

# A new diagnostic method for identifying working conditions of submersible reciprocating pumping systems

Yu Deliang<sup>1</sup>, Zhang Yongming<sup>1</sup>, Bian Hongmei<sup>2</sup>, Wang Xinmin<sup>3</sup> and Qi Weigui<sup>1\*</sup>

<sup>1</sup> School of Electrical Engineering and Automation, Harbin Institute of Technology, Harbin, Heilongjiang 150001, China

<sup>2</sup> Research and Development Center, Daqing Oilfield Company, Daqing, Heilongjiang 163000, China

<sup>3</sup> Research Institute of Oil Production Engineering, Daqing Oilfield Company, Daqing, Heilongjiang 163000, China

© China University of Petroleum (Beijing) and Springer-Verlag Berlin Heidelberg 2013

**Abstract:** The submersible pumping unit is a new type of pumping system for lifting formation fluids from onshore oil wells, and the identification of its working condition has an important influence on oil production. In this paper we proposed a diagnostic method for identifying the working condition of the submersible pumping system. Based on analyzing the working principle of the pumping unit and the pump structure, different characteristics in loading and unloading processes of the submersible linear motor were obtained at different working conditions. The characteristic quantities were extracted from operation data of the submersible linear motor. A diagnostic model based on the support vector machine (SVM) method was proposed for identifying the working condition of the submersible pumping unit, where the inputs of the SVM classifier were the characteristic quantities. The performance and the misjudgment rate of this method were analyzed and validated by the data acquired from an experimental simulation platform. The model proposed had an excellent performance in failure diagnosis of the submersible pumping system. The SVM classifier had higher diagnostic accuracy than the learning vector quantization (LVQ) classifier.

**Key words:** Submersible reciprocating pump, working condition, failure diagnosis, linear motor, characteristic quantity, support vector machine, misjudgment rate

## 1 Introduction

The submersible reciprocating pumping system is a new type of pumping unit used in oil fields. The submersible reciprocating pump is driven by a linear motor to lift formation liquids to the surface directly through oil tubing. A control device at the surface regulates the linear motor, which is powered by a submersible cable, to control the working conditions of the pumping unit through changing the operating time, stroke and the rate of the linear motor stator.

Different from conventional pumping systems, in this pumping system a submersible plunger pump is directly connected to a linear permanent magnet synchronous motor at the well bottom to lift crude oil (Yu et al, 2011a; 2011b). Since the output power of the linear motor acts directly on the pump for lifting crude oil, the submersible pump has high energy conversion efficiency (Rossini et al, 2008). Furthermore, no sucker rod and other equipment are introduced into the submersible pumping system compared

with conventional pumping units, thus completely avoiding eccentric wear between a sucker rod and oil tubing. However, the linear submersible motor and the plunger pump are exposed to severe well bottom conditions, failures such as linear motor damage may occur, which seriously shortens the work period of the pump (Fu et al, 2006; Yu et al, 2011b). Therefore, it is important to find an effective way to detect early abnormal working conditions before serious breakdown of the submersible pumping system. With more and more oilfields becoming maturing fields, the submersible pumping units have been widely used for this type of pumping system. This has major significance for energy saving and efficient production (Yu et al, 2011a; 2011b).

The abnormal working conditions of a pumping system, such as paraffin deposits in oil tubing, sand production, leakage, liquid shortage in the pump and gas obstruction, may increase the work load of the linear motor. If the linear motor works under abnormal conditions for a long period of time, its in-service life, as well as the continuous running time, will be shortened. Therefore, it is necessary to seek a way to dynamically process the abnormal conditions of the pumping system and to perform online diagnosis, thus providing early

\*Corresponding author. email: qwg1944@sina.com

Received March 7, 2012

failure warning.

For conventional beam pumping units, an intelligent fault diagnosis is used to recognize the dynamometer card patterns (Gibbs and Neely, 1966; Foley and Svinos, 1989; Nazi et al, 1994; Wu et al, 2004). However, the submersible reciprocating pump is a new type of rodless pumping unit, so it is impossible to measure polished rod load and stroke (or displacement) to obtain a surface dynamometer card. Meanwhile, as a new type of pumping unit, there have been few papers on the analysis of its typical failure information. The pumping system would not be running under abnormal conditions for a long time in practice, so it is difficult to obtain the samples of failure signals. The acquired failure information and running state have an observably nonlinear relation (Zhou and Bennett, 1998; Polycarpou and Trunov, 2000; Polycarpou, 2001; Zhang et al, 2005), especially in the transitory stage of starting or stopping. For these reasons, the diagnosis of the working conditions of the submersible reciprocating pumping unit is characterized by small sample size, nonlinear relation and dynamic change, thus the conventional diagnostic methods are not well suited (Pazzani, 1987; Youn and Hammen, 1998; Vemuri et al, 1999; Carrasco, 1999).

In order to solve the above problems, we propose a new method for the diagnosis of working conditions of submersible pumping systems based on a support vector machine (SVM) method. On the basis of the failure diagnosis of monodrome SVM (Muller et al, 2001; Zhang, 2000; Du and Hou, 2007), a diagnostic model is established to analyze multiple working conditions. Meanwhile the characteristics of the loading/unloading process of a submersible linear motor are investigated and several characteristic quantities are extracted. Then the sample data of the characteristic quantities are used to train and test the diagnostic machine based on SVM. The online diagnosis of the submersible pumping systems is determined by observing and identifying those abnormal working conditions. Diagnostic methods based on SVM and learning vector quantization (LVQ) are compared.

## 2 Failures and their characteristics of the submersible pumping system

We have tracked and investigated working conditions of several submersible pumping systems in several oil production units in the Daqing Oilfield, China. The results showed that abnormal working conditions like paraffin deposits in oil tubing, sand production, leakage, liquid shortage in the pump and gas obstruction may occur after the submersible pumping system has been working for a period of time. If these problems are not checked, they would probably cause damage to submersible linear motors and pumps, affecting the oil production.

The submersible reciprocating pumping system is different from the conventional beam pumping units. A linear motor is under the pump body, while its standing valve is above the pump body. Furthermore, the submersible pumping system is a type of rodless pumping unit and so its dynamometer card cannot be measured. In this paper, the load changes under

abnormal conditions in the cycle of pumping are summarized by analyzing the structural characteristics of the submersible pumping system.

**Influence of gas** If the gas content in downhole fluids is relatively high, some gas may be separated from the fluids and remain in the pump at the start of the upstroke. As the gas may be compressed and the exhaust valve cannot open quickly, the loading rate slows down. On the other hand, some solution gas and compressed gas are left in the clearance space at the end of the upstroke, and the pressure in the pump cannot decrease quickly due to gas expansion at the start of the downstroke, so the suction valve cannot open quickly. This results in unloading hysteresis. Although there are differences in the working time of the suction and discharge valves between the beam pumping unit and the submersible pumping unit, the gas obstruction for the motor load is the same. The more the gas there is, the longer the loading and unloading time is, and the less the work is done in a cycle. Air lock may occur in extreme cases. Therefore, the characteristic quantities (loading time during the upstroke  $T_{UU}$ , loading time during the downstroke  $T_{DU}$ , unloading time during the upstroke  $T_{UD}$ , unloading time during the downstroke  $T_{DD}$ , average power during the upstroke  $P_{UE}$ , average power during the downstroke  $P_{DE}$ ) are extracted from loading time, unloading time and power of the linear motor, which serve as the base to analyze the influence of gas.

**Influence of liquid shortage in the pump** If the submergence depth in a well is too small or the fluid supply from the reservoir is insufficient, the liquid cannot completely fill the pump. When the upstroke of the submersible pump is started, the suction valve cannot immediately open, so unloading hysteresis occurs and the load begins to fluctuate when the traveling valve bumps to the liquid in the pump. Therefore, the characteristic quantities (average motor current during the upstroke  $I_{UE}$ , average motor current during the downstroke  $I_{DE}$ , variance of the motor current during the upstroke  $I_{UV}$ , variance of the motor current during the downstroke  $I_{DV}$ , unloading time during the upstroke  $T_{UD}$ , and unloading time during the downstroke  $T_{DD}$ ) are extracted from the load fluctuation of the linear motor to analyze the influence of liquid shortage in the pump.

**Influence of leakage** 1) If the traveling valve is leaky, the pressure between the suction and exhaust valves gradually goes up, and the pressure underneath the suction valve gradually goes down during the upstroke of the pump. At the same time, the liquid may leak underneath through the suction valve, supporting the plunger of the reciprocating pump and then slowing the load rate. However, with the opening of the exhaust valve and the upward movement of the plunger, the pressure between the suction and exhaust valves will gradually reduce, so the support effect will be weakened and the load will gradually become stabilized. The movement of the plunger decelerates during the latter half of the upstroke. When the liquid velocity is greater than the plunger rising velocity, it will reproduce the support effect and lead to premature unload. 2) When standing valve leakage occurs, the liquid will flow into the space (or gap) between the suction and exhaust valves where the exhaust valve is not tight during the pump downstroke, which results in a delay of the opening

of the suction valve. Therefore, the unload rate becomes slow and the movement of the plunger decelerates during the latter half of the downstroke. The fluids continue to leak through the standing valve and the support effect of the liquid in the pump leads to premature loading. 3) When the travelling and standing valves are both leaky, the load change is equal to the superimposed effect of these cases in isolation. 4) When the leakage occurs at the oil tubing, although the loading and unloading processes are not hysteretic, the maximum load cannot reach the theoretical maximum. Therefore, the characteristic quantities ( $T_{UU}$ ,  $T_{DU}$ ,  $T_{UD}$ ,  $T_{DD}$ ,  $I_{UE}$ ,  $I_{DE}$ ,  $P_{UE}$ ,  $P_{DE}$ ) are extracted from loading time, unloading time and maximum load of the linear motor to analyze the influence of leakage.

**Influence of sand production** When the submersible pump is installed at a location of the tubing string which is surrounded with unconsolidated formation or the production pressure is too high, sand will move into the wellbore from the reservoir. If some fine sand gets into the pump, it will generate additional resistance to the plunger in a certain region or the whole region. As a result, the load of the pumping unit increases during the upstroke and downstroke. In general, the fine sand is randomly distributed in the pump, thus the resistance is different in each cycle stroke and may induce a drastic change in the pump load in a short time. Therefore, the characteristic quantities ( $I_{UV}$ ,  $I_{DV}$ ) are extracted from load fluctuation in a cycle of the linear motor to analyze the influence of sand production.

**Influence of paraffin deposits** Paraffin deposited in the oil tubing creates an additional resistance to the plunger during operation. This will increase the load during the upstroke and downstroke. In this case, as the linear motor is located at the lower part of the pump, the load change of the submersible pumping unit is quite different from that of a beam pumping unit. Therefore, the characteristic quantities ( $I_{UE}$ ,  $I_{DE}$ ,  $P_{UE}$ ,  $P_{DE}$ ) are extracted from load fluctuation in a cycle of the linear motor to analyze the influence of paraffin deposits.

The analyses of the pump structure and the operation principle of the linear motor, characteristics of eight working conditions are obtained. Although the characteristics of different working conditions may be the same, the value range and the variation of the characteristic quantities are different. Therefore, by using of SVM classifier, characteristic quantities are extracted to identify the working conditions of the submersible reciprocating pumping system.

### 3 Extraction of characteristic quantities

The submersible reciprocating pump is operated in an intermittent mode. Each running cycle includes four states “upstroke-upstroke interval-downstroke-downstroke interval”. Changes of the motor voltage and current in a cycle are shown in Fig. 1, the upstroke/downstroke statuses are the research focus.

As shown in Fig. 1, in the initial course of the upstroke/downstroke of the submersible pumping unit, sharp peaks are observed in the motor current curve, which are the starting current of the linear motor. The starting current may

influence diagnosis of working conditions based on SVM when extracting characteristic parameters. Therefore, it is necessary to eliminate the peak value of the starting current before processing data. The method is shown in Fig. 2, which uses the first minimum value (at time  $t_2$ ) after the peak value on the curve as a reference to find a point (at time  $t_1$ ) before the peak value (the motor current at this point is equal to the first minimum current). All current values between  $t_1$  and  $t_2$  are substituted with the first minimum value.

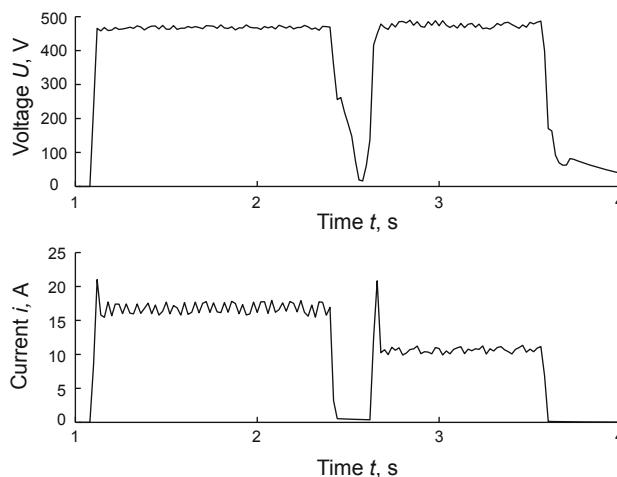


Fig. 1 Voltage and current of the linear motor

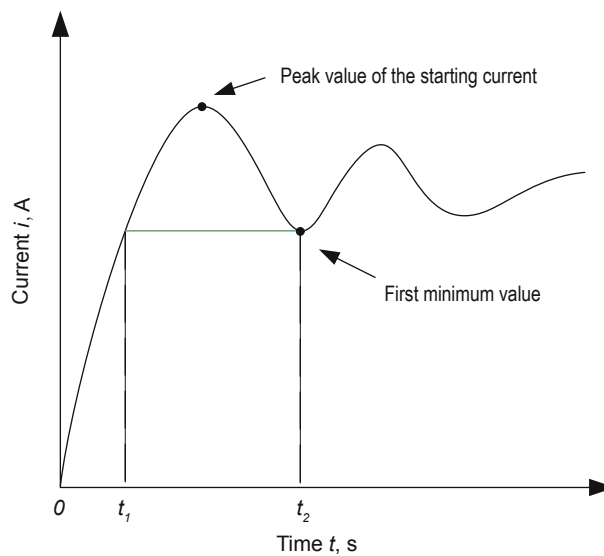


Fig. 2 Starting current of the linear motor

Above-mentioned analysis (Section 2) indicates that there is a close relationship between the working condition of the pumping system and the load of the submersible motor, and the data such as operating voltage, current and power may mirror the change of the load of the linear motor. Therefore, it is necessary to extract the characteristic parameters of the submersible linear motor for dynamic diagnosis of working conditions.

In an operational cycle of the linear motor, the motor current during the upstroke is defined as follows:

$$I_{UE} = \frac{1}{t_U} \sum_{t=1}^{t_U} i_U(t) \quad (1)$$

where  $i_U$  is the motor current during the upstroke;  $t_U$  is the number of samples during the upstroke;  $I_{UE}$  is the average motor current during the upstroke.

$$I_{UU} = k_{UU} \cdot I_{UE} \quad (0 < k_{UU} < 1) \quad (2)$$

$$I_{UD} = k_{UD} \cdot I_{UE} \quad (0 < k_{UD} < 1) \quad (3)$$

where  $k_{UU}$  and  $k_{UD}$  are the loading and unloading coefficients of the motor current during the upstroke, respectively;  $I_{UU}$  and  $I_{UD}$  are the threshold values of the loading current and the unloading current during the upstroke.

$$T_{UU} = \min S \quad (4)$$

$$\text{s.t.} \quad I_S \geq I_{UU} \quad S \in \{1, 2, \dots, t_U\} \quad (5)$$

$$T_{UD} = \max Q + 1 \quad (6)$$

$$\text{s.t.} \quad I_Q \geq I_{UD} \quad Q \in \{1, 2, \dots, t_U\} \quad (7)$$

where  $T_{UU}$ ,  $T_{UD}$  in Eqs. (4) and (6) are the loading time and the unloading time during the upstroke, respectively.

Similarly, during the downstroke of the submersible linear motor, we define:

$$I_{DE} = \frac{1}{t_D} \sum_{t=1}^{t_D} i_D(t) \quad (8)$$

where  $i_D$  is the motor current during the downstroke;  $t_D$  is the number of samples during the downstroke;  $I_{DE}$  is the average motor current during the downstroke.

$I_{DU}$ ,  $I_{DD}$  and  $k_{DU}$ ,  $k_{DD}$  can be obtained from Eqs. (2) and (3) ( $0 < k_{DU} < 1$ ,  $0 < k_{DD} < 1$ ).  $k_{DU}$  and  $k_{DD}$  are the loading and unloading coefficients of the motor current during the downstroke.  $I_{DU}$  and  $I_{DD}$  are the threshold of loading current and threshold of unloading current during the downstroke. In a similar way,  $T_{DU}$  and  $T_{DD}$  are the loading time and the unloading time during the downstroke, they can be obtained from Eqs. (4)-(7). Besides, other characteristic quantities should be defined as:

$$P_{UE} = \frac{1}{t_U} \sum_{t=1}^{t_U} P_U(t) \quad (9)$$

$$P_{DE} = \frac{1}{t_D} \sum_{t=1}^{t_D} P_D(t) \quad (10)$$

$$I_{UV} = \frac{1}{t_U} \sum_{t=1}^{t_U} (i_U(t) - I_{UE})^2 \quad (11)$$

$$I_{DV} = \frac{1}{t_D} \sum_{t=1}^{t_D} (i_D(t) - I_{DE})^2 \quad (12)$$

where  $P_{UE}$  and  $P_{DE}$  are the average motor power values during the upstroke and downstroke, respectively;  $I_{UV}$  and  $I_{DV}$  are the variances of the motor current during the upstroke and

downstroke, respectively.

## 4 Diagnosis of working conditions based on SVM

### 4.1 Bayes decision rule based on SVM

In actual oil production, downhole conditions of production wells are different, each submersible reciprocating pumping system works under different and changing conditions, so it is necessary to propose a method for identifying nonlinear dynamic working conditions on small size samples. The support vector machine (SVM) technique is a widely used method with great adaptability (Schölkopf et al, 1999; Vapnik, 1999; Luts et al, 2010). It is well suited to the diagnosis of the working condition of the submersible reciprocating pumping system. The nature of the SVM is based on the principle of structural risk minimization. It is a statistical learning machine which maps low-dimensional, linearly non-separable data to a high dimensional space and constructs an optimal hyperplane to classify the data (Chen and Huang, 2006; Luo et al, 2010).

For a set of non-linearly separable samples,  $\Omega = \{(x_i, y_i)\}_{i=1}^N$ ,  $y_i = +1$  or  $y_i = -1$ ,  $N$  samples in all, if there is a nonlinear mapping.

$$\phi(x), \mathbb{R}^{n_0} \rightarrow \mathbb{R}^{m_0}, (m_0 \geq n_0) \quad (13)$$

After mapping the samples to a space with some characteristic quantities from the original input space, the hyperplane can be expressed as follows:

$$f(x) = \omega \cdot \phi(x) + b \quad (14)$$

where  $\omega$  and  $\phi(x)$  are  $m_0$  dimensional vector;  $b$  is offset.

On the optimal hyperplane,  $2 / \|\omega\|$  should be the largest. So quadratic programming problems may be solved with the Lagrange method.

$$\min_{\omega, b, \xi} J = \frac{1}{2} \omega^T \cdot \omega + C \sum_{i=1}^N \xi_i \quad (15)$$

$$\text{s.t.} \quad y_i(\omega \cdot x_i + b) \geq 1 - \xi_i \quad (16)$$

where  $\xi_i$  is the slack variable and,  $\xi_i \geq 0$  ( $i = 1, 2, \dots, N$ );  $C$  is the penalty coefficient. When  $C$  increases, the degree of penalty increases accordingly. A dual problem can be obtained:

$$\max \bar{J} = \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j y_i y_j \phi(x_i) \phi(x_j) \quad (17)$$

Lagrange's multiplier,  $\{\alpha_i\}_{i=1}^N$ , meets the constraints:

$$\text{s.t.} \quad 0 \leq \alpha_i \leq C, \quad \sum_{i=1}^N \alpha_i y_i = 0 \quad (i = 1, 2, \dots, N) \quad (18)$$

As Eq. (17) only relates to a transvection operation, if a

kernel function,  $k(x_i \cdot x_j) = \phi(x_i)^T \phi(x_j)$ , satisfies the Mercer condition, it does not need to know the specific form of the nonlinear function. And the following discriminant function can be obtained from the original input space:

$$f(x) = \sum_{i=1}^N \alpha_i y_i k(x_i, x) + b \quad (19)$$

Working condition identification of the submersible reciprocating pumping system is mostly multi-valued. Multi-valued classification usually can be obtained by using a double-valued classifier structure, such as 1-a-r method, 1-a-1 method or DAG-SVM method (Casdagh, 1989; Bottou et al, 1994; Krebel, 1999). A multi-valued diagnostic method for working conditions is proposed, which is based on Bayes decision rules and single-valued SVM diagnosis method (Wu et al, 2005; Lu et al, 2009). The single-valued SVM diagnostic method is to construct a super-sphere in high dimensional space, the super-sphere contain the required samples, and refuses other samples.

For an  $h$  classification problem, each category is seen as a positive class, that is  $y_i = +1$ , and a single-valued SVM model is trained. Assuming that the training set of each category is  $S_i$ , and the corresponding category is  $\gamma_i$ , then

$$f_i(x) = \sum_i \alpha_i k(x, x_i), \quad x_i \in S_i \quad i = 1, 2, \dots, h \quad (20)$$

Then we obtain  $h$  classifiers.  $f_i(x)$  is the SVM output of the sample category  $i$  ( $b$  is neglected for it is a constant and does not affect the following discussion), which can be normalized as follows:

$$f'_i(x) = \frac{\sum_i \alpha_i k(x, x_i)}{\int \sum_i \alpha_i k(x, x_i) dx} = \frac{f_i(x)}{\int f_i(x) dx} \quad (21)$$

Because  $f'_i(x) \geq 0$  and  $\int f'_i(x) dx = 1$ ,  $f'_i(x)$  can be sed as a density function of some probability distribution to construct the Bayes classifier. Conditional probability is estimated as follows:

$$P(x | \gamma_i) = \frac{f'_i(x)}{\int f'_i(x) dx} \quad (22)$$

Prior probability estimation can be expressed as:

$$P(\gamma_i) = \frac{\int f'_i(x) dx}{\int f'_1(x) dx + \int f'_2(x) dx + \dots + \int f'_h(x) dx} \quad (23)$$

And the total probability density can be expressed as:

$$P(x) = P(x | \gamma_1)P(\gamma_1) + P(x | \gamma_2)P(\gamma_2) + \dots + P(x | \gamma_h)P(\gamma_h) \quad (24)$$

According to the Bayes theorem, the posteriori probability estimation can be expressed as:

$$P(\gamma_i | x) = \frac{P(x | \gamma_i)P(\gamma_i)}{P(x)} = \frac{f_i(x)}{\int f_1(x) dx + \int f_2(x) dx + \dots + \int f_h(x) dx} \times \frac{1}{P(x)} \quad (25)$$

Because the total probability density does not affect classification results, then  $P(x)=1$ , Eq. (25) can be expressed as follows:

$$P(\gamma_i | x) = \frac{f_i(x)}{\int f_1(x) dx + \int f_2(x) dx + \dots + \int f_h(x) dx} \quad (26)$$

$(i = 1, 2, \dots, h)$

Based on the minimum error rate criteria, the unknown sample categories can be judged.

$$\text{If } f_i(x) = \max_{j=1,2,\dots,h} \{f_j(x)\}, \text{ then } x \in \gamma_i \quad (27)$$

The classifier structure is shown in Fig. 3,  $h$  single-valued SVMs are built in this way, without repetition training in various types of samples (Chua, 2003; Cherkassky and Ma, 2004; Du and Hou, 2007).

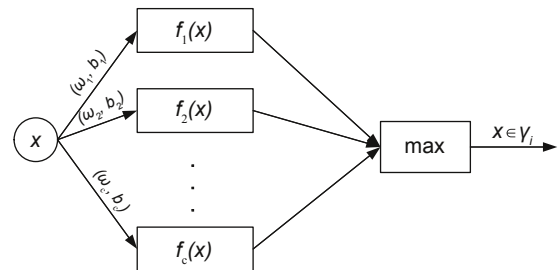


Fig. 3 Classifier structure

### 4.2 Constructing an input space vector

In a cycle of operation of the pumping unit, several characteristic quantities defined in the previous section are extracted. Because the orders of magnitude of these characteristic quantities are quite different, it is necessary to normalize them:

$$\hat{x}_i = 2 \frac{x_i - x_{\min}}{x_{\max} - x_{\min}} - 1 \quad (28)$$

where  $\hat{x}_i$  is the normalized data ( $\hat{x}_i \in [-1, 1]$ ), and  $x_i$  is the original sample data of the characteristic quantity;  $x_{\max}$  and  $x_{\min}$  are the maximum and minimum values of the characteristic quantity, respectively.

The input vector for identifying working conditions of the submersible pumping system is obtained from Eq. (28):

$$G = [\hat{T}_{UU}, \hat{T}_{UD}, \hat{T}_{DU}, \hat{T}_{DD}, \hat{I}_{UV}, \hat{I}_{DV}, \hat{I}_{UE}, \hat{I}_{DE}, \hat{P}_{UE}, \hat{P}_{DE}]^T \quad (29)$$

### 4.3 Diagnosis of working conditions of the submersible pumping system

One of the important problems in the diagnosis of the working condition of the submersible pumping system is how to choose an SVM kernel. In the SVM literature, most of attention has been paid to the polynomial kernel function, RBF kernel function and the sigmoid kernel function. Therefore, these kernel functions are tested one by one and the RBF function is selected as the kernel function. The process of tests is not described here. The RBF kernel function is expressed as follows:

$$k(x_k, x_l) = \exp(-\|x_k - x_l\|^2 / \sigma^2) \quad (30)$$

where  $\sigma$  is a parameter that controls the kernel width.

In these characteristic quantities,  $I_{UE}$ ,  $I_{DE}$ ,  $P_{UE}$  and  $P_{DE}$  are inherent in the running of the submersible pumping system, and they only vary with the sample data. But  $T_{UD}$ ,  $T_{DU}$ ,  $T_{DD}$ ,  $I_{UV}$  and  $I_{DV}$  are not only related with sample data but also related with  $k_{UU}$ ,  $k_{UD}$ ,  $k_{DU}$  and  $k_{DD}$ , so the selection of the loading/unloading threshold will affect the accuracy of the working condition diagnosis.

Under 9 kinds of working conditions, which are gas obstruction, liquid shortage in the pump, valve leakage, sand production, paraffin deposits and normal running, original operational parameters are measured. After pretreatment (normalization), training samples corresponding to each working condition are obtained. The parameters  $C$  and  $\sigma$  of the model are defined through those training samples. After being trained, single-valued SVM classifiers under each working condition are obtained and constitute the diagnostic machine for multiple conditions of the submersible pumping system, as shown in Fig. 3.

Another part of the running data measured under each working condition is normalized and then used as test samples for working condition diagnosis. The misjudgment rate under each SVM parameter and threshold is used to judge the working condition classifier of the submersible pumping system.

## 5 Simulation experiments

### 5.1 Simulation of working condition diagnosis based on SVM

An experimental platform was developed to simulate fluid circulation in an oil well. The platform consisted of a simulated circulation system, a flow control device, a submersible reciprocating pumping system and its control device. A simplified schematic and a photo of the experimental platform are shown in Figs. 4 and 5, respectively. Electrical parameters of the cylinder-shaped linear motor used were as follows: outer diameter 110 mm, length 6 m, rated voltage 660 V, the maximum input current 50 A, the maximum output power 30 kW, rated thrust 24,000 N, maximum thrust 35,000 N, maximum travel of the slider 123 cm. Fig. 6 shows the cylinder-shaped linear motor and the reciprocating pump, the red one is the linear motor, and the blue one is the reciprocating pump. Fig. 7 shows the control cabinet of the submersible reciprocating pumping system.

In experiments, working conditions were simulated by controlling the circulation pressure and introducing artificial failures. Meanwhile real-time recording of the running state of the linear motor and its current and power and other parameters were acquired. The linear submersible motor was rated at 8 strokes per minute. When the system was stabilized, data of 300 full-cycles were recorded under each working condition. Every full-cycle data created an input vector, the first 230 input vectors were used to construct training samples for the classifier, and the remaining 70

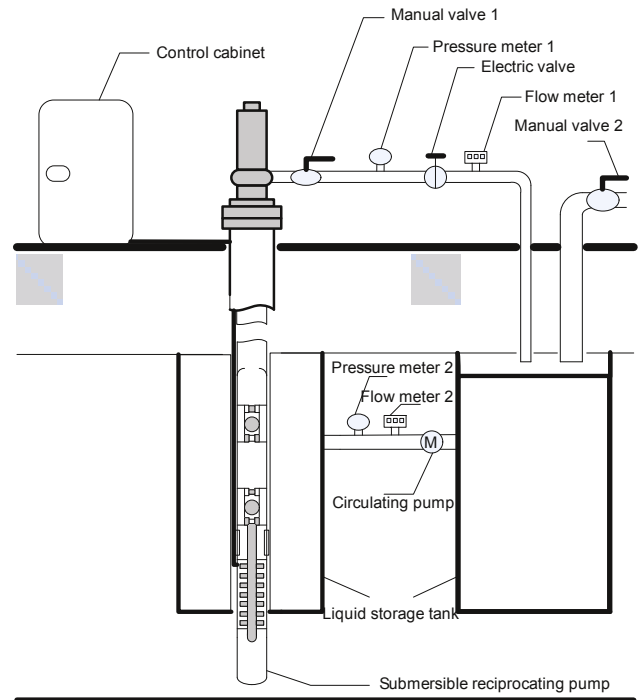


Fig. 4 Schematic of the simulation platform



Fig. 5 Photo of the simulation platform

were used to construct test samples. The diagnosis classifier for the submersible pumping system was trained by the 230 samples and tested by the 70 samples.  $k_{UU}$ ,  $k_{UD}$ ,  $k_{DU}$  and  $k_{DD}$  were manually-set parameters, if they changed, the training samples would change, and  $C$  and  $\sigma$  would also change. The misjudgment rate of each failure with different parameters is shown in Table 1,  $F_1$ : gas obstruction,  $F_2$ : liquid shortage in the pump,  $F_3$ : travelling valve leakage,  $F_4$ : standing valve leakage,  $F_5$ : leakage in both valves,  $F_6$ : tubing leakage,  $F_7$ : sand production,  $F_8$ : paraffin deposits,  $F_9$ : normal running (when the output of  $F_1$  to  $F_8$  classifiers was less than a setting value, the running state may be determined as  $F_9$ ).

The values of relevant parameters shown in Table 1 are only part of the experimental data. It was clear that when  $k_{UU}=0.85$ ,  $k_{UD}=0.80$ ,  $k_{DU}=0.80$ ,  $k_{DD}=0.75$ , the misjudgment rate of each failure could be controlled in an expected



**Fig. 6** Linear motor and reciprocating pump



**Fig. 7** Control cabinet of the submersible reciprocating pumping system

range. It should be noted that the normal running state  $F_9$  was not an individual classifier, its output came from other classifiers. In our study, only 8 failures were simulated, and the corresponding classifiers were trained and tested. So the misjudgment rate of  $F_9$  was higher than those of others. However, in practice production, this may lead to an increase in the misjudgment rate because of some unknown failure.

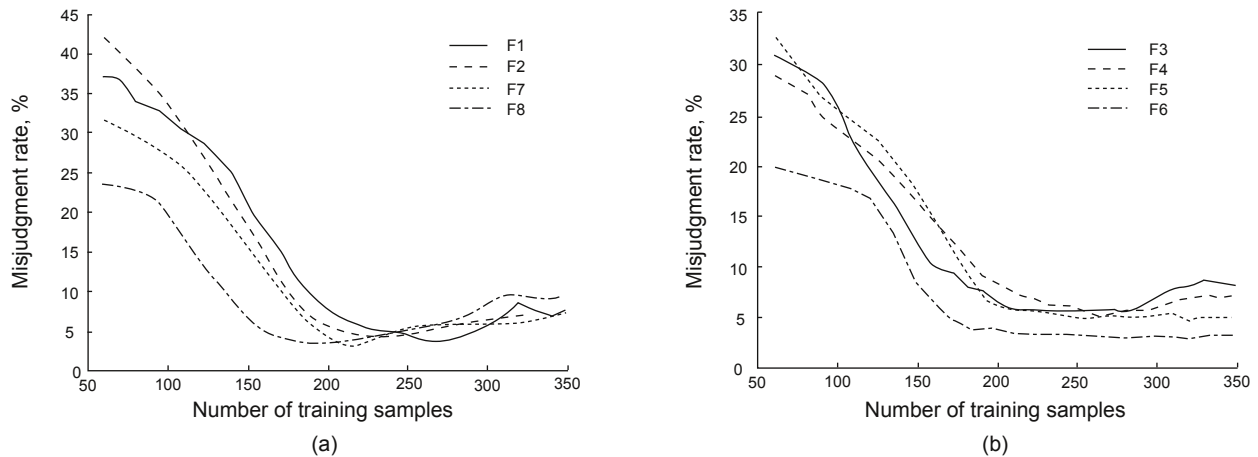
The misjudgment rate of the SVM-based method for identifying the working condition of the submersible pumping system is not only related to the model parameters, but also related to the number of training samples. This paper also described the effect of sample size (60-350 samples) on the misjudgment rate at the same model parameters. Different training sample sizes, ranging from 60 to 350 by an increment of 10, were used to estimate the dependence of

the misjudgment rate on the sample size (Fig. 8). But the test sample size was kept at 70.

As shown in Fig. 8, the misjudgment rate of the working condition diagnosis decreases with an increase in the training sample size. It can be seen from the curve that the misjudgment rate decreases from about 40% to 5% with an increase in the training sample size when the free gas has an influence on the working condition. Under the condition of liquid shortage in the pump, the misjudgment rate decreases sharply with an increase in the training sample size. When the size of training samples is about 200, the misjudgment rate drops to 6%, and then the rate fluctuation is small with an increase in the training sample size. Under the condition of sand production, the misjudgment rate decreases at an almost constant rate. The misjudgment rate may slip below 5% as

**Table 1** Misjudgment rate based on the SVM method

	Misjudgment rate ( $k_{UU}, k_{UD}, k_{DU}, k_{DD}$ ), %				
	(0.70, 0.80, 0.75, 0.70)	(0.75, 0.80, 0.75, 0.75)	(0.80, 0.85, 0.75, 0.80)	(0.85, 0.80, 0.80, 0.75)	(0.90, 0.85, 0.90, 0.80)
$F_1$	11.7	8.4	6.9	4.9	5.7
$F_2$	10.4	7.3	6.6	4.9	5.4
$F_3$	8.5	9.5	7.2	5.8	5.9
$F_4$	9.6	9.8	7.7	6.0	6.4
$F_5$	9.1	9.7	7.1	5.6	6.1
$F_6$	7.3	7.7	6.4	3.1	4.0
$F_7$	10.1	8.1	6.2	4.4	6.8
$F_8$	9.9	8.8	6.7	4.3	6.1
$F_9$	6.5	6.6	5.4	4.1	4.9
Average	9.2	8.4	6.7	4.8	5.7



**Fig. 8** Misjudgment rate versus the number of training samples

the number of training samples reaches 200 or so, and then the misjudgment rate changes slightly. Under the condition of paraffin deposits, the misjudgment rate decreases sharply when the number of training samples is less than 150. The rate decreases below 5% when the sample size is about 200, and then increases slightly with an increase in the sample size. Under the leakage condition of the travelling valve, standing valve and the both valves, the change of the misjudgment rate are similar. The misjudgment rates are 7%-8% when there are 250 training samples or so, and then fluctuate a little. Under the condition of the tubing leakage, the misjudgment rate decreases slowly when the number of training samples is below 120. The rate decreases to below 5% when 120-180 samples are used to train the classifier. After that, the training sample size has little effect on the misjudgment rate. Therefore, the misjudgment rate can reach a desired value when the training sample size is about 200 and after that the misjudgment rate is hardly affected by the sample size.

## 5.2 Comparative analysis of SVM- and LVQ-based methods

Using the same experimental data, the misjudgment rates of the diagnostic method based on the learning vector quantization (LVQ) were calculated. The training samples for classifier included the first 230 input vectors, and the test samples included the last 70 ones. The neural network structure of LVQ was 10-50-8: the number of the neurons in the input layer was the vector dimension of characteristic quantity; the number of the neurons in the output layer was the number of working conditions; and the number of the neurons in the competitive layer was 50 (the design and optimization of network structure is another issue which is not discussed here). The classifier based on LVQ for identifying the working condition of the submersible pumping system was trained and tested, respectively. The misjudgment rate (listed in Table 2) was also related to the training sample size. Taking the influence of paraffin deposits as an example (because of results are almost the same, other working

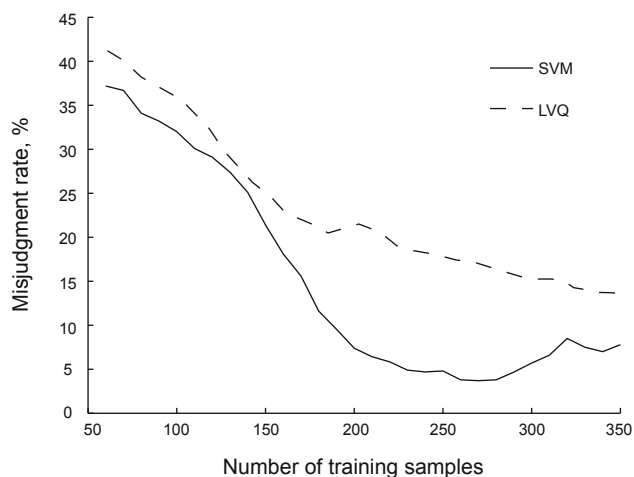
**Table 2** Misjudgment rate of the LVQ-based method

Fault	Misjudgment rate at different training sample size, %				
	50	100	150	200	230
F <sub>1</sub>	43.7	35.1	23.1	19.2	15.7
F <sub>2</sub>	40.1	33.2	22.3	16.8	14.1
F <sub>3</sub>	38.9	31.6	22.6	17.2	14.5
F <sub>4</sub>	42.1	37.2	23.7	16.3	13.3
F <sub>5</sub>	39.1	32.8	22.9	16.1	12.7
F <sub>6</sub>	37.4	30.3	21.8	17.1	14.3
F <sub>7</sub>	42.5	36.3	23.2	20.2	18.1
F <sub>8</sub>	41.3	34.9	22.1	19.3	17.2
Average	40.6	33.9	22.7	17.8	14.9

conditions are not discussed here), the effect of the sample size (60-350 samples) on the misjudgment rate was discussed when the model parameters are the same. The training sample size was varied from 60 to 350, with an increment of 10, and the test sample size was 70. The curves of the misjudgment rate of the SVM- and LVQ-based methods are shown in Fig. 9.

As can be seen in Fig. 9, the misjudgment rate of working condition diagnosis based on LVQ decreases significantly at a constant rate when the sample number is below 170, and then decreases slowly with an increase in the training sample size. The misjudgment rate decreases to about 13% when the training samples size reaches 350. The misjudgment rate curve of working condition diagnosis based on SVM slips smoothly when the sample size is below 120; and the misjudgment rate decreases quickly from 25% to 7% when the sample size is 120-180, then it decreases slowly. The misjudgment rate reaches a minimum, 5%, when the training sample set includes 280 samples. A comparison of the misjudgment rate calculated from the LVQ- and SVM-methods indicates that the misjudgment rate based on SVM





**Fig. 9** Misjudgment rates of the SVM- and LVQ-based methods

decreases at a faster rate and finally reaches a minimum value, which indicates that the diagnosis based on SVM is more suited for solving problems encountered in working conditions of the submersible reciprocating pumping system.

## 6 Conclusions

We have analyzed the principle and working conditions of the submersible reciprocation pumping. Based on the structural characteristics of the pumping unit, a new way to diagnosis the working condition of the submersible reciprocating pumping unit was proposed and then tested by the data obtained from the simulation platform.

1) Based on the structural and running characteristics of the submersible reciprocating pumping unit, the characteristics of the loading/unloading of the submersible linear motor under different working conditions were extracted. The characteristics of the working conditions listed in this paper could be accurately described by the loading/unloading characteristics.

2) The loading/unloading characteristics of the submersible motor under various working conditions were analyzed, and the loading/unloading thresholds of upstroke/downstroke, such as  $k_{UU}$ ,  $k_{UD}$ ,  $k_{DU}$ ,  $k_{DD}$  and other characteristic quantities, were defined. Extracting the characteristic quantities from the collected original data of the submersible linear motor, characteristic quantities  $T_{UU}$ ,  $T_{UD}$ ,  $T_{DU}$ ,  $T_{DD}$ ,  $I_{UV}$ ,  $I_{DV}$ ,  $I_{UE}$ ,  $I_{DE}$ ,  $P_{UE}$ ,  $P_{DE}$  were obtained, which were used as the input of the SVM working condition diagnosis of the submersible reciprocating pumping unit.

3) Training and test samples were acquired through simulating various working conditions of the submersible reciprocating pumping unit on the simulation platform. These samples were used for training and testing the SVM classifier for identifying the working condition of the submersible reciprocating pumping unit. The test results were compared with those from the LVQ classifier. The test results indicated that the misjudgment rate of the SVM classifier varied with

the training sample size and the working conditions. The SVM classifier had an excellent performance on pattern recognition, with a misjudgment rate of about 5%.

## References

- Bottou L, Cortes C, Denker J, et al. Comparison of classifier methods: A case study in handwritten digit recognition. Proceedings of the 12th International Conference on Pattern Recognition and Neural Networks, Jerusalem, 1994. 77-82
- Carrasco J A. An algorithm to find minimal cuts of coherent fault-trees with event-classes. IEEE Transactions on Reliability. 1999. 48(1): 31-41
- Casdagh M. Nonlinear prediction of chaotic time series. Physica D: Nonlinear Phenomena. 1989. 35(3): 335-356
- Chen L and Huang J. Motor broken rotor bar fault diagnosis with support vector machine. Transactions of China Electrotechnical Society. 2006. 21(8): 48-52 (in Chinese)
- Cherkassky V and Ma Y. Practical selection of SVM parameters and noise estimation for SVM regression. Neural Networks. 2004. 17(1): 113-126
- Chua K S. Efficient computations for large least square support vector machine classifiers. Pattern Recognition Letters. 2003. 24(3): 75-80
- Du J and Hou Y. Multi-fault classifier with kernels and its applications. Journal of China Coal Society. 2007. 32(1): 103-106 (in Chinese)
- Foley W L and Svinos J G. Expert advisor program for rod pumping. Journal of Petroleum Technology. 1989. 41(4): 394-400
- Fu G, Zhang B and Ruo H. Lifting technology for reciprocating magnetic drive plunger pump. Acta Petrolei Sinica. 2006. 27(4): 122-124 (in Chinese)
- Gibbs S G and Neely A B. Computer diagnosis of down-hole conditions in sucker rod pumping. Journal of Petroleum Technology. 1966. 18(1): 91-98
- Krebel U. Pairwise classification and support vector machines. In: Scholkopf B, Burges C J C and Smola A J, ed. Advances in Kernel Methods: Support Vector learning, Cambridge. MA: MIT Press. 1999. 255-268
- Lu J, Zhu Y, Zhao H, et al. Power system transient stability assessment based on boosting Bayesian classifier. Transactions of China Electrotechnical Society. 2009. 24(5): 179-181 (in Chinese)
- Luo H, Wang Y and Cui J. A method of analog circuit two-layer fault feature extraction based on a fractional Hilbert transform. Transactions of China Electrotechnical Society. 2010. 25(6): 150-154 (in Chinese)
- Luts J, Ojeda F, Van de Plas R, et al. A tutorial on support vector machine-based methods for classification problems in chemometrics. Analytica Chimica Acta. 2010. 665(2): 129-145
- Muller K, Mika S, Ratsch G, et al. An introduction to kernel based learning algorithms. IEEE Transactions on Neural Networks. 2001. 12(2): 181-201
- Nazi G M, Ashenayi K, Lea J F, et al. Application of neural artificial network to pump card diagnosis. SPE Computer Application. 1994. 6(6): 9-14 (paper SPE 25420)
- Pazzani M J. Failure driven learning of fault diagnosis heuristics. IEEE Transactions on Systems, Man and Cybernetics. 1987. 17(3): 380-384
- Polycarpou M and Trunov A. Learning approach to nonlinear fault diagnosis: Detectability analysis. IEEE Transactions on Automatic Control. 2000. 45(4): 806-812
- Polycarpou M M. Fault accommodation of a class of multivariable nonlinear dynamical systems using a learning approach. IEEE Transactions on Automatic Control. 2001. 46(5): 756-742

- Rossini W M, Alvarenga B, Chabu I E, et al. New concept for lifting in onshore oil wells. *IEEE Transactions on Industry Applications*. 2008. 44(4): 951-961
- Schölkopf B, Mika S, Burges C J C, et al. Input space versus feature space in kernel-based methods. *IEEE Transactions on Neural Networks*. 1999. 10(5): 1000-1017
- Vapnik V N. An overview of statistical learning theory. *IEEE Transactions on Neural Networks*. 1999. 10(5): 988-999
- Vemuri K K, Duggan J B and Sullivan K J. Automatic synthesis of fault trees for computer-based systems. *IEEE Transactions on Reliability*. 1999. 48(4): 394-402
- Wu H, Xing Z and Shu P. Theoretical and experimental study on indicator diagram of twin screw refrigeration compressor. *International Journal of Refrigeration*. 2004. 27(4): 331-338
- Wu L, Zhu Y and Yuan L. Novel method for integrated diagnosis of transformer faults based on Bayesian network classifier. *Transactions of China Electrotechnical Society*. 2005. 20(4): 45-51 (in Chinese)
- Youn W C and Hammen J M. Aiding the operation during novel fault diagnosis. *IEEE Transactions on Systems, Man and Cybernetics*. 1998. 18(1): 142-147
- Yu D, Qi W, Deng S, et al. Optimization frequency of stroke based on submergence depth SVM forecasting of submersible pump. *Proceedings of the CSEE*. 2011a. 31(27): 138-144 (in Chinese)
- Yu D, Qi W, Deng S, et al. Submergence forecasting of a submersible plunger pump based on the support vector machine. *Acta Petrolei Sinica*. 2011b. 32(3): 534-538 (in Chinese)
- Zhang X D, Parisini T and Polycarpou M M. Sensor bias fault isolation in a class of nonlinear systems. *IEEE Transactions on Automatic Control*. 2005. 50(3): 370-376
- Zhang X. Introduction to statistical learning theory and support vectormachines. *Acta Automatica Sinica*. 2000. 26(1): 32-44 (in Chinese)
- Zhou J and Bennett. Adaptive error compensation for fault detection. *International Journal of System Science*. 1998. 29(1): 57-64

(Edited by Sun Yanhua)