Research on support vector machine method for comprehensive evaluation after compression

Wang Chengwang¹, Chen Shuai¹, Chen Gaojie¹, Lu Han²*

¹PetroChina coalbed methane company limited, Beijing, 100020, China; ²School of science, southwest petroleum university, Chengdu, Sichuan, 610500, China.

Abstract. Hydraulic fracturing is one of the important measures to increase production of oil and gas reservoirs. Pressure after the assessment of the existing technology is usually divided into direct and indirect diagnosis of hydraulic fracture. The diversity of evaluation results is due to the uncertainty and immaturity of the technology itself. Up to now, we have not found an economical and accurate hydraulic fracture evaluation method in the development and application of fracturing technology. In this paper, the pressure treatment technology after the comprehensive evaluation of boundary conditions is studied. By introducing the support vector regression theory, an evaluation model and a solution for correcting fracturing parameters are proposed. Fracturing parameters include fracture length, fracture height, fracture width, integrated fracturing fluid leakage coefficient, fracture conductivity, fracture closure pressure, and so on. For the optimization of various parameters, objective and scientific comprehensive evaluation results can be obtained by selecting different kernel functions. The results show that the model and method based on support vector machine are effective and practical.

1 Introduction

With the development and application of fracturing techniques for tight gas reservoirs, there is an urgent need to evaluate methods and techniques for fracturing tight gas reservoirs. The purpose of post-pressure evaluation is to understand the extension of hydraulic fractures. The evaluation results are of great guiding significance for the deployment of well pattern, optimization of perforation, rational extraction and optimization of the construction scale, amount of liquid, amount of sand and concentration of coal seam fracturing. The existing post-fracturing techniques can be divided into two categories: direct diagnosis and indirect diagnosis. Direct diagnostic techniques include well temperature testing, radioactive tracer testing, micro seismic measurement, surface and bottom hole inclinometer, as well as downhole television testing, etc. Indirect diagnosis techniques generally include real-time monitoring and interpretation of fracturing, fracturing simulation, well test analysis of hydraulic fractures, fitting of post-pressure production history, and post-pressure drop analysis, etc. These technologies can be expensive, complex, or both. The evaluation results obtained by various methods are also quite different. How to get more accurate evaluation results under simple and feasible conditions is a matter that everyone in the well area wants to know.

In 2003, researchers such as He Decai et al. used single-well numerical simulation to predict the productivity of fractured Wells. The predicted productivity of fractured Wells can be used for economic benefit analysis of fractured Wells, as well as historical fitting of fractured Wells [1]. In 2011, a comprehensive evaluation model and solution method based on the uncertainty theory was proposed by Guo Dali et al. for relevant fracturing parameters (fracture length, fracture width, fracture height, comprehensive fluid loss coefficient, fracture diversion capacity, fracture closure pressure, etc.) [2]. In 2016, researchers such as Bian Xiaobing and others could qualitatively judge the embrittlement and plasticity of strata by counting the number of formation fractures, the average pressure drop and the average pressure drop during the pre-liquid injection stage. Comprehensive brittleness index can be calculated quantitatively according to the energy area in fracturing construction [3]. In 2018, Yang Xiantong et al. conducted technological and stimulation evaluation of fracturing effects on two core post-pressure evaluation technologies -- construction pressure analysis and numerical simulation technology [4].

This article will study pressure processing technology, comprehensive evaluation of boundary conditions by the support vector machine (SVM) theory is introduced to calculate the return of the fracturing parameters of the evaluation model and method, including fracture length, fracture height, fracture width, comprehensive fracturing fluid leak off coefficient, fracture diverting capacity, fracture closure pressure, etc., optimization of various parameters, choice of different kernel functions, in order to obtain objective and scientific comprehensive evaluation results.

*Corresponding author’s e-mail: luhan_1024@outlook.com
2 Establishment and solution of support vector regression model

Support vector machine was first proposed by Cortes et al in 1995. Its basic idea is that a nonlinear mapping \( \Phi \) is used to map data to a high-dimensional feature space, and then regression estimation function is constructed in the high-dimensional feature space and then mapped back to the original space. The nonlinear transformation is realized by defining appropriate kernel function. In addition, considering that some samples cannot be properly classified by hyperplane separation, we generally adopt relaxation variables to solve this problem, and then the optimization equation of support vector regression model can be expressed as follows:

\[
\begin{align*}
\max_{w,b} : & \quad \frac{1}{2} \|w\|^2 + C \sum_{i=1}^{N} (\xi_i^+ + \xi_i^-) \\
\text{s.t.} & : \quad y_i - f(X_i) \leq \varepsilon + \xi_i^+, \xi_i^- \geq 0 \\
& : \quad f(X) - y_i \leq \varepsilon + \xi_i^+, \xi_i^- \geq 0
\end{align*}
\]  
\tag{1}

Where \( \varepsilon \) is the super parameter that determines the width of the interval boundary. By introducing the Lagrangian multiplier \( \alpha, \alpha^+, \mu, \mu^+ \), the Lagrangian function and dual problem can be obtained:

\[
\begin{align*}
\max_{w,b} : & \quad \frac{1}{2} \|w\|^2 + C \sum_{i=1}^{N} (\xi_i^+ + \xi_i^-) \\
\text{s.t.} & : \quad y_i - f(X_i) \leq \varepsilon + \xi_i^+, \xi_i^- \geq 0 \\
& : \quad f(X) - y_i \leq \varepsilon + \xi_i^+, \xi_i^- \geq 0
\end{align*}
\]

The dual problem can be expressed as:

\[
\begin{align*}
\max_{\alpha} : & \quad \sum_{i=1}^{N} (\alpha_i^* - \alpha_i) - \varepsilon \sum_{i=1}^{N} \alpha_i^+ + \alpha_i^- \\
\text{s.t.} & : \quad \sum_{i=1}^{N} (\alpha_i^* - \alpha_i) = 0 \\
& : \quad 0 \leq \alpha_i^*, \alpha_i^- \leq C
\end{align*}
\]

\tag{3}

For solving the dual problem, we can obtain the function of SVR:

\[
f(X) = \sum_{i=1}^{m} (\alpha_i^* - \alpha_i) X_i^T X_j + b
\]

\tag{4}

SVM can obtain nonlinear regression results through kernel function. According to the relevant theory of functional, as long as a kind of kernel function \( K(X_i, X_j) \) satisfies mercer condition symmetric function can be regarded as kernel function.

The types of kernel functions are shown in the following table:

| category | expression | parameter |
|----------|------------|-----------|
| Linear   | \( K(X_i, X_j) = X_i^T X_j \) | none |
| polynomial | \( K(X_i, X_j) = (X_i^T X_j)^d \) | \( d \geq 1 \) is the degree of polynomial |
| Gaussian | \( K(X_i, X_j) = \exp \left(-\frac{\|X_i - X_j\|^2}{2\sigma^2}\right) \) | \( \sigma > 0 \) is the bandwidth of Gaussian kernel |
| Laplace  | \( K(X_i, X_j) = \exp \left(-\frac{\|X_i - X_j\|^2}{2\sigma}\right) \) | \( \sigma > 0 \) is the bandwidth of the Laplace kernel |
| Sigmoid  | \( K(X_i, X_j) = \tanh(\beta X_i^T X_j + \theta) \) | \( \text{Tanh is the hyperbolic tangent function, } \beta > 0, \theta < 0 \) |

For the above problems, Gaussian kernel function can be selected, because Gaussian kernel function only contains an unknown parameter \( \sigma \), which is easy to be optimized [5]. Gaussian kernel function can be expressed as:

\[
K(X_i, X_j) = \exp \left(-\frac{\|X_i - X_j\|^2}{2\sigma^2}\right)
\]

\tag{5}

in which \( \sigma > 0 \). At present, there are three unknown parameters that affect SVM learning performance, namely, the parameter \( \sigma \) in the kernel function, the parameter \( \varepsilon \) in the loss function and the regularized parameter \( C \) in the objective function. The performance of SVR largely depends on the selection of parameters, so Adam algorithm is adopted to optimize the parameters.

Final choice:

\[
C = 32, \sigma = 0.004, \varepsilon = 0.01
\]

Based on the above analysis, the support vector nonlinear model applicable to the comprehensive assessment of tight air pressure can be expressed as follows:

\[
f(X) = \sum_{i=1}^{m} (\alpha_i^* - \alpha_i) \exp \left(-\frac{\|X_i - X\|^2}{2\sigma^2}\right) + b
\]

\tag{6}

3 Application and analysis of the model

The support vector machine method is applied to 30 Wells in A certain block. After analyzing data, selecting factors, regression and classification, the model is trained and the training results are applied to another well A in the block. For reference and comparison, we simultaneously used six methods including support vector machine evaluation method for evaluation, and the evaluation results are shown in the following table:
Table 2: results and comparison of various evaluation methods and their comprehensive evaluation

| Evaluation methods                  | Seam width /m | Closure pressure /MPa | Fracture conductivity /D*cm | Comprehensive filtration loss coefficient / (m/min^0.5) |
|------------------------------------|---------------|-----------------------|-----------------------------|--------------------------------------------------------|
| Monitoring and interpretation      | 0.0032        | 44.5                  | 17.3                        | 0.00128                                                 |
| 3D fracturing simulation           | 0.0034        | **            | 22.1                        | 0.00074                                                 |
| Pressure drop after pressure analysis | 0.0038        | 43.6                  | **                          | 0.00166                                                 |
| Uncertainty assessment             | 0.0034        | 44.7                  | 16.4                        | 0.00114                                                 |
| Support vector evaluation          | 0.0033        | 44.8                  | 16.8                        | 0.00136                                                 |

Among them, the uncertainty method is the fuzzy comprehensive evaluation of many methods. This method avoids the boundary condition problem which is difficult to be solved, and integrates multiple evaluation values into one value by means of uncertainty theory. Support vector machine method is based on the evaluation of the object’s own impact factors for regression prediction. In addition, this method gets rid of the complicated evaluation process to some extent and is more simple and feasible than the traditional evaluation method.

4 Conclusion

Support vector machine (SVM) regression modeling is the mapping of low-dimensional nonlinear input to high-dimensional linear output. At the same time, the support vector machine algorithm and the corresponding statistical learning theory put forward the problem of small sample statistics, which provides a strong theoretical basis for solving the problem of machine learning and prediction in the case of multi-factor finite sample. The results of this paper show that the SVM method with Gaussian function as the core has high accuracy in the comprehensive evaluation of oil and gas well pressure.

Acknowledgment

This work was funded by the National Science and Technology Major Project of China (No. 2016ZX05042-003 & No. 2016ZX05065-003)

Reference

1. He Decai, Cao Shuyu, She Yuehui. Production capacity prediction and post-pressure evaluation method for oil-base fracturing in Qiquan oilfield [J]. Drilling and production technology, 2003, 26(1).
2. Guo Dali, Zhu kai, Chen Chaofeng, et al. Comprehensive assessment of fracturing pressure based on uncertainty theory [J]. Journal of southwest petroleum university (natural science edition), 2011, 33(4):173-176.
3. Bian Xiaobing, Jiang Yanxue, Jia Changgui, et al. A new method for post-pressure evaluation of shale gas Wells based on construction curve [J]. Natural gas industry, 2016, v.36; No. 268 (02): 67-72.
4. Yang Xiangtong, Teng Qi, Zhang Yang, et al. Post-pressure evaluation and productivity prediction of fractured tight sandstone gas reservoirs supported by integrated geological engineering [J]. China petroleum exploration, 2018.
5. Qi Hengnian. A review of support vector machines and their applications [J]. Computer engineering, 2004, 30(10):6-9.