ISSN: 2688-8645

Research on the Identification of Structural Size Parameters in Beam Pumping Units Using an Enhanced Particle Swarm Optimization Algorithm

Zhang Liu, Zhewei Ye

School of Mechanical and Electrical Engineering, Southwest Petroleum University, Chengdu Sichuan 610500, China

Abstract

The structural size parameters (SSP) of the beam pumping unit (BPU) are crucial for kinematic analysis and other related technologies. However, early equipment often lacked data, and measuring a large number of SSPs for the BPU is costly and inefficient. To enable the automatic recognition of SSPs and reduce labor and resource investments, this paper proposes a method for recognizing the SSPs of the BPU. A kinematic model for calculating the crank angle based on beam tilt angle and SSPs is established, along with an SSP recognition model. Based on the traditional particle swarm optimization (PSO) algorithm, a comprehensive enhanced particle swarm optimization (CEPSO) algorithm is designed to solve the model. Latin hypercube sampling (LHS) is used to obtain the initial values of the SSPs, which are then substituted into the kinematic model using the measured beam tilt angle. The algorithm aims to minimize the average absolute error between the measured and calculated crank angles. The CEPSO algorithm iteratively updates the SSP values, and when the error converges to a minimum, the SSPs are obtained. Experimental results demonstrate the effectiveness of the model and the algorithm. The parameter recognition error of the CEPSO algorithm is within 5%, significantly outperforming traditional PSO. This method provides a new technological approach for measuring SSPs.

Keywords

Beam Pumping Unit; Structural Parameters; Particle Swarm Optimization Algorithm; Parameter Identification.

1. Introduction

The BPU is one of the primary pieces of production equipment in oilfields, with a large installed capacity worldwide[1]. Real-time kinematic analysis of the BPU can be used to optimize design, fault detection, and motion balance analysis of the unit[2]. The four-bar linkage mechanism is a key component of the BPU, and the SSPs of its rods are crucial for calculating torque factors, displacement, speed, and acceleration of the sucker rod[3]. Additionally, many emerging technologies, such as calculating dynamometer cards (DC) based on measured electrical power[4], fault diagnosis[5], energy-saving variable-speed drives for pumping units[6], and beam balancing calculation models[7], also rely on precise SSPs. Therefore, accurately identifying SSPs is essential for implementing these new technologies.

Currently, SSPs are provided by manufacturers. However, due to size changes caused by equipment installation and long-term use, these parameters often fail to reflect the actual situation accurately[8]. This is particularly true in aging oilfields, where poor management or data loss often results in inaccurate or unavailable SSPs. Measuring the SSPs of a large number of BPUs is labor- and resource-intensive, and inefficient. Therefore, developing an effective soft measurement method for obtaining SSPs is of great importance.

ISSN: 2688-8645

In recent years, deep learning-based keypoint detection models have been used for real-time monitoring of the motion trajectory of BPUs. By acquiring real-time video streams via cameras and integrating deep learning techniques, the movement of key points, such as the sucker rod, can be accurately recognized[2][9][10]. However, due to limitations in camera installation locations and environmental conditions, these methods are limited in scattered development oilfields. This is particularly true for remote oil wells far from the oilfield's core area, where the high costs and maintenance requirements of existing monitoring equipment present challenges to widespread use. Based on the BPU's four-bar linkage motion model, Yin J[8] et al. proposed an SSP identification algorithm using the Kalman filter with nonlinear constraints. This algorithm calculates the sucker rod displacement equation using crank angle and rod length, and establishes a model for calculating crank angle based on electrical and motor nameplate parameters. Due to the highly nonlinear characteristics of the four-bar linkage motion and the need to identify multiple parameters, this algorithm has high computational complexity and is highly sensitive to initial parameter values. Therefore, there is a need for an algorithm better suited to identify the SSPs of the pumping unit. PSO has been widely applied[11][12][13] due to its global random search, simple algorithmic structure, robustness, and fast computation when solving nonlinear problems. However, PSO still faces challenges regarding local optima and requires further improvements in algorithm structure and optimization efficiency.

Moreover, in practical engineering, due to design, manufacturing, and installation errors, the initial crank angle when the horsehead reaches the top and bottom dead centers can be severely affected[14]. Additionally, belt slippage, causing disproportionate motor speed and crankshaft speed, further impacts the accuracy of sucker rod displacement calculations. These factors limit the effectiveness of dimension identification methods based solely on crank angle models. Zhao H[15] et al. established a mathematical model for calculating the torque factor using beam tilt angles. Results from practical engineering applications show that this method is more accurate than calculating torque factors based on crank angle. The method demonstrates good stability and practical engineering applicability. By measuring the beam tilt angle to calculate sucker rod displacement, higher accuracy is achieved. This method is not affected by stroke, stroke cycles, or other process parameters, making it more accurate in determining a complete stroke cycle. This method is applicable to various types and models of beam pumping units[16]. Therefore, measuring the beam angle improves the accuracy of sucker rod displacement calculations and directly provides movement information for the sucker rod.

1.1. Contributions

The core objective of this study is to identify SSPs for the BPU. The shortcomings of the aforementioned methods are outlined below: Deep learning-based key point detection methods are unsuitable for scattered oil wells. Additionally, single-well SSP identification algorithms based on the BPU's four-bar linkage motion model are complex and prone to errors in the geometric model. Furthermore, the application of PSO in parameter identification needs further improvement. Engineering practice has shown that measuring the beam tilt angle is highly effective. Based on these backgrounds and requirements, this paper proposes a method using CEPSO algorithms for the fast and accurate identification of SSPs for the BPU. The contributions of this paper are outlined below:

(1) A kinematic model for calculating the crank angle based on beam tilt angles and SSPs is proposed.

(2) An SSP identification method based on the kinematic model is developed.

(3) The PSO algorithm is improved, and a CEPSO algorithm for SSP identification is proposed.

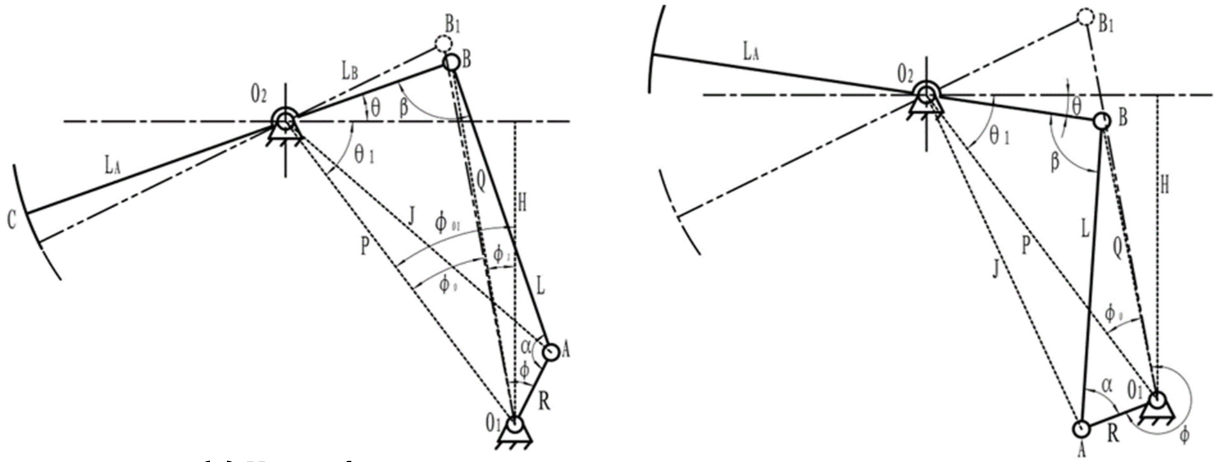
(4) The practicality of this method for SSP identification in BPU is explored.

2. Materials and Methods

2.1. Kinematic Model of BPU

Parameter identification combines theoretical models with experimental data to determine model parameters, minimizing the error between the model output and actual observed data. Essentially, it is an optimization problem within the parameter space[17]. Thus, the first step is to establish the mathematical model of the BPU and convert it into an optimization problem.

Under the assumption of generality, the beam pumping unit can be simplified as the four-bar mechanism geometric model shown in Fig 1. The meanings of the parameters are as follows: R: Crank radius(m), L: Link length(m), L_A: Front arm length of the beam(m), L_B: Rear arm length of the beam(m), P: Base rod length(m), Q: Length of the line connecting the crank's rotation center with the connection point between the link and the beam(m), J: Length of the line connecting the beam's rotation center with the connection point between the crank and the link(m), α: Angle between the crank and the link(°), β: Angle between the link and the beam(°), ϕ: Crank angle(°), θ: Angle between the rear arm of the beam and the horizontal direction. It is considered positive above the horizontal line and negative below the horizontal line(°), θ_1: Angle between the base rod and the horizontal line passing through the rotation center of the beam(°), ϕ_0: Angle between the base rod and the line O_1B_1 (where the crank and the link are collinear, and the horsehead is at the bottom dead point) (°), ϕ_1: Angle between O_1B_1 and the vertical direction(°), ϕ_{01}: Angle between the base rod and the vertical direction(°).



(a) Up stroke (b) Down stroke

Fig 1. Four-bar linkage mechanism of beam pumping unit

The following geometric relationship formula can be obtained through analysis:

$$Q^2 = R^2 + L^2 - 2 \cdot R \cdot L \cdot \cos\alpha \tag{1}$$

$$Q^2 = P^2 + L_B^2 - 2 \cdot P \cdot L_B \cdot \cos(\theta + \theta_1) \tag{2}$$

$$J^2 = L_B^2 + L^2 - 2 \cdot L_B \cdot L \cdot \cos\beta \tag{3}$$

$$J^2 = P^2 + R^2 - 2 \cdot P \cdot R \cdot \cos(\phi_0 + \phi) \tag{4}$$

In the formula: $\phi_0 = ar\cos\left(\dfrac{P^2 + (R+L)^2 - L_B^2}{2P(R+L)}\right)$, $\theta_1 = \arcsin(H/P)$

As shown in Fig 1(a), during the up stroke, the following geometric relationship is satisfied:

$$\phi_0 + \phi + \alpha + \beta + \theta + \theta_1 = 2\pi \tag{5}$$

As shown in Fig 1(b), during the up stroke, the following geometric relationship is satisfied:

$$\phi_0 + \phi + \beta + \theta + \theta_1 = 2\pi + \alpha \tag{6}$$

By combining equations (1) and (2):

$$\alpha = ar\cos\frac{R^2 + L^2 - P^2 - L_B^2 + 2\cdot P\cdot L_B\cdot\cos(\theta+\theta_1)}{2RL} \tag{7}$$

During the up stroke, it can be obtained from equations (3), (4)and (5)that:

$$\phi_{up} = ar\cos\left(\frac{\frac{t}{m}\left[\frac{R}{L}\sin\phi_1 + \sin\left(\alpha-\phi_1\right)\right] - \frac{H + L_B\sin\theta}{L}}{\frac{n}{m}\left[\frac{R}{L}\sin\phi_1 + \sin\left(\alpha-\phi_1\right)\right] - \left[\frac{R}{L}cos\phi_1 - cos\left(\alpha-\phi_1\right)\right]}\right) \tag{8}$$

In the formula, $t = L_B^2 + L^2 - P^2 - R^2$;

$$\phi_1 = \frac{\pi}{2} - \theta_1 - \phi_0;$$

$$m = 2L_B L\sin\left(2\pi - \phi_0 - \theta - \theta_1 - \alpha\right) + 2RP\sin\phi_0;$$

$$n = 2L_B L\sin\left(2\pi - \phi_0 - \theta - \theta_1 - \alpha\right) - 2RP\sin\phi_0.$$

During the down stroke, it can be obtained from equations (3), (4)and (6)that:

$$\phi_{down} = ar\cos\left(\frac{\frac{t}{M}\left[\frac{R}{L}\sin\phi_1 - \sin\left(\alpha+\phi_1\right)\right] - \frac{H + L_B\sin\theta}{L}}{\frac{N}{M}\left[\frac{R}{L}\sin\phi_1 - \sin\left(\alpha+\phi_1\right)\right] - \left[\frac{R}{L}cos\phi_1 - cos\left(\alpha+\phi_1\right)\right]}\right) \tag{9}$$

In the formula, $M = 2L_B L\sin(2\pi - \phi_0 - \theta - \theta_1 + \alpha) + 2RP\sin\phi_0$;

$$N = 2L_B L\sin(2\pi - \phi_0 - \theta - \theta_1 + \alpha) - 2RP\sin\phi_0.$$

Thus, a mathematical model for calculating the crank angle of a pumping unit based on the inclination angle of the beam can be obtained:

$$\phi = \begin{cases} ar\cos\left(\frac{\frac{t}{m}\left[\frac{R}{L}\sin\phi_1 + \sin\left(\alpha-\phi_1\right)\right] - \frac{H + L_B\sin\theta}{L}}{\frac{n}{m}\left[\frac{R}{L}\sin\phi_1 + \sin\left(\alpha-\phi_1\right)\right] - \left[\frac{R}{L}cos\phi_1 - cos\left(\alpha-\phi_1\right)\right]}\right), up \quad stroke \\ ar\cos\left(\frac{\frac{t}{M}\left[\frac{R}{L}\sin\phi_1 - \sin\left(\alpha+\phi_1\right)\right] - \frac{H + L_B\sin\theta}{L}}{\frac{N}{M}\left[\frac{R}{L}\sin\phi_1 - \sin\left(\alpha+\phi_1\right)\right] - \left[\frac{R}{L}cos\phi_1 - cos\left(\alpha+\phi_1\right)\right]}\right), down \quad stroke \end{cases} \tag{10}$$

2.2. SSP Identification Model

The objective function established based on the above mathematical model is as follows:

$$f\left(\boldsymbol{\xi}\right) = \frac{1}{n}\sum_{i=1}^{n}\left|\phi_i^{ref} - \phi_i\right| \tag{11}$$

In the equation, $\boldsymbol{\xi} = [\boldsymbol{\xi_1}, \dots, \boldsymbol{\xi_5}] = [R, L, P, H, L_B]$ represents the solution vector of the parameters to be determined; ϕ_i^{ref} and ϕ_i represent the measured and calculated values of the crank angle, respectively; n denotes the number of data points measured within one cycle.

Each parameter has specific bounds, and the range for the oil pump dimensions can be easily determined. To simplify the process, each parameter's value range is assumed to be consistent, i.e., $\xi_{min} \le \boldsymbol{\xi} \le \xi_{max}$. However, implicit constraints exist within this range. Since the inverse trigonometric functions $arcsin(\Omega)$ and $arccos(\Omega)$ are used in the model, where Ω is the argument with a valid range of $[-1,1]$, it is crucial to ensure Ω stays within this range during the calculation. To avoid adding complexity by treating all inverse trigonometric functions as constraints, the following approach is adopted: First, check if Ω exceeds the defined domain. If it does, restrict it to the boundary values of the domain, as shown in equation (12).

$$\Omega = \begin{cases} -1, & -1 \ge \Omega \\ \Omega, & -1 \le \Omega \le 1 \\ 1, & 1 \ge \Omega \end{cases} \tag{12}$$

Based on the above, the high-dimensional optimization problem for identifying the SSP of the BPU can be formulated as follows:

$$\begin{aligned} &\min f(\boldsymbol{\xi}) = \frac{1}{n}\sum\nolimits_{i=1}^{n}\left|\phi_i^{ref} - \phi_i\right| \\ &s.t. \quad \xi_{\min} \le \boldsymbol{\xi} \le \xi_{\max} \end{aligned} \tag{13}$$

The goal of parameter identification is to find a set of parameters $\boldsymbol{\xi}_{if}$ that minimizes $f(\boldsymbol{\xi})$. Next, a CEPSO algorithm is established to solve this problem.

The model shown in equation (10) cannot calculate the length of the beam's forearm L_A. However, the swing angle of the beam is proportional to the displacement of the suspension point. According to the geometric model shown in Fig 1, the corresponding relationship is given by the following equation[16].

$$L_A = \frac{S_i}{\theta_m - \theta_i} \tag{14}$$

In the equation, θ_m represents the angle between the beam's rear arm and the horizontal direction when the horsehead is at the bottom dead center, in radians; θ represents the angle between the beam's rear arm and the horizontal direction at a given moment, in radians.

The length of the beam's front arm, L_A, can be calculated using Equation (14), but this formula is invalid when the horsehead is at the bottom dead center. To avoid significant deviations due to sensor measurement errors, L_A is calculated as follows: First, data corresponding to the bottom dead center are excluded from the dataset. The remaining data are then used to compute a vector of L_A values, denoted as $L_A^P = \left[L_{A1}, \dots, L_{A(n-j)}\right]$, where j represents the number of data points corresponding to the bottom dead center within one cycle. Next, Shawnee data selection criteria[18] are used to decide whether to accept or reject anomalous outliers in L_A^P. Finally, the average of the remaining data is calculated to obtain the beam arm length, L_A.

2.3. The Method for IdentifyingSSP of BPU

Fig 1 shows the schematic diagram of the beam pumping unit SSP identification process. The procedure begins by obtaining the measured data: one cycle of beam angle $\boldsymbol{\theta}$, crank angle $\boldsymbol{\phi^{ref}}$, sucker rod displacement $\boldsymbol{S}$, and crank radius R_m. Next, LHS is used to randomly generate a dataset, $\boldsymbol{\tau} = (\boldsymbol{\xi_1}, \dots, \boldsymbol{\xi_m})$, containing m parameter vectors within the range $[\xi_{min}, \xi_{max}]$. These parameter vectors serve as the initial values for the CEPSO iteration, accounting for the algorithm's randomness. The condition $f(\boldsymbol{\xi_{if}}) \le \Delta$ is set during the iteration to determine the optimal parameter vector $\boldsymbol{\xi_{if}} = (R_{if}, L_{if}, P_{if}, H_{if}, L_{Bif})$. Finally, the identified parameter vector $\boldsymbol{\xi_{if}}$ is adjusted based on R_m to obtain the SSP identification values $\boldsymbol{\xi} = (R, L, P, H, L_B)$. The beam front arm length, L_A, is then calculated using the beam angle and the displacement.

LHS is a stratified sampling technique that ensures good uniformity in high-dimensional space, with initial values generated by LHS[19]. In the five-dimensional optimization problem discussed in this paper, LHS improves the global search capability and convergence speed of the CEPSO algorithm. The dimensions of the pumping unit remain valid for the mathematical model shown in Equation (10) even when scaled arbitrarily, resulting in many optimal

ISSN: 2688-8645

solutions that satisfy Equation (10).This characteristic allows CEPSO to obtain the global optimal solution. However, the optimal parameter vector $\boldsymbol{\xi_{if}}$ may not represent the actual values and requires correction using real measurements. The crank radius, R_m, is relatively easy to measure, and the ratio k between R_m and R_{if} is used for correction.

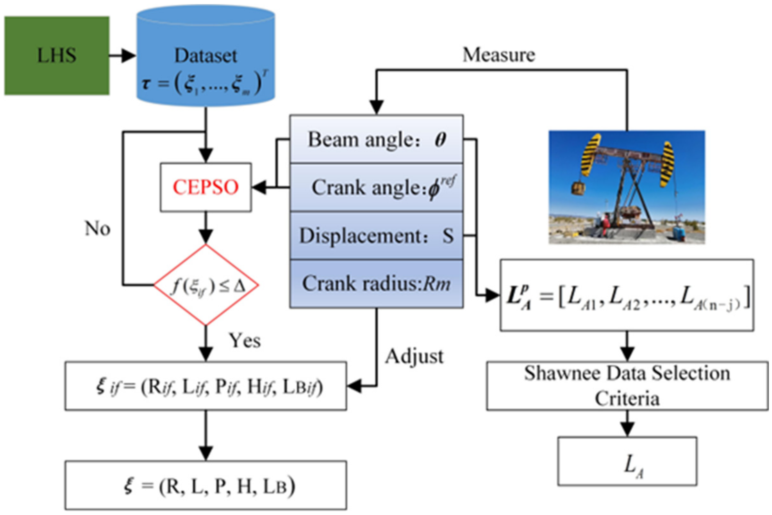


Fig 2. SSP Identification Process for Beam Pumping Units

$$\xi = (R, L, P, H, L_B) = \frac{R_m}{R_{if}} \cdot \xi_{if} = k \cdot \xi_{if} \tag{15}$$

2.4. Design of the CEPSO Algorithm

PSO is a swarm intelligence algorithm inspired by the foraging behavior of birds. In this study, PSO is applied to a five-dimensional optimization problem, where the parameter vector $\boldsymbol{\tau} = (\boldsymbol{\xi_1}, \dots, \boldsymbol{\xi_m})$ represents the positions of the particles, and the velocity of each particle is denoted as $\boldsymbol{V} = (v_1, \dots, v_m)$. The algorithm searches for the optimal solution by continuously updating the velocity V_{IJ} and position ξ_{IJ} of each particle[20].

$$V_{IJ}^{t+1} = w \cdot V_{IJ}^{t} + c_1 r_{(0,1)} \left(\xi_{IJ,pbest}^{t} - X_{IJ}^{t} \right) + c_2 r_{(0,1)} \left(\xi_{Gbest}^{t} - X_{IJ}^{t} \right) \tag{16}$$

$$\xi_{IJ}^{t+1} = \xi_{IJ}^{t} + V_{IJ}^{t+1} \tag{17}$$

In the equations, I ranges from 1 to 5, J ranges from 1 to m, and t represents the current iteration step.w is the inertia weight, $r_{(0,1)}$ is a random number between 0 and 1, c_1 is the cognitive learning factor, and c_2 is the social learning factor. $\xi_{IJ,pbest}^{t}$ represents the historical best position of the Ith particle, and ξ_{gbest}^{t} represents the global best position of the entire swarm.

Metaheuristic algorithms need to balance exploration and exploitation. In the initial stages, strong exploration is needed to locate the global optimum region, while in the later stages, exploitation should be enhanced to fine-tune the solutions and achieve a more precise global optimum[21]. As shown in Equation (16), in PSO, the exploration and exploitation capabilities depend on three parameters, w, c_1, and c_2:(1)w controls the influence of the previous velocity on the current one. A larger w strengthens exploration but weakens exploitation. (2) c_1 increases the attract(2)ion of the particle to its best position. A larger c_1 strengthens exploration but weakens exploitation. (3)c_2 increases the attraction of the particle to the global best position. A larger c_2 weakens exploration but strengthens exploitation.

Thus, during algorithm operation, w, c_1, and c_2 are dynamically adjusted in a non-linear manner to optimize the balance between exploration and exploitation, improving the algorithm's performance. The adjustment strategy is as follows:

$$w_t = w_{\max} - (w_{\max} - w_{\min})\sin\left(\frac{\pi}{2}\cdot\sqrt[4]{t/t_{\max}}\right) \tag{18}$$

$$c_{1t} = c_{1\max} - (c_{1\max} - w_{1\min})\sin\left(\frac{\pi}{2}\cdot\sqrt[4]{t/t_{\max}}\right) \tag{19}$$

$$c_{2t} = c_{2\min} + (c_{2\max} - w_{2\min})\sin\left(\frac{\pi}{2}\cdot\sqrt[4]{t/t_{\max}}\right) \tag{20}$$

In the above equation, w_{max} and w_{min} represent the maximum and minimum values of the inertia weight; c_{1max} and c_{1min} represent the maximum and minimum values of the cognitive learning factor; c_{2max} and c_{2min} represent the maximum and minimum values of the social learning factor; and t_{max} is the maximum number of iterations.

Increasing the diversity of the population can further enhance the global search ability[22]. To achieve this, the concept of Opposition-based Learning (OBL) is introduced. This method enhances the diversity of the population by applying an opposition transformation to the existing population[23]. Specifically, let ξ_{ij} be a point in $\boldsymbol{\tau} = (\boldsymbol{\xi_1}, \dots, \boldsymbol{\xi_m})$, then its opposite value ξ_{ij}^{OBL} is given by:

$$\xi_{IJ}^{OBL} = \xi_{\min} + \xi_{\max} - \xi_{IJ}, \qquad I = 1,\dots,\mathrm{m}; J = 1,\dots,5 \tag{21}$$

For each point in the current population, apply the operation in Equation (21) to generate an opposite population, $\boldsymbol{\tau^{OBL}} = \left(\boldsymbol{\xi_1^{OBL}}, \dots, \boldsymbol{\xi_m^{OBL}}\right)$. Next, substitute each value from $\boldsymbol{\tau^{OBL}}$ into the objective function. Select the top m parameter vectors with the smallest objective function values from both populations for the next iteration. Incorporating OBL allows the algorithm to explore a broader solution space, enhancing its global search capability.

In traditional PSO algorithms, when a particle exceeds the value range, its value is usually set to the upper or lower bound. This can cause significant population clustering, reducing diversity and increasing the likelihood of the algorithm converging to a local optimum[24]. To address this, a boundary rebound strategy is introduced:

$$\xi_{IJ}^{'} = \begin{cases} \xi_{\max} - r(0,1)\cdot(\xi_{IJ} - \xi_{\max}), & \xi_{IJ} > \xi_{\max} \\ \xi_{\min} + r(0,1)\cdot(\xi_{\min} - \xi_{IJ}), & \xi_{IJ} < \xi_{\min} \end{cases} \tag{22}$$

Similarly, when a particle's velocity exceeds the value range, the same operation is applied to adjust its velocity:

$$v_{IJ}^{'} = \begin{cases} v_{\max} - r(0,1)\cdot(v_{IJ} - v_{\max}), & v_{IJ} > v_{\max} \\ v_{\min} + r(0,1)\cdot(v_{\min} - v_{IJ}), & v_{IJ} < v_{\min} \end{cases} \tag{23}$$

3. Experimental Results

3.1. Experimental Setup

To verify the effectiveness of the above dimension identification method, the CEPSO algorithm was implemented using MATLAB. The measured data for one cycle of the pumping unit included the beam angle $\boldsymbol{\theta}$, crank angle $\boldsymbol{\phi^{ref}}$, and plunger displacement $\boldsymbol{S}$. The measured sequence length was $n = 511$. The curve of the measured data is shown in Fig 3. Table 1 lists the structural dimension parameters of the test pumping unit, and Table 2 shows the settings of the CEPSO algorithm parameters.

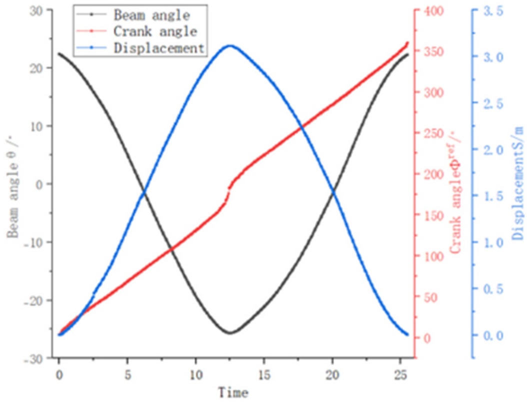


Fig 3. The experimental data

Table 1. SSP of beam pumping unit

Structural size parameters	L_A	R	L	L_B	P	H
Measurement value(m)	3.640	1.120	3.400	2.745	4.424	3.470

Table 2. The setting values of intermediate parameters in CEPSO algorithm

Parameter	Value	Parameter	Value	Parameter	Value
ξ_{min}	0	c_{1max}	3	m	100
ξ_{max}	10	c_{2min}	1	t_{max}	200
w_{min}	0.5	c_{2max}	2	Δ	1
w_{max}	0.8	v_{min}	-1		
c_{1min}	2	v_{max}	1		

3.2. Results

First, the real SSP values are substituted into the mathematical model shown in Equation (10) to validate the model by calculating the crank angle. Fig 4 compares the calculated crank angle with the measured values. The results show that the calculated values align well with the measured values, confirming that the model accurately represents the geometric relationship between the crank angle and the beam tilt angle.

Next, the effectiveness of the parameter identification process is validated. A comparison between the identified and actual parameter values is made, and the relative errors are calculated, as shown in Table 3. Overall, the identified values closely match the actual values, with small errors. This indicates that the CEPSO algorithm is highly accurate and effective. The identification error for each parameter is within 5%, with the maximum relative error being 4.54%. Thus, the proposed algorithm is reliable for identifying dimensional parameters and accurately recognizing the SSP of the BPU.

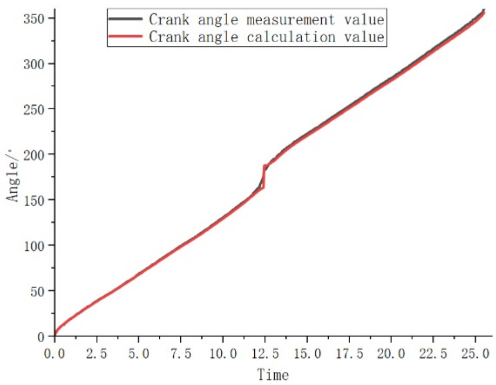


Fig 4. Measured and calculated values of crank angle

Table 3. Measured Values and Identified Values of SSP

SSP	L_A	R	L	L_B	P	H
Identification value(m)	3.686	1.12	3.308	2.713	4.223	3.384
relative error(%)	1.26	0	2.71	1.80	4.54	2.48

Fig 5 presents the variation curve of the calculated beam arm length L_A^P throughout the entire cycle, both before and after removing outliers, using Equation (14). The figure shows significant calculation errors at the beginning and end of the cycle. If these points with large errors are not excluded and all values are averaged, the calculation error of the beam arm length will increase. To improve accuracy, the Shawnee data selection criteria are applied to eliminate outliers with large deviations.

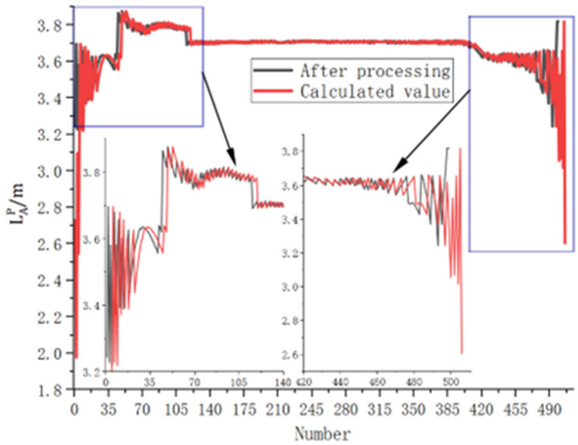


Fig 5. Calculated Value of Beam Front Arm Length

A comparative experiment was conducted to validate the advantages of the CEPSO algorithm over the traditional PSO algorithm, particularly in avoiding local optima and premature convergence. In the experiment, the condition $f(\xi_{if}) \leq \Delta$ was removed, and 50 repeated experiments were conducted, considering the algorithm's inherent randomness. In these experiments, the average value curve, the worst curve (corresponding to the maximum objective function value), and the best curve (corresponding to the minimum objective function value) were recorded, as shown in Fig 6. The obtained parameter averages and their relative errors compared to the actual values are presented in Fig 7 and Table 4. Since the crank radius requires direct measurement, the recognition error for this parameter was not considered.

Fig 6 shows that, overall, the CEPSO algorithm outperforms the traditional PSO algorithm in terms of efficiency and reliability. Although each curve stabilizes at a constant value after a certain number of iterations, the randomness of the algorithm causes a significant gap between the best and worst curves. However, the parameter identification method proposed in this paper includes the condition $f(\xi_{if}) \leq \Delta$, which enables the CEPSO algorithm to perform multiple calculations and achieve the desired accuracy, thereby validating the necessity of this condition.

Further analysis of the mean curve of the objective function shows that, in 50 repeated experiments, the traditional PSO algorithm often prematurely converges after getting stuck in a local optimum, halting the search for the global optimum. In contrast, the CEPSO algorithm exhibits stronger exploration capabilities, continuously searching and eventually converging to the global optimum. The data in Fig 7 and Table 4 further validate this: compared to the traditional PSO algorithm, the parameter values obtained by the CEPSO algorithm are closer to the actual values, with higher accuracy.

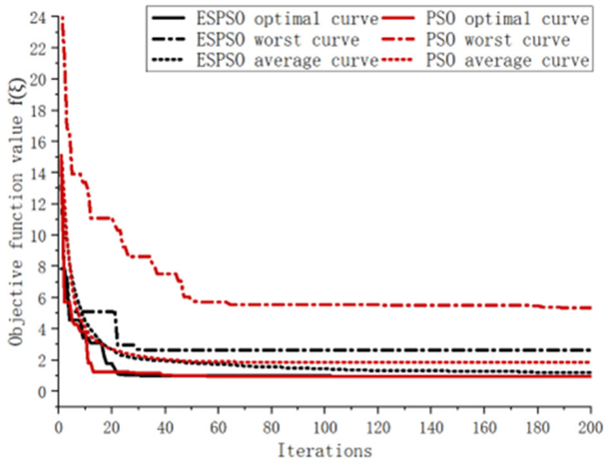


Fig 6. Objective Function Curve of 50 Repeated Experiments

Table 4. The average value of the SSPs from 50 repeated experiments

SSP	Average value of CEPSO algorithm(m)	Average value of PSO algorithm(m)
L	3.514	2.815
L_B	2.711	2.493
P	4.282	3.783
H	3.589	2.782

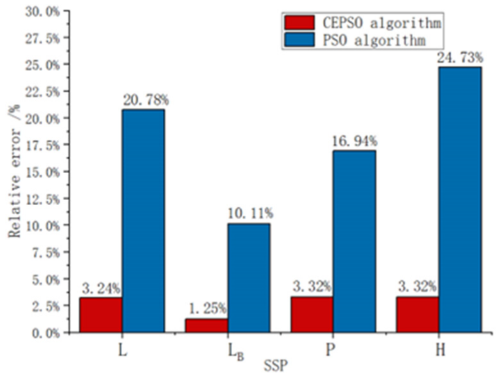


Fig 7. The relative error of the average value of the SSPs from 50 repeated experiments

In summary, by improving the particle swarm optimization algorithm, both the convergence speed and stability were significantly enhanced, yielding remarkable optimization results. The improved CEPSO algorithm not only maintains the stability of the optimization process but also enhances its exploration capability, making it more efficient and reliable for identifying the SSP of the BPU.

4. Conclusion

To address the challenge of directly measuring the structural parameters of a beam pumping unit in practical production, this study proposes a mathematical model that calculates the crank angle based on beam angle and rod lengths, combined with the CEPSO algorithm for SSP identification. This method can accurately identify the pumping unit's SSP. Compared to methods that estimate size parameters using real-time video captured by intelligent patrol robots, this approach is more cost-effective. Moreover, the algorithm is relatively simple, iterating over the initial particle swarm to automatically search for the optimal solution within the parameter range.

Building upon traditional PSO, this paper introduces concepts such as nonlinear dynamic adjustment of inertia and learning factors, opposition-based learning, and boundary rebound strategies, resulting in the CEPSO algorithm. This algorithm outperforms traditional PSO in terms of performance and effectively avoids the premature convergence issue caused by limited exploration in the later stages. The improved CEPSO algorithm not only enhances identification accuracy but also improves stability and global search capabilities in complex environments.

References

[1] Zhang C.; Wang L.; Wu X.; Gao X. Performance analysis and design of a new-type wind-motor hybrid power pumping unit. Electric Power Systems Research. 2022, 208, 107931. https:// doi.org/ 10. 1016/j.epsr.2022.107931.

[2] Wang Q.; Zhang K.; Zhao H.; Zhang H.; Zhang L.;Yan X.; Liu P.; Fan L.; Yang Y. Yao J. A novel method for trajectory recognition and working condition diagnosis of sucker rod pumping systems based on high-resolution representation learning. Journal of Petroleum Science and Engineering. 2022, 218, 110931. https://doi.org/10.1016/j.petrol.2022.110931.

[3] Feng, Z.; Tan J.; Li Q.; Fang X. A review of beam pumping energy-saving technologies. Journal of Petroleum Exploration & Pro-duction Technology.2018, 8, 299-311. https://doi.org/ 10.1007/ s1 3202-017-0383-6.

[4] Zou J.; Wu Y.; Wang Z.; Dong S. A Novel Hybrid Method for Indirect Measurement Dynamometer Card Using Measured Motor Power in Sucker Rod Pumping System. IEEE Sensors Journal. 2022, 22,13971-80. https://doi.org/10.1109/JSEN.2022.3181621.

[5] Hao D.; Gao X. Unsupervised Fault Diagnosis of Sucker Rod Pump Using Domain Adaptation with Generated Motor Power Curves. Mathematics. 2022, 10, 1224. https://doi.org/ 10.3390/ math 10081224.

[6] Tan C.; Feng Z.; Liu X.; Fan J.; Cui W.; Sun R.; Ma Q. Review of variable speed drive technology in beam pumping units for energy-saving. Energy Reports. 2020, 6, 2676-88.https://doi.org/ 10. 1016/j.egyr.2020.09.018

[7] Cen X.; Wu X.; Wang L.; Zhen L.; Ge L. A New Model for Calculatin the Ideal Beam Counterbalance Weight for aPumping Unit. Petroleum Drilling Techniques. 2016, 44, 82-86. https://doi. org/doi: 10. 11911/syztjs.201602014.

[8] Yin J.; Sun D.; Ma H. Identification of the Four-Bar Linkage Size in a Beam Pumping Unit Based on Cubature Kalman Filter. Machines. 2022, 10, 1133. https://doi.org/10.3390/machines10121133.

[9] Zhang, K.; Xia X.; Song Z.; Zhang L.; Yang Y.; Wang J.; Yao J.; Zhang H.; Zhang Y.; Feng G; Liu C. Trajectory Recognition and Working Condition Analysis of Rod Pumping Systems Based on Pose Estimation Method with Heatmap-Free Joint Detec-tion. SPE JOURNAL. 2024, 29, 5521-5537. https://doi.org/10.2118/223095-PA.

[10] Sun J.; Huang Z.; Zhu Y.; Zhang Y. Real-time kinematic analysis of beam pumping unit: a deep learning approach. Neural Computing and Applications. 2022, 34, 7157-71. https://doi.org/ 10. 1007/ s00521-021-06783-0.

[11] Yang L.; Wang H.; Yan M. Particle Swarm Optimization-Based Model Abstraction and Explanation Generation for a Recurrent Neural Network. Algorithms. 2024, 17, 210. https://doi.org/ 10.3390/ a17050210.

[12] Bai Z.; Yu P.; Liu P.; Guo J. Linear System Identification-Oriented Optimal Tampering Attack Strategy and Implementation Based on Information Entropy with Multiple Binary Observations. Algorithms. 2024, 17, 239. https://doi.org/10.3390/a17060239.

[13] Ishaq H.; Rached D. Identification of Mechanical Parameters in Flexible Drive Systems Using Hybrid Particle Swarm Opti-mization Based on the Quasi-Newton Method. Algorithms. 2023, 16, 371. https://doi.org/ 10.3390/a16080371.

[14] Zhang J.; Li X.; Shi H. Design calculation of beam pumping unit[M]. Beijing: Petroleum Industry Press. 2005： 9-11.

[15] Zhao H.; He K.; Hu D.; Lu M. Research on the soft-sensing method of polished rod load of beam pumping unit. Chinese Journal of Scientific Instrument. 2021, 42, 160-171. https://doi.org/ 10. 19650/ j.cnki.cjsi.J2107962.

[16] Wang H.; Zheng W.; Yan J.; Wang J. On-line test of polished rod displacement of pumping unit. Oil-Gasfield Surface Engineering. 2014, 33, 41-42. https://doi.org/doi:10.3969/j.issn.1006-6896.2014. 11. 020.

[17] Wang B.; He Y.; Liu J.; Luo B. Fast parameter identification of lithium-ion batteries via classification model-assisted Bayesian optimization. Energy. 2024, 288,129667. https://doi.org/ 10.1016/ j. energy. 2023.129667.

[18] Wang L.; Li Y.; Liang Q.; Dong S.; Tang L. Greenhouse environmental data collection based on improved Chauvenet's criterion. Transactions of the Chinese Society of Agricultural Engineering. 2015, 31, 212-217. https://doi.org/ doi: 10.3969/j.issn.1002-6819.2015.05.030.

[19] Ali R.A.; Mohammad M. Probabilistic load flow in distribution networks: An updated and comprehensive review with a new classification proposal. Electric Power Systems Research. 2023, 222, 109497. https://doi.org/10.1016/j.epsr.2023.109497.

[20] Juan R.V.; Zhang M.; Seah W. A performance study on synchronicity and neighborhood size in particle swarm optimization. Soft Computing. 2013, 17, 1019-1030. https://doi.org/ 10.1007/ s00 500-013-1015-9.

[21] Jordehi A.R. Time varying acceleration coefficients particle swarm optimisation (TVACPSO): A new optimisation algorithm for estimating parameters of PV cells and modules. Energy Conversion and Management. 2016, 129, 262-74. https://doi.org/10.1016/j.enconman.2016.09.085.

[22] Wu Z.; Luo Y.; Hu S. Optimization of jamming formation of USV offboard active decoy clusters based on an improved PSO algorithm. Defence Technology. 2024, 32, 529-540. https://doi.org/ 10.1016/ j.dt. 2023.03.017.

[23] Mahdavi S.; Rahnamayan S.; Deb K. Opposition based learning: A literature review. Swarm and Evolutionary Computation. 2018, 39, 1-23. https://doi.org/ 10.1016/j.swevo.2017.09.010.

[24] Song M.; Gu G. Research on particle swarm optimization: a review. Machine Learning and Cybernetics, 2004. Proceedings of 2004 International Conference on 2004. https://doi.org/ 10.1109/ ICMLC.2004.1382171.