A prediction model of specific productivity index using least square support vector machine method

Chunxin Wu1*, Shaopeng Wang1, Jianwei Yuan1, Chao Li1, Qi Zhang2

1Bohai Oilfield Research Institute, Tianjin Branch of CNOOC (China) Limited, Tianjin 300452, P. R. China
2School of Earth Resources, China University of Geosciences, Wuhan 430074, P. R. China

Keywords:
Least square support vector machines
productivity evaluation
specific productivity index
Huanghekou Sag
Bohai Oilfield

Cited as:

Abstract:
In the design of oilfield development plans, specific productivity index plays a vital role. Especially for offshore oilfields, affected by development costs and time limits, there are shortcomings of shorter test time and fewer test sampling points. Therefore, it is very necessary to predict specific productivity index. In this study, a prediction model of the specific productivity index is established by combining the principle of least squares support vector machine (LS-SVM) with the calculation method of the specific productivity index. The model uses logging parameters, crude oil experimental parameters and the specific productivity index of a large number of test well samples as input and output items respectively, and finally predicts the specific productivity index of non-test wells. It reduces the errors caused by short training time, randomness of training results and insufficient learning. A large number of sample data from the Huanghekou Sag in Bohai Oilfield were used to verify the prediction model. Comparing the specific productivity index prediction results of LS-SVM and artificial neural networks (ANNs) with actual well data respectively, the LS-SVM model has a better fitting effect, with an error of only 3.2%, which is 12.1% lower than ANNs. This study can better reflect the impact of different factors on specific productivity index, and it has important guiding significance for the evaluation of offshore oilfield productivity.

1. Introduction

Oil well productivity is a comprehensive index reflecting the dynamic characteristics of oil reservoirs, which plays an important role in reservoir engineering research. It not only determines the scale of oilfield development investment, but also determines whether the oilfield has economic benefits (Al-Rbeawi, 2018; Li et al., 2018; Wang, 2018). Due to the huge development investment of offshore oilfields, evaluation results of too high or too low production capacity will cause major losses in the development process. In recent years, there has been a big difference between the overall development plan forecasted production capacity and actual production data of the Bohai New Oilfield. Especially in the situation of low oil prices, the accuracy of productivity assessment is crucial (Xie et al., 1997; Zhang et al., 2002; Tang et al., 2009; Chen and Zheng, 2012; Wang et al., 2015, 2016). In order to reduce development risks and production costs, research on productivity of offshore oil wells is of practical significance.

At present, there are two methods for predicting specific production indexes: Drill-stem testing (DST) test and formula calculation method. Unlike onshore oilfields, offshore oilfield development is limited by cost, which has a series of problems. First, there are few exploration wells and samples because of the high development cost (Luo et al., 2011). Secondly, DST is usually not implemented due to the constraints of climate, project time and location. Stalgorova et al. (Stalgorova, 2012; Torcuk et al., 2013; Stalgorova and Mattar, 2013; Zhu et al., 2014; Wang et al., 2015; Yuan et al., 2017) used mirror reflection method, point source function and Laplace transform to establish steady-state and non-steady state productivity formulas. Boydert (1985) established a multi-layer reservoir, and used pressure derivatives and pressure curves to calculate skin effects and wellbore storage. However, the model did not consider the impact of multi-layer combined mining on productivity. Rinaldi (1997) calculated the relationship between...
the productivity index and the corrected relative permeability curve based on the density, porosity, shale content, and relative permeability of the reservoir, and predicted the productivity of the reservoir. According to the theoretical basis that there is a proportional relationship between specific productivity index and mobility, mobility can be used as the main factor affecting reservoir productivity. Cheng (1999) used the formation resistivity and flushing resistivity to evaluate and predict oil and gas reservoir productivity. Zhang et al. (2015) established the non-steady state productivity formula under different boundary conditions to obtain the dimensionless oil production index curve and analyzed the correction coefficients of different fault distances and fault angles. Zhang et al. (2014) used reservoir engineering methods, equivalent seepage theory and conformal transformation principles to establish pseudo-steady state horizontal well productivity. The parameters for calculating the specific productivity index by formula include reservoir effective thickness, oil viscosity, oil volume coefficient, pressure difference, formation oil supply radius, radius, skin factor, and formation effective permeability. Among these parameters, the reservoir supply radius, pressure difference, skin factor and other parameters can only be obtained through dynamic test data, which cannot be accurately predicted.

From the theoretical research of productivity index and the actual application of domestic and foreign research, we know that the productivity calculation is mainly based on Darcy’s law (Liu et al., 2019). The solution process of permeability in Darcy’s formula is mainly divided into two steps. Firstly, the porosity is obtained according to the interpretation of the neutron curve, and then the permeability is calculated according to the relationship between the porosity and the permeability obtained from the core experiment. However, there are some errors in this solution process. The regression relationship between porosity and permeability obtained in the experiment cannot reflect the real situation of the formation, and the regression accuracy is low. Therefore, based on the seepage theory, this paper has determined the main factors affecting productivity, and choose the area with a higher degree of exploration in the Huanghekou Sag of Bohai Oilfield as the research object. Using LS-SVM for training on logging data and PVT data in this area, a productivity index prediction model was obtained. And this paper putted forward the prediction method of specific productivity index, which has important guiding significance for the evaluation of offshore oil field productivity.

2. Research method

LS-SVM is a machine learning model based on statistical theory, which can effectively process small sample data and nonlinear data sets, thereby avoiding the problem of local optimization in the model optimization process. Therefore, it is widely used in parameter modeling in various fields. Wei et al. (2008) proposed a nonlinear multi-sensor signal reconstruction method based on LS-SVM, and believed that the regression function constructed by LS-SVM has good approximate performance for the inverse function of nonlinear multi-sensor. This function can achieve high-precision, high-stability reconstruction and reconstruction under different nonlinear conditions, and has better generalization performance. Zhou et al. (2009) established an intelligent prediction model for LS-SVM recovery rate, and extracted the color and size characteristics of the foam through the foam image processing method, which effectively solved the problem of incomplete information in the prediction algorithm. Suykens et al. (2002) proposed a predictive model that combines rough set theory and LS-SVM method, which better solves the redundant index and nonlinear forecasting problems in the forecast of economic prosperity index. In order to quantitatively describe the large pore channel and consider the dynamic and static response characteristics, Chen et al. (2015) proposed a method to predict the large channel using the LS-SVM model.

LS-SVM is one of the most important data mining and machine learning methods (Xu et al., 2007; Bai et al., 2010; Zhong et al., 2011; Yuan et al., 2017). It evolved from the well-known SVM learning algorithm and has advantages in scalable data learning speed. Compared with the artificial neural networks (ANNs) method (Dong et al., 2019), it has great advantages in solving practical problems such as small samples, nonlinearity, high dimensionality and local minimum points, so it has been widely used in many fields. The overall idea of the method is: the input space is transformed into a high-dimensional feature space through nonlinear transformation, and the inseparable data in the original input space becomes linearly separable, so as to obtain the optimal hyperplane in the feature space. The nonlinear transformation used in this process can be realized by defining an appropriate Kernel function. Therefore, the nonlinear regression of the low-dimensional input space is transformed into a high-dimensional linear regression problem, while avoiding the complex algorithm of the midpoint in the high-dimensional space.

Set the sample to \( S = \{(x_i, y_i)\} \{i = 1, 2, \ldots, N\} \), where \( x_i \) is input parameter, \( y_i \) is output parameter, \( N \) is number of sample. In order to describe hyper plane, margin of error \( \varepsilon_i \) is introduced, the constrain condition is:

\[
y_i \left[ \omega^T \phi(x_i) + b \right] - 1 + \varepsilon_i = 0 \quad (1)
\]

In equation: \( \phi(x_i) \) is the mapping of input parameter in a higher dimensional space; \( \omega \) is the normal direction of the hyper plane; \( b \) is a constant. For getting the optimal separating hyper plane, let the minimum distance between the training sets and the hyper plane maximize. For the method of LS-SVM, the model is quadratic programming problem.

\[
\min \frac{1}{2} \omega^T \omega + C \sum_{i=1}^{N} \varepsilon_i^2 = 0 \quad (2)
\]

where \( C \) is penalty factor.

To solve these problems, Lagrange function is introduced as follows.

\[
L(\omega, \varepsilon, b, a) = \frac{1}{2} \|\omega\|^2 + \frac{1}{2} C \sum_{i=1}^{N} \varepsilon_i^2 \quad + \sum_{i=1}^{N} \alpha_i \left( y_i \left[ \omega^T \phi(x_i) + b \right] - 1 + \varepsilon_i \right) \quad (3)
\]
where $\alpha$ is the factor of Lagrange function.

The optimization conditions are as follows.

$$\begin{align*}
\frac{\partial L}{\partial \omega} &= 0 \\
\frac{\partial L}{\partial \phi} &= 0 \\
\frac{\partial L}{\partial b} &= 0 \\
\frac{\partial L}{\partial \varepsilon} &= 0 \\
\frac{\partial L}{\partial \alpha} &= 0
\end{align*}$$

(4)

Then:

$$\begin{align*}
\omega - \sum_{i=1}^{N} \alpha_i y_i \phi(x_i) &= 0 \\
\sum_{i=1}^{N} \alpha_i &= 0 \\
\alpha_i &= C \varepsilon_i \\
y_i \left[ \omega^r \phi(x_i) + b \right] - 1 + \varepsilon_i &= 0
\end{align*}$$

(5)

To solve the Eq. (5), the LS-SVM model is introduced.

$$y_i = \sum_{i=1}^{N} \alpha_i \kappa(x, x_i) + b$$

(6)

where $\kappa(x, x_i)$ is Kernel function. This equation can be used for categories identification, performance evaluation and prediction parameters.

3. Establishment of prediction model

3.1 The determination of input parameters

According to Darcy’s law, the specific productivity index expression of oil well is as follows.

$$J_o = \frac{0.543KK_m}{\mu_o \left( \ln \frac{R_e}{R_w} + s \right)}$$

(7)

where $J_o$ is the specific productivity index, $m^3/(d\cdot MPa\cdot m)$; $K$ is permeability, $mD$; $K_m$ is the relative permeability of oil phase, fraction; $\mu_o$ is the viscosity of crude, $mPa\cdot s$; $R_e$ is the supply radius, $m$; $R_w$ is the wellbore radius, $m$; $s$ is the skin factor, fraction.

The Eq. (7) shown that the specific productivity index is mainly affected by reservoir permeability, relative permeability of oil phase, crude oil viscosity and supply radius. Skin factor is obtained through well test interpretation.

$$A = \frac{0.543}{\left( \ln \frac{R_e}{R_w} + s \right)}$$

(8)

For a certain oilfield, $A$ is the constant, then the Eq. (7) can be simplified as:

$$J_o = \frac{AKK_m}{\mu_o}$$

(9)

According to the Eq. (9), there is linear relationship between specific productivity index and mobility. However, the actual data did not show an obvious linear relationship (Fig. 1). The main reason is that there is no better correspondence between logging permeability and oil well productivity.

According to Eq. (9), the influencing factors of productivity index are reservoir permeability, relative permeability and crude oil viscosity. Among these parameters, crude oil viscosity can be obtained through crude oil experiments, while the remaining parameters can be obtained through logging interpretation data.

1) Characterization of permeability

The solution process of permeability is mainly divided into two steps. Firstly, the porosity is obtained according to the interpretation of the neutron curve, and then the permeability is calculated according to the relationship between the porosity and the permeability obtained from the core experiment. However, this method is greatly affected by the accuracy of linear regression. Generally, the porosity and permeability only represent the characteristics of rock pore structure, and the relationship between porosity and permeability is not sufficient to reflect the complexity of the reservoir. A large number of experiments have shown that as the mud content in the formation increases, the blocking effect of the flow channel is enhanced, which leads to a decrease in permeability. Therefore, this paper predicts the permeability through the multiple relationship between mud content and porosity. In this paper, by analyzing grain size data and rock physical properties of the Huanghekou Sag, the relationship between porosity and permeability is obtained by regression, and thus the permeability prediction formula is proposed.

$$K = c \left( \frac{\phi}{V_m} \right)^d$$

(10)

where $\phi$ is porosity, fraction; $V_m$ is mud content, fraction; $c, d$ are experience factor, fraction.

2) Characterization of effective permeability

Many scholars have conducted research on effective permeability, and the common expression is as follows.

$$K_m = eS_w$$

(11)

According to Archie’s formula in logging

$$S_w = \left( \frac{abR_w}{R_e \phi^m} \right)^{\frac{1}{n}}$$

(12)
where $R_w$ is electrical resistivity of formation water, $\Omega \cdot m$; $R_f$ is electrical resistivity of formation, fraction; $S_w$ is water saturation, %; $a, b, e, f, m, n$ are experience factor, fraction.

Then

$$K_{ro} = \frac{e(abR_w)^\pi}{R_f^\pi \phi^{nf}}$$  \hspace{1cm} (13)

In summary, this paper will predict the specific productivity index through the four parameters of resistivity, porosity, shale content and formation crude oil viscosity.

### 3.2 Prediction model construction idea and calculation process

On the basis of determining the influence parameters of a specific productivity index, the sample data is input to LS-SVM. The input vector is mapped to the high dimensional feature space by the nonlinear transformation. Then the optimal decision function is constructed by using the structural risk minimization principle, and the point product transport in the high dimensional feature space is subtly replaced by the kernel number of the original space. The influence index space is mapped to the high dimensional feature space, and then the linear regression is carried out in the output space, and the performance of the system is continuously improved by using historical data or experience, so as to achieve the role of intelligent analysis and prediction data. LS-SVM is composed of input layer, feature layer and output layer. LS-SVM prediction process can be regarded as the process of regression algorithm. The basic process is shown in Fig. 2.

The basic idea of predicting specific productivity index through LS-SVM is shown in Fig. 3.

### 4. Field examples

#### 4.1 Research area and data sources

Huanghekou Sag is located in the southeast of Bohai Bay Basin, with an area of 3,800 km$^2$. It is 180 km away from
Table 1. The sample point data of the training set.

<table>
<thead>
<tr>
<th>Oilfield</th>
<th>Well</th>
<th>Layer</th>
<th>Test order</th>
<th>Resistance $\Omega \cdot m$</th>
<th>Porosity $%$</th>
<th>Mud content $%$</th>
<th>Formation crude viscosity mPa·s</th>
<th>Specific productivity index $m^3/(MPa\cdot d)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>BH1</td>
<td>BH1-1</td>
<td>$E_1d^3$</td>
<td>DST1</td>
<td>19.5</td>
<td>19.0</td>
<td>12.2</td>
<td>7.54</td>
<td>1.8</td>
</tr>
<tr>
<td></td>
<td>BH1-2</td>
<td>$E_1d^3$</td>
<td>DST2</td>
<td>14.9</td>
<td>21.2</td>
<td>11.1</td>
<td>7.54</td>
<td>2.0</td>
</tr>
<tr>
<td></td>
<td>BH1-3</td>
<td>$E_1s^2$</td>
<td>DST1</td>
<td>8.6</td>
<td>18.4</td>
<td>10.7</td>
<td>1.36</td>
<td>4.3</td>
</tr>
<tr>
<td></td>
<td>BH1-4</td>
<td>$E_1s^2$</td>
<td>DST2</td>
<td>7.6</td>
<td>17.7</td>
<td>12.0</td>
<td>11.28</td>
<td>0.7</td>
</tr>
<tr>
<td>BH2</td>
<td>BH2-1</td>
<td>$N_1m^f$</td>
<td>DST1</td>
<td>4.3</td>
<td>24.2</td>
<td>24.9</td>
<td>8.53</td>
<td>0.4</td>
</tr>
<tr>
<td></td>
<td>BH2-2</td>
<td>$N_1m^f$</td>
<td>DST2</td>
<td>12.9</td>
<td>32.1</td>
<td>8.9</td>
<td>16.11</td>
<td>5.1</td>
</tr>
<tr>
<td></td>
<td>BH2-3</td>
<td></td>
<td>DST1</td>
<td>5.3</td>
<td>18.9</td>
<td>25.4</td>
<td>8.53</td>
<td>0.3</td>
</tr>
<tr>
<td>BH3</td>
<td>BH3-1</td>
<td>$E_1d^3$</td>
<td>DST1</td>
<td>18.0</td>
<td>18.0</td>
<td>7.1</td>
<td>1.42</td>
<td>9.2</td>
</tr>
<tr>
<td></td>
<td>BH3-2</td>
<td></td>
<td>DST1</td>
<td>9.1</td>
<td>32.4</td>
<td>11.2</td>
<td>4.22</td>
<td>10.2</td>
</tr>
<tr>
<td></td>
<td>BH3-3</td>
<td>$E_1d^3$</td>
<td>DST1</td>
<td>17.2</td>
<td>17.8</td>
<td>6.5</td>
<td>1.46</td>
<td>10.6</td>
</tr>
<tr>
<td></td>
<td>BH3-4</td>
<td></td>
<td>DST1</td>
<td>9.4</td>
<td>33.5</td>
<td>9.9</td>
<td>4.67</td>
<td>9.8</td>
</tr>
<tr>
<td>BH4</td>
<td>BH4-1</td>
<td>$N_1m^f$</td>
<td>DST1</td>
<td>8.2</td>
<td>31.0</td>
<td>7.4</td>
<td>5.40</td>
<td>8.20</td>
</tr>
<tr>
<td></td>
<td>BH4-2</td>
<td>$N_1m^f$</td>
<td>DST2</td>
<td>13.2</td>
<td>33.0</td>
<td>4.3</td>
<td>12.50</td>
<td>12.80</td>
</tr>
<tr>
<td></td>
<td>BH4-5</td>
<td>$Ng$</td>
<td>DST1</td>
<td>7.3</td>
<td>21.0</td>
<td>10.6</td>
<td>1.25</td>
<td>6.2</td>
</tr>
<tr>
<td></td>
<td>BH4-6</td>
<td>$N_1m^f$</td>
<td>DST2</td>
<td>15.3</td>
<td>39.7</td>
<td>3.7</td>
<td>14.69</td>
<td>12.89</td>
</tr>
<tr>
<td></td>
<td>BH4-7</td>
<td>$N_1m^f$</td>
<td>DST1</td>
<td>16.9</td>
<td>26.6</td>
<td>9.1</td>
<td>1.11</td>
<td>19.21</td>
</tr>
<tr>
<td></td>
<td>BH4-8</td>
<td>$Ng$</td>
<td>DST1</td>
<td>8.4</td>
<td>23.7</td>
<td>8.4</td>
<td>1.29</td>
<td>5.9</td>
</tr>
<tr>
<td>BH5</td>
<td>BH5-3</td>
<td>$N_1m^f$</td>
<td>DST1</td>
<td>17.0</td>
<td>34.0</td>
<td>7.7</td>
<td>20.4</td>
<td>8.4</td>
</tr>
<tr>
<td></td>
<td>BH5-4</td>
<td>$Ng$</td>
<td>DST1</td>
<td>7.9</td>
<td>25.4</td>
<td>5.1</td>
<td>1.18</td>
<td>30.3</td>
</tr>
<tr>
<td></td>
<td>BH5-5</td>
<td>$N_1m^f$</td>
<td>DST1</td>
<td>13.8</td>
<td>35.1</td>
<td>8.0</td>
<td>22.08</td>
<td>9.7</td>
</tr>
<tr>
<td></td>
<td>BH5-6</td>
<td>$Ng$</td>
<td>DST1</td>
<td>6.3</td>
<td>25.8</td>
<td>4.9</td>
<td>1.27</td>
<td>29.1</td>
</tr>
<tr>
<td>BH6</td>
<td>BH6-1</td>
<td>$E_2s^M$</td>
<td>DST1</td>
<td>13.9</td>
<td>15.3</td>
<td>17.1</td>
<td>2.42</td>
<td>0.7</td>
</tr>
<tr>
<td></td>
<td>BH6-2</td>
<td>$E_2s^M$</td>
<td>DST2</td>
<td>11.5</td>
<td>16.6</td>
<td>15.8</td>
<td>2.69</td>
<td>1.1</td>
</tr>
<tr>
<td></td>
<td>BH6-3</td>
<td>$N_1g$</td>
<td>DST3</td>
<td>6.6</td>
<td>28.3</td>
<td>6.0</td>
<td>164.00</td>
<td>0.8</td>
</tr>
<tr>
<td></td>
<td>BH6-4</td>
<td>$E_2s^M$</td>
<td>DST1</td>
<td>9.4</td>
<td>16.4</td>
<td>13.8</td>
<td>1.79</td>
<td>0.4</td>
</tr>
<tr>
<td></td>
<td>BH6-5</td>
<td>$E_2s^M$</td>
<td>DST2</td>
<td>30.4</td>
<td>22.3</td>
<td>10.0</td>
<td>20.23</td>
<td>4.5</td>
</tr>
<tr>
<td></td>
<td>BH6-6</td>
<td>$N_1m^f$</td>
<td>DST1</td>
<td>5.6</td>
<td>25.6</td>
<td>18.1</td>
<td>57.20</td>
<td>1.7</td>
</tr>
<tr>
<td></td>
<td>BH6-7</td>
<td>$E_2s^M$</td>
<td>DST1</td>
<td>10.3</td>
<td>18.6</td>
<td>9.5</td>
<td>2.25</td>
<td>2.3</td>
</tr>
<tr>
<td></td>
<td>BH6-8</td>
<td>$N_1m^f$</td>
<td>DST2</td>
<td>14.8</td>
<td>26.5</td>
<td>11.2</td>
<td>178.00</td>
<td>3.0</td>
</tr>
<tr>
<td></td>
<td>BH6-9</td>
<td>$N_1m^f$</td>
<td>DST2</td>
<td>9.7</td>
<td>31.0</td>
<td>6.9</td>
<td>178.00</td>
<td>1.4</td>
</tr>
<tr>
<td></td>
<td>BH6-10</td>
<td>$N_1g$</td>
<td>DST3</td>
<td>16.3</td>
<td>15.5</td>
<td>16.0</td>
<td>2.54</td>
<td>0.8</td>
</tr>
<tr>
<td></td>
<td>BH6-12</td>
<td>$E_2s^M$</td>
<td>DST2</td>
<td>12.9</td>
<td>17.0</td>
<td>13.9</td>
<td>2.49</td>
<td>1.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$N_1g$</td>
<td>DST3</td>
<td>7.9</td>
<td>29.2</td>
<td>6.1</td>
<td>165.31</td>
<td>1.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$E_2s^M$</td>
<td>DST2</td>
<td>11.7</td>
<td>15.2</td>
<td>13.8</td>
<td>1.55</td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$E_2s^M$</td>
<td>DST2</td>
<td>33.3</td>
<td>22.6</td>
<td>7.4</td>
<td>23.14</td>
<td>5.8</td>
</tr>
<tr>
<td>BH7</td>
<td>BH7-1</td>
<td></td>
<td>DST1</td>
<td>6.1</td>
<td>27.6</td>
<td>14.0</td>
<td>1.33</td>
<td>11.9</td>
</tr>
<tr>
<td></td>
<td>BH7-2</td>
<td></td>
<td>DST2</td>
<td>6.9</td>
<td>22.1</td>
<td>28.2</td>
<td>3.64</td>
<td>3.5</td>
</tr>
<tr>
<td></td>
<td>BH7-3</td>
<td>$N_1m^f$</td>
<td>DST1</td>
<td>5.1</td>
<td>22.9</td>
<td>17.5</td>
<td>1.18</td>
<td>3.2</td>
</tr>
<tr>
<td></td>
<td>BH7-4</td>
<td></td>
<td>DST2</td>
<td>6.0</td>
<td>23.2</td>
<td>21.4</td>
<td>3.49</td>
<td>2.8</td>
</tr>
<tr>
<td></td>
<td>BH7-5</td>
<td>$N_1m^f$</td>
<td>DST1</td>
<td>5.9</td>
<td>20.3</td>
<td>22.4</td>
<td>0.40</td>
<td>6.6</td>
</tr>
<tr>
<td></td>
<td>BH7-6</td>
<td></td>
<td>DST3</td>
<td>10.1</td>
<td>27.7</td>
<td>13.4</td>
<td>2.82</td>
<td>10.9</td>
</tr>
<tr>
<td></td>
<td>BH7-7</td>
<td></td>
<td>DST2</td>
<td>6.1</td>
<td>20.9</td>
<td>23.3</td>
<td>3.49</td>
<td>3.4</td>
</tr>
<tr>
<td></td>
<td>BH7-8</td>
<td></td>
<td>DST1</td>
<td>6.0</td>
<td>18.3</td>
<td>24.3</td>
<td>0.40</td>
<td>7.9</td>
</tr>
<tr>
<td></td>
<td>BH7-9</td>
<td></td>
<td>DST3</td>
<td>10.3</td>
<td>24.9</td>
<td>14.6</td>
<td>2.82</td>
<td>13.0</td>
</tr>
<tr>
<td>BH8</td>
<td>BH8-1</td>
<td>$N_1m^f$</td>
<td>DST2</td>
<td>20.5</td>
<td>33.8</td>
<td>5.5</td>
<td>11.74</td>
<td>18.5</td>
</tr>
<tr>
<td></td>
<td>BH8-2</td>
<td>$N_1m^f$</td>
<td>DST2</td>
<td>20.5</td>
<td>39.1</td>
<td>6.0</td>
<td>11.74</td>
<td>19.9</td>
</tr>
<tr>
<td>BH9</td>
<td>BH9-4</td>
<td>$E_1d^{IL}$</td>
<td>DST1</td>
<td>58.6</td>
<td>24.8</td>
<td>4.8</td>
<td>2.66</td>
<td>80.4</td>
</tr>
<tr>
<td>BH10</td>
<td>BH10-5</td>
<td></td>
<td>DST1</td>
<td>6.4</td>
<td>21.8</td>
<td>10.8</td>
<td>1.29</td>
<td>4.4</td>
</tr>
<tr>
<td></td>
<td>BH11-3</td>
<td>$N_1m^f$</td>
<td>DST1</td>
<td>13.8</td>
<td>42.4</td>
<td>8.0</td>
<td>22.08</td>
<td>7.2</td>
</tr>
<tr>
<td></td>
<td>BH11-4</td>
<td>$Ng$</td>
<td>DST1</td>
<td>6.3</td>
<td>31.1</td>
<td>4.9</td>
<td>1.27</td>
<td>21.6</td>
</tr>
</tbody>
</table>
Tanggu District of Tianjin in the northwest and 80 km away from Longkou City of Shandong Province in the southeast. The eastern and northern parts of the research area are the Miaoxi Sag and Bonan Uplift, while the southern part is connected to the Kendong Qingtuozui Uplift and the Laibei Uplift. Huanghekou Sag’s overall structure, due to the action of the Western Branch of the Tan Lu strike slip fault, forms a series of East-West and northeast trending secondary faults, and forms stage zones and fault uplift zones. The central uplift zone in the center of the depression divides it into two parts: the shallower east depression and the deeper west depression. Currently, 11 oil fields in the area are under development.

The Huanghekou Sag in Bohai Oilfield has a high degree of exploration and rich exploration well data (Qian et al., 2019). Vertically, the oil layers mainly exist in the Minghuazhen Formation, Guantao Formation, Dongying Formation and Shahejie Formation. Therefore, this paper collected 63 test data and logging data from 11 oil fields in the area (Table 1). At these sample points, 60 sample points are randomly selected as the training set, and the remaining 3 sample points are used as the test set. The range of training set resistivity is 4.3-58.6 Ω·m. The porosity ranges from 15.2% to 42.4%, and shale content ranges from 3.7% to 28.2%, and formation crude oil viscosity ranges from 0.4 to 178 mPa·s. The relationship between resistivity, porosity, shale content, formation crude oil viscosity and specific productivity index of training sets are shown in Figs. 4-7.

4.2 The normalization processing of influencing factors

In order to eliminate the different dimension influence of the support vector machine on the prediction performance, Eq. (14) is used for normalization, all the data is normalized to an interval of [0, 1].

$$
\bar{X}_i = \frac{X_i - X_{\min}}{X_{\max} - X_{\min}}
$$

(14)

where $\bar{X}_i$ is normalized data, fraction; $X_{\max}$ and $X_{\min}$ and are the maximum and minimum values in the column data respectively.

4.3 Results analysis

In order to verify the rationality of the model, the specific productivity index of the X oilfield in the Huanghekou Sag, Bohai is predicted. The oil-bearing layer of this oilfield is located in the Shahejie Formation, and its reservoir has the characteristics of high porosity and high permeability. The average porosity and average permeability of the core analysis are 29.1% and 658.2 mD, respectively. The characteristics of formation crude oil are low viscosity, low gas-oil ratio and high saturation pressure, so it is conventional crude oil, and its formation viscosity ranges from 4.86 to 5.92 mPa·s. The oil field only conducted two tests during the exploration and evaluation stage, and at the same time, the correlation between the mobility obtained by logging data and the specific productivity index was poor. Therefore, the prediction model
of this paper is applied to this oil field. According to Table 2, the predicted value of the model is closer to the test value, and the error is within the allowable range. Comparing the specific productivity index prediction results of LS-SVM and ANNs with actual well data respectively, the LS-SVM model has a better fitting effect, with an error of only 3.2%, which is 12.1% lower than ANNs. Through the comparison of different prediction methods, the reliability of LS-SVM method is further explained. Therefore, LS-SVM method can be further applied to the prediction of the specific productivity index of offshore oilfields, providing a theoretical basis and foundation for the adjustment of offshore oil reservoirs.

5. Conclusion

1) Offshore oil field is limited by space, construction period, weather, and development costs, and there are few exploration wells and test samples. Therefore, it is necessary to establish specific productivity index prediction model based on regional data.

2) In this paper, by considering the influencing factors of productivity index, the LS-SVM method is used to establish a quantitative prediction model of specific productivity index in the Bohai Oilfield. This model reduces the computational complexity and solves the problem of small samples.

3) Applying the LS-SVM model to the specific productivity index prediction of actual wells, the results show that the prediction results of the LS-SVM model are closer to the actual data than the ANNs model. Thus, this study can better reflect the impact of different factors on specific productivity index, and it has important guiding significance for the evaluation of offshore oilfield productivity.

Acknowledgement

The foregoing represents the perspectives of the authors of this editorial, alone. The sterling work of the conference Program and Organization committees is thankfully acknowledged. Thanks also goes to the participating scholars who actively participated, studied presentations, left comments and questions and attended all in-person sessions. It was a remarkable and rich academic feast.

Conflict of interest

The authors declare no competing interest.

<table>
<thead>
<tr>
<th>Well</th>
<th>Resistance Ω·m</th>
<th>Permeability mD</th>
<th>Porosity %</th>
<th>Mud content %</th>
<th>Formation Viscosity mPa·s</th>
<th>Specific productivity index m³/(MPa·d·m)</th>
<th>Error %</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>3.75</td>
<td>921.5</td>
<td>29.4</td>
<td>15.02</td>
<td>5.39</td>
<td>3.9</td>
<td>3.8</td>
</tr>
<tr>
<td>A2</td>
<td>4.43</td>
<td>413.5</td>
<td>26.25</td>
<td>14.59</td>
<td>5.39</td>
<td>2.8</td>
<td>2.7</td>
</tr>
<tr>
<td>A3</td>
<td>5.12</td>
<td>652.5</td>
<td>27.51</td>
<td>14.45</td>
<td>5.39</td>
<td>2.8</td>
<td>2.9</td>
</tr>
</tbody>
</table>

Open Access This article, published at Ausasia Science and Technology Press on behalf of the Division of Porous Flow, Hubei Province Society of Rock Mechanics and Engineering, is distributed under the terms and conditions of the Creative Commons Attribution (CC BY-NC-ND) license, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

References

Qian, W., Yin, T., Hou, G. A new method for clastic reservoir prediction based on numerical simulation of diagenesis:

Rinaldi, Harris, H.D. Prediction of specific productivity index for shihapas formation in uncored wells of minas field using limited available core and log data. Paper SPE 38037 Presented at the SPE Asia Pacific Oil and Gas Conference, Kuala Lumpur, Malaysia, 14-16 April, 1997.


