CSCanada
**Advances in Petroleum Exploration and Development**
Vol. 10, No. 1, 2015, pp. 61-65
**DOI:**10.3968/7501

ISSN 1925-542X [Print]
ISSN 1925-5438 [Online]
www.cscanada.net
www.cscanada.org

# Application of Support Vector Machine in Friction Coefficient Prediction for Extended-Reach Well

LI Yun[a],*

[a] Technical equipment bureau, Shengli petroleum engineering Co., Ltd, Sinopec, Dongying, China.
*Corresponding author.

Received 21 August 2015; accepted 10 September 2015
Published online 28 September 2015

## Abstract

Torque and drag is a major problem in the drilling process for extended-reach well due to the great well depth and large displacement. The value of friction & torque is mainly determined by friction coefficient value, and there are many factors affecting the coefficient friction, reasonable and correct determination of the friction coefficient is an issue that must be addressed in the friction & torque analysis and prediction. On the basis of the friction coefficient calculation, the prediction model of friction coefficient for designing well was established based on support vector machine, the results show its prediction accuracy is over 90%, the limitation of using experiences to determine friction coefficient was broken down in the process of well designing.

**Key words:** Torque and drag; Extended-reach well; Support vector machine; Friction coefficient; Prediction

Li, Y. (2015). Research and application of the technology to increase viscosity and viscosity retention for polymer solution prepared with oilfield wastewater. *Advances in Petroleum Exploration and Development*, *10*(1), 61-65. Available from: URL: http://www.cscanada.net/index.php/aped/article/view/7501
DOI: http://dx.doi.org/10.3968/7501

## INTRODUCTION

With the development of drilling technology, the well's depth and horizontal and apace curvature become more and more large. So it often happens that the drag and torque overtakes the bearing capacity maximum of the installation and drilling string. During ERD (Extended-Reach Drilling) well planning, torque and drag analysis is recognized as an essential part of the risk management process[1]. Torque and drag is related to and affected by many things including the wellpath design, drill string design, hole size, drilling fluid, and so forth. High levels of torque &drag can lead to situations where the casing, liner and/or completion cannot be installed at the planned depth. Problems getting casing to bottom, getting weight on bit or trouble sliding could also result from this. All of these can limit the ultimate depth of the well[2]. However, the value of friction & torque is mainly determined by friction coefficient, friction coefficient is characterized by uncertainty, ambiguity and time-varying, and there are many factors affecting the coefficient friction. In this paper, artificial intelligence techniques have been used in friction coefficient prediction, the method of support vector machine is optimally selected to establish friction coefficient prediction model, which is a black system and particularly suitable to solve the small sample, nonlinear and high dimension problem.

## 1. MACHINE LEARNING AND SUPPORTING VECTOR MACHINE METHOD

### 1.1 Machine Learning Method

The machine learning techniques are sometimes called artificial intelligence methods since they fairly follow and mimic some of human properties and abilities such as learning, generalization, memorization, and prediction. However, scholars found that smart machines need to be improved and developed in such a way they can detect noisy data and faults, and also can plan despite uncertainties[3].

The main objective of seeking smart machine methods is to predict the occurrence of some problems based on previous experience with reasonable cost and time.

The reliability of the method depends on the accuracy of prediction and the error between the actual and the predicted class labels of the problem.

The most common machine learning techniques are Artificial Neural Networks (ANN) and Support Vector Machine (SVM). Both of these tools were designed mainly for classification or pattern recognition. These two computing methods are considered complimentary rather than competitive. The difference between these two methods is in their mathematical approach and construction[4]. Each technique would initiate a pattern that could produce outputs of a certain problem.

In this paper, the Support Vector Machine method is mainly introduced.

### 1.2 The Principle of Supporting Vector Machine

Supporting Vector Machine (SVM) is based on general machine learning method of statistical learning theory and it has better generalization ability especially for the case of limited samples of statistical learning problem than other methods. Basic idea of support vector machine is that by using the definition of inner product of the non-linear transform function to transform input space into a high-dimensional space, and then to find out the non-linear relationship between input variables and output variables in this high-dimensional space. Supporting Vector Machine has a strict theoretical basis, which uses the principle of structural risk minimization, and has good generalization ability. Supporting vector machine algorithm is a convex quadratic optimization problem, which makes sure that the solution is the global optimal solution, which is particularly suitable to solve the small sample, nonlinear and high dimension problem[5-6].

SVM method derived from the settlement of problem classification and SVR promotes the result of SVM from regression to a wide range. Linear regression and nonlinear regression are the main modes[7]. Linear regression can apply the following formula and match the data $\{(x_i, y_i), i = 1,2,\Lambda l\}$:

$$y = f(x) = w \cdot x + b. \tag{1}$$

$x_i \in R^d$, $y_i \in R$, $w$ and $b$ are the vector and offset of linear regression function.

Suppose all the data can precisely match with linear regression function at accuracy $\varepsilon$ and then the optimization can be realized with the introduction of slack variable $\xi_i$ and $\xi_i^*$. The optimization can be transformed to[8]:

$$\min \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{l}(\xi_i + \xi_i^*). \tag{2}$$

Constraints

$$\begin{cases} y_i - w \cdot x_i - b \le \varepsilon + \xi_i \\ w \cdot x_i + b - y_i \le \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \ge 0 \\ C > 0 \end{cases} \quad i = 1,2,\ldots,l. \tag{3}$$

The above optimization function is quadratic and constraints are linear so it is a typical quadratic programming problem. Lagrange multipliers can be used to settle it. Here Lagrange multipliers $\alpha_i$, $\alpha_i^*$, $\eta_i$, $\eta_i^*$ are introduced:

$$\begin{aligned} L\left(w, b, \xi_i, \xi_i^*\right) = \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{l}(\xi_i + \xi_i^*) - \sum_{i=1}^{l}\alpha_i(\varepsilon + \xi_i - y_i + w \cdot x_i + b) \\ - \sum_{i=1}^{l}\alpha_i^*(y_i + \varepsilon + \xi_i^* - w \cdot x_i - b) - \sum_{i=1}^{l}\left(\eta_i\xi_i + \eta_i^*\xi_i^*\right). \end{aligned} \tag{4}$$

Optimal solution

$$\begin{cases} \dfrac{\partial L}{\partial w} = w - \sum_{i=1}^{n}(\alpha_i - \alpha_i^*) \cdot x_i = 0 \\ \dfrac{\partial L}{\partial b} = \sum_{i=1}^{n}(\alpha_i - \alpha_i^*) = 0 \\ \dfrac{\partial L}{\partial \xi_i} = C - \alpha_i - \eta_i = 0 \\ \dfrac{\partial L}{\partial \xi_i^*} = C - \alpha_i^* - \eta_i^* = 0 \end{cases}. \tag{5}$$

$$\max \sum_{i=1}^{l} y_i(\alpha_i - \alpha_i^*) - \varepsilon\sum_{i=1}^{l}(\alpha_i + \alpha_i^*) - \frac{1}{2}\sum_{i,j=1}^{l}(\alpha_i - \alpha_i^*) \cdot (\alpha_j - \alpha_j^*)(x_i \cdot x_j). \tag{6}$$

Regression function can be get by applying Equation (5) to Equation (6):

$$f(x) = \sum_{i=1}^{l}(\alpha_i - \alpha_i^*)(x_i \cdot x) + b. \tag{7}$$

The above concepts can also be extended to the case of a nonlinear regression by a mapping of the input space onto a high dimensional space. The key property of this mapping is that the function $\varphi$ is subject to the condition that the dot product of the two functions $\varphi(X_i) \cdot \varphi(X_i)$ can be written as a kernel function $K(X_i, X)$, then the optimization is transformed to[9-10]:

$$\max \sum_{i=1}^{l} y_i(\alpha_i - \alpha_i^*) - \varepsilon\sum_{i=1}^{l}(\alpha_i + \alpha_i^*) - \frac{1}{2}\sum_{i,j=1}^{l}(\alpha_i - \alpha_i^*) \cdot (\alpha_j - \alpha_j^*)K(x_i \cdot x_j). \tag{8}$$

The corresponding regression function is:

$$f(x) = \sum_{i=1}^{l}(\alpha_i - \alpha_i^*)K(x_i \cdot x) + b. \tag{9}$$

### 1.3 Steps of Supporting Vector Machine Model

Specific SVM system prediction process is shown in Figure 1.

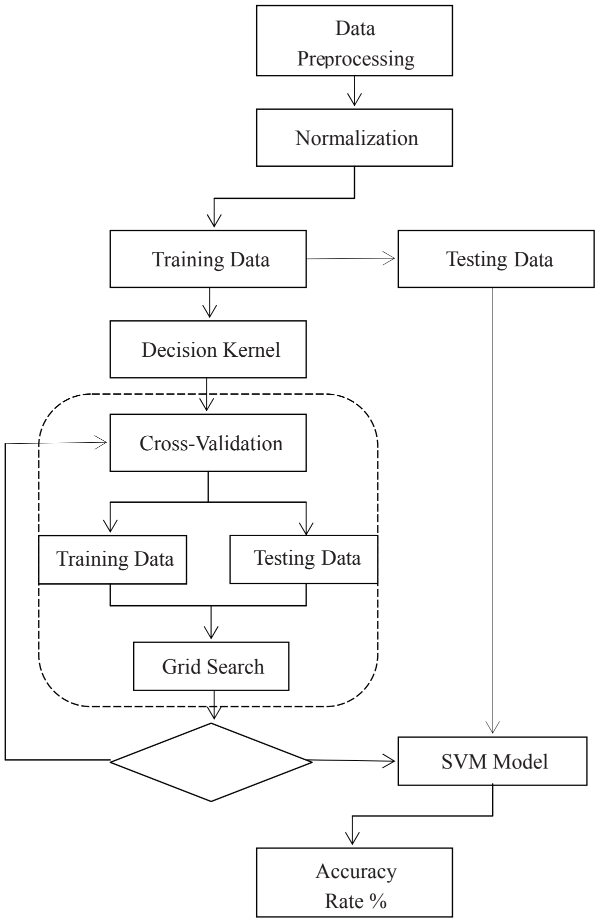


**Figure 1**
**Flowchart of the Proposed SVM System**

## 2. THE APPLICATION AND RESULTS ANALYSIS

### 2.1 The Selection of Friction Coefficient Influence Factors

Studies have shown that friction coefficient influence factors includes drilling fluid property, string structure, well trajectory parameters, rock property, and so on. In this paper, TD,TVD, displacement, average of overall angle change, drilling fluid type, density, viscosity, plastic viscosity, dynamic shear force, fluid loss is choose as input parameters, The inversion result of friction coefficient and input parameters is shown as Table 1.

To avoid negative impact on predicting results owing to differences of the relevant parameter dimensionless. This article normalizes the properties of the sample values. The properties of the samples are processed to interval [0, 1], assume the importation of samples number is $n$, then:

$$x_i' = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}} \qquad (i = 1,2,\cdots,n). \tag{10}$$

### 2.2 Selection of SVM Parameters

The parameters of SVM are mainly kernel parameters and the punishment parameter $C$, different parameters will have a direct influencing on learning efficiency and generalization ability of SVM. However, the SVM does not provide the general method of selecting inner product kernel function parameter which is easy to receive such as the width of radial basis function. For this reason, test algorithm is applied. That is to say the appropriate parameters are obtained by the sample studying of selected casing samples and testing of the kernel function and its parameters and the penalty parameter $C$.

### 2.3 Analysis of Prediction Results

The parameters of the previous 16 wells is adopted to establish learning model, the data of later 5 wells is used as prediction samples. Compared with the BP neural network method, the Prediction result based on SVM is shown as Table 2.

It can be drawn from Table 2 that Prediction result based on SVM has higher precision with an average of 91.6%, While the neural network method was 78.0%. The results demonstrated that friction coefficient prediction of designing well was reliable based on support vector machines, and it had broken the limitation of the traditional method which the friction coefficient was determined by experience relying on drilling fluid system, support vector machines method provide new insights into the prediction of friction coefficient during the design of extended reach wells. The prediction precision would be improved significantly, after the database was formed by continuously replenished prediction model samples, as the extended reach wells were more and more.

**Table 1**
**The Result of Friction Coefficient Inversion**

| Well number | Friction coefficient | Hole size (mm) | Well trajectory parameters: TD (m) | TVD (m) | Displacement (m) | Avg. rate of over all angle change(°/30m) | Drilling fluid parameters: Drilling fluid type | Density (g/cm$^3$) | Viscosity (s) | Plastic viscosity (mPa·s) | Dynamic shearforce (Pa) | Fluid loss (ml) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Bin 173-1HF | 0.43 | 215.9 | 3,983 | 2,892 | 1,366 | 1.19 | 2 | 1.25 | 80 | 23 | 14 | 2 |
| Bin 435-4HF | 0.27 | 215.9 | 5,097 | 3,699 | 1,786 | 1.4 | 2 | 1.4 | 66 | 30 | 14 | 2.8 |
| Boye-ping2 | 0.12 | 215.9 | 3,645 | 2,568 | 1,329 | 1.19 | 1 | 1.25 | 44 | 18 | 9 | 0.6 |
| Fan 116-1HF | 0.34 | 215.9 | 4,002 | 2,814 | 1,405 | 1.28 | 2 | 1.33 | 60 | 35 | 23 | 1.2 |
| Fan 116-3HF | 0.31 | 215.9 | 4,096 | 2,789 | 1,515 | 1.34 | 2 | 1.35 | 147 | 43 | 31 | 1.8 |
| Fan 116-4HF | 0.3 | 215.9 | 3,818 | 2,789 | 1,187 | 1.2 | 2 | 1.3 | 60 | 22 | 14 | 2 |
| Fan 154-8HF | 0.47 | 152.4 | 4,158 | 2,761 | 1,619 | 1.3 | 2 | 1.3 | 85 | 25 | 16 | 2.2 |
| Fan 162-2HF | 0.31 | 152.4 | 4,522 | 2,666 | 2,066 | 1.24 | 2 | 1.24 | 62 | 20 | 11.2 | 1.6 |
| Gao-ping1 | 0.47 | 215.9 | 4,535 | 948 | 3,814 | 1.43 | 2 | 1.14 | 60 | 22 | 8.67 | 2.8 |
| Niu871-X7 | 0.19 | 215.9 | 3,752 | 3,170 | 1,678 | 0.84 | 2 | 1.51 | 60 | 26 | 9 | 4 |
| Xin 14-8HF | 0.49 | 215.9 | 4,069 | 2,582 | 2,254 | 1.62 | 2 | 1.25 | 63 | 27 | 12.5 | 3.1 |
| Yan 227-2HF | 0.3 | 215.9 | 4,522 | 3,653 | 1,345 | 0.92 | 2 | 1.28 | 69 | 28 | 11.8 | 1.8 |
| Yan 227-7HF | 0.22 | 215.9 | 4,206 | 3,411 | 1,138 | 1.06 | 2 | 1.28 | 66 | 21 | 9.2 | 2.6 |
| Yan 227-4HF | 0.36 | 215.9 | 4,323 | 3,453 | 1,183 | 0.91 | 2 | 1.28 | 66 | 22 | 17 | 2.6 |
| Zhuang 129-1HF | 0.21 | 215.9 | 5,341 | 3,341 | 3,168 | 0.98 | 2 | 1.4 | 68 | 26 | 14 | 2.4 |
| Boye-ping 1 | 0.23 | 215.9 | 4,335 | 2,969 | 1,597 | 1.03 | 1 | 1.75 | 71 | 56 | 13 | 0.2 |
| Fan 116-2HF | 0.35 | 215.9 | 3,912 | 2,783 | 1,306 | 1.29 | 2 | 1.32 | 162 | 26 | 29 | 2 |
| Fan 154-7HF | 0.47 | 152.4 | 4,260 | 2,758 | 1,739 | 1.32 | 2 | 1.38 | 67 | 33 | 11 | 1.8 |
| Yan 227-5HF | 0.29 | 215.9 | 4,740 | 3,919 | 1,430 | 0.97 | 2 | 1.25 | 55 | 22 | 17 | 2 |
| Yan 227-9HF | 0.4 | 215.9 | 4,925 | 3,841 | 1,557 | 1.09 | 2 | 1.25 | 60 | 23 | 19 | 3.2 |
| Liangye 1HF | 0.26 | 152.4 | 3,969 | 3,206 | 1,010 | 0.91 | 1 | 1.54 | 80 | 36 | 10 | 2 |

*Note*. In the column of drilling fluid type, "1" means oil-based drilling fluid, "2" means water-based drilling fluid.

**Table2**
**Prediction Result Contrast of Friction Coefficient**

| Well Number | Inversion result | Prediction result based on SVM | Prediction accuracy based on SVM (%) | Prediction result based on BP | Prediction accuracy based on BP (%) |
|---|---|---|---|---|---|
| Fan 116-2HF | 0.35 | 0.32 | 91.4 | 0.27 | 77.1 |
| Fan 154-7HF | 0.47 | 0.43 | 91.5 | 0.4 | 85.1 |
| Yan 227-5HF | 0.29 | 0.28 | 96.6 | 0.22 | 75.9 |
| Yan 227-9HF | 0.4 | 0.41 | 97.5 | 0.30 | 75.0 |
| Liangye 1HF | 0.26 | 0.21 | 81.8 | 0.32 | 76.9 |

## CONCLUSION

(a) Supporting vector machine（SVM）is a new kind of learning machine based on statistical theory. In recent years, both the algorithm theory itself and the application in other areas have made great development.

(b) Due to the fact that multitude factors influence friction coefficient, it is difficult to build a certain relationship between friction coefficient and the relevant factors. Supporting vector machine is applied to the predicting of friction coefficient prediction in this paper. The SVM can take full advantage of the inversion cases of friction coefficient and establish the prediction model between friction coefficient and relevant factors.

(c) Friction coefficient is characterized by uncertainty, ambiguity and time-varying, the prediction model based on the SVM has the advantages of simplicity, convenience and high reliability, which open up a new way forfriction coefficient prediction in well designing.

## REFERENCES

[1] Johancsik, C. A., Friesen, D. B., & Dawson, R. (1984). Torque and drag in directional wells–prediction and measurement. *Journal of Petroleum Technology*, *36*(6), 987-992.

[2] Zhang, L. Q. (2008). Calculating model of torque and drag in extended reach well. *Fault-Block Oil & Gas Field*, *15*(2), 88-90.

 

[3] Fan, X. W., & Du, S. X. (2004). Rough support vector machine and its application to waste water treatment process. *Control and Decision*, *19*(5), 573-576.
[4] Wei, X. G. (2003). *Research on handwritten numerical digits recognition based on kernel-based learning algorithm* (Doctoral dissertation). Nanjing University of Science and Technology, Nanjing.
[5] Burbidge, R., Trotter, M., & Buxton, B. (2001). Drug design by machine learning: Support vector machines for pharmaceutical data analysis. *Computer and Chemistry*, *26*(1), 5-14.
[6] Trotter, M. W. B., Buxton, B. F., & Holden, S. B. (2001). Support vector machines in combinatorial chemistry. *Measurement and Control*, *34*(8), 235-239.
[7] Van Gestel, T., Suykens, J. A. K., & Baestaens, D. E. (2001). Financial time series prediction using least squares support vector machines within the evidence framework. *IEEE Transactions on Neural Networks*, *12*(4), 809-821.
[8] Burges, C. A. (1998). Tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery*, *2*(2), 1-47.
[9] Suykens, J. A. K., Vandewalle, J., & De Moor, B. (2001). Optimal control by least squares support vector machines. *Neural Networks*, *14*(1), 23-35.
[10] Platt, J. (1998). *Sequential minimal optimization*: *A fast algorithm for training support vector machines* (pp.169-182). Cambridge: The MIT Press.