

Research Article

Electric-Parameter-Based Inversion of Dynamometer Card Using Hybrid Modeling for Beam Pumping System

Jingliang Wei  and Xianwen Gao 

College of Information Science and Engineering, Northeastern University, Shenyang, China

Correspondence should be addressed to Xianwen Gao; gaoxianwen@mail.neu.edu.cn

Received 11 March 2018; Revised 25 May 2018; Accepted 30 May 2018; Published 27 June 2018

Academic Editor: Akhil Garg

Copyright © 2018 Jingliang Wei and Xianwen Gao. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Conventional fault diagnosis and production calculation of an oil well can be conducted with the surface dynamometer cards, which are obtained by load sensor installed on the horse head. This method to measure the dynamometer cards is limited by the sensor maintenance and calibration, battery replacement, and safety hazards for staff. As the basic parameter of the oil extraction industry, electric parameters have the advantages of low cost and high efficiency. So the inversions of dynamometer card with electric parameters are attracting more and more attention. In order to solve the problem of insufficient data and consider the real-time performance in the actual oil extraction process, this paper proposes a novel hybrid model which consists of two parts: the mechanism model of polished rod load and the suspension displacement calculated with the space vector equations of motor and a data-dependent kernel online sequential extreme learning machine (DDKOS-ELM) model proposed to correct the output error of the mechanism model, which improves the kernel function selection and makes it real-time. Thus, the highlights of this paper can be summed up in two points: (1) under the circumstance of the bottom dead point detection without sensors, the mechanism modeling of the polished rod load and suspension displacement has been carried out from the perspective of mathematical model of AC motor; (2) a novel data-driven model based on data-dependent kernel online sequential extreme learning machine (DDKOS-ELM) has been proposed to improve the kernel functions selection. The coefficients in the data-dependent kernel function are optimized with improved free search algorithm (IFSA). The proposed hybrid model has been applied to a normal working oil well and the prediction results show better accuracy than the pure data-driven model and mechanism model.

1. Introduction

The beam pumping unit plays a predominant part in oil extraction, which is the major artificial lift method and employed by virtually more than 80% oil wells all over the world [1]. As one of the key data to identify the various downhole operation conditions, dynamometer card puts forward an immense influence on the oil well fault diagnosis. It describes the relationship between displacement and load of the polished rod in each stroke and fully reflects the oil exploitation status of the pumping unit. Accordingly, the real-time measurement of dynamometer card has invariably been the major topic of discussion in oil well diagnosis. The commonly used methods of measuring the dynamometer card in actual production are realized by directly mounting

load sensor. The existing detection methods of dynamometer card have the following defects:

(1) The installation of the measuring equipment of the dynamometer card requires the unloading of the suspension point, the operation is cumbersome, and there are safety risks.

(2) Generally, the overall dynamometer card measuring instrument uses battery power which requires large amount of daily maintenance. Because the production will be affected when the well is forced to stop working by the replacement of the battery and the installation of the load sensor, so the measured intervals are usually relatively long, and the downhole conditions cannot be reflected in real time. Therefore, the fault diagnosis of the pumping well may not be timely.

(3) The load of the suspension point is measured by the pressure strain gauge which must be calibrated periodically.

Besides, the ground dynamometer card tester mainly reflects the underground working condition. It is difficult to fully reflect many important operation indicators such as the balance of pumping unit, the system efficiency, energy consumption, and motor running status. This leads to a single measurement function and low cost performance.

As the significant reference variables which are the direct and real-time reflection of the working condition of the beam pumping system, the induction motor electronic parameters are often ignored by the diagnostic staff. There are some calculation methods for the transformation from motor parameters to dynamometer card in recent years. Li et al. [2] presented a method for drawing the dynamometer card based on the actual input power curve of the motor, but it needs to estimate the location of the bottom dead point and thus decreases the accuracy of the calculation. Qi [3] realized the indirect measurement of the dynamometer card of the pumping unit by combining the rotating speed method with the mechanical characteristic curve of the pumping unit, but the modified engineering formula is used to calculate the electromagnetic torque, which affects the accuracy of the method. Besides, the mechanical characteristic curve of the pumping unit is obtained from the load sensor installed on the polished rod. Zhang et al. [4] detected the dead point position of polished rod by installing two microswitches near the link of the pumping beam and bracket and then got the corresponding relation between the power of the motor and the displacement of the polished rod. Because of the mechanical friction between the beam and the microswitches, the accuracy will also be affected. Guo and Ma [5] established a mathematical model containing useful motor power, polished rod load, and polished rod displacement, which can figure out the relationship between the beam pumping unit output and the load. Luo et al. [6] developed a multiple factor nonlinear mathematical model to analyze the operating performance of the induction motor system and introduced the field-circuit coupling equation of induction motors and the mechanical model of the four-bar linkage system. The mechanism models above all have some uncertain environment factors that are hard to precisely describe because of the real-time changes with undulations of the pumping units' operating conditions. Consequently, the data-driven methods have captured increasingly researchers' attention. Li and Han [7] proposed a serial hybrid model of motor load torque by establishing a data-driven model of underground friction. Considering the torque factor is zero when the polished rod is at the top and bottom dead centers, Zhang et al. [8] proposed a power-dynamometer card inversion model based on big data technologies by collecting the measured power curves and corresponding dynamometer cards of typical wells under different working conditions within a period of time as the training data for deep belief network. However, in the case of insufficient data, the data-drive model is difficult to achieve excellent performance. After the above discussion, we adopt the hybrid modeling by combining the mechanism model with the data-driven model in order to solve the problem of dynamometer card inversion. Nowadays, hybrid modeling algorithms have been successfully put into use in a multitude of complex industry

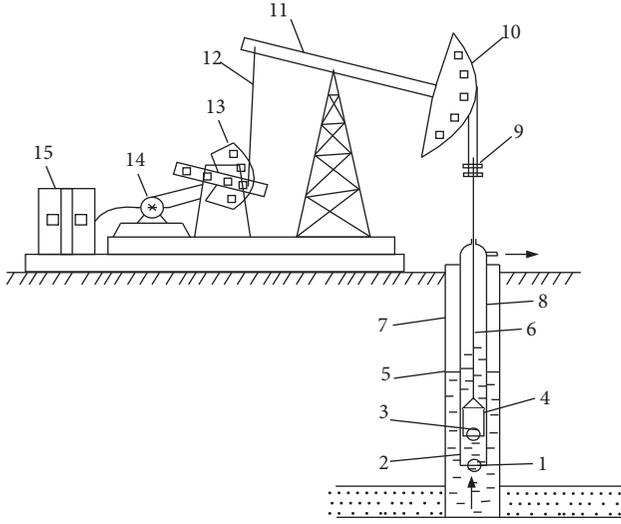
processes [9–22], and they mostly consist of serial structure and parallel structure. The serial hybrid model firstly performs data-driven modeling for the uncertain parameters, and then the outputs of data-driven model are taken as the input of the mechanism model, while the outputs of the parallel hybrid model are the sum of the outputs of data-driven model and mechanism model, and the data-driven model is built to compensate the error of the mechanism model. The latter is adopted as the method in this paper.

In consideration and summary of above discussion, our main contributions in this paper can be summed up as follows. For the part of the mechanism model, the relationship among voltage, current, and torque in the stationary second rectangular coordinate system was derived based on the mathematical model of AC induction motor. The instantaneous torque of AC induction motor was identified according to the stator line voltage and the line current of the motor, and then the polished rod load was calculated from the motor torque in consideration of the crank weight. Finally, the corresponding relation between the polished rod load and suspension displacement is obtained from the signal of the crank rotation angle. For the part of the data-driven model, the inputs of it are the mechanism and electric motor parameters including stroke, the crank radius, stroke frequency, pulley diameter, unbalanced weight, crank quality, balance block weight, two line voltages and line currents of the stator, motor speed, and other technical parameters, and the outputs are the error estimation between the actual measured values and the outputs of the mechanism model.

2. Process Description

The typical structure diagram of beam pumping units is shown as Figure 1, according to the working principle of oil extraction, the oil pump barrel is placed under the oil surface through tubing, with a standing valve installed on the bottom of it; piston is placed into the oil pump barrel through tubing from rod and with a travelling valve installed on the top of it. The sucker rod is connected to a polished rod suspended from the “horse head”. Under the influence of the crankshaft and four connecting rod of the pumping unit, the motor's rotational motion is turned into a reciprocating movement up and down, and then the pump is driven to extract oil. The load of the beam hanger is called polished rod load, which reflects the load of sucker rod column and the liquid column above the piston. In the upper stroke process, the sucker rod is pulled up to pull the plunger, the travelling valve is closed, the standing valve is opened, the oil is pumped into the pump, and the polished rod load is equal to the sum of the weight of sucker rod and the weight of liquid above the plunger, while in the lower stroke process, the sucker rod pushes down the plunger, the standing valve is closed first, then the travelling valve is pushed open under the pressure, and the pump discharges the oil.

In a stroke, the closed curve mapped out by the load and displacement of the suspension point is known as the dynamometer card, and the area of the curve is the size of the work done by the pump in a single stroke.



1-Standing valve, 2-Pump barrel, 3-Travelling valve, 4-Plunger, 5-Dynamic fluid level, 6-Sucker rod, 7-Casing, 8-Tubing, 9-Beam hanger, 10-Horse head, 11-Beam, 12-Link rod, 13-Crank, 14-Induction motor, 15-Electronic control cabinet

FIGURE 1: Typical structure diagram of the beam pumping units.

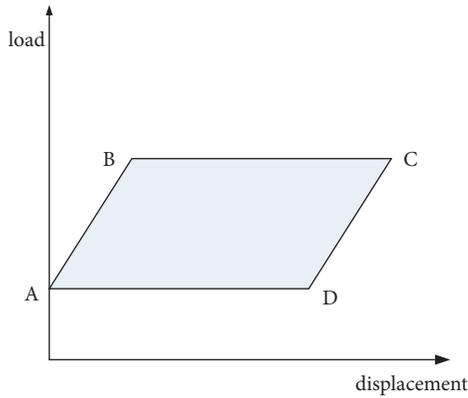


FIGURE 2: Theoretical dynamometer card.

When only considering the liquid column static load above the rod string and the plunger surface and the elastic deformation of sucker rod and tubing, regardless of the sand, wax, gas, and interference of other environmental factors, the curve is obtained as the theoretical dynamometer card, which is similar to parallelogram. As shown in Figure 2, line AB is the loading section, which starts at the moment when the plunger reaches the bottom dead point after ascending, while the travelling valve is closed, the plunger pushes the liquid column above it move upward, and the weight is loaded on the polished rod through the sucker rod, elastic deformation of tubing and sucker rod occurs, the tubing is shortened, and the sucker rod elongates. During this period, although the polished rod is moving upward, the plunger relative to the pump barrel does not produce relative changes.

When the load on the polished rod increases enough to pull the plunger, the elastic deformation of tubing and sucker

rod ends, the high load BC section starts drawing, at this time the load of the polished rod stays the same, the plunger is pulled upward, standing valve is opened, and the pumped oil takes up the pump space until the top dead point.

In unloading section CD, the plunger reaches the top dead point and starts to go down, the travelling valve is opened and the standing valve is closed, elastic deformation of tubing and sucker rod occurs again, the tubing is elongated, and the sucker rod is shortened. During this time, although the polished rod is moving downward, the plunger relative to the pump barrel does not produce relative changes either.

When the load on the polished rod is reduced to the same as the weight of the plunger and the polished rod, the elastic deformation of tubing and sucker rod ends, the low load DA section starts drawing, at this time the load of the polished rod remains the same, the plunger and the polished rod move downward in a constant speed, travelling valve is opened, and the pumped oil flowed into the plunger from the pump barrel.

It is not hard to conclude from the oil pumping process introduced above that different shapes of the dynamometer cards may cause corresponding changes because of the combined variation of gas, liquid, and machinery.

3. Mechanism Modeling of Beam Pumping Units

3.1. Pumping Motor. In the actual oil production field, the induction motor is employed as the main equipment for the oil pumping motor, and the space vector equation of ac motor can be described as follows [23]:

$$\begin{aligned} u_s &= R_s i_s + \frac{d\psi_s}{dt} \\ u_r &= R_r i_r + \frac{d\psi_r}{dt} - j\omega_r \psi_r \\ \psi_s &= L_s i_s + L_m i_r \\ \psi_r &= L_m i_s + L_r i_r \end{aligned} \quad (1)$$

where u_s is the stator voltage space vector, u_r is the rotor voltage space vector, i_s is the stator current space vector, i_r is the rotor current space vector, ψ_s is the stator magnetic chain space vector, ψ_r is the rotor magnetic chain space vector, R_s is the resistance of motor stator, R_r is the resistance of motor rotor, L_s is the stator self-inductance, L_r is the rotor self-inductance, L_m is the mutual inductance of rotor to stator winding, and ω_r is the angular velocity of rotor.

Then the α axis of two phase static coordinates (α - β coordinate system) is overlapped to the A axis of three-phase static coordinate system (ABC coordinate system) and the β axis of two phase static coordinates lag α axis 90° , and then the electromagnetic torque, synchronous angular velocity, and electromagnetic power in the α - β coordinate system can be obtained by coordinate transformation as follows:

$$T_e = p_n (\vec{\psi}_s \times \mathbf{i}) = p_n (\psi_{s\alpha} i_{s\beta} - \psi_{s\beta} i_{s\alpha}) \quad (2)$$

$$\omega_s = \frac{(u_{s\beta} - R_s i_{s\beta}) \psi_{s\alpha} - (u_{s\alpha} - R_s i_{s\alpha}) \psi_{s\beta}}{\psi_s^2} \quad (3)$$

$$P_e = T_e \Omega_s = \frac{T_e \omega_s}{p_n} \quad (4)$$

where p_n is the pole logarithm, ω_s is the synchronous motor angular velocity, Ω_s is the mechanical angular velocity, T_e is the electromagnetic torque, and P_e is the electromagnetic power. The line voltage and line current component of the α - β coordinate system are obtained by the phase transformation of the stator line voltage and line current directly measured in the electrical parameter acquisition process.

Considering the mechanical loss of the motor, the output torque and output power of the motor can be obtained from the following formulas:

$$\omega = \frac{R_r T_e}{p_n \Psi_s^2} \quad (5)$$

$$T_m = \frac{(1-s)P_e - p_0}{(\omega_s - \omega)/p_n} \quad (6)$$

where ω is the angular velocity of the rotor slip, T_m is the output torque of motor, s is the slip ratio, and p_0 is the mechanical loss of the motor.

3.2. The Movement Rule of Pumping Units. The surface structure of the beam pumping unit consists of drive motor, reduction gearbox, and four-bar linkage. For the four-bar linkage mechanism, the fixed rod is the connection between the beam support point and the crankshaft center, and the three movable rods are crank, connecting rod, and backward beam. As shown in Figure 3, R is the radius of crank, L_c is the length of the connection rod, L_b is the length of backward beam, L_f is the length of forward beam, K is the length of the fixed rod, Ω denotes the mechanical angular velocity, O and O_1 are reference points, and H is the vertical height of the support point to the center of the gearbox. Setting the clockwise direction positive, the motor parameters are collected at 12 o'clock of crank, and θ_o denotes the angle of crank spinning clockwise.

The angle between the backward beam and connecting rod can be calculated by applying the cosine theorem to triangle $\Delta_{L_b L_c J}$ and Δ_{KRJ} :

$$\beta = \cos^{-1} \frac{L_b^2 + L_c^2 - K^2 - R^2 + 2KR \cos \theta}{2L_b L_c} \quad (7)$$

$$P = \frac{T_m \times i_{belt} \times i_{reduction\ gearbox} \times (\eta_{belt} \times \eta_{reduction\ gearbox})^\sigma - (W_{cb} r + W_c r_c) \sin \theta_0}{\overline{TF}}, \quad \sigma = \begin{cases} 1 & T_m > 0 \\ -1 & T_m < 0 \end{cases} \quad (12)$$

where $\overline{TF} = (L_f/L_b)r(\sin \alpha/\sin \beta)$ is the torque factor, i_{belt} and η_{belt} are transmission ratio and transmission efficiency of

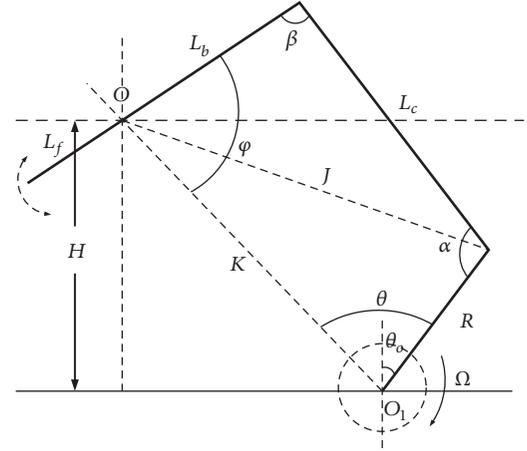


FIGURE 3: Four-bar linkage mechanism.

The angle between the crank and connecting rod can be calculated as follows:

$$\alpha = 360^\circ - (\beta + \varphi + \theta) \quad (8)$$

The angles of φ when the horse head is at the top and bottom dead points are obtained as follows, respectively:

$$\varphi_{bottom} = \cos^{-1} \frac{L_b^2 + K^2 - (L_c + R)^2}{2L_b L_c} \quad (9)$$

$$\varphi_{top} = \cos^{-1} \frac{L_b^2 + K^2 - (L_c - R)^2}{2L_b L_c} \quad (10)$$

Counting from the bottom dead point (12 o'clock of crank), the motor parameters are collected every 40ms, φ can be calculated by multiplying time by the motor speed, and then the displacement is expressed as a percentage of the stroke:

$$S = \frac{\varphi_{bottom} - \varphi}{\varphi_{bottom} - \varphi_{top}} S_C \quad (11)$$

where S is the displacement of suspension point and S_C is the stroke of the sucker rod pumping system.

When the beam pumping unit is working, the torque synthesized on the crankshaft by the suspension point load and the balance weight is in equilibrium with the output torque of motor.

the belt, $i_{reduction\ gearbox}$ and $\eta_{reduction\ gearbox}$ are transmission ratio and transmission efficiency of the reduction gearbox,

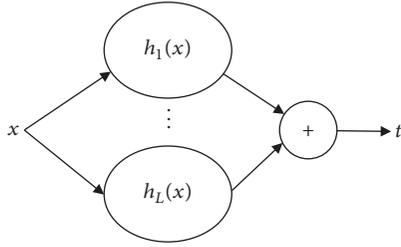


FIGURE 4: Structure diagram of basic ELM model with single input and output.

W_{cb} and W_c are the weight of the crank balance block and the crank, r_c is the crank radius, and P is the load of the suspension point.

Thus, the relation between the displacement and the crank angle θ_0 is obtained from (11), and the load of the suspension point can be calculated by (12). Then the curve of dynamometer card can be obtained with variable substitution according to the displacement and load curve of the crank angle θ_0 .

4. Data-Dependent Kernel Online Sequential Extreme Learning Machine

It can be observed from the description in Section 3 that there are many factors affecting the accuracy of the mechanism model such as the flux estimation, system power loss, sliding, and viscous frictions. Considering the related data information, we employ the data based regression algorithm to compensate the mechanism model by training the deviation between the output of the mechanism model and the measured value.

4.1. Kernel Extreme Learning Machine. Extreme Learning Machine (ELM) [24] was initially proposed for single-hidden-layer feedforward neural networks (SLFNs) by Huang and extended to the generalized single-hidden-layer feedforward neural networks where the hidden layer needs not to be neuron alike. As a supervised algorithm, its parameters are randomly generated, and then the output weights can be analytically determined by the generalized inverse operation. For a set of training samples $\{(x_i, t_i) \mid x_i \in \mathbf{R}^d, t_i \in \mathbf{R}^m, i = 1, 2, \dots, N\}$, there is a nonlinear relationship between the input \mathbf{x}_i and the output \mathbf{t}_i . This nonlinear projection relationship can be expressed as a linear projection in the feature space according to the model of ELM. The outputs of ELM with L hidden layer neuron nodes are mathematically formulated as

$$t_i = f_L(x_i) = \sum_{j=1}^L \beta_j G(a_j, b_j, x_i), \quad i = 1, 2, \dots, N \quad (13)$$

where $a_j \in \mathbf{R}^n$ and $b_j \in \mathbf{R}$ are the parameters of the hidden layer of SLFN and $\beta_j \in \mathbf{R}^m$ is the weight which connects the j th hidden layer node to the output layer node. $G(a_j, b_j, x_i) = g(a_j x_i + b_j)$ is the output of the j th hidden layer node with

respect to the input x_i . The activation functions $g(\cdot)$ can be sigmoid as well as the radial biases, sine, cosine, wavelet, and many other piecewise continuous computational functions. The graphic architecture of the basic ELM model with single input and single output is shown as Figure 4.

Then the hidden layer output matrix H is written as

$$H = \begin{bmatrix} h(x_1) \\ \vdots \\ h(x_N) \end{bmatrix} = \begin{bmatrix} G(a_1, b_1, x_1) & \cdots & G(a_L, b_L, x_1) \\ \vdots & \ddots & \vdots \\ G(a_1, b_1, x_N) & \cdots & G(a_L, b_L, x_N) \end{bmatrix} \quad (14)$$

The training purpose of ELM model is to find the optimal output weight matrix β with the following formula:

$$\beta = \mathbf{H}^+ \mathbf{T} \quad (15)$$

where $\mathbf{T} = [\mathbf{t}_1, \mathbf{t}_2, \dots, \mathbf{t}_N]^T$ denotes the output matrix and \mathbf{H}^+ is the Moore-Penrose generalized inverse of matrix. In order to calculate the norm of the output weights, we minimize the training error as follows:

$$\min_{\beta \in \mathbf{R}^m} \frac{1}{2} \|\beta\|^2 + \frac{C}{2} \sum_{i=1}^N \|\varepsilon_i\|^2 \quad (16)$$

$$\text{s.t. } \varepsilon_i^T = t_i^T - h(x_i) \beta, \quad i = 1, \dots, N \quad (17)$$

After solving the quadratic optimization problem above, the output of the ELM is obtained as follows:

$$f(x) = \begin{cases} h(x) H^T \left(\frac{I}{C} + H H^T \right)^{-1} T, & N < L \\ h(x) \left(\frac{I}{C} + H^T H \right)^{-1} H^T T, & N \geq L \end{cases} \quad (18)$$

If the nonlinear mapping function $h(\cdot)$ is unknown to users, the corresponding ELM kernel matrix is given to users as replacement. The kernel matrix for ELM is defined as follows:

$$K_{ELM, i, j} = H H^T = h(x_i) \cdot h(x_j) = K(x_i, x_j) \quad (19)$$

Then the output function of kernel-based ELM can be expressed as follows:

$$y = f(x) = \begin{bmatrix} K(x, x_1) \\ \vdots \\ K(x, x_N) \end{bmatrix}^T \left(\frac{I}{C} + K_{ELM} \right)^{-1} T \quad (20)$$

Apparently, the selection and construction of kernel functions greatly influence the performance of K-ELM. Besides, as the volume of datasets increases, the training time complexity of $O(N^3)$ and the kernel matrix size of $O(N^2)$ become an inevitable problem to concern.

4.2. Online Sequential Extreme Learning Machine with Kernels. In the actual oil extraction process, the measured dynamometer cards data is arriving in a streaming fashion in the actual oil extraction process. Therefore, in order to ease the computation complexity of the inverse matrix in (18) and extend the kernel extreme learning machine to the online version, Simone *et al.* [25] proposed a kernel online sequential extreme learning machine(KOS-ELM) by straightforwardly extending the kernel recursive least squares to the online sequential extreme learning machine(OS-ELM) framework.

According to the OS-ELM, the dataset is divided into successive minibatches X_i , $i = 1, 2, \dots, B$, so that $X = \cup_{i=1}^B X_i$. Similar to OS-ELM, the first approximation of the weight from the first minibatch is computed as follows for KOS-ELM:

$$P'_1 = (C^{-1} + k(x_1, x_1))^{-1} \quad (21)$$

$$\beta'_1 = P'_1 y_1 \quad (22)$$

In combination with recursive least squares (RLS), at each new minibatch k , the output weights are updated as

$$k_{k+1} = [k(x_{k+1}, x_1), \dots, k(x_{k+1}, x_k)]^T \quad (23)$$

$$z_{k+1} = P_k k_{k+1} \quad (24)$$

$$r_{k+1} = C^{-1} + k(x_{k+1}, x_{k+1}) - z_{k+1}^T k_{k+1} \quad (25)$$

$$P_{k+1} = r_{k+1}^{-1} \begin{bmatrix} P_k r_{k+1} + z_{k+1} z_{k+1}^T & -z_{k+1} \\ -z_{k+1}^T & 1 \end{bmatrix} \quad (26)$$

$$\beta_{k+1} = \begin{bmatrix} \beta_k - z_{k+1} r_{k+1}^{-1} e_{k+1} \\ r_{k+1}^{-1} e_{k+1} \end{bmatrix} \quad (27)$$

$$e_{k+1} = y_{k+1} - k_{k+1}^T \beta_k \quad (28)$$

For each sample (x_i, y_i) , e_i denotes the error of the filter before the update and $-r_i^{-1} e_i$ is the corresponding coefficient of the center x_i . And all the previous coefficients are updated by a factor $-z_i r_i^{-1} e_i$.

4.3. Data-Dependent Kernel. In the traditional kernel ELM for regression, the selection of kernel function mostly depends on experience and the optimization of kernel parameters generally depends on intelligent optimization algorithms such as genetic algorithm, clonal selection algorithm, and particle swarm optimization algorithm. To a certain extent, these methods improve the performance of kernel based learning algorithms. However, the impact of data is not taken into account, the distribution structure of data in the kernel mapping space does not change and the performance of kernel learning algorithm cannot be fundamentally raised to higher level because the kernel function is fixed in different training samples' structures. Thus, a novel data-dependent kernel learning method is proposed to improve the performance by replacing the basic kernel with a data-dependent

kernel function [26–28]. Through conformal transformation of the basic kernel, the data-dependent kernel is defined as

$$k(x, x') = c(x) k_0(x, x') c(x') \quad (29)$$

where $K_0(x, x')$ is a basic kernel function such as the sigmoid kernel or the Gaussian kernel and the positive value function of x , $c(x)$ is defined as

$$c(x) = b_0 + \sum_{i=1}^N b_i k'(x, a_i) \quad (30)$$

where the whole samples are chosen as “empirical cores” $\{a_i = x_i \in R^d, i = 1, 2, \dots, N\}$, N is the number of samples, b_i represents the combination expansion coefficients of a_i , $k(x, a_i) = e^{-\delta \|x - a_i\|^2}$, and δ is a free parameter. Thus, the geometrical structure of the data in the kernel mapping space is determined by the expansion coefficients b_i ($i = 1, 2, \dots, N$). The kernel matrices corresponding to k and k_0 are denoted as K and K_0 , respectively. Consequently, (29) can be rewritten as

$$K = CK_0C \quad (31)$$

where the diagonal matrix $C = \text{diag}\{c(x_1), c(x_2), \dots, c(x_N)\}$. After defining $c = [c(x_1), c(x_2), \dots, c(x_N)]^T$ and $b = [b_1, b_2, \dots, b_N]^T$, we have

$$c = \begin{bmatrix} 1 & k'(x_1, a_1) & \dots & k'(x_1, a_N) \\ 1 & k'(x_2, a_1) & \dots & k'(x_2, a_N) \\ \vdots & \vdots & \ddots & \vdots \\ 1 & k'(x_N, a_1) & \dots & k'(x_N, a_N) \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_N \end{bmatrix} = K_N b \quad (32)$$

4.4. Kernel Optimization. There are many methods to optimize the combination expansion coefficients and the Gaussian kernel parameter σ^2 . For two key uncertain parameters, the improved free search algorithm (IFSA) [29] is adopted to address the optimal solution, which is put forward to find the optimal objective function value with the following equations:

$$f_{tj} = f(x_{tji}), \quad (33)$$

$$f_j = \min(f_{tj})$$

$$Ph_j = \frac{\min(f_j)}{f_j} \quad (34)$$

f denotes the position where an individual completes a search step; t denotes the search step; j denotes the number of individuals; and i denotes the search spatial dimension.

As a major parameter of the algorithm, the search radius R_{ji} is adjusted in real time. The initial value of it is set to be 1, and with the growth of the search step, the radius shrinks as (35). In addition, the sensitivity parameter determines the scope of the pheromone value and also confirms the initial

Input: The initial population size M , max algebra G , search step length T , the parameters σ^2 and b ;
Output: The optimal parameter σ^2 and b ;
Step 1: Initialization
 1.1 Normalize the input parameters σ^2 and b with the range of $[-1, 1]$;
 1.2 The initial pheromone is obtained with the initial value, which is released at the same time;
 1.3 The initial search results are obtained.
Step 2: Searching process
 2.1 Calculate the sensitivity parameter with (36) and (37);
 2.2 Confirm the new search point of the next search round using the sensitivity parameter;
 2.3 Record the RMSE value of the error between the predicted value and the actual value as the fitness function. Fitness function is defined as follows;

$$f_j = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$$

 2.4 The new search results can be obtained by (33) and (34);
 2.5 Preserve the best individuals;
 2.6 Adjust the search radius according to (35).
Step 3: Termination condition judgement
 When the termination condition is satisfied, output the optimal parameter σ^2 and b .

ALGORITHM 1: Improved free search algorithm.

Input: Training samples $\{(x_i, t_i) \mid x_i \in \mathbf{R}^d, t_i \in \mathbf{R}^m, i = 1, 2, \dots, N\}$, the Gaussian kernel function is selected as the basic kernel function $K_0(x, x') = \exp(-\|x - x'\|^2/2\sigma^2)$, the nonzero constant C , the number of hidden nodes L and the number of data blocks B in a data set.
Output: The predictive value of the polished rod load and suspension point displacement
 (1) Randomly generate the input layer parameters (a_j, b_j) ($j = 1, 2, \dots, L$);
 (2) Construct the data-dependent kernel function $K(x, x')$ by modifying the initial basic kernel function according to (29), whose parameters are optimized by Algorithm 1;
 (3) **for** each $k \in [1, B]$ **do**
 (4) Update the output weight β by (27);
 (5) **end for**
 (6) Compute the predictive value according to (20).

ALGORITHM 2: IFS-DDKOS-ELM.

search point of the next search round, so the modification of sensitivity parameters is shown as (36) and (37).

$$R_{ji} = \begin{cases} \rho R_j(t - g), & R_j(t) \geq R_{\min} \\ R_{\min}, & \text{else} \end{cases} \quad (35)$$

$$Se_j = Se_{\min} + \Delta Se_j \quad (36)$$

$$\Delta Se_j = (Se_{\max} - Se_{\min}) \times \frac{t - G}{G} \quad (37)$$

Thus, the implementation steps of improved free search algorithm can be summarized as Algorithm 1.

By applying the data-dependent kernel to the kernel online sequential extreme learning machine for regression, the training framework of the proposed IFS-DDKOS-ELM error model is summarized as Algorithm 2.

5. Hybrid Model Structure

Based on the above discussion, the parallel hybrid model is constructed in Figure 5, which not only considers the

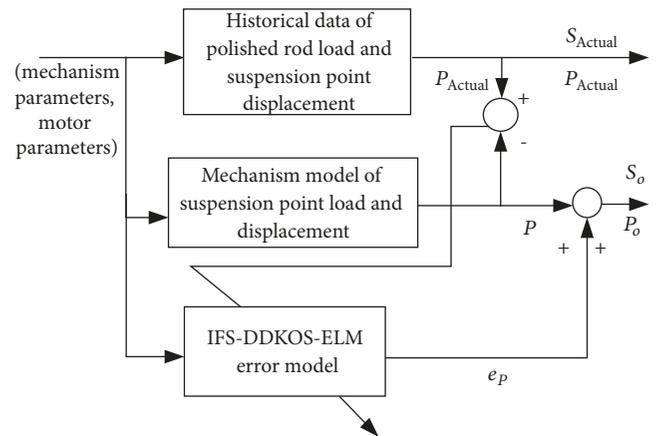


FIGURE 5: The prediction process of hybrid model.

processing technology but also makes full use of relevant data information.

TABLE 1: Technical parameters of the beam pumping unit.

	CYJ10-3-53HB
Number of oil wells	5
Stroke number per minute (min^{-1})	150.82
Transmission ratio of belt-pulley gear box	1.134
Crank radius (m)	3.60
Length of the link rod (m)	3.8
Length of the beam forearm (m)	2.5
Length of the beam backward arm (m)	3.2
Vertical height H (m)	4.53
Distance between the reference points K (m)	0.95
Transmission efficiency	-1.12
Unbalanced weight (kN)	1.40
Balance block weight (kN)	

TABLE 2: Electronic parameters of the induction motor.

	CYJ10-3-53HB
Number of oil wells	25
Rated power (kW)	380
Rated voltage (V)	585
Rated rotor speed ($\text{rad} \cdot \text{min}^{-1}$)	0.36
Stator resistance (Ω)	0.19
Rotor resistance (Ω)	33
Stator inductance (mH)	35
Rotor inductance (mH)	41
Peak value of the mutual inductance between rotor and stator (mH)	1.71
Moment of inertia ($\text{kg} \cdot \text{m}^2$)	0.019
Rotational resistance coefficient	3
Pole-pairs number	0.98
Motor transmission efficiency	

The inputs of the hybrid model are the mechanism and electric motor parameters including stroke, the crank radius, stroke frequency, pulley diameter, unbalanced weight, crank quality, balance block weight, three line voltages and line currents of the stator, motor speed, and other technical parameters. The outputs of the hybrid model are expressed as follows:

$$P_o = P + e_p \quad (38)$$

where load P and displacement S are the outputs of the mechanism model and e_p is the output of the IFS optimized data-dependent kernel extreme learning machine (IFS-DDKOS-ELM), which is the error estimation between the actual measured value and the output of the mechanism model.

6. Case Research and Experimental Results

In this section, the proposed hybrid model is firstly employed to predict the load and the displacement of the suspension point of the beam pumping unit. Secondly, the curves of polished rod load and suspension point displacement with respect to crank angle θ_0 are drawn by the outputs of the hybrid model respectively. Finally, the dynamometer card is drawn according to the curves of the polished load and suspension point through variable substitution. Taking a

normal oil well in Liaohe oil field of China for case research, the technical parameters of the beam pumping unit and the parameters of the induction motor are shown in Tables 1 and 2, respectively.

The data to establish the hybrid model is collected with the indicator instrument, the electric parameters measurement instrument, and two proximity switches. The indicator instrument installed at the suspension point collects the surface dynamometer cards periodically when the beam pumping units are working. The collected dynamometer cards are taken as the reference data and the historical training data for the IFS-DDK-ELM error predictive model. The electric parameters measurement instrument is installed in the electric control cabinet connecting to the induction motor with the sampling interval of 47.8ms, and the collected current, voltage, and power of the induction motor are taken as the input of the hybrid model. One proximity switch is installed at 12 o'clock on the pedestal of the crank shaft to measure the cycle of one stroke, when the crank is rotated to the 12 o'clock position, triggering the proximity switch action, and the timer of the electric parameters measurement instrument starts ticking until the crank is rotated again to 12 o'clock. Therefore, a periodic displacement curve can be obtained. The other one is installed on the motor shaft to measure the motor speed.

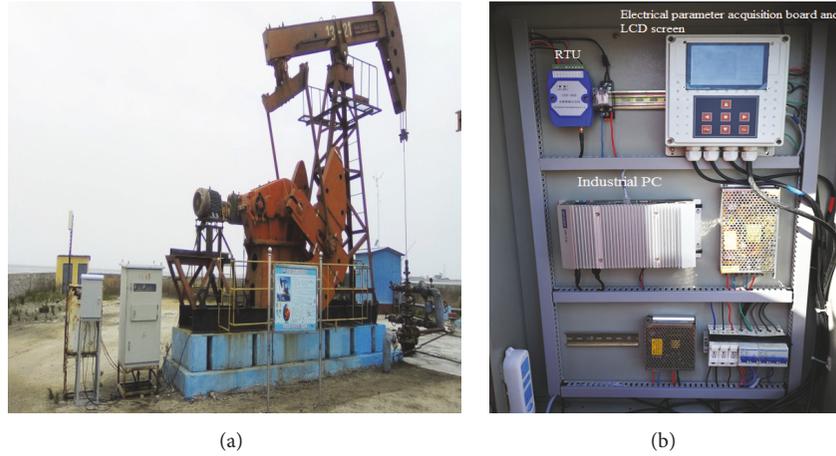


FIGURE 6: The data collection instruments installed behind the electronic control cabinet.

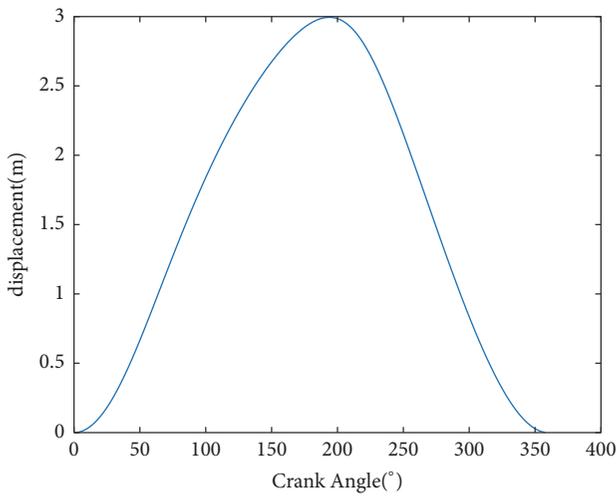


FIGURE 7: The suspension point displacement curve.

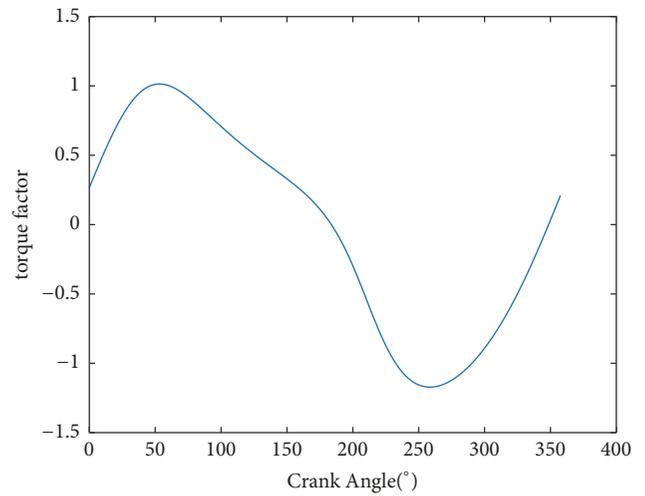


FIGURE 8: The torque factor curve.

The data collection instruments are installed behind the electronic control cabinet (as shown in Figure 6(a)), and it mainly includes the RTU (Remote Terminal Unit), the electrical parameter acquisition board with LCD screen, and the industrial PC (as shown in Figure 6(b)). The RTU is used to receive the surface dynamometer cards collected with the wireless integrated indicator instrument installed at the suspension point of the horse head. The function of the electrical parameter acquisition board is data collection and dynamometer cards inversion; the industrial PC is used to preserve the work data and generate work reports. And the communication mode among the three modules is TCP/IP.

1800 groups of measurement data in total three months from Aug 9, 2017, to Nov 10, 2017, including the electric parameters and the dynamometer cards are collected to establish the data-driven model. Randomly select 1200 groups of the collected data as the training samples and the other 600 groups of data as the test samples. Figure 7 shows the suspension point displacement curve in one stroke, Figure 8 is the torque factor curve, and Figure 9 is the

predictive curve of the polished rod load obtained by the hybrid model. Consequently, the predictive result of dynamometer card is obtained as Figure 10. The test error of maximum load, minimum load, and the area of the dynamometer card obtained by the hybrid model are shown as Table 3. It can be concluded from Figure 10 and Table 3 that the predictive dynamometer card of the proposed hybrid model is basically consistent with the measured dynamometer card and can reflect the actual working operation vividly.

In order to approve the effectiveness of the proposed hybrid prediction method, the mechanism predictive model without data-driven model and pure IFS-DDKOS-ELM model are chosen to compare the prediction performance, and the prediction performances are shown as Figure 11(a). In addition, combination kernel function LSSVM and the KOS-ELM in [25] are selected as the comparison algorithms to verify the effectiveness of the IFS-DDKOS-ELM. The prediction effect of the KOS-ELM in [25] is proved to be better than ELM, K-ELM, and OS-ELM. So this paper directly chooses these two methods to compare the prediction performance

TABLE 3: Comparison of performance indicators between predictive value and actual value.

	Predictive value	Actual value	Test error (%)
Maximum load (kN)	82.87	82.22	0.7905
Minimum load (kN)	44.49	44.32	0.3835
Area of dynamometer card	67.0445	67.2565	0.3152

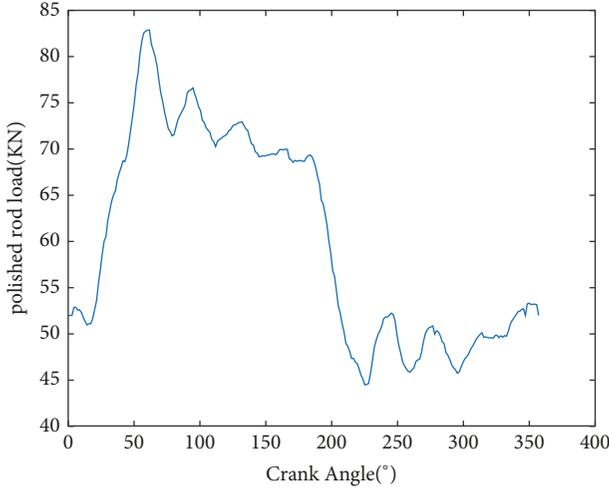


FIGURE 9: Polished rod load curve.

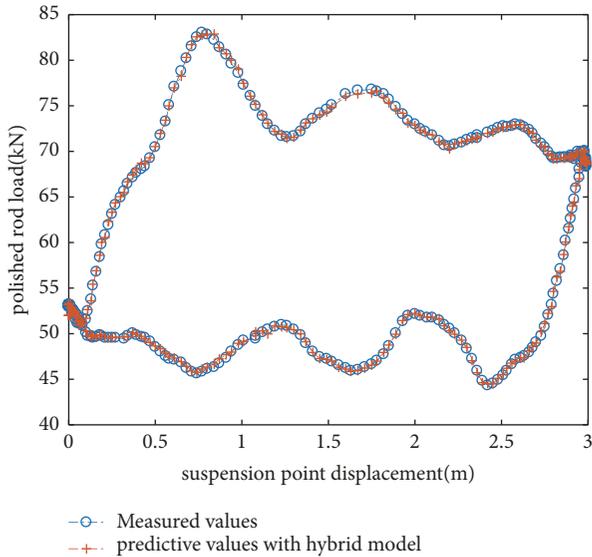


FIGURE 10: The prediction result of dynamometer card.

of the IFS-DDKOS-ELM model. The comparison results are shown in Figure 11(b).

Root mean square error (RMSE) and mean absolute error (MAE) are used as the prediction evaluation indicators, which are defined as follows:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (39)$$

TABLE 4: Comparison of the prediction results with the three models by RMSE and MAE.

Predictive methods	RMSE	MAE
Hybrid model	0.2383	0.1731
Mechanism model	1.4829	1.1775
Pure IFS-DDKOS-ELM	1.4365	1.1780

TABLE 5: Comparison of the prediction results with the three data-driven models by RMSE and MAE.

Predictive methods	RMSE	MAE
IFS-DDKOS-ELM model	1.4365	1.1780
LS-SVM model	3.4873	2.9140
KOS-ELM model	2.5195	2.1492

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (40)$$

where n denotes the number of test data, y_i is the real value of test data, and \hat{y}_i is the prediction value. After operating the algorithms introduced above 50 times, the average is selected as the experimental result, which are shown as Tables 4 and 5.

From the prediction performance of Figure 11(b) and Table 5, it can be known that the IFS-DDKOS-ELM model of this paper has the better prediction effect for dynamometer card inversion than the other two data-driven models. Although the LS-SVM with combination kernel function and the KOS-ELM in [25] show good prediction effect, the prediction accuracy is further improved by employing the optimized data-dependent kernel function. From the prediction performance of Figure 11(a), it can be observed that the proposed hybrid model has better prediction accuracy than the pure IFS-DDKOS-ELM and mechanism models, which also can be concluded from the corresponding RMSE and MAE with different models shown in Table 4. Notably, because of the stable operating condition of the motor, the fluctuation of the prediction dynamometer card with pure mechanism model is relatively weak.

In summary, the experiment results demonstrate that the proposed hybrid model can perform better dynamometer cards inversion effect than the pure mechanism model and the pure data-driven model. And it can produce precise dynamometer cards for fault diagnosis of the beam pumping unit.

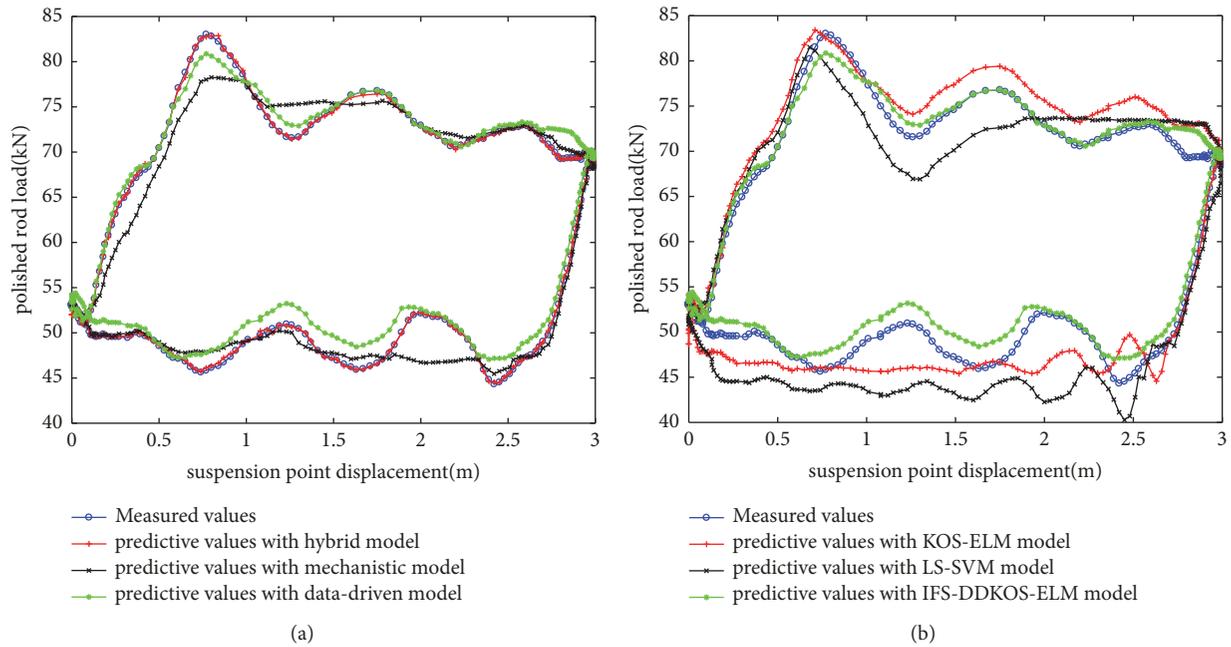


FIGURE 11: The predictive contrast results of dynamometer card.

7. Conclusion

In actual oil production, real-time and accurate collection of dynamometer cards has far-reaching influence on the oil well fault diagnosis. In view of the insufficient data in the actual oilfield, the data collection is not timely. This paper proposes a novel hybrid modeling for dynamometer cards inversion based on the electronic parameters collected from the motor. The simulation results can basically match with the practical measurement dynamometer cards. The advantages of the proposed hybrid model are summarized as follows.

(1) The mechanism model is carried out to indirectly measure the polished rod load and suspension displacement from the perspective of mathematical model of AC motor. Besides, in the process of data collection, the proximity switch is used to realize the cycle identification without sensor.

(2) In order to compensate the error of the mechanism model output, an online learning with a data-dependent kernel-based ELM (DDKOS-ELM) is firstly presented for regression. The comparison results show that the performance of the data-dependent kernel function is proved to be better than the combination kernel function.

Data Availability

As the project will not be finished until December 2019 and we have confidentiality agreement with Liaohe Oilfield, the data could not be released so far. For any information about the article, please contact us via weijingl_neu@163.com.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

Financial support from the National Natural Science Foundation of China under Grants 61573088, 61573087, and 61433004 is acknowledged. The authors are also grateful for the support from the Liaohe Oilfield of China National Petroleum Corporation, providing them with research and experimental conditions.

References

- [1] B. M. Wilamowski and O. Kaynak, "Oil well diagnosis by sensing terminal characteristics of the induction motor," *IEEE Transactions on Industrial Electronics*, vol. 47, no. 5, pp. 1100–1107, 2000.
- [2] L. Hujun and Z. Lianyou, "Predicting dynamometer cards by actual motor power curves," *Petroleum Geology & Oilfield Development in Daqing*, vol. 10, no. 4, pp. 63–67, 1991.
- [3] Q. Xuehao, *The indicator diagram remote monitoring system based on indirect measurement*, Dalian: Dalian University of Technology, 2008.
- [4] Z. Shirong and L. Changxi, "Indirect measurement of dynamometer card of beam pumping unit," *Journal Huazhong University of Science and Technology: Nature Science Edition*, vol. 32, no. 11, pp. 62–64, 2004.
- [5] Y. Guo and L. Ma, "The indicator diagram collection based on beam pumping unit motor effective power measurement," in *Proceedings of the International Petroleum Technology Conference 2013: Challenging Technology and Economic Limits to Meet the Global Energy Demand, IPTC 2013*, Beijing, China, March 2013.
- [6] W. Luo, B. Wang, H. Zhao, and Y. Luo, "Modeling and simulation of non-linear dynamic process of the induction motor system with fluctuating potential loads," *Science China Technological Sciences*, vol. 57, no. 9, pp. 1729–1737, 2014.

- [7] L. Kun and Y. Han, "Modeling for motor load torque with dynamic load changes of beam pumping units based on a serial hybrid model," *Transactions of the Institute of Measurement and Control*, vol. 40, no. 3, Article ID 0142331216670454, 2016.
- [8] X. Zhang, J. Shi, R. Zhao et al., "Research and Application of Electric Power Curve Inverting Dynamometer Diagram Technology Using Big Data Approach," in *Proceedings of the SPE Symposium: Production Enhancement and Cost Optimisation*, Society of Petroleum Engineers, Kuala Lumpur, Malaysia, 2017.
- [9] Y. Chu, F. You, and J. M. Wassick, "Hybrid method integrating agent-based modeling and heuristic tree search for scheduling of complex batch processes," *Computers & Chemical Engineering*, vol. 60, pp. 277–296, 2014.
- [10] Y. R. Lee, H. S. Choi, H. C. Park, and J. E. Lee, "A numerical study on biomass fast pyrolysis process: A comparison between full lumped modeling and hybrid modeling combined with CFD," *Computers & Chemical Engineering*, vol. 82, pp. 202–215, 2015.
- [11] P. Setoodeh, A. Jahanmiri, and R. Eslamloueyan, "Hybrid neural modeling framework for simulation and optimization of diauxic-involved fed-batch fermentative succinate production," *Chemical Engineering Science*, vol. 81, pp. 57–76, 2012.
- [12] J.-S. Wang and N.-N. Shen, "Hybrid multiple soft-sensor models of grinding granularity based on cuckoo searching algorithm and hysteresis switching strategy," *Scientific Programming*, vol. 2015, Article ID 146410, 11 pages, 2015.
- [13] J. Pian and Y. Zhu, "A hybrid soft sensor for measuring hot-rolled strip temperature in the laminar cooling process," *Neurocomputing*, vol. 169, pp. 457–465, 2015.
- [14] J. Zhang, Z.-Z. Mao, R.-D. Jia, and D.-K. He, "Real time optimization based on a serial hybrid model for gold cyanidation leaching process," *Minerals Engineering*, vol. 70, pp. 250–263, 2015.
- [15] J. Zhang, Z.-Z. Mao, R.-D. Jia, and D.-K. He, "Serial hybrid modelling for a gold cyanidation leaching plant," *The Canadian Journal of Chemical Engineering*, vol. 93, no. 9, pp. 1624–1634, 2015.
- [16] X. Wang, Y. Wang, C. Yang, D. Xu, and W. Gui, "Hybrid modeling of an industrial grinding-classification process," *Powder Technology*, vol. 279, pp. 75–85, 2015.
- [17] T. Liu, X.-W. Gao, and L.-N. Wang, "Hybrid modeling method of comprehensive energy consumption for oil and gas production process," *Journal of Northeastern University (Natural Science)*, vol. 34, no. 11, pp. 1525–1528, 2013.
- [18] B. M. Brentan, E. Jr Luvizotto, M. Herrera, J. Izquierdo, and R. Pérez-García, "Hybrid regression model for near real-time urban water demand forecasting," *Journal of Computational and Applied Mathematics*, vol. 309, pp. 532–541, 2017.
- [19] M. Hasanipanah, A. Shahnazari, H. Bakhshandeh Amnieh, and D. Jahed Armaghani, "Prediction of air-overpressure caused by mine blasting using a new hybrid PSO–SVR model," *Engineering with Computers*, vol. 33, no. 1, pp. 23–31, 2017.
- [20] Z. Yang, L. Ce, and L. Lian, "Electricity price forecasting by a hybrid model, combining wavelet transform, ARMA and kernel-based extreme learning machine methods," *Applied Energy*, vol. 190, pp. 291–305, 2017.
- [21] C. M. Chew, M. K. Aroua, and M. A. Hussain, "A practical hybrid modelling approach for the prediction of potential fouling parameters in ultrafiltration membrane water treatment plant," *Journal of Industrial and Engineering Chemistry*, vol. 45, pp. 145–155, 2017.
- [22] S. Salcedo-Sanz, R. C. Deo, L. Cornejo-Bueno, C. Camacho-Gómez, and S. Ghimire, "An efficient neuro-evolutionary hybrid modelling mechanism for the estimation of daily global solar radiation in the Sunshine State of Australia," *Applied Energy*, vol. 209, pp. 79–94, 2018.
- [23] A. K. Gupta and A. M. Khambadkone, "A space vector PWM scheme for multilevel inverters based on two-level space vector PWM," *IEEE Transactions on Industrial Electronics*, vol. 53, no. 5, pp. 1631–1639, 2006.
- [24] G.-B. Huang, H. Zhou, X. Ding, and R. Zhang, "Extreme learning machine for regression and multiclass classification," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 42, no. 2, pp. 513–529, 2012.
- [25] S. Scardapane, D. Comminiello, M. Scarpiniti, and A. Uncini, "Online sequential extreme learning machine with kernels," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 26, no. 9, pp. 2214–2220, 2015.
- [26] X. I. E. Jianhong, "Impact damage detection by LS-SVM with data-dependent kernel for piezoelectric smart structures," *Chinese Journal of Sensors and Actuators*, vol. 25, no. 6, pp. 864–868, 2012.
- [27] W. Dingcheng, N. Yujia, and C. Beijing, "A wind speed forecasting model based on support vector regression with data dependent kernel," *Journal of Nanjing Normal University (Natural Science Edition)*, vol. 37, no. 3, pp. 15–20, 2014.
- [28] W. Dingcheng, Z. Youzhi, C. Beijing, and L. Yiyi, "Fast learning algorithm of multi-output support vector regression with data-dependent kernel," *Journal of Computer Applications*, vol. 37, no. 3, pp. 746–749, 2017.
- [29] T. Zhongda, L. Shujiang, W. Yanhong, and S. Yi, "A prediction method based on wavelet transform and multiple models fusion for chaotic time series," *Chaos, Solitons & Fractals*, vol. 98, pp. 158–172, 2017.



Hindawi

Submit your manuscripts at
www.hindawi.com

