SPE 99261

Novel Approach To Predict Potentiality of Enhanced Oil Recovery

H. Jian, China U. of Petroleum, and H. Wenfen, Imperial College London

This paper was prepared for presentation at the 2006 SPE Intelligent Energy Conference and Exhibition held in Amsterdam, The Netherlands, 11–13 April 2006.

Abstract

The potentiality prediction of enhanced oil recovery (EOR) is the basis of EOR potentiality analysis as well as the robust guarantee of the reliability of analysis results. In the light of statistical learning theory, establishing an EOR predictive model substantially falls within the problem of function approximation. According to Vapnik's structural risk minimization principle, one should improve the generalization ability of learning machine, that is, a small error from an effective training set can guarantee a small error for the corresponding independent testing set. The up-to-date results from studies on statistical theory in recent decades even recent years are firstly applied to EOR potentiality analysis. The applications of group method of data handling (GMDH), Improved BP artificial neural network, and support vector machine (SVM) are discussed. The comparison of the results from three methods indicates that SVM can pay more attention to both the universality and extendibility of a model when the samples are very limited, which shows a good prospect of its application. A method used to generate a set of sample theoretically is developed in this research by combining quadrate designing, reservoir simulation, and economical evaluation.

Introduction

The analysis of EOR potentiality can guide the implementations of EOR methods and lower the risks in tertiary recovery in China, where a good potentiality of EOR has been seen. Potentiality prediction is the basis of potentiality analysis as well as the guarantee of the reliability of results. At present, one popular method used to predict the potentiality of EOR is numerical simulation[1,2], in which a huge amount of work has to be done due to varieties and heterogeneities of interested areas.

If various factors that affect performances and profits of EOR can be identified and the relationships among these factors can be recognized by a statistical method, one can employ these relationships to predict EOR potentiality with statistical samples obtained from either fields or the results of simulations. In fact, since the basic principle of statistic method is drawing a universal law from a complex system and applying it back into the same system, in which there are some principles beyond the descriptions of mathematic models, one can always achieve a good result. For example, there are four sub-models of steam-flooding potentiality prediction used by the United States Department of Energy (USDE). All of them were built by using mathematical modeling or statistic analysis/empirical formulation; one of them, Gomma — a statistical model based on numerical simulation, performances better than others. In addition, statistics-based prediction models are also popular for faster calculating and simpler styles.

Statistical learning Theory (SLT) is a new theory about studying laws of learning machine aimed to small sample sets. Vapnik, *et al*.[3] started the research in this area from 1960s. In the mid-1990s, SLT had gained increasingly popular attention because of its increasingly mature theory and slow developing in other learning theories (for instance, the method of neural network). In the same period, support vector machine (SVM) was developed on the basis of SLT as a new versatile learning method and showed its superiority at the very beginning of its birth. According to Vapnik's structural risk minimization principle, SVM has powerful generalization ability. Moreover, since its algorithm falls within the problem of convex optimization, its global optimal solution equals its local optimal solution.

There are three sorts of statistical learning questions, namely, pattern recognition, function approximation and density estimation. EOR potentiality prediction is in the category of function approximation. In this study on the application of SLT in EOR, a polymer-flooding theoretical sample is taken as an example.

Methodology

Basic question in Statistical Learning. Statistical learning aims to estimate the relationships of input and output in a system, according to a specific training set, and to predict unknown output on the basis of these relationships as accurately as possible. In the case of limited samples, the smallest empirical risk does not necessarily mean the smallest

expected risk (the actual risk); the complexity of studying machine is not only related to the studied system but also adapted to the limited samples.

Compared with traditional gression analysis and netrual network, the method of function matching using SLT can not only achieve a satisfying learning effect with certain complexities, namely satisfying a certain learning precision, but also emphasize an extensive application, which enables so-built models to give a good forecasting effect.

In order to study the rate of convergence in learning process and the extendibility of the method, SLT defines a series of indices related to the properties of function learning. The most important one is Vapnik—Chervonenkis Dimension (VC), which reflects the learning ability of a set of function. The larger a VC, the more complicated the learning machine, the more powerful its study ability and the larger capacity of samples.

For various sets of function, the relationship between empirical risk and actual risk, so-called the circle of extension, is defined as

$$R_{act}(\omega) \le R_{emp}(\omega) + \Phi\left(h/n\right) \tag{1}$$

where ω is a generalized parameter and can be a free real number or a set of plural number. h is the Vapnik — Chervonenkis Dimension (VC) and n is the number of samples.

Eq. (1) shows that actual risk $R_{act}(\omega)$ consists of two parts: one is empirical risk $R_{emp}(\omega)$ (training error); the other is fiducial range $\Phi\left(h/n\right)$, which is related to VC and n. With limited training samples, the higher the VC (the more complicated the learning machine), the larger the fiducial range and then the larger a possible difference between actual risk and empirical risk. Therefore, SLT follows Structural Risk Minimization (SRM) principle and balances empirical risk and fiducial range to gain the smallest actual risk.

Support Vector Machine. Support vector machine (SVM) is the latest and most practicable part of SLT, the core content of which was put forward during 1992-1995. According to Vapnik, minor representative samples, so-called support vectors, were chosen from a set of sample to form matching functions, whereby the problems of universality and extendibility can be solved[4]. SVM is an ideal method to perform nonlinear function matching.

The definition of inner product kernel. In order to solve the problem of universality, SVM adopts the technique of kernel function. The kernel function is undependent on the style of nonlinear matching function. We estimate an unknown function $f(x)$ using Fourier polynomial in order to find a universal nonlinear matching function:

$$f(x) = \frac{a_0}{\sqrt{2}} + \sum_{k=1}^{N}\left(a_k \sin kx + b_k \cos kx\right) \tag{2}$$

If a basic function system is

$$U(x) = \left(1/2, \sin x, \cdots, \sin Nx, \cos x, \cdots, \cos Nx\right)^T,$$

$$C = \left(a_0, a_1, \cdots, a_N, b_1, \cdots, b_N\right),$$

then

$$f(x) = \sum_{l=1}^{M} c_l u_l(x) = \langle C, U(x)\rangle \tag{3}$$

where $M = 2N + 1$. $\langle\ \rangle$ stands for inner products.

Improving the accuracy of function approximation depends on increasing the number of function (N). However, the number of basic functions increases too rapidly as the dimensions of variable x increase. When $n << M$, there is no point to solve C using underdetermined equation group. Define $C\left(C \in R^M\right)$ with $\left\{U\left(x_j\right) \middle| U \in R^M, j = 1,2,\cdots,n\right\}$, then we have

$$C = \sum_{j=1}^{n} \beta_j U\left(x_j\right) \tag{4}$$

where $\beta_1, \beta_2, \cdots, \beta_n$ are parameters that need to be defined.

Eq.(3) can be transformed as

$$f(x) = \langle C, U(x)\rangle = \sum_{j=1}^{n} \beta_j K_N\left(x, x_j\right) \tag{5}$$

Define kernel function with inner products as

$$\begin{aligned} K_N\left(x, x_j\right) &= \left\langle U(x), U\left(x_j\right)\right\rangle \\ &= \frac{1}{2} + \sum_{k=1}^{N}\left(\sin(kx)\sin\left(kx_j\right) + \cos(kx)\cos\left(kx_j\right)\right) \\ &= \frac{1}{2} + \sum_{k=1}^{N}\cos\left[k\left(x - x_j\right)\right] \\ &= \sin\left[\frac{2N+1}{2}\left(x - x_j\right)\right] \Big/ \sin\left(\frac{x - x_j}{2}\right) \end{aligned} \tag{6}$$

When n becomes bigger, non-zero β_j is just a very small part of Eq.(5). Its corresponding sample vectors are support vectors and can reflect main changes in the function. It is not difficult to see that the matching function Eq.(5), defined by Eq.(6), is appliable to any nonlinear functions satisfying Fourier series expansion and knowing specific forms of matched functions is unneccessary.

Generally speaking, support vector machine implements the following idea: it tranforms the input space R^n into a high-dimensional feature space through a nonlinear transformation defined by the inner product function. In this space, it matches functions. Since only inner product operations $\left(x_i \cdot x_j\right)$ among training samples are involved in the process of optimizing function, the only calculation in high-dimensional feature space is the inner product operation, which can be done in the original space with no need to know the form of transformation.

According to functional theory, so long as one kernel function satisfies the Mercer condition, there is a corresponding inner product operation in feature space. In other word, a function Ψ exists to make

$\Psi(x)\cdot\Psi(x_j)=K(x,x_j)$. At present, ordinary kernel functions used in SVM include polynomial kernel function, radial basis kernel function (RBF), and Sigmoid function.

SVM in function matching. Numerical method for underdetermined problems is employed to solve the problem of the extendibility of SVM. With a certain kernel function $K(x,x_j)$, the matching function can be written as

$$y(x)=\sum_{j=1}^{n}\beta_j K(x,x_j)+b_0 \tag{7}$$

For a given set of sample $\{(x_i,y_i)|i=1,2,\cdots,n\}$, $b_0,\beta_1,\cdots,\beta_n$ are solutions for the following underdetermined equation group:

$$\begin{cases} y_i=\sum_{j=1}^{n}\beta_j K(x_i,x_j)+b_0, i=1,\cdots,n \\ b_0=\frac{1}{n}\sum_{i=1}^{n}y_i-\sum_{j=1}^{n}\beta_j\left(\frac{1}{n}\sum_{i=1}^{n}K(x_i,x_j)\right) \end{cases} \tag{8}$$

Normally, Eq.(8) is a ill-posed equation group, whose solutions, given by the conventional least-squares method, are often dependence on dataset discontinually. Previded (x_i,y_i) is a sample. When the input is $x_i+\Delta x$ and the output is $y_i+\Delta y$ for a matching function, $|\Delta y|$ might be very large even with a very small $|\Delta x|$. This means Eq. (7) is not applicable to the outside of the set of sample. In the 1960s, Ivanov and Phillips, *et al.*[5,6] introduced regularized function $\Omega(f)$ into learning theory and proposed several regularization methods to solve underdetermined problems.

Solving the equation group (8) using the Phillips Method, the solutions ($b_0,\beta_1,\cdots,\beta_n$) are the solutions for the following quadratic programming problem:

$$\min\Omega(\beta)=\frac{1}{2}\sum_{j=1}^{n}\beta_j^2 \tag{9}$$

$$s.t.\ y_i-\sum_{j=1}^{n}\beta_j K(x_i,x_j)-b_0\le\varepsilon \quad i=1,\cdots,n \tag{10}$$

$$\sum_{j=1}^{n}\beta_j K(x_i,x_j)+b_0-y_i\le\varepsilon \quad i=1,\cdots,n \tag{11}$$

The equations above reflect an application of Vapnik's structural risk minimization principle in support vector machine (SVM). The aim of minimizing $\Omega(\beta)$is to make a matching function as smooth as possible. In Eq.(7), when $\beta_1=\beta_2\cdots=\beta_n=0$, $y(x)=b_0$; the matching function is in its smoothest state. A smooth matching function has a smaller VC dimension, and consequently, a smaller fiducial range $\Phi(h/n)$. The constraints of inequality assume that all training samples can be matched correctly under an accuracy of ε, whereby a small empirical risk $R_{emp}(\omega)$is guaranteed. To solve a quadratic programming problem is to lessen $R_{emp}(\omega)+\Phi(h/n)$. Therefore, the accuracy ε is a positive value used to coordinate matching precision (universality) and forecasting precision (extendibility) in a matching function. The smaller the value, the higher the matching accuracy and the worse the extendibility.

According to Wolf's dual programming theory, the quadratic programming problem above can be transformed into a dual programming problem with simple constraints and simple calculations. The following equation can be derived:

$$\max L(\alpha',\alpha'')=-\frac{1}{2}\sum_{i,j=1}^{n}(\alpha_i''-\alpha_i')(\alpha_j''-\alpha_j')K(x_i,x_j)+\sum_{i=1}^{n}(\alpha_i''-\alpha_i')y_i-\varepsilon\sum_{i=1}^{n}(\alpha_i'+\alpha_i''),$$

$$s.t.\quad \sum_{i=1}^{n}\alpha_i'-\sum_{i=1}^{n}\alpha_i''=0,\quad \alpha_i'>0,\alpha_i''>0\ ,\quad i=1,\cdots,n \tag{12}$$

Then Eq. (7) is transformed to

$$y(x)=\sum_{i=1}^{n}(\alpha_i''-\alpha_i')K(x,x_i)+b_0 \tag{13}$$

$$b_0=\frac{1}{n}\sum_{i=1}^{n}y_i-\frac{1}{n}\sum_{i,j=1}^{n}(\alpha_j''-\alpha_j')K(x_i,x_j) \tag{14}$$

After the solutions of Eq.(12), $\{\alpha_i',\alpha_i''|i=1,2,\cdots,n\}$, have been obtained using nonlinear programming method, $y(x)$ can be calculated from Eq.(13) and Eq.(14) with any x value.

Constructing of a set of sample. Although a rapid development of the technique of polymer flooding has been seen, it is still very difficult to construct a set of learning sample from a polymer-flooded field. The following method is proposed to build learning samples by combining orthogonal design, reservoir numerical simulation and economic estimation[7]:

· Screen sensitive parameters. According to the conditions of an interested area, parameters sensitive to the technique of polymer-flooding and economics are selected empirically and theoretically;

· Generate sample projects using orthogonal design. Make sure that variances of sensitive parameters can be represented by minimum projects;

· Perform polymer-flooding simulation and economic estimation on each project;

· Sort samples according to the indices of economic estimation.

Orthogonal design of sample projects. Fundamental research and field pilots show the factors that affect actual performance of polymer-flooding include reservoir and fluid properities, the indices of polymer performance, field injecting conditions and so on. Here, potentiality estimation is performed mainly according to the properties of reservoir and fluids with determined polymer and suitable injecting

conditions. We include the size of reservoir, initial oil saturation, permeability, variation coefficient, viscosity of reservoir oil, reservoir temperature, salinity of reservoir water, concerntration of Ca^{2+}、Mg^{2+} and the start time into the process of orthogonal designing, under the consideration that the scale and the start time of polymer-flooding might affect the profits of polymer flooding. In addition, the effects caused by formation rhythm and wettability are not taken into accout because the formations suit for polymer flooding in Shengli oilfield have been known mainly as positive rhythm and water-wetting.

By setting 10 values for the size of reservoir and 5 for other parameters, a mixed problem with 1-parameter-10- level and 8-parameter-5-level is formed. An orthogonal designing table $L_{50}\left(10^1 \times 5^{10}\right)$ is used, whereby 50 sample projects are generated. In **Table 1**, the values of orthogonal designing parameters are set by referring those from pilot fields in Shengli oilfield, China. For parameter A, the size of reservoir (block reserve, $\times 10^4$t), given the values of 200, 500, 800, 1100, 1400, 1700, 2000, 2300, 2600, and 3000.

Simulation of Polymer-flooding. We emply a 3D-2phase-5composition chemical model in the simulation, where various phsical and chemical phenomena in the process of polymer-flooding are taken into acount systmetically. Geologically, the model is isotropic in plane and anisotropic in vertical with five layers and 15m in total thickness. A five-point well pattern with 250m injector-producer spacing is placed in it (see **Figure 1**). The size of grid is $21 \times 21 \times 5$ with anisotropy in plane. The heterogeneity of permeability is represented by the variation coefficient (V_k). The permeabilities for each layer are calculated from average permeabilities and the variation coefficient.

Economic estimation of Polymer-flooding. Incremental method is adopted in the economic estimation of polymer-flooding, whereby increased incomes, saved costs and increased investments are analyzed comprehensively to determine the profits and costs of the increased production in the polymer-flooding. Polymer-flooding projects are judged to be applicable or not on the basis of appraising indices.

Forming a set of sample. **Table 2** gives the technological and economic indices obtained from the simulation and economic estimation on each polymer-flooding project. We use the performance indices of Polymer 3530S in the simulations. Injection conditions are set as: 1500mg/L injection concentration, 450mg/L.PV injection dose and 0.1PV per year injection rate. Other parameters, such as parameters of economic estimation, are borrowed from polymer-flooded pilots in Shengli oilfield.

In the whole set of sample, sample 4, 14, 30, 38 and 39 are selected randomly to form a post-appraisal set to test the extrapolation ability of matching function, while the rest 45 samples are used as a matching set to determine the parameters of matching function.

Results

On the basis of statistic learning theory (SLT), the applications of group method of data handling (GMDH), Improved BP artificial neural network and support vector machine (SVM) in EOR potentiality estimation are discussed here.

Group method of data handling (GMDH). In the 1970s, on the basis of K-G polynomial, Ivakhnenko, a former UUSR scholar, proposed a heuristic recognizing method for complex nonlinear system, Group Method of Data Handling (GMDH), using the auto-organization principle in biological control theory. Based on limitied information of the system, this method (also called auto-organization method) can automatically organize the structure of model, select input variables and propose grouping methods and selecting criteria of intermediate variable for different data to meet the different needs of recognizing model in various systems. Since being proposed, the GMDH method has gained universal attention and application as well as continuos improvement[8].

In this paper, a multi-layer prediction model of GMDH is generated after testing calculation and comparing, where enhanced recovery, utilization rate of polymer and internal rate of return are output variables and the other nine parameters (such as the size of reservoir, et al.) are input variables.

Improved BP artificial neural network. Artificial Neural Network (ANN), being a branch of Artificial Intelligence, was proposed on the basis of modern neural science. It is a comlpex network system formed by a large amount of widely-connected simple processing units which are similar to neurons in human's brain. Several dozens of ANN models have been proposed, the most famous and widest used of which is Error Back-Propagation (BP network). BP network is a learning algorithm of multilayer neural network proposed by Rumelhart in 1986 on the basis of mutilayer neural network models. It has been applied to nonlinear regression, pattern recognition and sorting, and also to the exploitation and development of oil[9].

In traditional BP network model, the problem of input/output for a set of sample is changed to a problem of nonlinear optimization. Since adopting an ordinary gradient method for optimization, this model has a good recognizing function and, theoretically, can satisfy any small error in the simulations of complex nonlinear models. However, some flaws exist in this model, such as local minimum problem, slower converging rate, so-called error platform, over-interfering in network structure, shortage of inductive rules, poor extraplolation ability, and so on. Some improvements, such as momentum factor rectification, learning factor self-adaptation, limitation exportation, pre- and post-process of data transformation, batch processing of sample, and so on, have been seen recently[7].

In this paper, we set the main factors that affect the potentiality of EOR as the inputs of the system, in which 9 nodes are included, and the indices of prediction as the outputs of the system, in which 3 nodes are included. A 3-layer neural network with 21 optimising implicit nodes is adopted. After performing a pseudo-linear transforming on the input and the output parameters, we set the initial learning factor is 0.9, momentum factor is 0.9, network iterations are 3000 times, and training system error is 0.0025.

Support vector machine. Regularized Fourier is used to expand kernel functions in function matching. The following is the specific form of Fourier expanding:

$$f(x)=\frac{a_0}{\sqrt{2}}+\sum_{k=1}^{\infty}q^k\left(a_k\sin kx+b_k\cos kx\right)\quad(0<q<1) \tag{15}$$

This expanding is different with the traditional Fourier expanding in that it has an extra factor q^k for regularization. For the problem of function approximation, the role of expanding items in improving the matching accuracy weakens as the value of k increases. Its corresponding kernel function is

$$K(x,x_i)=\frac{1}{2}+\sum_{k=1}^{\infty}q^k\left(\sin(kx)\sin(kx_i)+\cos(kx)\cos(kx_i)\right)$$

$$=\frac{1}{2}+\sum_{k=1}^{\infty}q^k\cos\left[k(x-x_i)\right]=\frac{1-q^2}{2\left[1-2q\cos(x-x_i)+q^2\right]} \tag{16}$$

We give q =0.5 in the above equation and build the matching functions for enhanced recovery, utilization rate of polymer and internal rate of return using SVM. The vectors of β coefficient are 45 on account of 45 matching samples. The vector of i^{th} sample whose coefficient is non-zero ($\beta_i \neq 0$) forms its own support vector.

Comparisons. **Figure 2**, **Figure 3**, and **Figure 4** show deviations between predicted values and actual values for enhanced oil recovery (EOR), polymer utility rate (PUR), and internal rate of return (IRR). From the comparison, we can obtain the following points:

(1) Group method of data handling (GMDH) gives lower accuracy in predication for both matching samples and prediction samples owing to using a multilayer dual quadratic complete polynomial to access the complicated nonlinear system. However, it has met basic needs of engineering calculation. The extrapolatibility of regression model in this method is improved because the data is grouped and, therefore, matching data and checking data are used separately in the process of modelling. In addition, it has superiorities over other methods by its simple principle and easily-realized model.

(2) Improved BP artificial neural network (Improved BP) gives higher accuracy for matching samples but lower accuracy for prediction samples. Therefore, its extendibility is poor, which indicates that the complexity of learning machine is unadapted to the limited samples. This problem can be solved by increasing the number of matching samples. However, that how to choose a network structure depends on users' skills might be an obstacle for promoting its application.

(3) Support vector machine (SVM) gives lower accuracy for matching samples but higher accuracy for prediction samples. It has a robust theoretical basis and relatively complicated principles and considers both universality and extendibility of a model in the application when only limited samples are available, which is advantageous to the application in engineering calculation.

Conclusions

A new method for EOR potentiality prediction, support vector machine (SVM), is proposed on the basis of statistical learning theory. Compared with the method of simulation, it is faster in calculating and easier to use. Compared with Group method of data handling(GMDH) and Improved BP artificial neural network (Improved BP), it overcomes the limitations appeared in conventional statistic methods and has better versatility and extendibility with limited sample sets. Under the guidance of Vapnik's structural risk minimization rule, the small errors obtained from an effective training set guarantee the errors from an independent testing set are small. Applying SVM into polymer-flooding potentiality prediction obtains good results and shows a good future of the method.

Nomenclature

a_k = *Fourier polynomial coefficient*
b_k = *Fourier polynomial coefficient*
C = *basic function*
f = *an unknown function*
h = *Vapnik—Chervonenkis Dimension*
K = *kernel function*
L = *Equation for a dual programming problem*
$M = 2N+1$
N = *number of functions*
n = *number of samples*
q^k = *regularization factor*
R = *real space*
R_{act} = *actual risk*
R_{emp} = *empirical risk*
U = *basic function*
V_k = *variation coefficient of permeability*
x = *variable*
y = *matching function*
α_i' = *coefficient to be defined*
α_i'' = *coefficient to be defined*
Φ = *fiducial range*
ω = *a generalized parameter*
Ψ = *any function*
β_j = *parameter to be defined*
Ω = *regularized function*
ε = *error accuracy*

Subscripts and superscripts

i = *sample index*
j = *sample index*
k = *function item index*
l = *function item index*

References

1. Jones Jr., R. S., Pope, G. A., and Lake, L. W.: *"A Predictive Model for Water and Polymer Flooding"*, paper SPE 12653 presented at the SPE/DOE Fourth Symposium on Enhanced Oil Recovery, Tulsa, OK, April 15-18 1984.
2. Aydelotte, S. R., and Pope, G. A.: *"A Simplified Predictive Model for Steamdrive Performance"*, J. Pet. Tech. 1983, 35(5): 991-1002.
3. Vapnik, V.: *"The Nature of Statistical Learning Theory"*, Springer –Verlay, New York, 1995.
4. Cortes, C., and Vapnik, V.: *"Support-vector Networks"*, Machine Learning. 1995, 20(3): 273-297.
5. Ivanov, V. V.: *"On Linear Problems Which Are Not Well-posed"*, Soviet Math. Docl. 1962, 3(4): 981-983.
6. Phillips, D. Z.: *"A Technique for Numerical Solution of Certain Integral Equation of the First Kind"*, J. Assoc. Compute Math. 1962, 9(1): 84-96.
7. Hou Jian.: *"A Basic Study on Potentiality Analysis of Enhanced Oil Recovery"*, Ph.D. dissertation, China University of Petroleum, 2002.
8. Liu, D. H., and Chai, E. F.: *"Analyzing Reservoir Performance Using Group Method of Data Handling"*, Acta Petrolei Sinica China. 1994, 15(2): 58-63.
9. Boomer, R. J.: *"Predicting Production Using a Neural Network (Artificial Intelligence Beats Human Intelligence)"*, paper SPE 30202 presented at the SPE Petroleum Computer Conference, Houston, Texas, June 11-14 1995.

Table 1. Parameters for orthogonal designing

The number of designings	1	2	3	4	5
B initial oil saturation (%)	55	60	65	70	75
C permeability($\times10^{-3}\mu m^2$)	100	750	1500	2000	3000
D variation coefficient	0.3	0.5	0.6	0.7	0.8
E viscosity of reservoir oil(mPa.s)	10	40	60	80	100
F temperature (□)	50	60	70	80	90
G salinity of reservoir water(mg/L)	2000	5000	10000	20000	30000
H concentrations of Ca^{2+}、Mg^{2+} (mg/L)	50	100	150	200	300
I water-cut at the beginning of polymer-flooding (%)	50	70	80	90	96

Table 2. Quadrature design table

No.	A 1	B 2	C 3	D 4	E 5	F 6	G 7	H 8	I 9	10	11	Enhanced oil recovery (%)	Polymer utilization rate (t/t)	Internal rate of return (%)
1	1	1	1	1	1	1	1	1	1	1	1	7.9	83.2	14.9
2	1	2	2	2	2	2	2	2	2	2	2	9.5	108.5	21.9
3	1	3	3	3	3	3	3	3	3	3	3	8.5	105.2	21.0
4	1	4	4	4	4	4	4	4	4	4	4	5.1	67.7	10.5
5	1	5	5	5	5	5	5	5	5	5	5	3.2	45.8	4.4
6	2	1	2	3	4	5	1	2	3	4	5	7.8	81.4	14.6
7	2	2	3	4	5	1	2	3	4	5	1	7.3	82.9	15.0
8	2	3	4	5	1	2	3	4	5	1	2	5.5	68.0	10.8
9	2	4	5	1	2	3	4	5	1	2	3	5.9	79.3	13.9
10	2	5	1	2	3	4	5	1	2	3	4	5.7	81.0	14.4
11	3	1	3	5	2	4	4	1	3	5	2	6.5	68.0	10.8
12	3	2	4	1	3	5	5	2	4	1	3	4.7	54.2	6.9
13	3	3	5	2	4	1	1	3	5	2	4	8.9	110.4	22.9
14	3	4	1	3	5	2	2	4	1	3	5	6.7	89.3	16.9
15	3	5	2	4	1	3	3	5	2	4	1	5.7	81.2	14.6
16	4	1	4	2	5	3	5	3	1	4	2	4.1	42.6	3.6
17	4	2	5	3	1	4	1	4	2	5	3	7.0	79.9	14.3
18	4	3	1	4	2	5	2	5	3	1	4	7.5	92.5	17.9
19	4	4	2	5	3	1	3	1	4	2	5	9.3	123.4	26.8
20	4	5	3	1	4	2	4	2	5	3	1	5.3	76.0	13.2
21	5	1	5	4	3	2	4	3	2	1	5	7.4	77.8	13.7
22	5	2	1	5	4	3	5	4	3	2	1	4.3	48.8	5.4
23	5	3	2	1	5	4	1	5	4	3	2	5.9	72.8	12.3
24	5	4	3	2	1	5	2	1	5	4	3	6.0	79.9	14.3
25	5	5	4	3	2	1	3	2	1	5	4	9.3	133.2	29.7
26	6	1	1	4	5	4	3	2	5	2	3	5.3	55.5	7.3
27	6	2	2	5	1	5	4	3	1	3	4	4.2	48.5	5.3
28	6	3	3	1	2	1	5	4	2	4	5	5.5	67.6	10.8
29	6	4	4	2	3	2	1	5	3	5	1	10.3	137.0	30.9
30	6	5	5	3	4	3	2	1	4	1	2	8.3	118.3	25.5
31	7	1	2	1	3	3	2	4	5	5	4	7.9	82.3	15.1
32	7	2	3	2	4	4	3	5	1	1	5	5.6	64.1	9.8
33	7	3	4	3	5	5	4	1	2	2	1	4.8	59.7	8.6
34	7	4	5	4	1	1	5	2	3	3	2	4.9	65.0	10.1
35	7	5	1	5	2	2	1	3	4	4	3	10.1	143.9	33.0
36	8	1	3	3	1	2	5	5	4	2	4	3.8	40.1	2.9
37	8	2	4	4	2	3	1	1	5	3	5	11.0	126.0	27.8
38	8	3	5	5	3	4	2	2	1	4	1	9.2	114.0	24.3
39	8	4	1	1	4	5	3	3	2	5	2	5.3	70.5	11.7
40	8	5	2	2	5	1	4	4	3	1	3	5.1	72.6	12.3
41	9	1	4	5	4	1	2	5	2	3	3	7.4	78.0	13.9
42	9	2	5	1	5	2	3	1	3	4	4	6.0	68.6	11.2
43	9	3	1	2	1	3	4	2	4	5	5	4.8	59.7	8.6
44	9	4	2	3	2	4	5	3	5	1	1	5.0	67.3	10.8
45	9	5	3	4	3	5	1	4	1	2	2	9.6	137.3	31.2
46	10	1	5	2	2	5	3	4	4	3	1	6.4	66.6	10.6
47	10	2	1	3	3	1	4	5	5	4	2	6.6	75.0	13.1
48	10	3	2	4	4	2	5	1	1	5	3	5.4	66.7	10.7
49	10	4	3	5	5	3	1	2	2	1	4	8.0	106.2	22.2
50	10	5	4	1	1	4	2	3	3	2	5	5.8	83.3	15.5

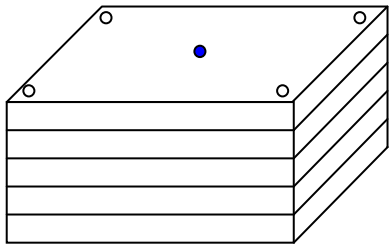
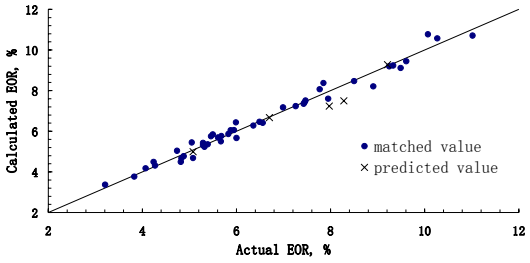
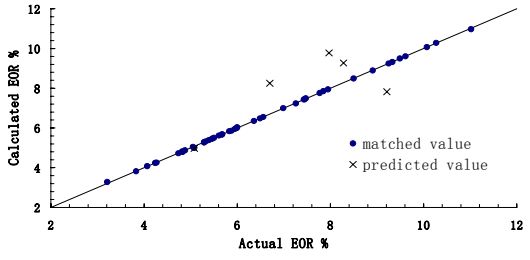


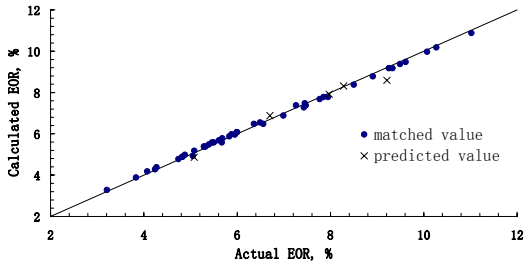
Figure 1. The reservoir simulation model.



(a) GMDH

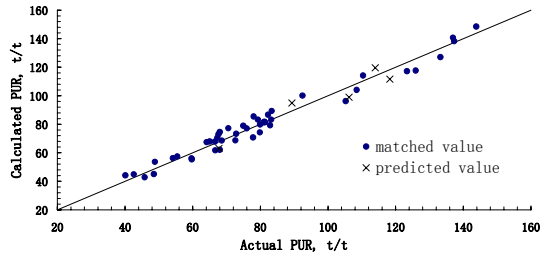


(b) Improved BP

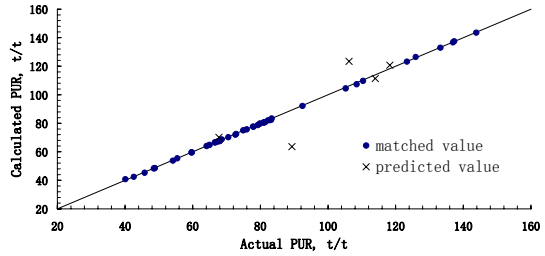


(c) SVM

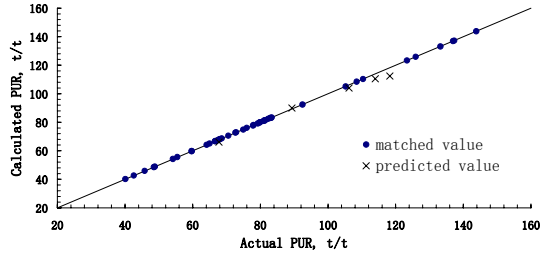
Figure 2. Comparison of enhanced oil recoveries in different methods.



(a) GMDH

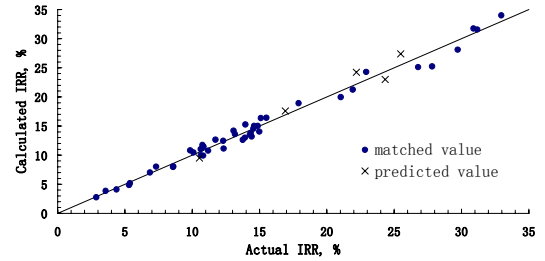


(b) Improved BP

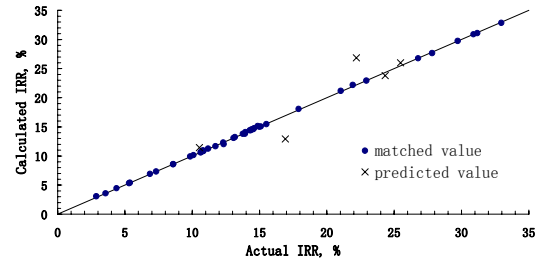


(c) SVM

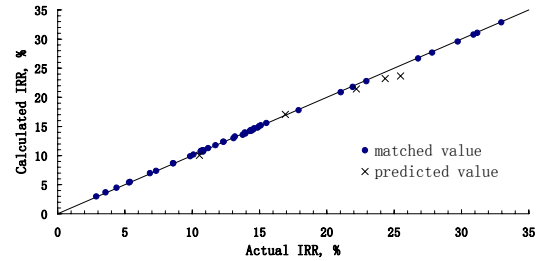
Figure 3. Comparison of polymer utility rates in different methods.



(a) GMDH



(b) Improved BP



(c) SVM

Figure 4. Comparison of internal rates of return in different methods.