed in Table 4-2. These kernels are two types: local kernels and global kernels (Smola and Schölkopf, 1998). A global kernel allows data points that are far away from each other to have an influence on the kernel values, as shown in Figure 4-2a. The polynomial kernel in Eq. (5) is a typical example of a global kernel (Smits and Jordaan, 2002; Zheng et al., 2004). The higher the degree of polynomial kernel, the better its interpolation ability will be. The lower the degree of polynomial kernel function, the better its extrapolation ability will be. In a local kernel only the data that are close to each other have an influence on the kernel values (Figure 4-2b). The typical local kernel is the radial basis function kernel (RBF) in Eq. (6). The smaller value of γ is, the worse its interpolation ability will be and vice versa. In other word, no single value of kernel parameter γ or polynomial degree will provide a model with both interpolation and extrapolation properties (Smits and Jordaan, 2002).

$$K_{ploy}(x_i, x_j) = (\langle x_i, x_j \rangle + 1)^d \qquad \text{Eq. (5),}$$

$$K_{rbf}(x_i, x_j) = \exp(-\gamma \left\| x_i - x_j \right\|^2) \qquad \text{Eq. (6),}$$

Based on the Mercer's conditions, there is one corollary. Assuming $K_1(\vec{x},\vec{x}^*)$, $K_2(\vec{x},\vec{x}^*)$ are admissible support vector kernels and $C_1, C_2 \geq 0$,

$$K(\vec{x},\vec{x}^*) = C_1 K_1(\vec{x},\vec{x}^*) + C_2 K_2(\vec{x},\vec{x}^*) \qquad \text{Eq. (7),}$$

$K(\vec{x},\vec{x}^*)$ is an admissible kernel (Smola and Schölkopf, 2004). The mixed kernels function $K(\vec{x},\vec{x}^*)$, which combine the good characteristics of both $K_1(\vec{x},\vec{x}^*)$ and $K_2(\vec{x},\vec{x}^*)$, has been

proposed by numerous investigators (Huang et al., 2012; Lian et al., 2013; Smits and Jordaan, 2002; Smola and Schölkopf, 2004; Yang et al., 2013; Zhu et al., 2005).

The mixtures of the RBF and polynomial kernels can be defined as Eq. (8, 9)

$$K_{mix} = mK_{poly} + (1-m)K_{rbf} \qquad 0 \le m \le 1 \qquad \text{Eq. (8),}$$

$$K_{mix}(x_i, x_j) = m(\gamma_1 \langle x_i, x_j \rangle + 1)^d + (1-m)\exp(-\gamma_2 \left\| x_i - x_j \right\|^2) \quad 0 \le m \le 1 \quad \text{Eq. (9),}$$

Where K_{poly} is a polynomial kernel and K_{rbf} is a radial basis function kernel, m is the mixing coefficient. If m=0, then the mixed kernels function will equal the RBF kernel, which means $K_{mix} = K_{rbf}$. If m=1, then the mixed kernels function will become the polynomial kernel, which means $K_{mix} = K_{ploy}$. The effect of the mixing a polynomial kernel with a RBF kernel is shown in Figure 4-3, where $x_i = 0.2, \gamma = 0.1, d = 1$. The mixed kernels function has characteristics of both the polynomial kernel and the RBF kernel, and improves fitting and generalizing ability.

4.2.3 PSO (Particle Swarm Optimization) Algorithm

Based on stochastic search and optimization processes, the particle swarm optimization algorithm (PSO) was developed (Eberhart and Kennedy, 1995), which is an evolutionary algorithm and imitated human (or insects) social behaviour. Individuals interact with one another while learning from their own experience, and gradually the individual population members' move into better regions of the problem space (Eberhart and Kennedy, 1995). Particles, indicating the potential solution, randomly locate in the architecture space (like birds or fish randomly distributed in a specific open environment), and are utilized to calculate the global optima of the fitness function. Assuming D-dimensional architecture space, population $x = \{x_1, x_2, x_3, ..., x_n\}$ is composed by n particles (x_i). Each $x_i = \{x_{i1}, x_{i2}, x_{i3}, ..., x_{iD}\}$ is indicating its position, and is also representing a potential solution. A global best value is $p_g = \{g_1, g_2, g_3, ..., g_D\}$ and personal best value is $p_i = \{p_{i1}, p_{i2}, p_{i3}, ... p_{iD}\}$. The velocity for particle x_i, representing the rate of position change, is written as $v_i = \{v_{i1}, v_{i2}, v_{i3}, ..., v_{iD}\}$. In PSO,

each particle updates its position by velocity adjustment and fluctuates between p_i and p_g. The process for velocity adjustment is presented in Figure 4-4. When the particle swarm algorithm is running without restraining velocities, it rapidly increases to unacceptable levels within a few iterations. Some form of damping of the dynamics of particles (e.g., $v_{\max}$) is necessary. To better allow an elegant and well-explained method for preventing explosion, ensuring convergence, reducing and eliminating the importance of $v_{\max}$, Shi et al.,(Shi and Eberhart, 1998) introduced the inertial weight.

$$\vec{v}_i^{\,new} = \omega_t \vec{v}_i^{\,old} + \phi_1 \otimes (\vec{p}_i - \vec{x}_i) + \phi_2 \otimes (\vec{p}_g - \vec{x}_i) \qquad \text{Eq. (10),}$$

$$\vec{x}_i^{\,new} = \vec{x}_i^{\,old} + \vec{v}_i^{\,new} \qquad \text{Eq. (11),}$$

$$\omega_t = \omega_{\max} - (\omega_{\max} - \omega_{\min}) * t_{ite} / t_{\max} \qquad \text{Eq. (12),}$$

Where ω_{max} is initial inertia weight, ω_{min} is the finial inertia weight, φ_1 is cognitive coefficient and φ_2 is social coefficient.

4.3 Model Development

SVR is an extended case of SVM, which was modified from the machine-learning community. Because of strong interpolation and extrapolation capability, the new MKF-SVR was chosen to build the model for CO_2-oil MMP prediction. This model was optimized by PSO. The existing SVR for comparison was also chosen because of its many advantages over the traditional methods and empirical correlations; these include fast convergence to the global optimum, high generalization performance, and less probability for over-fitting. The construction of the MKF-SVR model was undertaken with the LIBSVM toolbox developed by Chang et al. (Chang and Lin, 2011); the optimized process was based on the MATLAB software. The design process for this model involves three parts: data acquisition, data normalization and model optimization and validation.

4.3.1 Experimental Dataset

The accuracy and stability of the developed model is largely dependent on the training data's reliability and comprehensiveness. Based on a review of recent research, reservoir temperature, oil composition, purity of injected gas play an important role in CO_2-oil MMP values (Alston et al., 1985; Sebastian et al., 1985; Shokir, 2007a; Yuan et al., 2004; Zuo et al., 1993). Thus, input data sets used in this model include average critical temperature of the drive gas ($T_{cm} = \sum Z_i \times T_{ci}$), reservoir temperature (T), molecular weight of the C_{5+} fraction (MC_{5+}), and the ratio of volatile (C_1 and N_2) to intermediate (C_2-C_4, H_2S and CO_2) components in crude oil, Pure and impure CO_2-oil MMP is the desired output. In this work, all of the experimental data used were collected from existing published literature (Al-Ajmi et al., 2009; Alston et al., 1985; Bon et al., 2006; Bon et al., 2005; Cao, 2012; Chaback et al., 1989; Dicharry et al., 1973; Dong, 1999; Dong et al., 2001; Eakin and Mitch, 1988; Frimodig et al., 1983; Graue and Zana, 1981; Harmon and Grigg, 1988; Henry and Metcalfe, 1983; Holm and Josendal, 1974; Jacobson, 1972; Khan et al., 1992; Li et al., 2012; Metcalfe, 1982; Rathmell et al., 1971; Sebastian et al., 1985; Shelton and Yarborough, 1977; Spence Jr and Watkins, 1980; Srivastava et al., 2000; Sun et al., 2006; Thakur et al., 1984; Zhang et al., 2015; Zhou, 2008; Zuo et al., 1993). Details of datasets are presented in Table 4-1.

Normally, all data for the training of a supervised machine-learning algorithm are divided into three parts: training data, validation data and testing data. This paper combines the validation data and training data. The combination is used as training dataset. An N-fold cross-validation method was used as resampling method during the training process. The purpose of training and validation process is to find the best parameters for MKF-SVR model and to ensure accurate prediction performance. The goal of the testing process is to evaluate the accuracy and stability of the well-trained MKF-SVR model (Chen et al., 2014). Of the 147 data samples collected, 90% (133) were used for training and validation, and the remaining 10% (14) were used for testing.

4.3.2 Parameters Optimization Based on Improved PSO

As discussed, the final performance of the MKF-SVR is strongly dependent on the parameters of the mixed kernels function. By applying the PSO algorithm and n-fold cross-validation

resampling methods, key parameters such as mixing coefficient (m), penalizing coefficient (C), RBF kernel parameter (ϒ), polynomial degree (d) and (ε) in ε-insensitivity function are optimized. Particles update their positions by changing velocity and converge finally at a global optimum within the searching space. In this study, the particles' positions are the vectors of m, ϒ, d, C, ε, which are denominated as P (m, ϒ, d, C, ε). In order to search the global optimum reasonably and to convergent quickly, the parameter m is limited in [0, 1], ϒ in $[2^{-4}, 2^{4}]$, d in [1, 3], C in $[2^{-5}, 2^{5}]$ and ε in [0.001, 0.01]. To evaluate the performance of training process, mean square error (MSE) is chosen as the fitness function, which is formulated as:

$$F_{Fitness} = \frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2 \qquad \text{Eq. (13),}$$

where $\hat{y}_i$ is the predicted value, y_i is the true value and n is the number of samples. Figure 4-5 shows the workflow to find the optimum values of each parameter in the mixed kernels function.

4.3.3 Testing of the MKF-SVR Model

After the training and validation processes were completed, all of the parameters were optimized and the expected regression model was established. In order to verify its generalization capability and stability, the well-trained model is tested by using the testing data set (10% data), which were not used in the training and validation process.

4.4 Results and Discussion

Asoodeh et al. (Asoodeh et al., 2014) used the original SVR model to estimate the CO_2-oil MMP; Shokrollahi et al. (Shokrollahi et al., 2013) applied the adapted LS-SVM model to estimate the CO_2-oil MMP. Although both models have relatively good prediction performance, the generalization ability of these models can be improved by the MKF-SVR model, because of the limitation of applied kernel function. In this study, the result of the original SVR is compared with different basic kernel functions and that of the adapted SVR with mixed kernels function in order to demonstrate improved performance of the MKF-SVR model than the traditional correlations. The results are compared to those produced by four other proposed numerical

correlations (Alston et al., 1985; Sebastian et al., 1985; Yellig and Metcalfe, 1980; and (Emera and Sarma, 2005).

4.4.1 Results of SVR and MKF-SVR

There are four different kernel functions as listed in Table 4-2: each has different parameters that should be optimized separately. The grid searching algorithm was applied to find the best parameters of each basic kernel function. PSO algorithm is applied to optimize the five parameters for the mixed kernels function. Table 4-3 shows the range of parameters for PSO algorithms.

The well-trained MKF-SVR model is applied to estimate the MMP value of crude oil, in order to verify the reliability and accuracy of the model (Al-Anazi and Gates, 2010a; Jin et al., 2001). Five statistical parameters were introduced as evaluation indexes to test the performance of the model. These parameters include coefficient of determination (R^2), correlation coefficient (r), root mean error (RMSE), average absolute error (AAE), and maximum absolute error (MAE), as defined in

. An R^2 (Coefficient of Determination) of 1 indicates a perfect model, while an R^2 of 0 indicates an inaccurate model (Al-Anazi and Gates, 2010c; Oyerokun et al., 2002). Moreover, R and RMSE are used to evaluate overall performance, while AAE and MAE are used to determine the error range of the predicted results (Chen et al., 2014).

The particle swarm optimization algorithm was applied to search for the optimal parameter sets. The optimal parameters in the MKF-SVR model were determined to be C=955.0476, ϒ=144.3693, m=0.3435, d=3.935, ε=0.0804.

The training and testing performance of the SVR model with mixed kernels function and other basic kernel functions are listed in Table 4-4. The performance of the MKF-SVR model improves the accuracy and stability as indicated by the decreased RMSE and AAE, and increases in the correlation coefficient and coefficient of determination. The process of searching for best gamma and cost value is shown in Figure 4-6, using grid searching algorithm for the SVR with linear kernel, polynomial kernel, RBF kernel, sigmoid kernel and n-fold of 4. The cross-validation mean square errors are 8.4967, 9.2179, 6.1250 and 8.4766 respectively. The process of searching for best parameters in MKF-SVR model by particle swarm optimization (PSO) is shown in Figure 4-7. A comparison of Figures 4-6 and 4-7 shows that the MKF-SVR model has the smallest cross-validation mean square error (4.0012). The MFK-SVR model has the highest training and testing coefficient of determination (0.8767) as shown in Figure 4-8. Compared with other kernel functions, the MKF-SVR model provides more accurate predictions as demonstrated by Figure 4-9 showing the comparison of actual values and forecasted values.

4.4.2 Comparison of Mixed Kernel Based SVR Model with Correlations from Literatures

In order to further evaluate the model accuracy and stability, the comparison between some well-known empirical correlations with the predicted results calculated from the newly MKF-SVR model is necessary. The empirical correlations selected for the comparisons include correlations by Yelling et al. (Yellig and Metcalfe, 1980)(corrected with Sebastian et al. (Sebastian et al., 1985)impurity correction factor), Yelling et al. (Yellig and Metcalfe, 1980)(corrected with Alston et al. (Alston et al., 1985) impurity correction factor), Emera et al. (Emera and Sarma, 2005)(corrected with Sebastian et al. (Sebastian et al., 1985)impurity correction factor), Emera et al. (Emera and Sarma, 2005)(corrected with Alston et al. (Alston et al., 1985) impurity correction

factor), Alston et al. (Alston et al., 1985), and Sebastian et al. (Sebastian et al., 1985). Detailed descriptions of each correlation are given in Appendix. The correlations consider the pure CO_2 injection case and also investigate the impact of an impure CO_2 drive gas stream. The Alston et al. (Alston et al., 1985) and Sebastian et al. (Sebastian et al., 1985) correlations were developed for both pure and impure CO_2-drive stream, while the Yelling et al. (Yellig and Metcalfe, 1980) and Emera et al. (Emera and Sarma, 2005) correlations were developed only for a pure CO_2-drive stream; thus they need impurities correction factors to correct for impure CO_2 data samples. The Alston et al. (Alston et al., 1985) and Sebastian et al. (Sebastian et al., 1985) impurity correction factors are summarized in the Appendix.

A scatter diagram (Figure 4-10) illustrates the comparison between the laboratory experimental MMP values and the predicted values. The cross-plots visualize the accuracy of proposed model and available correlations. The diagonal line in each figure represents the theoretical equality between predicted and experimental values. Tightness of the points to the diagonal line directly expresses the agreement level between the simulated and experimental values. The MKF-SVR results show a total coefficient of determination of 0.87, and the majority of data points for the training and testing process lie on a 45° line, indicating that the newly developed MKF-SVR model is robust and stable (Figure 4-10a). There is an excellent agreement between the experimental data and the predictions of newly developed MKF-SVR model indicating that this model can be successfully applied to predicting MMP for both pure and impure CO_2 injection cases. As Figure 4-10b-f illustrates, the best fit lines of predicted results of correlations deviate from the diagonal line, which indicate that the agreement between the experimental data and value predicted by mixed kernel based SVR model is improved among the selected correlations. The simulation results of classic correlations and newly developed model (in terms of R, RMSE, AAE, MAE) were also compared and presented in Table 4-6. The MFK-SVR model has the smallest RMES (of 1.9151), AAE (of 1.1140) and MAE (of 4.6291); and highest R (of 0.9381). In terms of results, it is clearly seen that MFK-SVR model have high performance compared with other correlations.

4.4.3 Sensitivity Analysis

In order to further investigate the relationship between the influence factor (i.e. reservoir temperature, average critical temperature, molecular weight of C_{5+} fraction and volatile to intermediate ratio) and CO_2-oil MMP, a sensitivity analysis was conducted. This process is executed by changing one studied variable gradually, while holding the other variables constant (Chen et al., 2014). As Figure 4-11a shows, the increase of reservoir temperature will increase MMP, as the solubility of CO_2 into oil is decreases (Dodds et al., 1956; Duan and Sun, 2003; Mungan, 1981), which means the CO_2 is less miscible with oil, an increase in CO_2-oil MMP. When reservoir temperature decreases, the interfacial tension between oil and CO_2 also declined (Rao, 1997; Rao and Lee, 2002), therefore the MMP decreased correspondingly. This model results are in agreement with previous work (Shokir, 2007b; Shokrollahi et al., 2013; Yuan and Johns, 2005).

The purity of the injection gas stream has significant and complex impacts on CO_2-oil MMP. Generally, the presence of H_2S or intermediate hydrocarbon components will decrease the CO_2-oil MMP, while the presence of C_1 or N_2 will increase the CO_2-oil MMP (Alston et al., 1985; Eakin and Mitch, 1988; Sebastian et al., 1985). In order to describe the impact of purity of injection gas stream on CO_2-oil MMP simply and concisely, the average critical temperature is the best reasonable parameter. As Figure 4-11b shows, the increase of average critical temperature will decrease the CO_2-oil MMP.

As Figure 4-11c illustrates, the MMP increases as the average molecular weight of pentane plus fraction increases. Pentane plus fraction, the major component of crude oil indicates an increase of the average molecular weight of C_{5+} and the percentage of heavy oil molecules (Clark et al., 1958; Turek et al., 1984). The increase in average molecular weight of C_{5+} is less miscible with CO_2, thus MMP increases.

The evaporation property of volatile components may increase the gas phase during the multi-contact of CO_2 and oil, decreasing the oil miscibility with CO_2 (Al-Wahaibi, 2010; Rathmell et al., 1971), while the intermediate components are easy to mix with CO_2 due to their inter-

solubility into each other (Metcalfe, 1982; Rathmell et al., 1971), thus the increase of volatile to intermediate ratio will increase the MMP (Alston et al., 1985), as shown in Figure 4-11d.

4.5 Conclusions

A hybrid model of mixed kernels function (MKF) based support vector regression (SVR) model was developed to predict pure and impure CO_2-oil minimum miscibility pressure (MMP) during a CO_2-EOR process. In this MKF-SVR model, four factors (i.e. reservoir temperature, average critical temperature, molecular weight of C_{5+} fraction of crude oil, and the ration of volatile components to intermediate components in crude oil) representing the most comprehensive and robust set were selected as the input variables while MMP was considered as output variable. Through the foregoing analysis and discussions of the simulation results, several conclusions are drawn:

(1) The mixed kernel function based support vector regression (SVR) model was successfully applied to predict the CO_2-oil MMP value for both pure and impure CO_2 gas.

(2) Different kernel functions affect the final performance of SVR significantly. Mixed kernel function, which combines the advantages of global radial basis function (RBF) and local (Polynomial) kernel functions, increases the applicability of SVR dramatically.

(3) The comparison of traditional correlations with the mixture kernel based SVR model shows excellent performance and greater generalization ability with higher correlation coefficient (R = 0.9381), smaller average absolute error (AAE = 1.1406), maximum absolute error (MAE = 4.6291) and the reduced root mean square error (RMSE =1.9151).

(4) Based on the sensitivity analysis, reservoir temperature, molecular weight of C_{5+} fraction and volatile/intermediate ratio of crude oil positively influence the CO_2-oil MMP, while average critical temperature has a negative impact on CO_2-oil MMP. Moreover the impacts of all factors towards the CO_2-oil MMP is nonlinear.

Acknowledgements

We are grateful to Dr. Kathy Bruner for their helpful comments in improving the early version of the manuscript. This work was funded by the US-China Clean Energy Research Center, Advanced Coal Technology Consortium, under grant DE-PI0000017 from the National Energy Technology Laboratory of the US Department of Energy.

Nomenclature

AAE	average absolute error	φ_2	social coefficient
BOPD	Barrels of oil per day	p_g	global best value
C	penalizing parameter in SVM	p_i	personal best value
CO_2	carbon dioxide	P_b	bubble pressure
C_{5+}	pentane plus fraction	$P_{m,\min,pure}$	MMP value in pure CO_2-oil systems
d	polynomial degree	$P_{m,\min,impure}$	MMP value in impure CO_2-oil systems
D	architecture space Dimension	$\vec{v}_i^{new}$	i^{th} particle's new velocity
ε	insensitivity loss parameter	$\vec{v}_i^{old}$	i^{th} particle's old velocity
ξi *and* ξi^*	slack variables	PSO	particle swarm optimization
EOR	enhanced oil recovery	r	correlation coefficient
F_{imp}	correction factor for impure CO_2 stream	R	coefficient of determination

GSA	grid searching algorithm	RBF	radial basis function
Υ	RBF kernel parameter	RMSE	root mean error
$K(x_i, x_j)$	kernel function	SVR	support vector regression
m	mixing coefficient	SVM	support vector machine
MAE	maximum absolute error	T_{ck}	critical temperature of component k (k = CO2, C1, N2,H2S, and C2–C4), ℃
MKF	mixed kernels function	T_{cm}	weight average pseudocritical temperature of the drive gas, ℃
MKF-SVR	mixed kernels function based support vector regression	T	reservoir temperature, ℃
MMP	minimum miscibility pressure, MPa	x_i	i^{th} particle's position
MSE	mean squared error, %	$x_{i,j}$	j^{th} value of i^{th} particle
MW_{C5+}	molecular weight of C5+ oil fraction, g/mol	$\vec{x}_i^{\,old}$	i^{th} particle's old position
N	sample number	$\vec{x}_i^{\,new}$	i^{th} particle's new position
$\omega_{\max}$	initial inertia weight	x_{MED}	mole fraction of intermediate oil components including C2–C4, CO2, and H2S, %
$\omega_{\min}$	finial inertia weight	x_k	mole fraction of component k (k = CO2, C1, N2, H2S, and C2–C4), %

φ_1	cognitive coefficient	x_{VOL}	mole fraction of volatile oil components including N2 and CH4, %

Appendix 4-A

In this Appendix, totally nine correlations used for MMP prediction in pure CO_2-oil systems are summarized and presented. All of the correlations are converted to consistent units.

4-A.1. Correlation Proposed by Yelling et al.

Yelling (Yellig and Metcalfe, 1980) correlation correlates CO_2-oil MMP only with the reservoir temperature. This correlation only can be used when T (reservoir temperature) ranges from 35.8 to 88.9 °C, which is given by:

$$P_{m,\min,pure} = 12.6472 + 0.01553(1.8T + 32) + 1.24192\times10^{-4}(1.8T + 32)^2 - \frac{716.9427}{(1.8T + 32)} \qquad \text{Eq. (A1)}$$

If MMP <7 P_b (P_b is the bubble pressure), then P_b is taken as MMP.

4-A.2. Correlation Proposed by Emera et al.

A new correlation modified from the Alston correlation by Emera et al. (Emera and Sarma, 2005). When $P_b \geq 0.345MPa$, this correlation is presented as follows:

$$P_{m,\min,pure} = 5.0093\times10^{-5}(1.8T + 32)^{1.164}(M_{C_{5+}})^{1.2785}(\frac{x_{VOL}}{x_{MED}})^{0.1073}, \qquad \text{Eq. (A2)}$$

otherwise, when $P_b < 0.345MPa$

$$P_{m,\min,pure} = 5.0093\times10^{-5}(1.8T + 32)^{1.164}(M_{C_{5+}})^{1.2785}, \qquad \text{Eq. (A3)}$$

where T is reservoir temperature, P_b is bubble pressure, MC_{5+} is the fraction of C_{5+} molecular weight; x_{vol} is the volatile components and x_{med} is the intermediate components.

However, those correlations can be used only when special conditions are satisfied: (i) the test temperature should range from 40.8 to 112.2°C; (ii) the MMP ranges between 8.28 and 30.2 MPa; (3) the C_{5+} molecular weight ranges from 166.2 g/mol to 267.5 g/mol. If the calculated MMP is smaller than P_b (bubble pressure), then P_b is taken as MMP.

4-A.3. Correlation Proposed by Alston et al.

The correlation developed by Alston et al. (Alston et al., 1985) for the pure CO_2 injection case is represented as a function of reservoir temperature, oil pentane and heavier molecular weight, and the ratio of volatile to intermediate mole fractions in the reservoir oil. When the bubble pressure of reservoir oil is greater than 50 psia (0.345MPa), the influence of the volatile/intermediate ratios is important, and the correlation can be written as:

$$P_{m,\min,pure} = 6.0536\times10^{-5}(1.8T+32)^{1.06}(M_{C_{5+}})^{1.78}(\frac{x_{VOL}}{x_{MED}})^{0.136}. \quad \text{Eq. (A4)}$$

However, when the bubble pressure of reservoir oil is smaller than 50 psia (0.345MPa), it contains relatively small or zero quantities of volatile and intermediate components, and the MMP is not affected significantly by these components. Thus the influence of the volatile/intermediate ratios is not significant, and the correlation is:

$$P_{m,\min,pure} = 6.0536\times10^{-5}(1.8T+32)^{1.06}(M_{C_{5+}})^{1.78}, \quad \text{Eq. (A5)}$$

where T is reservoir temperature, M_{C5+} is the fraction of C_{5+} molecular weight; x_{vol} is the volatile components and x_{med} is the intermediate components.

For the impure CO_2 stream injection case, contaminants have the adverse or positive effects on MMP. Thus a correction factor F_{imp} is introduced. It was calculated based on the critical temperature of each component in the stream. The pseudo-critical temperature (T_{cm}) of the solvent stream was calculated using a weight-fraction mixing rule as follows:

$$F_{imp} = (87.8/T_{cm})^{1.935\times87.8/T_{cm}} \quad \text{Eq. (A6)}$$

$$T_{cm} = \sum_{k=1}^{n} w_k (1.8T_{ck} + 32) \qquad \text{Eq. (A7)}$$

where T_{cm} is the weight average critical temperature of the solvent stream, w_k is the weight fraction of component k (k = CO_2, C_1, N_2, H_2S, and C_2–C_4), and T_{ck} is the critical temperature of component k. Then the impure CO_2-oil MMP was calculated as:

$$P_{m,mim,impure} = P_{m,\min,pure} \times F_{imp} . \qquad \text{Eq. (A8)}$$

4-A.4. Correlation Proposed by Sebastian et al.

Based on the Sebastian and co-worker's research (Sebastian et al., 1985), combined with previous study, Sebastian et al. proposed a new correlation to predict the change in MMP value resulting from the impurities (i.e. C_1, H_2, O_2, CO_2, N_2, H_2S, and C_2–C_5 hydrocarbons) in the CO_2 drive gas. The relationship between impure and pure CO2-oil MMP was described as:

$$F_{imp} = 1.0 - 0.0213(1.8\mathrm{T}_{cm} - 31) + 2.51\times10^{-4}(1.8\mathrm{T}_{cm} - 31)^2 - 2.35\times10^{-7}(1.8\mathrm{T}_{cm} - 31)^3 \ , \quad \text{Eq. (A9)}$$

where T_{cm} is the mole average pseudocritical temperature of the gas stream, and can be expressed as:

$$T_{cm} = \sum_{k=1}^{n} x_k T_{ck} \qquad \text{Eq. (A10)}$$

where x_k is the mole fraction of component k (k = CO_2, C_1, N_2, H_2S, and C_2–C_4), and T_{ck} is the supercritical temperature of each component k.

The MMP pure was calculated use the following formula:

$$P_{m,\min,pure} = 12.6472 + 0.01553(1.8\mathrm{T} + 32) + 1.24192\times10^{-4}(1.8\mathrm{T} + 32)^2 - \frac{716.9427}{1.8\mathrm{T}_R + 32}, \quad \text{Eq. (A11)}$$

and the impure CO_2-oil MMP was calculated as:

$$P_{m,mim,impure} = P_{m,\min,pure} \times F_{imp} .$$ Eq. (A12)

References

Al-Ajmi, M. F., O. A. Alomair, and A. M. Elsharkawy, 2009, Planning miscibility tests and gas injection projects for four major Kuwaiti reservoirs: Kuwait International Petroleum Conference and Exhibition.

Al-Anazi, A. F., and I. D. Gates, 2010a, Support-vector regression for permeability prediction in a heterogeneous reservoir: A comparative study: SPE Reservoir Evaluation & Engineering, v. 13, p. 485-495.

Al-Anazi, A. F., and I. D. Gates, 2010b, A support vector machine algorithm to classify lithofacies and model permeability in heterogeneous reservoirs: Engineering Geology, v. 114, p. 267-277.

Al-Anazi, A. F., and I. D. Gates, 2010c, Support vector regression for porosity prediction in a heterogeneous reservoir: A comparative study: Computers & Geosciences, v. 36, p. 1494-1503.

Al-Anazi, A. F., I. D. Gates, and J. Azaiez, 2011, Support vector machines for petrophysical modelling and lithoclassification.

Al-Wahaibi, Y. M., 2010, First-contact-miscible and multicontact-miscible gas injection within a channeling heterogeneity system: Energy & Fuels, v. 24, p. 1813-1821.

Alston, R. B., G. P. Kokolis, and C. F. James, 1985, CO_2 minimum miscibility pressure: A correlation for impure CO_2 streams and live oil systems: Society of Petroleum Engineers Journal, v. 25, p. 268-274.

Asoodeh, M., A. Gholami, and P. Bagheripour, 2014, Oil-CO_2 MMP determination in competition of neural network, support vector regression, and committee machine: Journal of Dispersion Science and Technology, v. 35, p. 564-571.

Bennett, K. P., and O. L. Mangasarian, 1992, Robust linear programming discrimination of two linearly inseparable sets: Optimization methods and software, v. 1, p. 23-34.

Bon, J., M. K. Emera, and H. K. Sarma, 2006, An experimental study and genetic algorithm (GA) correlation to explore the effect of nC_5 on impure CO_2 minimum miscibility pressure (MMP): SPE Asia Pacific Oil & Gas Conference and Exhibition.

Bon, J., H. K. Sarma, and A. M. Theophilos, 2005, An investigation of minimum miscibility pressure for CO_2-rich injection gases with pentanes-plus fraction: SPE International Improved Oil Recovery Conference in Asia Pacific.

Boser, B. E., I. M. Guyon, and V. N. Vapnik, 1992a, A training algorithm for optimal margin classifiers: Proceedings of the fifth annual workshop on Computational learning theory, p. 144-152.

Boser, B. E., I. M. Guyong, and V. N. Vapnik, 1992b, A training algorithm for optimal margin classifiers.

Burges, C. J. C., 1998, A tutorial on support vector machines for pattern recognition: Data mining and knowledge discovery, v. 2, p. 121-167.

Cao, P., 2012, Feasibility assessment on CO_2 miscible flooding for enhancing oil recovery in Gbeibe oil reservoir, Southwest Petroleum University.

Chaback, J. J., R. A. Harmon, and R. B. Grigg, 1989, Discussion of vapor-density measurement for estimating minimum miscibility pressure: SPE reservoir engineering, v. 4, p. 253-254.

Chang, C.-C., and C.-J. Lin, 2011, LIBSVM: A library for support vector machines: ACM Transactions on Intelligent Systems and Technology (TIST), v. 2, p. 27.

Chen, G., K. Fu, Z. Liang, T. Sema, C. Li, P. Tontiwachwuthikul, and R. Idem, 2014, The genetic algorithm based back propagation neural network for MMP prediction in CO_2-EOR process: Fuel, v. 126, p. 202-212.

Christiansen, R. L., and H. K. Haines, 1986, Apparatus and method for determining the minimum miscibility pressure of a gas in a liquid, USA.

Christiansen, R. L., and H. K. Haines, 1987, Rapid measurement of minimum miscibility pressure with the rising-bubble apparatus: SPE Reservoir Engineering, v. 2, p. 523-527.

Clark, N. J., H. Shearin, W. Schultz, K. Garms, and J. Moore, 1958, Miscible drive-Its theory and application: Journal of Petroleum Technology, v. 10, p. 11-20.

Cortes, C., and V. Vapnik, 1995, Support-vector networks: Machine learning, v. 20, p. 273-297.

Cronquist, C., 1978, Carbon dioxide dynamic miscibility with light reservoir oils: Proc. Fourth Annual US DOE Symposium, Tulsa, p. 28-30.

Dicharry, R. M., T. L. Perryman, and J. D. Ronquille, 1973, Evaluation and design of a CO_2 miscible flood project-SACROC unit, Kelly-Snyder field: Journal of Petroleum Technology, v. 25, p. 1,309-1,318.

Dodds, W. S., L. F. Stutzman, and B. J. Sollami, 1956, Carbon dioxide solubility in water: Industrial & Engineering Chemistry Chemical & Engineering Data Series, v. 1, p. 92-95.

Dong, M., 1999, Task 3- minimum miscibility pressure (MMP) studies, technical report: potential of greenhouse storage and utilization through enhanced oil recovery, Petroleum Research Center, Saskatchewan Research Council Saskatchewan,, Canada.

Dong, M., S. Huang, S. B. Dyer, and F. M. Mourits, 2001, A comparison of CO_2 minimum miscibility pressure determinations for Weyburn crude oil: Journal of Petroleum Science and Engineering, v. 31, p. 13-22.

Drucker, H., C. J. C. Burges, L. Kaufman, A. Smola, and V. Vapnik, 1997, Support vector regression machines: Advances in neural information processing systems, v. 9, p. 155-161.

Duan, Z., and R. Sun, 2003, An improved model calculating CO_2 solubility in pure water and aqueous NaCl solutions from 273 to 533 K and from 0 to 2000 bar: Chemical geology, v. 193, p. 257-271.

Eakin, B., and F. Mitch, 1988, Measurement and correlation of miscibility pressures of reservoir oils: SPE annual technical conference and exhibition.

Eberhart, R. C., and J. Kennedy, 1995, A new optimizer using particle swarm theory: Sixth international symposium on micro machine and human science, p. 39-43.

Egwuenu, A. M., 2004, Improved fluid characterization for miscible gas floods, University of Texas at Austin.

Emera, M. K., and H. K. Sarma, 2005, Use of genetic algorithm to estimate CO_2–oil minimum miscibility pressure-a key parameter in design of CO_2 miscible flood: Journal of petroleum science and engineering, v. 46, p. 37-52.

Fazlali, A., M. Nikookar, A. Agha-Aminiha, and A. H. Mohammadi, 2013, Prediction of minimum miscibility pressure in oil reservoirs using a modified SAFT equation of state: Fuel, v. 108, p. 675-681.

Frimodig, J. P., N. A. Reese, and C. A. Williams, 1983, Carbon dioxide flooding evaluation of high pour-point, paraffinic red wash reservoir oil: Society of Petroleum Engineers Journal, v. 23, p. 587-594.

Fu, Y., and Y. Cheng, 2011, Application of an integrated support vector regression method in prediction of financial returns: International Journal of Information Engineering and Electronic Business (IJIEEB), v. 3, p. 37.

Glass, O., 1985, Generalized minimum miscibility pressure correlation Society of Petroleum Engineers Journal, v. 25, p. 927-934.

Graue, D. J., and E. T. Zana, 1981, Study of a possible CO_2 flood in Rangely Field: Journal of Petroleum Technology, v. 33, p. 1,312-1,318.

Harmon, R. A., and R. B. Grigg, 1988, Vapor-density measurement for estimating minimum miscibility pressure SPE reservoir engineering, v. 3, p. 1,215-1,220.

Hemmati-Sarapardeh, A., S. Ayatollahi, M.-H. Ghazanfari, and M. Masihi, 2013, Experimental determination of interfacial tension and miscibility of the CO_2–crude oil system; temperature, pressure, and composition effects: Journal of Chemical & Engineering Data, v. 59, p. 61-69.

Henry, R. L., and R. S. Metcalfe, 1983, Multiple-phase generation during carbon dioxide flooding: Society of Petroleum Engineers Journal, v. 23, p. 595-601.

Holm, L. W., and V. A. Josendal, 1974, Mechanisms of oil displacement by carbon dioxide: Journal of petroleum Technology, v. 26, p. 1,427-1,438.

Huang, H., S. Ding, F. Jin, J. Yu, and Y. Han, 2012, A novel granular support vector machine based on mixed kernel function: International Journal of Digital Content Technology and its Applications, v. 6, p. 484-492.

Huang, Y. F., G. H. Huang, M. Z. Dong, and G. M. Feng, 2003, Development of an artificial neural network model for predicting minimum miscibility pressure in CO_2 flooding: Journal of Petroleum science and Engineering, v. 37, p. 83-95.

Hutchinson, C. A., and P. H. Braun, 1961, Phase relations of miscible displacement in oil recovery: AIChE Journal, v. 7, p. 64-72.

Jacobson, H. A., 1972, Acid gases and their contribution to miscibility: Journal of Canadian Petroleum Technology, v. 11.

Jaubert, J.-N., L. Avaullee, and J.-F. Souvay, 2002, A crude oil data bank containing more than 5000 PVT and gas injection data: Journal of Petroleum Science and Engineering, v. 34, p. 65-107.

Jin, R., W. Chen, and T. W. Simpson, 2001, Comparative studies of metamodelling techniques under multiple modeling criteria: Structural and Multidisciplinary Optimization, v. 23, p. 1-13.

Khan, S. A., G. A. Pope, and K. Sepehrnoori, 1992, Fluid characterization of three-phase CO_2/oil mixtures: SPE/DOE Enhanced Oil Recovery Symposium.

Kuuskraa, V., and M. Wallace, 2014, CO_2-EOR set for growth as new CO_2 supplies emerge: Oil & Gas Journal, v. 112, p. 92-92.

Lee, J. I., 1979, Effectiveness of carbon dioxide displacement under miscible and immiscible conditions, Report RR-40, Petroleum Recovery Inst., Calgary.

Li, H., J. Qin, and D. Yang, 2012, An improved CO_2–oil minimum miscibility pressure correlation for live and dead crude oils: Industrial & Engineering Chemistry Research, v. 51, p. 3516-3523.

Lian, C., Z. Zeng, W. Yao, and H. Tang, 2013, Displacement prediction of landslide based on PSOGSA-ELM with mixed kernel: Advanced Computational Intelligence (ICACI), 2013 Sixth International Conference on, p. 52-57.

Metcalfe, R. S., 1982, Effects of impurities on minimum miscibility pressures and minimum enrichment levels for CO_2 and rich-gas displacements: Society of Petroleum Engineers Journal, v. 22, p. 219-225.

Müller, K.-R., A. J. Smola, G. Rätsch, B. Schölkopf, J. Kohlmorgen, and V. Vapnik, 1997, Predicting time series with support vector machines, Artificial Neural Networks, Springer, p. 999-1004.

Mungan, N., 1981, Carbon dioxide flooding-fundamentals: Journal of Canadian Petroleum Technology, v. 20.

NETL, 2010, Carbon dioxide enhanced oil recovery-untapped domestic energy supply and long term carbon storage solution: The Energy Lab.

Noble, W. S., 2006, What is a support vector machine?: Nature biotechnology, v. 24, p. 1565-1567.

Norman, J. H., 2001, Non technical Guide to Petroleum Geology Exploration: Drilling and Production. 2nd edition, printed in USA, p. 1-15.

Orr, F. M., and K. Jessen, 2007, An analysis of the vanishing interfacial tension technique for determination of minimum miscibility pressure: Fluid phase equilibria, v. 255, p. 99-109.

Orr Jr, F. M., and C. M. Jensen, 1984, Interpretation of pressure-composition phase diagrams for CO_2/crude-oil systems: Society of Petroleum Engineers Journal, v. 24, p. 485-497.

Oyerokun, A. A., K. Aminian, S. Ameri, H. I. Bilgesu, and D. Della-Giustina, 2002, A new approach for training and testing artificial neural networks for permeability prediction, West Virginia University Libraries.

Rao, D. N., 1997, A new technique of vanishing interfacial tension for miscibility determination: Fluid phase equilibria, v. 139, p. 311-324.

Rao, D. N., and J. I. Lee, 2002, Application of the new vanishing interfacial tension technique to evaluate miscibility conditions for the Terra Nova Offshore Project: Journal of Petroleum Science and Engineering, v. 35, p. 247-262.

Rathmell, J. J., F. I. Stalkup, and R. C. Hassinger, 1971, A laboratory investigation of miscible displacement by carbon dioxide: Fall meeting of the society of petroleum engineers of AIME.

Sebastian, H. M., R. S. Wenger, and T. A. Renner, 1985, Correlation of minimum miscibility pressure for impure CO_2 streams: Journal of Petroleum Technology, v. 37, p. 2,076-2,082.

Shelton, J. L., and L. Yarborough, 1977, Multiple phase behavior in porous media during CO_2 or rich-gas flooding: Journal of Petroleum Technology, v. 29, p. 1,171-1,178.

Shi, Y., and R. Eberhart, 1998, A modified particle swarm optimizer: Evolutionary Computation Proceedings, 1998. IEEE World Congress on Computational Intelligence., The 1998 IEEE International Conference on, p. 69-73.

Shokir, E. M. E.-M., 2007a, CO_2–oil minimum miscibility pressure model for impure and pure CO_2 streams: Journal of Petroleum Science and Engineering, v. 58, p. 173-185.

Shokir, E. M. E.-M., 2007b, Precise model for estimating CO_2—oil minimum miscibility pressure: Petroleum Chemistry, v. 47, p. 368-376.

Shokrollahi, A., M. Arabloo, F. Gharagheizi, and A. H. Mohammadi, 2013, Intelligent model for prediction of CO_2–reservoir oil minimum miscibility pressure: Fuel, v. 112, p. 375-384.

Smits, G. F., and E. M. Jordaan, 2002, Improved SVM regression using mixtures of kernels: In Neural Networks, 2002. IJCNN'02. Proceedings of the 2002 International Joint Conference on, p. 2785-2790.

Smola, A. J., 1996, Regression estimation with support vector learning machines: Master's thesis, Technische Universit at M unchen.

Smola, A. J., and B. Schölkopf, 1998, Learning with kernels, MIT.

Smola, A. J., and B. Schölkopf, 2004, A tutorial on support vector regression: Statistics and computing, v. 14, p. 199-222.

Smola, A. J., B. Schölkopf, and K.-R. Müller, 1998, General cost functions for support vector regression: In Proceedings of the 8th International Conference on Artificial Neural Networks.

Spence Jr, A. P., and R. W. Watkins, 1980, The effect of microscopic core heterogeneity on miscible flood residual oil saturation: SPE Annual Technical Conference and Exhibition.

Srivastava, R. K., S. S. Huang, and M. Dong, 2000, Laboratory investigation of Weyburn CO_2 miscible flooding: Journal of Canadian Petroleum Technology, v. 39.

Stalkup Jr, F. I., 1983a, Miscible displacement, Society of Petroleum Engineers,Richardson, TX.

Stalkup Jr, F. I., 1983b, Status of miscible displacement: Journal of Petroleum Technology, v. 35, p. 815-826.

Sun, Y. H., G. Z. Lv, Y. F. Wang, and A. Q. Dong, 2006, A method of state equation for determining minimum miscible pressure of CO_2: Pettroleum Geology and Recovery Efficiency, v. 13, p. 82-84.

Thakur, G. C., C. J. Lin, and Y. R. Patel, 1984, CO_2 minitest, little knife field, ND: a case history: SPE enhanced oil recovery symposium.

Turek, E. A., R. S. Metcalfe, L. Yarborough, and R. L. Robinson Jr, 1984, Phase equilibria in CO_2-multicomponent hydrocarbon systems: experimental data and an improved prediction technique: Society of petroleum engineers journal, v. 24, p. 308-324.

Vapnik, V. N., 1963, Pattern recognition using generalized portrait method: Automation and remote control, v. 24, p. 774-780.

Vapnik, V. N., and A. J. Chervonenkis, 1964, On the one class of the algorithms of pattern recognition: Automation and Remote Control, v. 25.

Vapnik, V. N., S. E. Golowich, and A. Smola, 1997, Support vector method for function approximation, regression estimation, and signal processing: Advances in neural information processing systems, p. 281-287.

Vapnik, V. N., and S. Kotz, 1982, Estimation of dependences based on empirical data, v. 40, Springer-Verlag New York.

Wang, G. C., T. R. Carr, Y. W. Ju, and C. F. Li, 2014, Identifying organic-rich Marcellus Shale lithofacies by support vector machine classifier in the Appalachian basin: Computers & Geosciences, v. 64, p. 52-60.

Wang, J., M. Dong, Y. Li, and H. Gong, 2015, Prediction of nitrogen diluted CO_2 minimum miscibility pressure for EOR and storage in depleted oil reservoirs: Fuel, v. 162, p. 55-64.

Yang, X., H. Peng, and M. Shi, 2013, SVM with multiple kernels based on manifold learning for breast cancer diagnosis, Information and Automation (ICIA), 2013 IEEE International Conference on, Yinchuang, China, IEEE, p. 396-399.

Yellig, W. F., and R. S. Metcalfe, 1980, Determination and prediction of CO_2 minimum miscibility pressures: Journal of Petroleum Technology, v. 32, p. 160-168.

Yuan, H., and R. T. Johns, 2005, Simplified method for calculation of minimum miscibility pressure or enrichment: SPE Journal, v. 10, p. 416-425.

Yuan, H., R. T. Johns, A. M. Egwuenu, and B. Dindoruk, 2004, Improved MMP correlations for CO_2 floods using analytical gas flooding theory: SPE/DOE symposium on improved oil recovery.

Yuan, H., R. T. Johns, A. M. Egwuenu, and B. Dindoruk, 2005, Improved MMP correlation for CO_2 floods using analytical theory: SPE Reservoir Evaluation & Engineering, v. 8, p. 418-425.

Zhang, H., D. Hou, and K. Li, 2015, An improved CO_2-crude oil minimum miscibility pressure correlation: Journal of Chemistry, v. 2015.

Zheng, S., J. Liu, and J. Tian, 2004, An SVM-based small target segmentation and clustering approach: Machine Learning and Cybernetics, 2004. Proceedings of 2004 International Conference on, p. 3318-3323.

Zhou, H., 2008, Experimental study on CO_2 miscible flooding in ultralow permeability reservoir Ph.D thesis thesis, Daqing Petroleum Institute, Heilongjiang.

Zhu, Y., L. Tian, Z. Mao, and L. Wei, 2005, Mixtures of kernels for SVM modeling, Advances in Natural Computation, Springer, p. 601-607.

Zuo, Y., J. Chu, S. Ke, and T. Guo, 1993, A study on the minimum miscibility pressure for miscible flooding systems: Journal of Petroleum Science and Engineering, v. 8, p. 315-328.

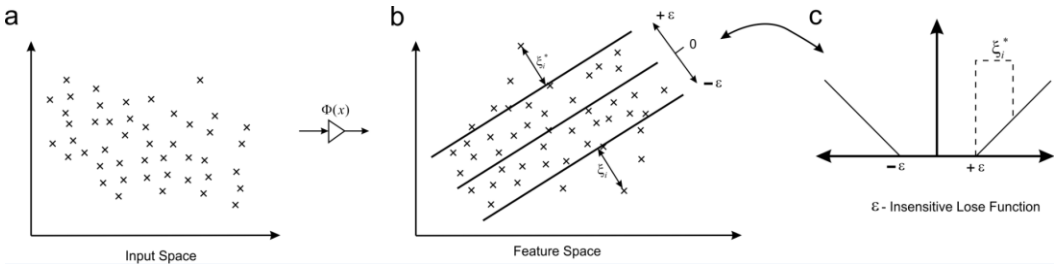


Figure 4-1. Transformation process illustration of a SVR model. A nonlinear mapping function ψ(x) defined to convert a nonlinear problem in the original (low dimensional) data input space (a) to linear problem in a (higher dimensional) feature space (b). The points

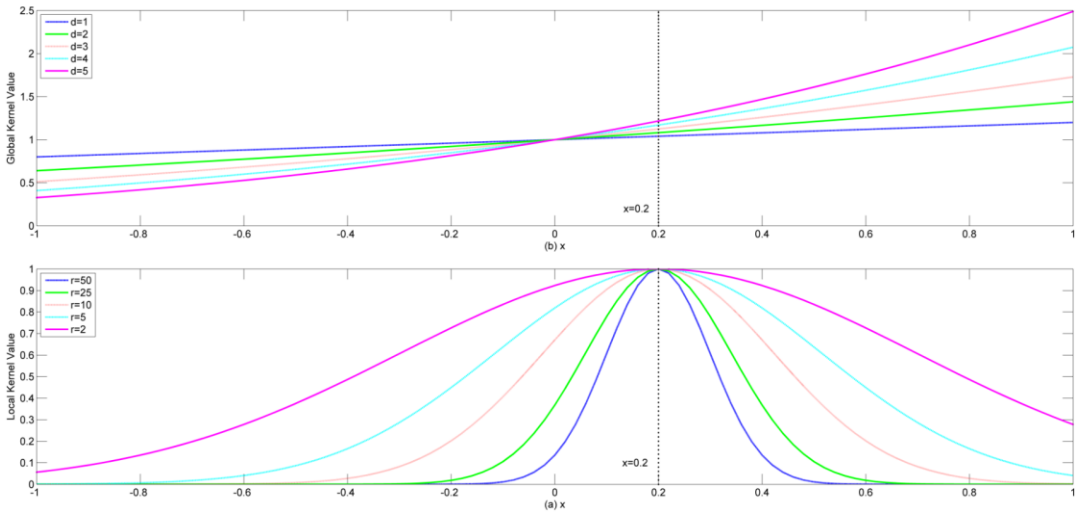


Figure 4-2.Mapping features of polynomial and RBF kernel. (a) is polynomial kernel, d is the operation degree, x=0.2 is test point. Various values of d was selected, only the points which are far enough from test point will have an effective influence on the kernel value; (b) is radial basis function kernel, x=0.2 is test point. Various values of ϒ was selected, the points adjacent to the test point have a great influence on the kernel values.

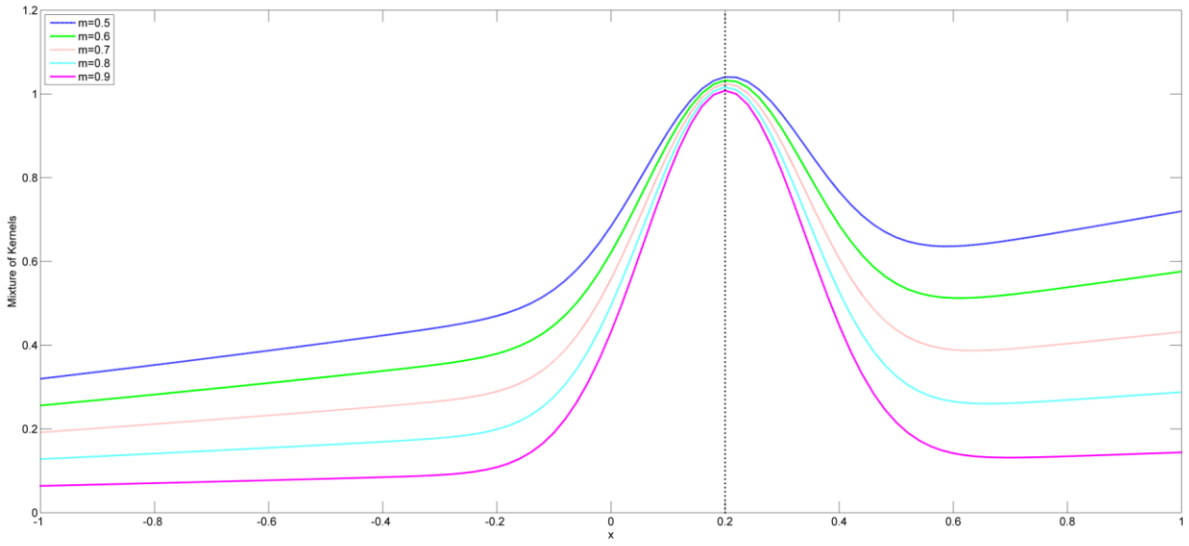


Figure 4-3.Curves of mixed kernel function. x=0.2 is the test point, ϒ is 0.1 and d=1. Various value of mixing coefficient (m) was selected, data points which are both far away from the test point and adjacent to the test point have a great influence on the kernel.

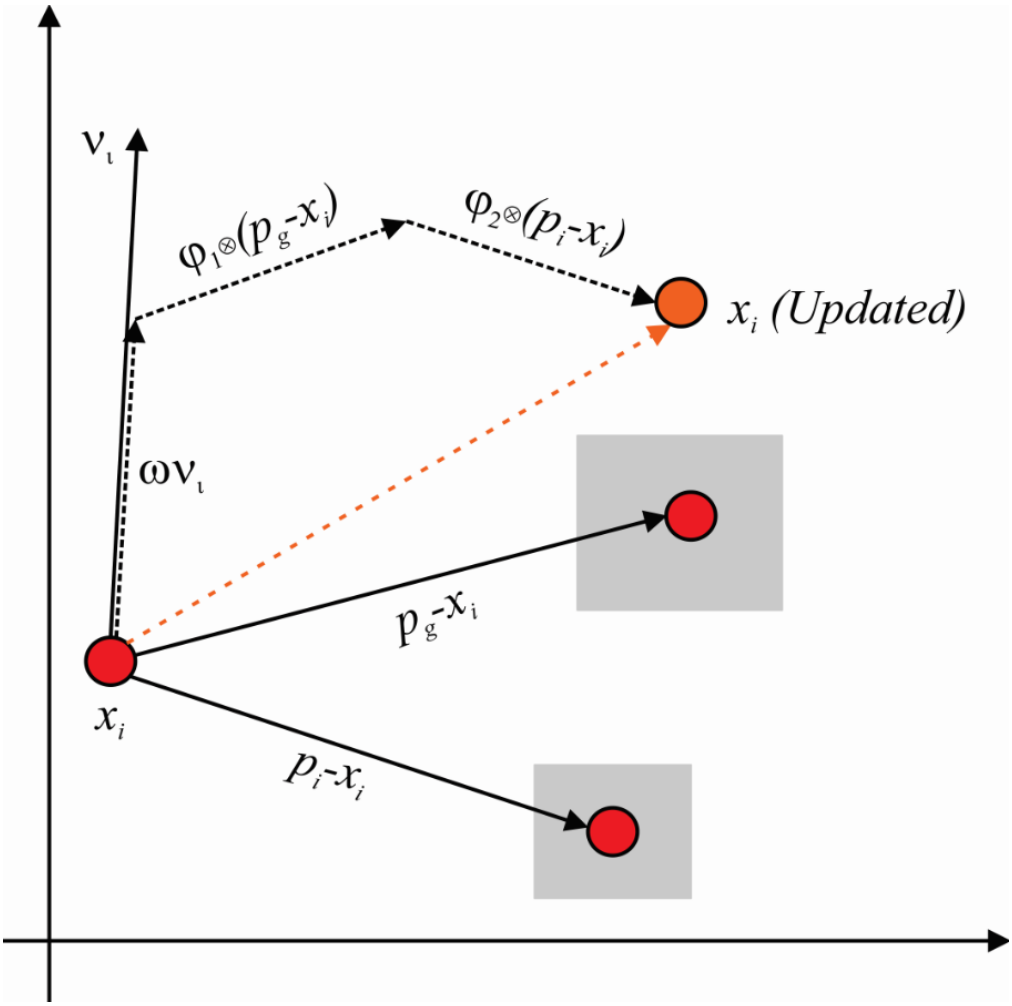


Figure 4-4. Illustration of PSO velocity and particle position update for particle xi in a two-dimensional search space.

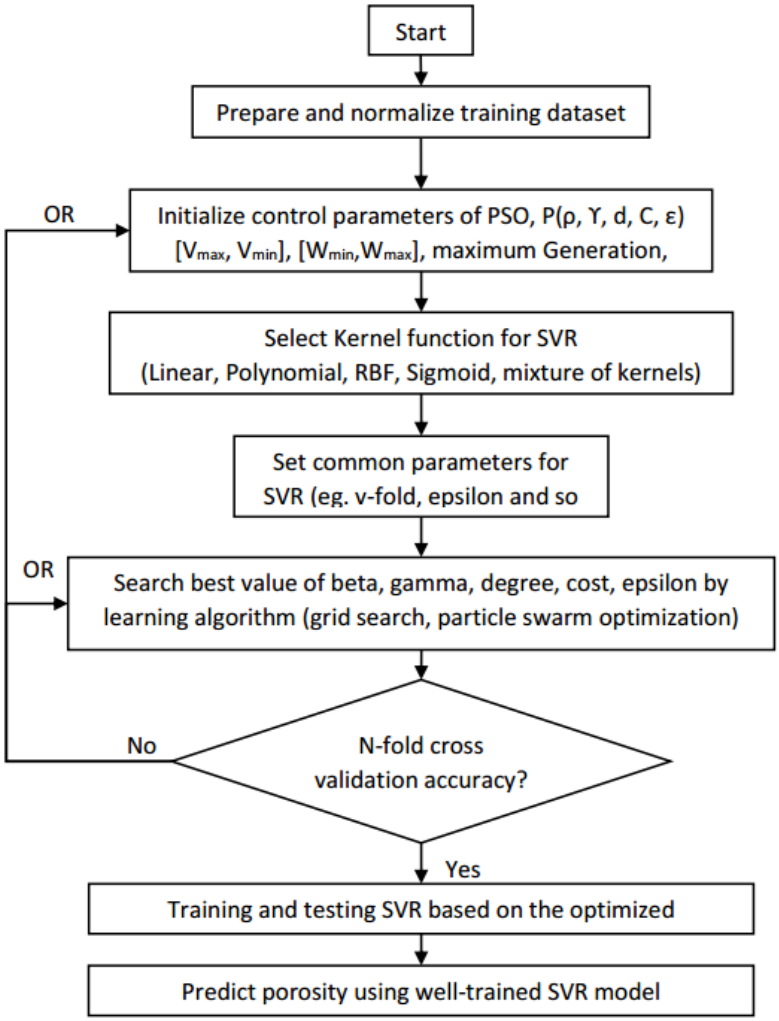


Figure 4-5. Workflow of PSO to optimize parameters of mixture kernel function.

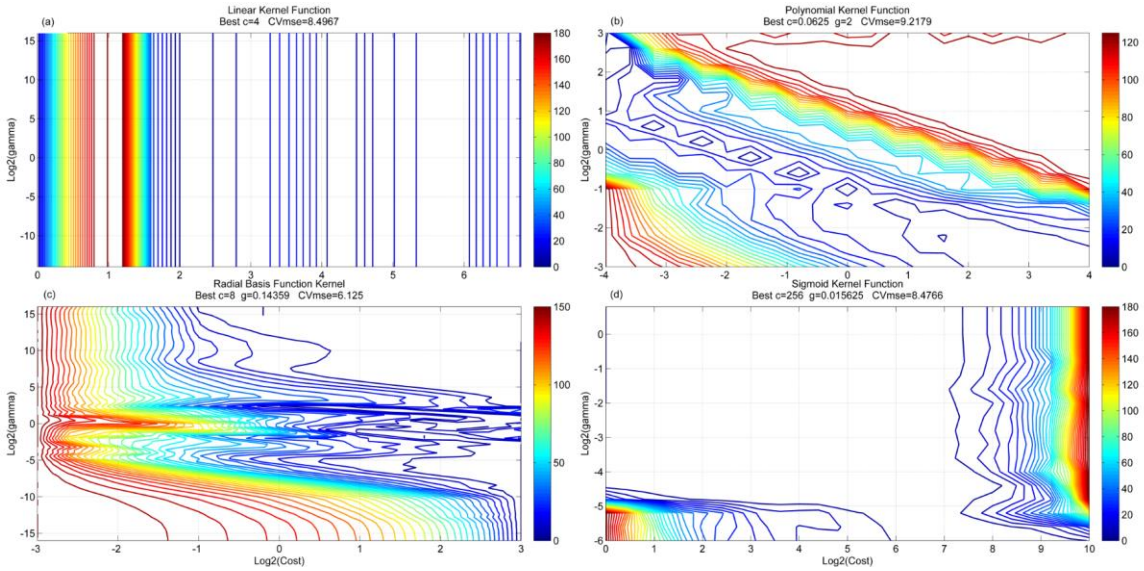


Figure 4-6. The process of searching best gamma and cost parameters by grid searching algorithm for SVR with linear kernel (a), polynomial kernel (b), RBF kernel (c), sigmoid kernel (d) and n-fold of 4. The color of the contour lines in the figure indicated the associated cross-validation mean square error.

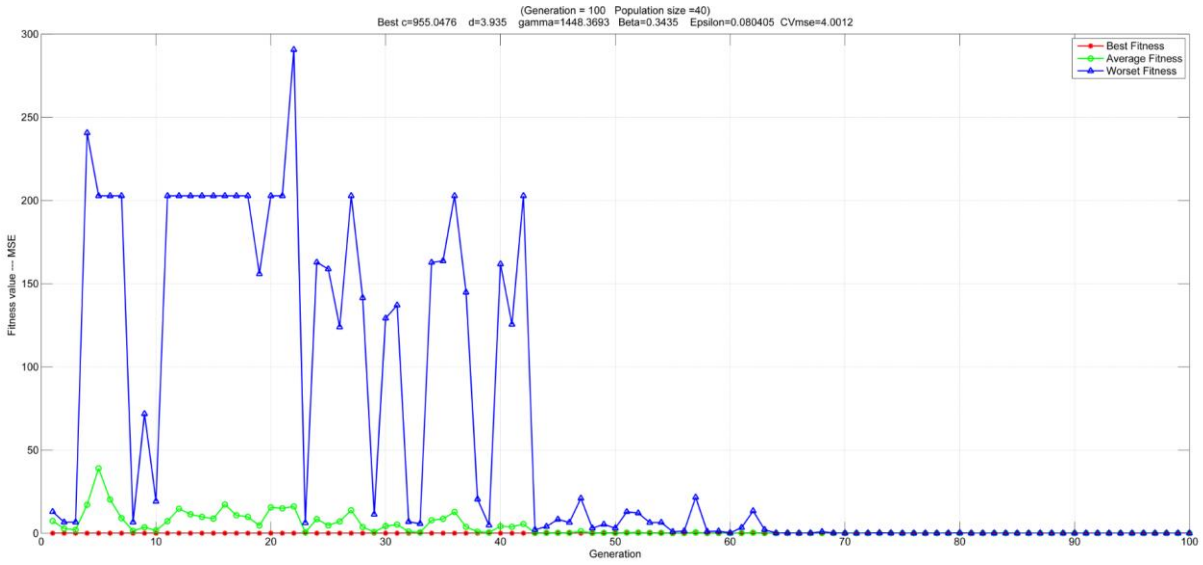


Figure 4-7. This plot shows the process of searching for best gamma and cost value by particle swarm optimization (PSO) for MKF-SVR model and n-fold of 4.

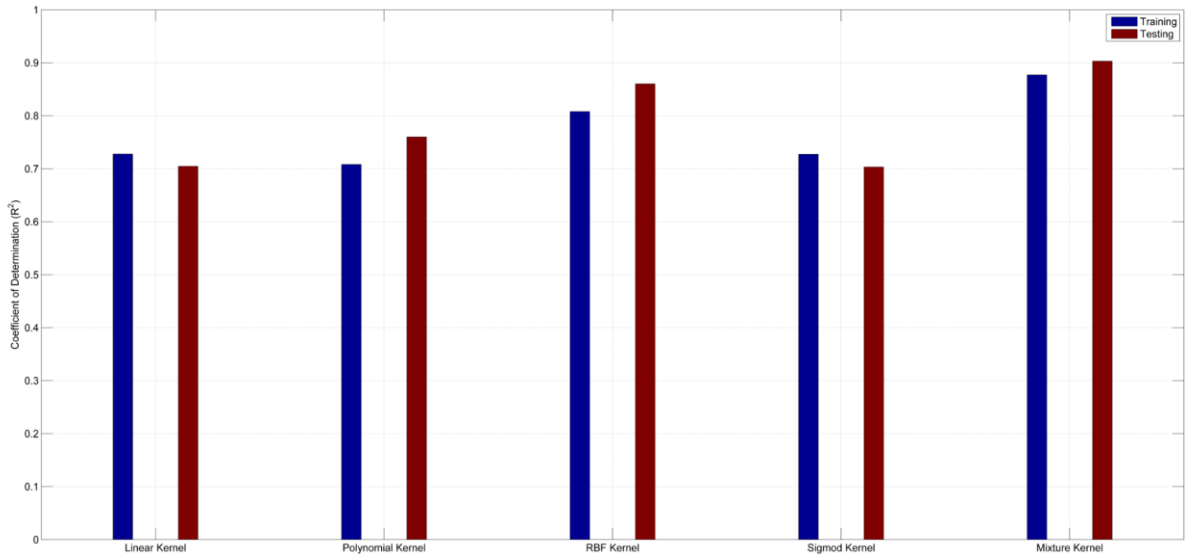


Figure 4-8. Determination of the correlation coefficient of training SVR with different Kernel function

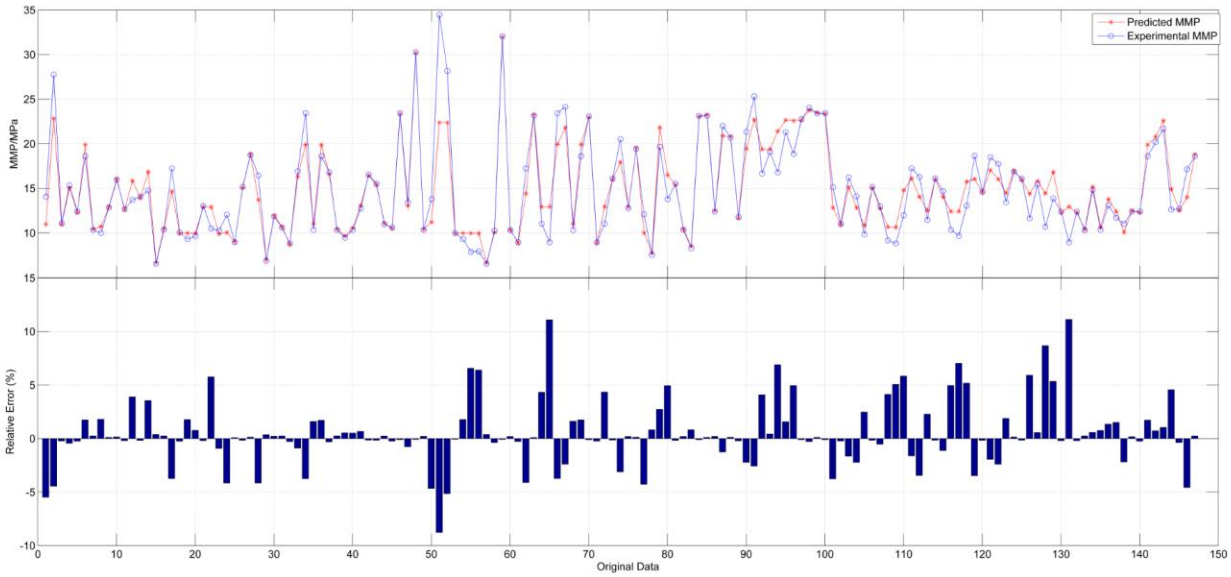


Figure 4-9. Comparison of actual values and forecasted values by mixed kernels function based SVR model.

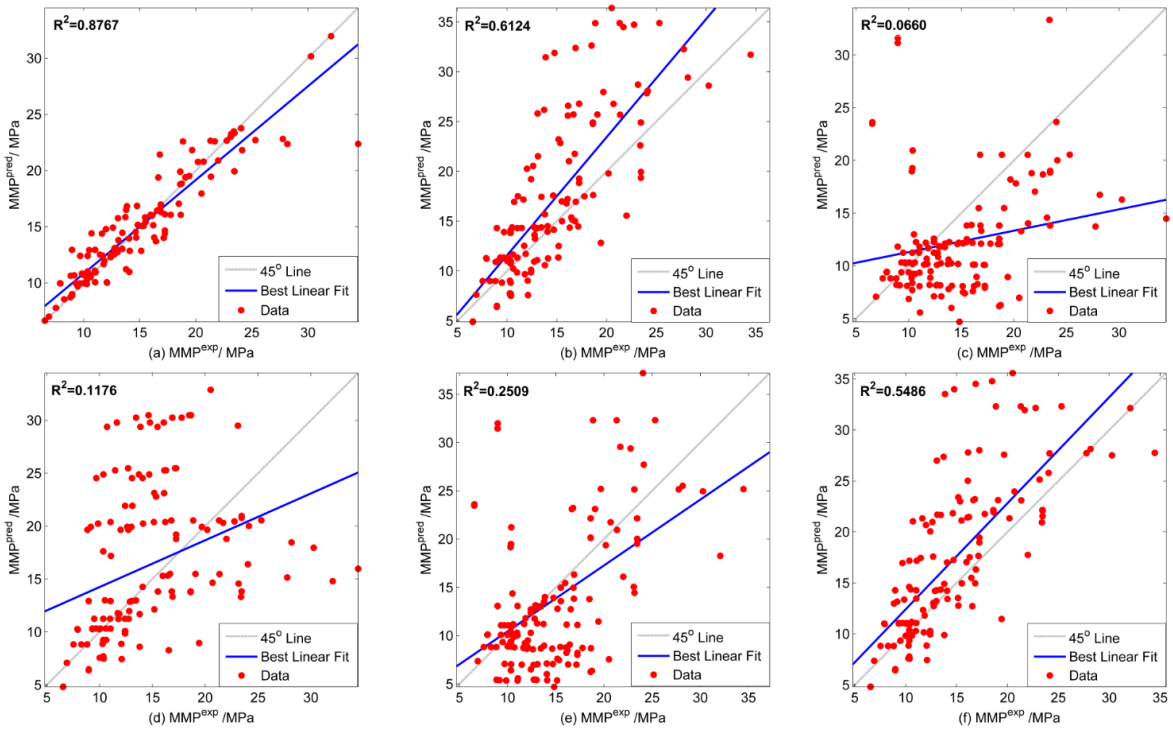


Figure 4-10. Comparison between the results of the developed model and other well-known correlations. (a) mixed kernels function based SVR (b) Alston et al. (1985) correlation, (c) Yellig and Metcalfe (1980) correlation (corrected with Sebastian et al. (1985) impurity correction factor), (e) Emera and Sarma (2005) correlation (corrected with Sebastian et al. (1985) impurity correction factor), and (f) Emera and Sarma (2005) correlation (corrected with Alston et al. (1985) impurity correction factor). MMP^{exp} is the MMP value measured by experiments, and MMP^{pred} is the MMP value predicted based on the correlation models.

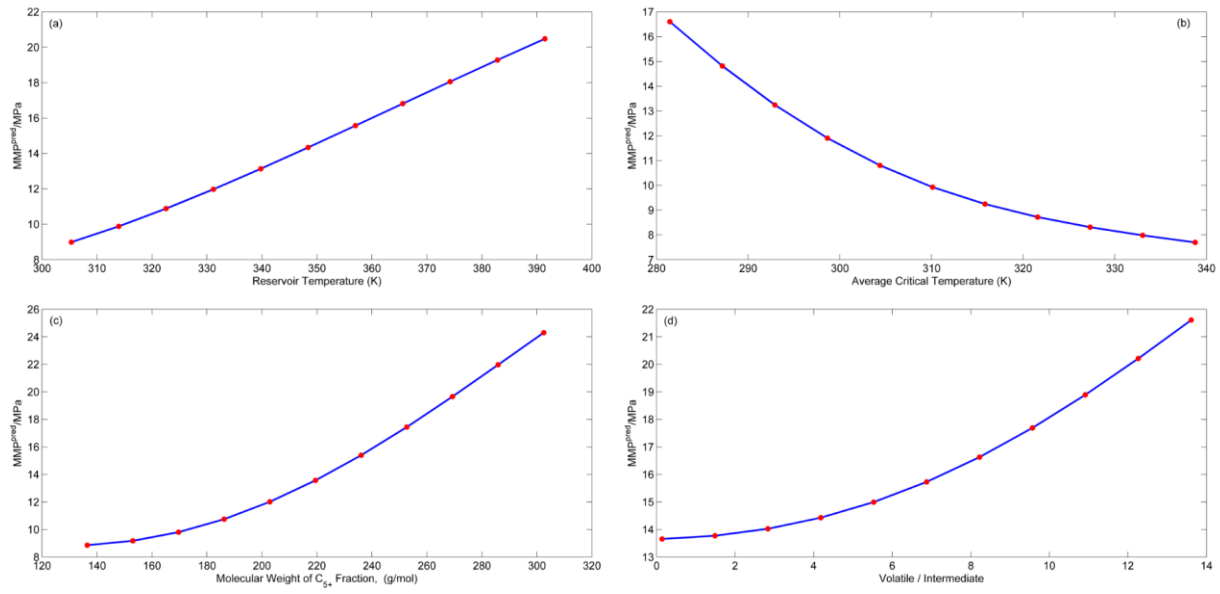


Figure 4- 11. Sensitivity analysis of the proposed model to vary input parameters. (a) reservoir temperature; (b) average critical temperature; (c) molecular weight of pentane plus fraction; and (d) volatile component to intermediate component ratio.

Table 4-1. Literature experimental data that used for MMP prediction.

Reference	Training Dataset		Testing Dataset	
	Pure	Impure	Pure	Impure
Rathmell et al. (1971)	3			
Jacobson, (1972)	1	3		
Dicharry et al. (1973)	1			
Holm et al. (1974)			1	
Shelton et al. (1977)			1	
Spence et al. (1980)	1		1	
Graue et al. (1981)		2		
Gardner et al. (1981)	1			
Metcalfe (1982)	4	20	1	2
Frimodig (1983)	3	18		2
Henry et al. (1983)	1			
Thakur et al. (1984)	2			
Alston et al. (1985)	11	12	1	1
Sebastian et al. (1985)	1	6		
Harmon et al. (1988)	1			
Eakin et al. (1988)	1	7	1	1
Chaback (1989)	4			
Khan et al. (1992)	2			
Zuo et al. (1993)	1	1		
Dong (1999)		2		
Srivastava et al. (2000)				
Dong et al. (2001)	1	1		1
Jaubert et al. (2002)	1			
Bon et al. (2005)	3			
Bon et al. (2006)	1			
Sun et al. (2006)	2			
Zhou et al. (2008)	1			
Al-Ajmi, et al. (2009)	1	2		
Li et al. (2012)	4			
Peng (2012)	1			
Zhang et al. (2015)	6		1	
Total	59	74	7	7

Table 4-2. Common Kernel function, corresponding mathematical expressions and parameters ranges.

Kernel Function		C	ϒ	d	ε	m
Linear Kernel	$K(x_i, x_j) = \langle x_i, x_j \rangle + c$	$2^0 \sim 2^7$	×	×	×	×
Polynomial Kernel	$K(x_i, x_j) = (\gamma \langle x_i, x_j \rangle + c)^d$	$2^{-4} \sim 2^4$	$2^{-3} \sim 2^3$	2	×	×
Radial Basis Function Kernel	$K(x_i, x_j) = \exp(-\gamma \left\| x_i - x_j \right\|^2)$	$2^{-3} \sim 2^3$	$2^{-15} \sim 2^{15}$	×	×	×
Sigmoid Kernel	$K(x_i, x_j) = \tanh(\gamma \langle x_i, x_j \rangle + c)$	$2^{-0} \sim 2^{10}$	$2^{-6} \sim 2^1$	×	×	×
Mixture Kernel	$K_{mix} = \rho K_{poly} + (1-\rho) K_{rbf}$	$2^{-2} \sim 2^5$	$2^{-3} \sim 2^4$	1~4.5	$10^{-4} \sim 10^{-1}$	0~1

Table 4-3. Typical parameters used for training SVR model with PSO algorithms.

Parameter	**Value**	**Parameter**	**Value**
Maximum Generation	150	Population size	50
Cognitive efficient (ω_1)	1.5	Social efficient (ω_2)	1.5
Initial inertia weights $\omega_{\max}$	0.9	Final inertia weight $\omega_{\min}$	0.4

Table 4-4. Error measures for accuracy assessment (Al-Anazi and Gates, 2010a)

Accuracy measure	Mathematical expression
Coefficient of Determination, R^2	$R^2 = 1 - \frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)}{\sum_{i=1}^{m}(y_i - averg(y_i))}$
Correlation coefficient, r	$\frac{\sum_{i=1}^{m}(y_i - \bar{y}_i)(\hat{y}_i - \bar{\hat{y}}_i)}{\sum_{i=1}^{m}(y_i - \bar{y}_i)^2 \sum_{i=1}^{N_p}(\hat{y}_i - \bar{\hat{y}}_i)^2}$
Root mean square error, RMSE	$\sqrt{\frac{1}{l}\sum_{i=1}^{m}(y_i - \hat{y}_i)^2}$
Average absolute error, AAE	$\frac{1}{l}\sum_{i=1}^{m}\lvert y_i - \hat{y}_i \rvert$
Maximum absolute error, MAE	$\max\lvert y_i - \hat{y}_i \rvert, i = 1,2,...m$

Table 4-5. The training and testing performance of SVR model with mixture kernel and other basic kernel functions.

Data Set	Statistical Parameters	Kernel Function				
		Linear	Polynomial	RBF	Sigmoid	MKF
Training Set	RMSE	2.9220	3.0128	2.4558	2.9181	2.0111
	AAE	2.1244	2.0668	1.7852	2.2223	1.1659
	MAE	5.8374	6.3763	4.8271	5.9940	5.1228
	r	0.8531	0.8415	0.8989	0.8528	0.9365
	N^e	133				
Test Set	RMSE	2.2851	1.9760	1.4670	2.2667	1.204
	AAE	1.7873	1.3221	1.0862	1.8198	0.8814
	MAE	4.7201	3.9290	1.4040	2.5362	2.4607
	r	0.8395	0.8717	0.9274	0.8386	0.9503
	N^e	14				
Total	RMSE	2.8675	2.9299	2.3794	2.8625	1.9151
	AAE	2.0923	1.9959	1.7186	2.1840	1.1406
	MAE	5.8374	6.3763	4.8271	5.9940	4.6291
	r	0.8411	0.8695	0.9099	0.8420	0.9381
	N^e	147				

Table 4- 6. The comparison of five correlations statistical results.

Model	R	RMSE	AAE	MAE	N
MFK-SVR	0.9381	1.9151	1.1140	4.6291	147
Alston et al., (1985)	0.7826	5.5644	30.9630	309.2128	141
Yelling et al. [24](Sebastian et al. [83] correction)	0.2145	7.0807	50.1360	599.5285	147
Yelling et al. (1980), (Alston et al.,1985 correction)	0.3429	7.4034	54.8104	348.8954	143
Emera et al. (2005), (Sebastian et al. (1985)correction)	0.5001	6.6600	44.3556	529.8390	146
Emera et al. (2005), (Alston et al. (1985) correction)	0.7407	5.7220	32.7420	386.5963	142