



Original Paper

A novel optimization scheme for structure and balance of compound balanced beam pumping units using the PSO, GA, and GWO algorithms



Jie Wang^a, Quan-Ying Guo^a, Cheng-Long Fu^a, Gang Dai^c, Cheng-Yu Xia^{a, b},
Li-Qin Qian^{a, *}

^a School of Mechanical Engineering, Yangtze University, Jingzhou, 434023, Hubei, China

^b Cooperative Innovation Center of Unconventional Oil and Gas, Yangtze University (Ministry of Education & Hubei Province), Wuhan, 430100, Hubei, China

^c Jingzhou Mingde Technology Co., LTD, Jingzhou, 434000, Hubei, China

ARTICLE INFO

Article history:

Received 25 July 2024

Received in revised form

13 October 2024

Accepted 17 January 2025

Available online 24 January 2025

Edited by Teng Zhu

Keywords:

Compound balanced BPU

Dynamic model

Structural optimization

Balance optimization

Constraints

ABSTRACT

The beam pumping unit (BPU) remains the most stable and reliable equipment for crude oil lifting. Despite its simple four-link mechanism, the structural design of the BPU presents a constrained single-objective optimization problem. Currently, a comprehensive framework for the structural design and optimization of compound balanced BPUs is lacking. Therefore, this study proposes a novel structural design scheme for BPUs, aiming to meet the practical needs of designers and operators by sequentially optimizing both the dynamic characteristics and balance properties of the BPUs. A dynamic model of compound balanced BPU was established based on D'Alembert's principle. The constraints for structural dimensions were formulated based on the actual operational requirements and design experience with BPUs. To optimize the structure, three algorithms were employed: the particle swarm optimization (PSO) algorithm, the genetic algorithm (GA), and the gray wolf optimization (GWO) algorithm. Each newly generated individuals are regulated by constraints to ensure the rationality of the outcomes. Furthermore, the integration of three algorithms ensures the increased likelihood of attaining the global optimal solution. The polished rod acceleration of the optimized structure is significantly reduced, and the dynamic characteristics of the up and down strokes are essentially symmetrical. Additionally, these three algorithms are also applied to the balance optimization of BPUs based on the measured dynamometer card. The calculation results demonstrate that the GWO-based optimization method exhibits excellent robustness in terms of structural optimization by enhancing the operational smoothness of the BPU, as well as in balance optimization by achieving energy conservation. By applying the optimization scheme proposed in this paper, the CYJW7-3-23HF type of BPU was designed, achieving a maximum polished rod acceleration of $\pm 0.675 \text{ m/s}^2$ when operating at a stroke of 6 min^{-1} . When deployed in two wells, the root-mean-square (RMS) torque was minimized, reaching values of $7.539 \text{ kN}\cdot\text{m}$ and $12.921 \text{ kN}\cdot\text{m}$, respectively. The proposed design method not only contributes to the personalized customization but also improves the design efficiency of compound balanced BPUs.

© 2025 The Authors. Publishing services by Elsevier B.V. on behalf of KeAi Communications Co. Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

Currently, the main methods for crude oil extraction in oil fields include flowing production and artificial lift production. Flowing production is primarily deemed suitable for newly developed oil

wells that possess substantial crude oil reserves. It utilizes the formation's energy to lift crude oil to the surface, with a critical prerequisite being sufficiently high formation pressure. Conversely, if the formation pressure is too low for natural flow, the artificial-lift method must be employed to extract the crude oil. Approximately 87% of oil production wells use artificial lifting methods (Catalina et al., 2022). Rough statistical data indicate that 71% of these wells employ beam pumping system, 16% use electrical

* Corresponding author.

E-mail address: qlq1010@126.com (L.-Q. Qian).

submersible pumps, 7% use gas lift systems, and 6% rely on other lifting methods (Pineda et al., 2023). BPU are particularly favored in major oil fields due to their simple operation, convenient installation and maintenance, high adaptability to harsh conditions, and other advantages (Sun et al., 2022; He et al., 2023). Despite these benefits, BPUs often face criticism for energy wastage, akin to “using a sledgehammer to crack a nut”. From an economic benefit perspective, it is vital to redesign conventional BPUs for greater energy efficiency and reduced consumption. Ideally, these improvements should be feasible for direct implementation at oilfield sites. This holds significant importance for the current state of oilfields and their overall economic efficiency.

The BPU employs a classic 4-bar linkage mechanism. In kinematics, this type of mechanism has one degree of freedom, meaning that the motion can be fully described by the rotation at any rotary joint. The graphical approach and the famous Freudenstein equation (Ghosal, 2010) are the most traditional methods for analyzing such mechanisms. Despite advancements, designers still frequently use the graphical approach to determine the dimensions of 4-bar linkages in BPUs. Liu (1985) introduced a novel method for the geometric design of sucker rod pumping units, combining both graphical and analytical methods. However, the design typically requires predefining several parameters based on experience, such as the stroke, front arm of beam, height of frame, horizontal offset distance, crank angle between extreme positions, and starting angle of crank. These predefined parameters suggest a unique structure for the BPU but overlook its dynamic characteristics. As theoretical research on BPUs continues to evolve and practical experience accumulates, structural optimization has become increasingly prominent. Hu et al. (2006) provided a comprehensive summary and analysis of the issues concerning the structural optimization of BPUs, covering the variables to be optimized, objective functions, and constraints. Zhang et al. (2005) performed a kinematic analysis of the BPU and suggested an optimization strategy for structural dimensions, detailing the constraints involved. However, they did not provide a clear implementation process for their optimization strategy. Furthermore, the “Specification for Pumping Units” (API Spec 11E, 2013), issued by American Petroleum Institute (API), specifies the components of the BPU but does not explicitly define the methods for determining dimensions of key components like the beam, pitman, frame, and crank, etc. Therefore, developing a comprehensive structural design and optimization strategy for BPUs is of significant practical engineering value, providing crucial guidance for their design and maintenance, and enhancing energy efficiency as discussed earlier.

In recent years, the focus on energy conservation in BPUs has intensified within the global petroleum industry (Lv et al., 2016; Fakhier et al., 2021; Tian et al., 2021), driven by the pressing need to address low profit margins, energy crises, and environmental pollution in oilfield operations. Traditional methods to enhance energy efficiency in BPUs include regular adjustments to stroke length or speed, and correcting balance deviations. Moreover, advancements in control technology for BPUs have enabled energy savings through automated control systems. For instance, full cycle variable speed operation control of the pumping system has been developed, which allows for more precise management of energy use. Feng et al. (2020) developed a dynamic model and an optimization method of a variable speed drive BPU. The complex mechanical behavior of sucker-rod pump, a major focus for many scholars, is crucial for determining energy-saving control strategies. However, traditional approaches, such as modifying stroke length or speed and rectifying balance deviations, are more favored for their cost-effectiveness, reliability, and simplicity. Operators can adjust the net torque by altering the mass and/or position of counterweights. Stroke adjustments are facilitated by changing the

crank pin's position, utilizing pre-existing pin holes. Takacs and Kis (2021) performed an in-depth analysis on the arrangement of crank counterweights and proposed a PSO-based model for counterbalance optimization. This model uniquely allows adjustments to both the magnitude and the phase angle of the maximum counterbalance moment during the optimization process. Various balancing methods, including beam, crank, and compound balancing, have been designed and implemented in BPUs (Xu et al., 2022). However, the implementation of these strategies still lacks specific theoretical guidance. Additionally, these balance adjustment strategies often disregard the effects of polished rod load fluctuations, yet the actual measured suspension loads can vary greatly and directly dictate the torque of the gearbox. Therefore, conducting balance optimization based on dynamometer card data carries more practical significance in engineering contexts.

Based on the analysis presented, it can be concluded that enhancing the stability and energy-saving effect of the compound balanced BPU requires both structural optimization and balance performance optimization. In recent years, numerous optimization metaheuristic algorithms and their improved versions (Zakian and Kaveh, 2024) have been developed and successfully applied to complex optimization challenges. Kaveh and Zaeerza (2023) introduced the shuffled shepherd optimization algorithm (SSOA) and its enhanced variants. These algorithms have been effectively applied to the optimal design of castellated beams, curved roof frames, and the detection of structural damage. Additionally, Kaveh and Mahdavi (2015, 2017) developed the colliding bodies optimization (CBO) algorithm and its enhanced versions, specifically designed for the optimal design of skeletal structures. This algorithm is particularly effective for structural optimization tasks as it leverages the principles of momentum and energy conservation during collisions. Following these contributions, a general framework for population-based metaheuristic algorithms, grounded in set theory (Kaveh and Hamedani, 2022), was proposed. This framework enhances the balance between global exploration and local exploitation during the search process, ensuring more effective optimization results across diverse applications. Zakian et al. (2021) addressed the design optimization problem of minimizing the weight of steel pipe rack structures using three metaheuristic algorithms: the modified PSO, the GWO, and the improved GWO. Mirjalili et al. (2017) introduced the multi-verse optimization algorithm and further advanced it by developing the multi-objective multi-verse optimizer (MOMVO) (Kumar et al., 2023). This algorithm was rigorously tested on 80 case studies, encompassing both unconstrained and constrained test functions as well as the engineering problems. The results demonstrated that MOMVO is capable of effectively approximating the true Pareto optimal fronts across all case studies, showcasing its efficacy. Yang et al. (2022) proposed the judgment-rule-based evolutionary algorithm, called JREA, which leverages two critical features of structural balance: the determination of node attributes based on the node's degree and the attributes of its neighboring nodes. JREA is primarily designed to achieve structural balance in signed social networks, aiming to guide social systems towards a more stable and harmonious state. Ma et al. (2020) integrated two orthogonal design (OD) engines into the brain storm optimization (BSO) algorithm to enhance its learning mechanism. The exploration OD engine and the exploitation OD engine work in tandem to discover and utilize valuable search experiences. This enhanced algorithm was validated on a suite of benchmarks and was also applied to the quantitative association rule mining problem, considering support, confidence, comprehensibility, and net confidence, proving its potency in optimizing complex functions. Rao et al. (2011) introduced a teaching-learning based optimization (TLBO) algorithm, which was applied to constrained mechanical design optimization

problems. Tejani et al. (2017) improved this algorithm by incorporating multiple teachers' models, adaptive teaching factors, self-motivated learning, and tutorial-based learning. The modified TLBO was applied to simultaneous topology, shape, and size optimization of spatial and planar trusses, and its performance was comparable to that of the original TLBO and other state-of-the-art algorithms, such as a genetic algorithm (GA), improved GA, force method and GA, ant colony optimization, adaptive multi manipulation differential evolution, a firefly algorithm, group search optimization (GSO), improved GSO, and intelligent garbage can decision-making model evolution algorithm. Tejani et al. (2017b) proposed a modified heat transfer search (HTS) algorithm that incorporates subpopulation-based simultaneous heat transfer modes, including conduction, convection, and radiation, into the basic HTS algorithm. The modified HTS (MHTS) algorithm was tested on six standard truss problems, and the results indicated that MHTS outperforms the original HTS algorithm. Metaheuristic algorithms have played an irreplaceable role in structural optimization, including large-scale and complex designs (Kaveh and Ghazaan, 2018; Kaveh and Eslamlou, 2020). More algorithms and their engineering applications, as well as MATLAB programs for these algorithms, can be found in references (Kaveh, 2021; Kaveh and Bakhshpoori, 2019). Although these algorithms have demonstrated exceptional performance in addressing complex multi-objective and constrained optimization challenges, the constraints involved in optimizing the structure of the BPU are notably stringent. Additionally, the objective function tends to exhibit reduced sensitivity to design variables near the optimal solution. As a result, the algorithm selected for this study must not only offer high computational efficiency but also ensure consistent and reliable optimization outcomes. To achieve these objectives, and to ensure both high precision and minimal programming complexity, three widely used metaheuristic optimization algorithms have been chosen: PSO (Wang et al., 2021; Zakian et al., 2021b), GA, and GWO (Zamfirache et al., 2022). These algorithms have been carefully refined to meet the unique challenges presented by this research.

In summary, the existing literature lacks a comprehensive and systematic approach for optimizing the design of compound balanced BPUs, particularly in determining structural dimensions and adjusting balance performance. This paper addresses these gaps by presenting a novel optimization strategy that integrates advanced algorithms, such as PSO, GA, and GWO, to enhance both the structural design and balance performance of BPUs. The approach is characterized by its ability to navigate the complex constraints and objectives inherent in BPU optimization, marking a significant advancement over conventional methods. The main contributions of this research are multifaceted. Firstly, a systematic framework is introduced that unifies structural and balance optimization, offering a holistic solution to BPU design challenges. Secondly, empirical analysis demonstrates the superior performance of the modified GWO algorithm, highlighting its robustness and reliability in achieving optimal solutions. Thirdly, the paper advocates for a practical balance optimization method that leverages dynamometer card data, ensuring that the solutions are applicable in practical scenarios. Furthermore, this research bridges the theoretical and practical aspects of BPU optimization, providing valuable insights for both academic researchers and industry practitioners. The comprehensive approach presented is expected to significantly improve the energy efficiency and operational effectiveness of BPUs, contributing to the sustainability and economic viability of oil extraction processes.

The paper is organized as follows: Section 2 provides an in-depth introduction to the structural characteristics of the compound balanced BPU, establishing the correlation between the polished rod load and gear torque, grounded in the principle of

D'Alembert. Section 3 elaborates on the prevalent methodologies for ascertaining the polished rod load, thereby establishing a basis for the balance optimization approach that utilizes measured dynamometer card. Section 4 delineates the specific variables involved in the structural and balance optimization design of the BPU, setting forth rational constraints and objective functions that are informed by actual engineering applications. Subsequently, the detailed scheme for both structural and balance optimization is presented. Throughout these optimizations, three optimization algorithms, such as PSO, GA, and GWO are utilized. These algorithms have been refined to leverage their unique attributes, with the dual aim of enhancing computational efficiency and ensuring that the obtained parameters are suitable for practical engineering applications. Section 5 illustrates the specific implementation steps and optimization outcomes of the computational methods proposed in this study through practical examples, thereby validating the effectiveness of proposed approach. It also serves to demonstrate the applicability of the refined GWO algorithm in the structural and balance optimization of compound balanced BPUs.

2. Calculation of gearbox torques

2.1. Introduction to the structure of compound balanced BPU

The focus of this paper is on a compound balanced BPU that incorporates an offset section in the rear arm of the beam. This configuration is particularly representative and provides a mechanical model that is relevant to other conventional BPUs. Fig. 1 illustrates the structure of the BPU being analyzed. The compound balanced BPU consists of several components: a brake device (1), motor unit (2), frame (3), carrier bar (4), horsehead (5), samson post bearing (6), main beam (7), downward beam (8), pitman (9), beam counterweight (10), gearbox (11), crank counterweight (12), crank (13), and foundation (14). The inclusion of the downward beam is specifically designed to facilitate adjustments of the beam counterweight by operators.

2.2. Kinematic analysis of BPU

A schematic of the mechanism, as shown in Fig. 2, was established based on the structure of the BPU. The four members R , P , C , and K of the 4-bar linkage mechanism can be represented by four vectors: \vec{R} , \vec{P} , \vec{C} and \vec{K} . The load on the suspension point is denoted as Q . γ_1 is the angle between the downward beam and the main beam. γ_2 is the angle between the moment arm of the gravity center of beam counterweight center and the main beam. γ_3 is the

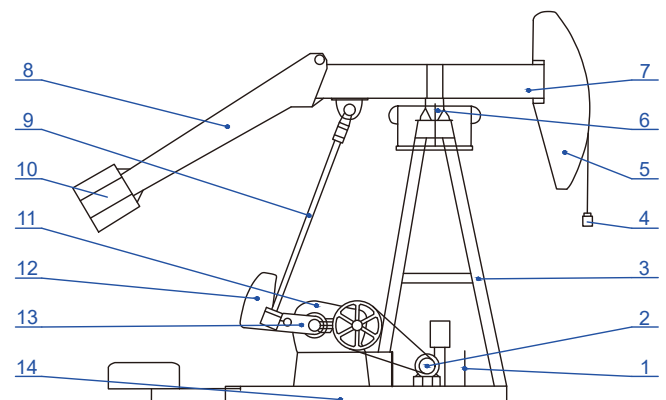


Fig. 1. The components of compound balanced BPU.

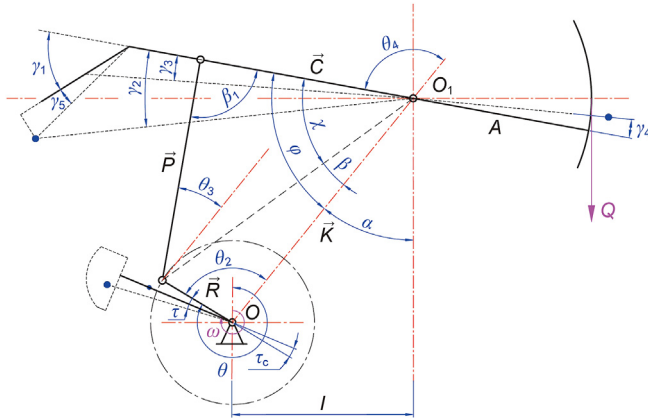


Fig. 2. Schematic of the mechanism of BPU.

angle between the moment arm of the gravity center of downward beam and the main beam. γ_4 is the angle between the moment arm of the gravity center of horsehead and the main beam. τ is the lag angle of the gravity center of the crank counterweight. τ_c is the lag angle of the gravity center of the crank. For the convenience of dynamic analysis, the conventions for the positive direction of each angle in Fig. 2 are as follows:

- 1) The crank's rotation angle θ is measured starting from the 12 o'clock position, with clockwise rotations considered positive.
- 2) The reference angles of the rods $\theta_2, \theta_3, \theta_4$ are measured from the base rod OO_1 , with counterclockwise direction considered positive.
- 3) The geometric dimensions of each member are specified as follows: R is the rotation radius of the crank pin, P is the length of the pitman, C is the length of the rear arm of the main beam, K is the length of the base rod, A is the length of the fore-arm of the main beam, and I is the horizontal projection length of the base rod.

The geometric relationship of the mechanism shown in the schematic (Fig. 2) can be described using the following expressions:

$$\alpha = \arcsin\left(\frac{I}{K}\right) \quad (1)$$

$$\theta_2 = 2\pi - \theta + \alpha \quad (2)$$

$$L = \sqrt{R^2 + K^2 - 2RK \cos \theta_2} \quad (3)$$

$$\beta = \arcsin\left(\frac{R}{L} \sin \theta_2\right) \quad (4)$$

$$\theta_3 = \arccos\left(\frac{P^2 + L^2 - C^2}{2PL}\right) - \beta \quad (5)$$

$$\theta_4 = \arccos\left(\frac{P^2 - L^2 - C^2}{2CL}\right) - \beta \quad (6)$$

$$\chi = \arccos\left(\frac{C^2 + L^2 - P^2}{2CL}\right) \quad (7)$$

$$\varphi = \chi + \beta \quad (8)$$

In the above equations, the virtual rod is an imaginary line that connects the crank pin to the beam's center of rotation, with its length L , dynamically changing as the pumping unit operates. The angles involved are defined as follows: α is the angle between the base rod K and the vertical direction; θ_2 is the angle between the crank and the base rod; θ_3 is the angle between the pitman and the base rod; θ_4 is the angle between the main beam and the base rod; β is the angle between the virtual rod and the base rod; χ is the angle between the main beam and the virtual rod; and φ , the angle between the main beam and the base rod, is the sum of β and χ . For a detailed understanding of the geometric meanings of these variables, Fig. 2 should be consulted. Each vector has the following relationship:

$$\vec{R} + \vec{P} = \vec{K} + \vec{C} \quad (9)$$

The above vector equation can be represented in complex form:

$$Re^{i\theta_2} + Pe^{i\theta_3} = K + Ce^{i\theta_4} \quad (10)$$

According to Euler's formula $e^{i\theta} = \cos \theta + i \sin \theta$, Eq. (10) can be transformed into:

$$R(\cos \theta_2 + i \sin \theta_2) + P(\cos \theta_3 + i \sin \theta_3) = K + C(\cos \theta_4 + i \sin \theta_4) \quad (11)$$

Taking the derivative of both sides of Eq. (11) with respect to time yields the following expression:

$$iR\dot{\theta}_2 \cos \theta_2 - R\dot{\theta}_2 \sin \theta_2 + iP\dot{\theta}_3 \cos \theta_3 - P\dot{\theta}_3 \sin \theta_3 = iB\dot{\theta}_4 \cos \theta_4 - B\dot{\theta}_4 \sin \theta_4 \quad (12)$$

If the real and imaginary parts on both sides of the equation are equal, the following system of equations can be obtained:

$$\begin{cases} R\dot{\theta}_2 \sin \theta_2 + P\dot{\theta}_3 \sin \theta_3 = B\dot{\theta}_4 \sin \theta_4 \\ iR\dot{\theta}_2 \cos \theta_2 + iP\dot{\theta}_3 \cos \theta_3 = iB\dot{\theta}_4 \cos \theta_4 \end{cases} \quad (13)$$

By solving the above equation system, the angular velocities $\dot{\theta}_3$ and $\dot{\theta}_4$ of the connecting rod and walking beam can be determined:

$$\begin{cases} \dot{\theta}_3 = \frac{R\dot{\theta}_2 \cdot \sin(\theta_4 - \theta_2)}{P \cdot \sin(\theta_3 - \theta_4)} \\ \dot{\theta}_4 = \frac{R\dot{\theta}_2 \cdot \sin(\theta_3 - \theta_2)}{B \cdot \sin(\theta_3 - \theta_4)} \end{cases} \quad (14)$$

By differentiating the above equation with respect to time, the angular accelerations $\ddot{\theta}_3$ and $\ddot{\theta}_4$ of the connecting rod and beam motion can be determined:

$$\begin{cases} \ddot{\theta}_3 = \dot{\theta}_3 \left[\frac{\ddot{\theta}_2}{\dot{\theta}_2} + (\dot{\theta}_4 - \dot{\theta}_2) \cot(\theta_4 - \theta_2) - (\dot{\theta}_3 - \dot{\theta}_4) \cot(\theta_3 - \theta_4) \right] \\ \ddot{\theta}_4 = \dot{\theta}_4 \left[\frac{\ddot{\theta}_2}{\dot{\theta}_2} + (\dot{\theta}_3 - \dot{\theta}_2) \cot(\theta_3 - \theta_2) - (\dot{\theta}_3 - \dot{\theta}_4) \cot(\theta_3 - \theta_4) \right] \end{cases} \quad (15)$$

where,

research, the practical applicability and ease of implementation of these findings continue to be debated among scholars. Several commonly used engineering methods for calculating polished rod load are introduced below, each addressing different aspects of these complex influences.

1) Constant mass method

Hanging test blocks on the carrier bar is a straightforward method commonly used to verify the structural strength of the BPU prototype under rated polished rod load. Assuming the mass of the

$$\begin{cases} Q_{\text{up}} = (m_{\text{rod}} + m_{\text{pump}} + m_{\text{oil}})(g + a) + p_o(A_{\text{plunger}} - A_{\text{rod}}) - (p_c + \rho_{\text{oil}}gh_s)A_{\text{plunger}} \\ Q_{\text{down}} = m_{\text{rod}} \frac{\rho_{\text{rod}} - \rho_{\text{oil}}}{\rho_{\text{rod}}} g + m_{\text{rod}} a \end{cases} \quad (25)$$

test block is m , the polished rod load can be expressed using Eq. (24). For optimizing the balance of the BPU, the mass of the test blocks can be estimated based on the average of the maximum and minimum polished rod loads.

$$Q = m(a + g) \quad (24)$$

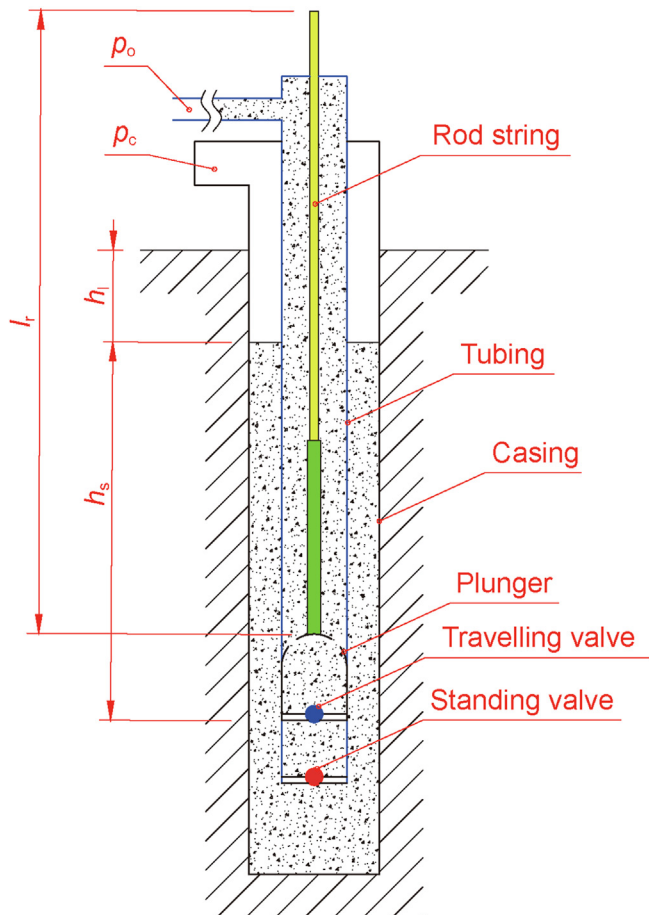


Fig. 4. Structural of the oil well.

2) Simplified dynamic load

The typical structural of an oil well is illustrated in Fig. 4. In this simplified analysis, no consideration is given to the deformation of the tubing and sucker rod, as well as the resistance of the fluid and piston during motion. The analysis focuses solely on the effects of gravity and inertia of the sucker rod and the lifted liquid. Under these assumptions, the corresponding polished rod loads for the upstroke and downstroke can be calculated as follows;

in which, Q_{up} is polished rod load corresponding to the upstroke; Q_{down} is polished rod load corresponding to the downstroke; m_{rod} is the mass of rod string; m_{pump} is the mass of pump; m_{oil} is the mass of lifted oil; p_o is the internal pressure of tubing; p_c is the inner pressure of casing; ρ_{oil} is the density of oil; ρ_{rod} is the density of rod string; h_s is submergence depth of pump. A_{plunger} is the cross-sectional area of plunger; A_{rod} is the cross-sectional area of rod string.

3) Dynamometer card data

As established in Section 2.3, all gearbox torque components are influenced by the polished rod load. The rod torque varies with the torque factor that changes with the crank angle, while the counterbalance torque is directly related to the crank angle and angular acceleration of the beam. Inertial torques, which are derived from the acceleration patterns of various components of the BPU, exhibit fluctuations that correspond to the time-varying history of the crank angle. Furthermore, understanding the dynamics of gearbox torque components requires detailed knowledge of the polished rod load and its temporal variations in position.

Modern dynamometer systems (Li et al., 2015) are essential for capturing the most important operational data of sucker rod pumping systems, namely the polished rod load and its position over time. The position of the polished rod varies with the crank angle. Following the kinematic analysis detailed in Section 2.2, the displacement of the polished rod is expressed by Eq. (26). With this data, the relationship between the crank angle and polished rod load can be accurately determined.

$$S = \left(\theta_4 - \arccos \left(\frac{K^2 + C^2 - (P + R)^2}{2KC} \right) \right) \cdot A \quad (26)$$

4. Structural design and optimization of BPU

4.1. Design variables and general optimization strategies

The kinematic and dynamic analyses of the BPU discussed in Section 2 form the theoretical foundation for its structural design. The parameters that influence the net torque of the crankshaft can be categorized into two main groups: structural and mass-related. These categories of parameters are illustrated in Fig. 5. This

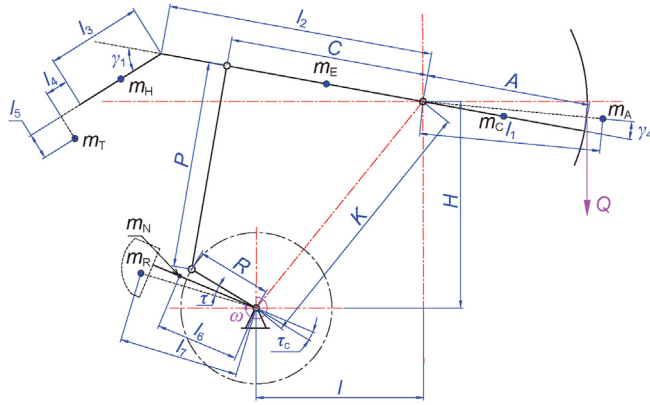


Fig. 5. Schematic of the structure and mass parameters of the BPU.

distinction is critical as it guides the selection of design variables and the application of optimization strategies to enhance the efficiency and functionality of the BPU.

The design of the BPU involves a total of 17 independent structural parameters, including variables such as γ_4 , l_1 , A , C , l_2 , l_3 , γ_1 , l_4 , l_5 , P , R , l_6 , τ_c , l_7 , τ , H , I . Additionally, there are 7 mass parameters which include variables such as m_A , m_C , m_E , m_H , m_T , m_N , m_R . Here is a more detailed breakdown of some critical parameters: m_A represents the mass of the horsehead, with γ_4 and l_1 determined by the horsehead's center of gravity. m_C represents the mass of the fore-arm of the main beam. m_E and l_2 respectively denote the mass and length of the beam's rear arm. Given that both the front and rear arms of the beam are box girders with identical cross-sectional dimensions, it's reasonable to assume their centers of gravity are situated at their midpoints. m_H and l_3 are, respectively, the mass and length of the downward beam. Unlike the front and rear arms, the cross-section of the downward beam may vary along its axis, necessitating a specific determination of its center of gravity based on its structure. l_4 and l_5 are the parameters utilized to ascertain the position of the beam counterweight's center of gravity relative to the end of the downward beam. m_T is the mass of the beam counterweight. m_N refers to the mass of the crank, with l_6 indicating the rotation radius of the crank's center of gravity. m_R is the mass of the crank counterweight, which shares l_7 as the rotation radius for its center of gravity. Some parameters are fixed and do not require redesign, which will be detailed in the subsequent sections.

- 1) In actual design of the BPU' structure, the length (A) of the fore-arm is approximately equal to the maximum stroke of the BPU. The swing angle of the walking beam is about 1 radian.
- 2) Typically, the length of the beam's rear arm and downward beam are not adjustable at the well site. To facilitate the adjustment of the balance block of the beam on the production site, it is crucial to minimize the ground clearance at the end of the lower beam. Additionally, interference between components must be avoided. The cross-sectional dimensions of the main beam and the downward beam can be designed based on existing related products. Once the size of the 4-bar linkage is determined, the structure (i.e. l_2 , l_3 and γ_1) and mass (i.e. m_E and m_H) of the beam's and downward beam can be determined.
- 3) Horsehead is a modular component that can be selected based on stroke requirements of the BPU. Parameters related to its mass (m_A) and center of gravity (i.e., γ_4 and l_1) can also be determined.
- 4) The total length of crank is typically based on experiential data. The mass (m_N) of the crank and the rotation radius (l_6) of its

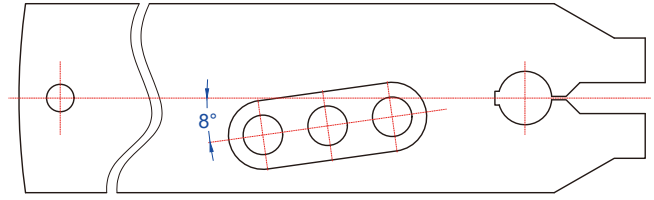


Fig. 6. Schematic of the structure of crank.

center of gravity can be accurately predicted. The crank features three pin holes for quick adjustment of the stroke, as shown in Fig. 6. Assuming the influence of pin holes on the uniformity of crank mass distribution is negligible, the lag angle (τ_c) of the center of gravity is equivalent to the angle between the rotation radius of the pin hole and the middle plane of the crank. After determining the size of the crank, the rotation radius (l_7) of the center of gravity of the crank counterweight is solely dependent on its shape, and its mass varies only with changes in thickness.

- 5) Similar to the counterweight of the crank, the counterweight of the beam is typically composed of stacked weight plates. Therefore, l_4 and l_5 are primarily determined by the shape of the weight plates.

Based on the analysis provided above, it can be concluded that the remaining variables to be designed include H , I , C , P , R , τ , m_R , m_T . When optimizing the structure of the BPU, size design and balance design can be conducted separately. Initially, the structural dimensions of the BPU are determined. This process aims to minimize the torque factor (TF) and the acceleration of suspension point (a_{max}), both of which are critical for improving the dynamic performance of BPU. The variables related to it include H , I , C , P , R . Subsequently, the remaining variables are optimized to minimize the RMS torque M_e , which influences the energy consumption of the BPU. The variables involved in this phase include τ , m_R , m_T . Takacs and Kis (2021) described the benefits of optimizing the counterweight of the BPU in detail, and recommended using the reduction of the cyclic load factor (CLF) as the optimization goal. The CLF is directly related to RMS torque. A comprehensive introduction to balance evaluation strategy of BPUs will be provided in the subsequent sections.

4.2. Constrains

The constraint conditions significantly impact the structural dimensions of the BPU. Only suitable and effective constraints can provide necessary support for quickly and accurately optimizing the structural dimensions of the BPU. Drawing on existing design experience with BPUs, several key constraints have been formulated to guide the optimization process.

- 1) Crank rocker mechanism (Ghosal, 2010)

The crank rocker mechanism forms an essential component of the BPU. It must be carefully designed to ensure efficiency and reliability. The following requirements must be met for the dimensions related to this mechanism:

$$R < K \quad (27a)$$

$$R < P \quad (27b)$$

$$R < C \quad (27c)$$

$$R + K < C + P \quad (27d)$$

2) Maximum stroke of polished rod

Typically, the stroke length (S_0) of the polished rod is a critical design parameter for a BPU. To simplify the calculation procedure, the stroke length (S) of the polished rod is defined within a specific interval, ensuring sufficient accuracy. In the equation below, “ e ” represents the relative error, which should not exceed 1%.

$$S \in [S_0 - e, S_0 + e] \quad (27e)$$

3) Transmission angle

The transmission angle (denoted as β_1 in Fig. 2) is the angle between the connecting rod and the beam. The recommended interval for the transmission angle β_1 is

$$\beta_1 \in [30^\circ, 150^\circ] \quad (27f)$$

4) Extreme position angle

There is no clear requirement for the extreme position angle. Zhang et al., (2005) provides several sets of recommended values based on the circulation characteristics of the BPU. Hu et al., (2006) recommends that the interval of the extreme position angle for a compound balanced BPU should be within specific ranges to ensure optimal performance and balance. The recommended interval for the extreme position angle is

$$\lambda \in [8^\circ, 12^\circ] \quad (27g)$$

5) The symmetry of the swing angle of the beam

In general, it is required that the swing angle of the beam is approximately symmetrically distributed with respect to the horizontal line. Assuming that the swing angle of the beam is divided into two parts by the horizontal line, the upper part is δ_{up} and the lower part is δ_{down} . To ensure symmetry, Hu et al., (2006) recommends the interval of the angle difference should be within a specific range. The recommended interval for the angle difference is

$$\delta_{up} - \delta_{down} \in [-10^\circ, 10^\circ] \quad (27h)$$

6) The angle α

The α represents the angle between the base rod and the perpendicular line. This parameter may affect the stability and structural dimensions of the BPU. Reference (Zhang et al., 2005), combined with design experience, provides a reference value for α as follows:

$$\alpha \in [32^\circ, 50^\circ] \quad (27i)$$

4.3. Summary of the optimization problem

Given all the information in the preceding sections, the entire optimization problem can be summarized as follows:

To minimize:

$$f_{obj1} = \min(|a|), a = f(H, I, C, P, R) \quad (28)$$

$$f_{obj2} = \min(M_e), M_e = f(\tau, M_R, M_H) \quad (29)$$

Subjected to:

$$\begin{cases} R < K \\ R < P \\ R < C \\ R + K < C + P \\ S \in [S_0 - e, S_0 + e], e/S_0 = 1\% \\ \beta_1 \in [30^\circ, 150^\circ] \\ \lambda \in [8^\circ, 12^\circ] \\ \delta_{up} - \delta_{down} \in [-10^\circ, 10^\circ] \\ \alpha \in [32^\circ, 50^\circ] \end{cases} \quad (30)$$

The formula for calculating the RMS torque is as follows:

$$M_e = \sqrt{\frac{1}{2\pi} \int_0^{2\pi} M^2 d\theta} = \sqrt{\frac{\sum (M_i^2 \Delta\theta_i)}{\sum \Delta\theta_i}} \quad (31)$$

in which, M_i is the instantaneous torque at the i -th specific moment; $\Delta\theta$ is the increment of the crank rotation angle; subscript “ i ” represents a specific moment in time.

4.4. Optimization scheme

The optimization of the relevant parameters of the BPU, as described in Section 5.1, is divided into two stages based on the objective function, as illustrated in Fig. 7. This division transforms each stage into a single-objective optimization problem, thereby reducing the overall complexity. From a programming perspective, the computational procedures for both stages are fundamentally identical. The primary difference lies in the imposition of constraints: in the first stage, constraints are explicitly applied, whereas in the second stage, these constraints are inherently satisfied. As a result, the core structure of the programming remains consistent across both stages; only the range of the design variables needs to be adjusted according to the specific objective function. Additionally, the Latin Hypercube Sampling (LHS) method is utilized for generating initial samples, which promotes the generation of more uniformly distributed samples. This method serves a dual purpose: it initially verifies the suitability of the design variable range in the first stage and, as experience has shown, the generation of qualified initial samples that satisfy the constraints can substantially expedite the subsequent optimization process. To achieve the global optimal solution, optimization schemes based on three different algorithms are implemented. A comprehensive explanation of each optimization scheme follows.

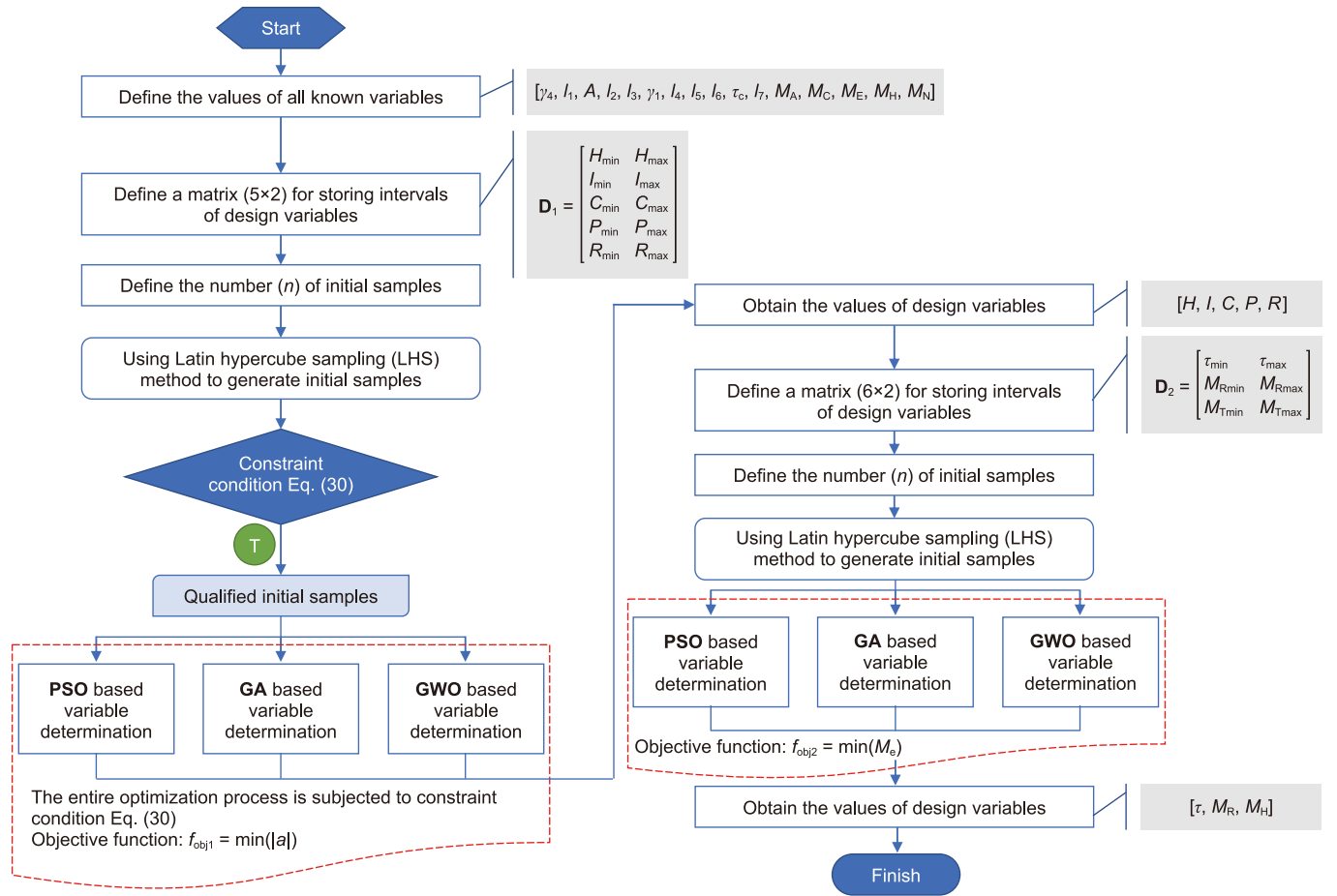


Fig. 7. Parameter optimization scheme flowchart.

4.4.1. The particle swarm optimization (PSO) (Liu et al., 2015; An et al., 2022) based variable determination method

A particle swarm is composed of N particles, each of which is a D -dimensional vector. The dimension D of a vector corresponds to the number of design variables. The spatial position of particles can be represented as

$$x_i = (x_{i1}, x_{i2}, \dots, x_{iD}), i = 1, 2, \dots, N \quad (32)$$

where, x_i denotes the position of the i -th particle in the D -dimensional design space. x_{ij} represents the value of the j -th design variable for the i -th particle. The spatial position of particles is a solution in the objective optimization problem. The fitness of particles can be evaluated through a fitness function, which allows the quality of particles to be measured based on the fitness value. The initial particle swarm consists of qualified initial samples after screening.

The flying speed of the i -th particle is also a D -dimensional vector, denoted as

$$v_i = (v_{i1}, v_{i2}, \dots, v_{iD}), i = 1, 2, \dots, N \quad (33)$$

where, v_i represents the velocity of the i -th particle, v_{ij} represents the speed component of the j -th dimension for the i -th particle. The initial values of velocity are randomly generated within the specified range. The spatial position where the i -th particle has the optimal fitness value is called the individual's historical optimal position p_{besti} , denoted as

$$p_{besti} = (p_{besti1}, p_{besti2}, \dots, p_{bestiD}), i = 1, 2, \dots, N \quad (34)$$

The optimal position of the entire particle swarm is called the global historical optimal position g_{besti} , denoted as

$$g_{besti} = (g_{besti1}, g_{besti2}, \dots, g_{bestiD}), i = 1, 2, \dots, N \quad (35)$$

The position update of particle swarm can be achieved through the update of both speed and position. The update of velocity can be expressed by the following formula.

$$v_{ij}(t+1) = v_{ij}(t) + c_1 r_1 (p_{bestij}(t) - x_{ij}(t)) + c_2 r_2 (g_{bestij}(t) - x_{ij}(t)); \\ i = 1, 2, \dots, N; j = 1, 2, \dots, D \quad (36)$$

Then, the position of the particles can be updated using the following formula:

$$x_{ij}(t+1) = x_{ij}(t) + v_{ij}(t+1) \quad (37)$$

where, subscript j represents the j -th dimension of the particle; subscript i represents the i -th particle; t represents the current number of iterations; c_1 and c_2 are individual and social learning factors, respectively, which determine the acceleration. The value range is usually $(0, 2)$; r_1 and r_2 are independent random numbers with values in the range of $[0, 1]$. Updating the spatial positions of particles means obtaining a new set of design variables that may not meet the constraint conditions, and even their fitness values

may even be singular. To avoid this problem, mutation operations from genetic algorithms are used. Each design variable is randomly updated within the specified interval until the particle meets the constraint conditions. In this paper, the values of c_1 and c_2 are both set to 2, while the maximum speeds of variables related to length, angle, and mass are set to 10, 1, and 5, respectively. This differentiated setting helps in fine-tuning the search process according to the specific characteristics of the problem at hand. It's worth noting that these parameter settings do introduce some complexity in programming, as they require a careful handling of the various elements involved in the particles before their velocities and positions are updated. However, this approach is justified as it allows for a more nuanced and effective optimization process. The values mentioned are not set in stone and are determined based on the parameter ranges, with the flexibility to be adjusted through a trial-and-error method. The detailed process is shown in Fig. 8.

4.4.2. The genetic algorithm (GA) based variable determination method

GA (Ongkunaruk et al., 2016; Chen et al., 2017) was first proposed by John Holland from the United States in the 1970s. The algorithm is designed based on the evolutionary principles observed in nature. It is a computational model that simulates the natural selection and genetic mechanisms of Darwin's theory of biological evolution. GA searches for optimal solutions by mimicking the natural evolution process. This algorithm uses mathematical methods and computer simulation operations to transform the problem-solving process into one that resembles the crossover and mutation of chromosome genes in biological evolution. In GA, both crossover and mutation operations can generate

new individuals. To avoid parameter optimization failures caused by singular fitness, the new individuals generated by each operation are tested against constraint conditions. If the constraint conditions are not met, the current operation is revoked. The detailed process is shown in Fig. 9. The crossover probability and mutation probability are set to 0.8 and 0.15, respectively. The genes (i.e., binary bits) are represented by 20 bits.

4.4.3. The gray wolf optimization (GWO) based variable determination method

The GWO is an intelligent optimization algorithm proposed by Mirjalili (2015) from Griffith university in Australia in 2014. The inspiration for GWO comes from the predatory behavior of gray wolf packs. The GWO algorithm simulates the leadership hierarchy and hunting mechanism of gray wolves in nature. Gray wolves are divided into four types to simulate hierarchical levels. Additionally, three main stages of hunting are simulated: searching for prey, surrounding prey, and attacking prey. In this algorithm, the best solution is considered the alpha (α) wolf, the second-best is the beta (β) wolf, and the third-best is the delta (δ) wolf (Meng et al., 2023). In the solution domain of optimization problems, the optimal solution (the location of prey) is often unclear. To simulate the hunting behavior of gray wolves, it is assumed that the alpha (α) wolf, beta (β) wolf, and delta (δ) wolf have strong abilities to identify potential prey positions. In each iteration process, the best 3 wolves (α , β , δ) in the current population are retained, and the positions of other search agents (ω) are updated based on their locations. The detailed process is shown in Fig. 10.

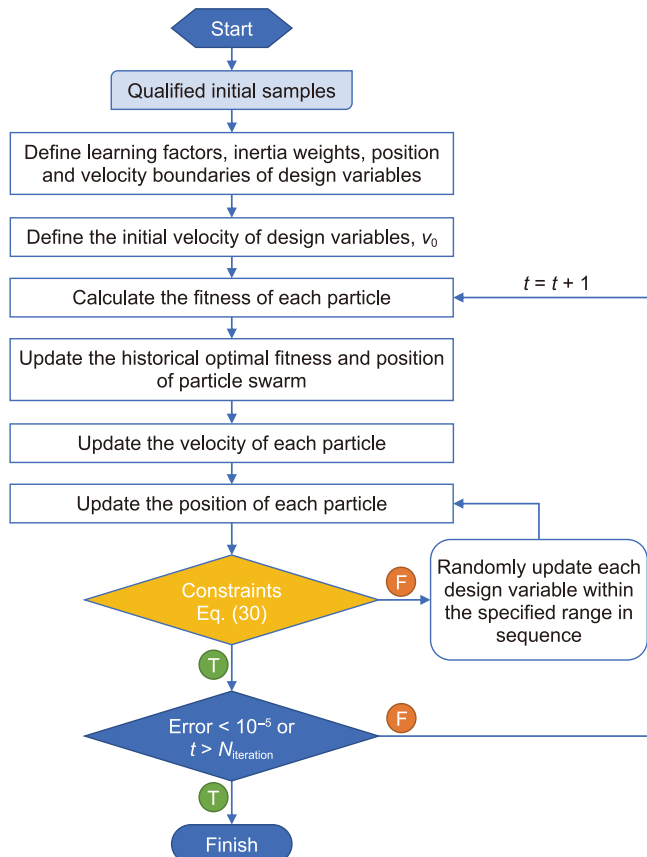


Fig. 8. Parameter optimization process based on the PSO algorithm.

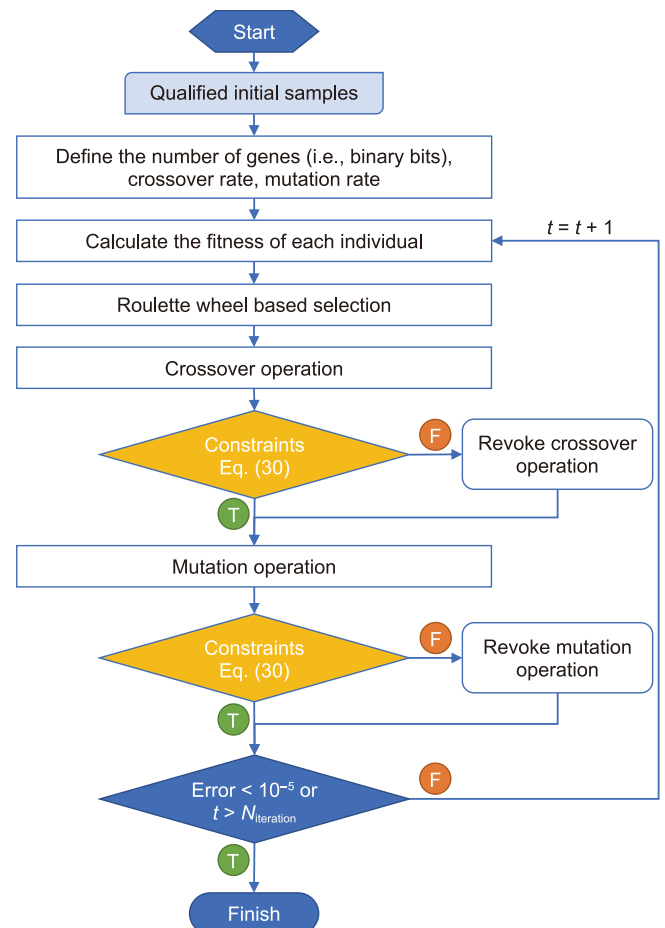


Fig. 9. Parameter optimization process based on the GA algorithm.

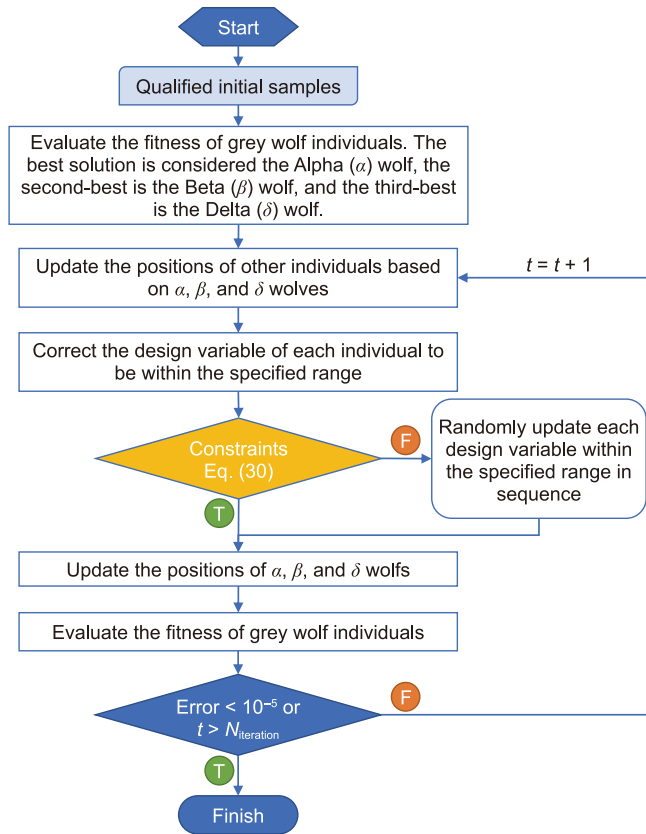


Fig. 10. Parameter optimization process based on the GWO algorithm.

The mathematical model of gray wolves tracking prey positions is described as follows:

$$\begin{cases} D_{\alpha} = |C_1 X_{\alpha}(t) - X(t)| \\ D_{\beta} = |C_2 X_{\beta}(t) - X(t)| \\ D_{\delta} = |C_3 X_{\delta}(t) - X(t)| \end{cases} \quad (38)$$

where, D_{α} , D_{β} and D_{δ} represent the distances between the α , β , δ , and other individuals, respectively; X_{α} , X_{β} and X_{δ} represent the current position of α , β and δ , respectively; C_1 , C_2 and C_3 are random vectors; X is the position of individual gray wolves. The mathematical model of hunting behavior is described as follows

$$\begin{cases} X_1 = X_{\alpha}(t) - A_1 D_{\alpha} \\ X_2 = X_{\beta}(t) - A_2 D_{\beta} \\ X_3 = X_{\delta}(t) - A_3 D_{\delta} \end{cases} \quad (39)$$

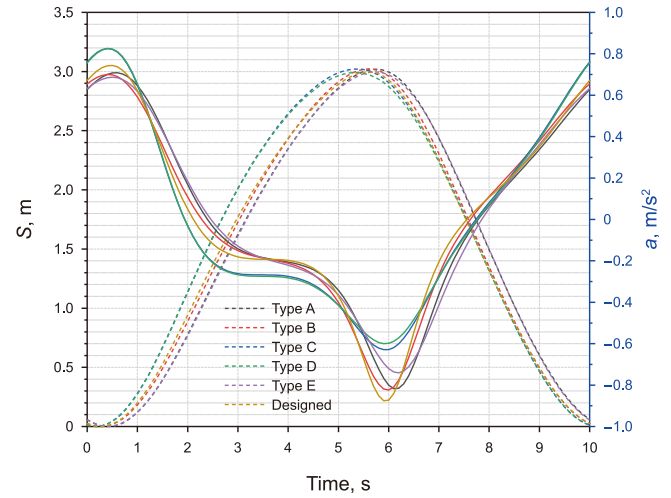


Fig. 11. The variation of stroke, polished rod acceleration, and displacement in single stroke.

$$X(t+1) = \frac{X_1 + X_2 + X_3}{3} \quad (40)$$

$$A = 2a \cdot r_1 - a \quad (41)$$

$$C = 2 \cdot r_2 \quad (42)$$

where, r_1 and r_2 are random numbers with values in the interval $[0, 1]$. The convergence factor a decrease linearly from 2 to 0 as the number of iterations increases. Similarly, an individual's position may not meet the constraint conditions after being updated. To address this, similar to the mutation operation in genetic algorithms, each design variable involved in the individual is updated to ensure compliance with the constraint condition.

5. Case study and analysis

5.1. Step 1: Optimization of basic structure

In the actual design of a BPU, the maximum stroke S_{\max} is a fundamental technical parameter, which is specifically requested by oilfield developers. To verify the effectiveness and advantages of the design method proposed in this paper, the results calculated using proposed method will be compared with the data from Hu et al., (2006). The references provide 5 sets of in-services and 1 set of designed BPU parameters, as detailed in Table 1. The maximum stroke S_{\max} listed on the nameplate of these BPUs is 3000 mm. However, there is a discrepancy between the actual values shown in the last row of Table 1 and the nameplate values.

Table 1
Basic structural parameters of various types of BPUs.

Structure parameter	In-service					Designed by Hu et al. (2006)	Proposed method
	Type A	Type B	Type C	Type D	Type E		
A, mm	3000	3000	3000	3000	3600	2520	3000
C, mm	2100	1860	2200	2200	2620	1620	[1500, 2700]
H, mm	3280	3250	3200	3200	3610	2950	[2800, 3700]
I, mm	2990	2400	2400	2400	3610	2030	[1900, 3800]
P, mm	3490	3200	3200	3240	3780	2950	[2800, 3900]
R, mm	950	840	1040	1040	990	860	[700, 1100]
S_{\max} , m	3.018	3.022	3.021	2.992	2.996	2.996	[2.970, 3.030]

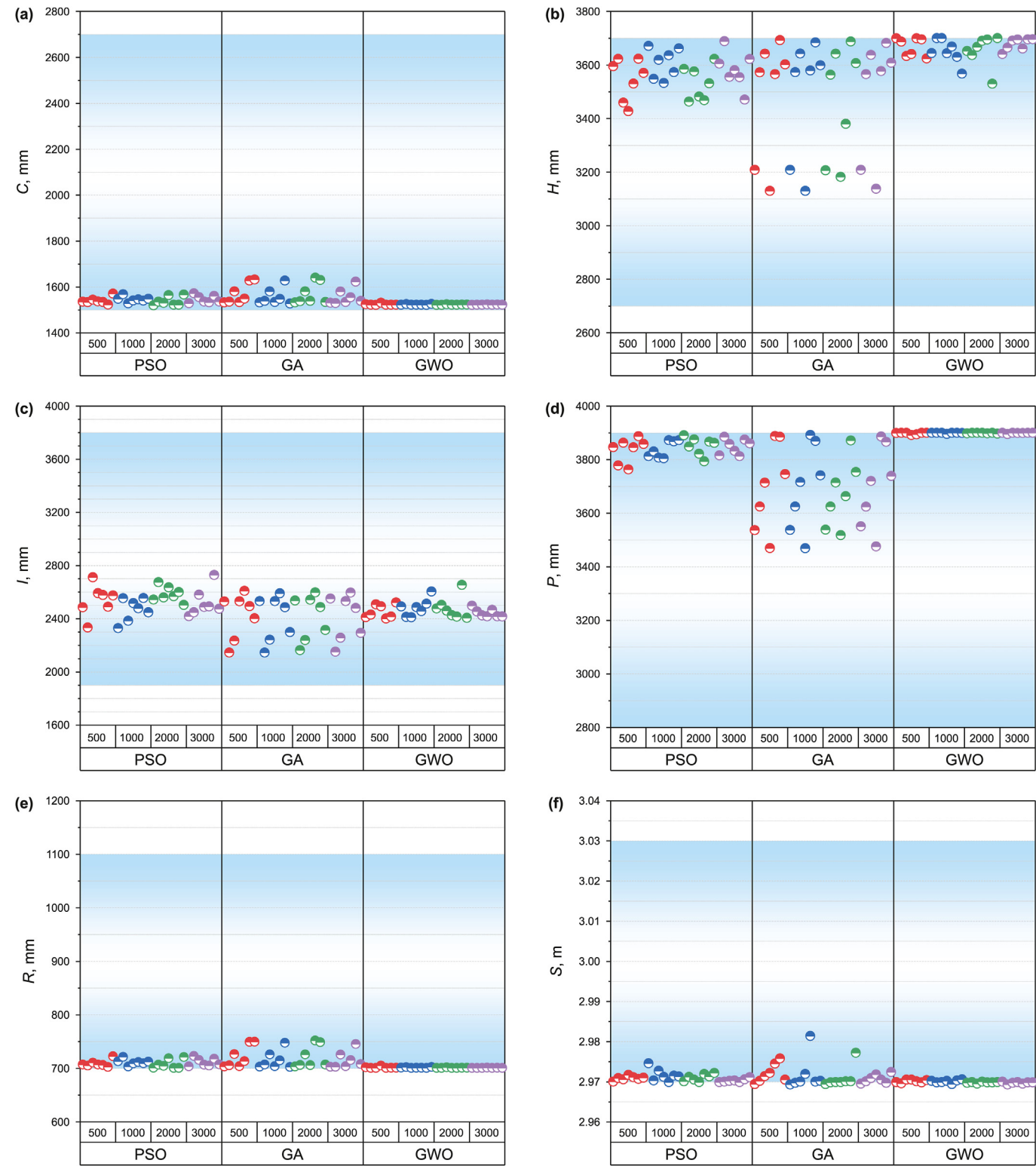


Fig. 12. (a)–(f) Optimization results for variables C , H , I , P , R , S .

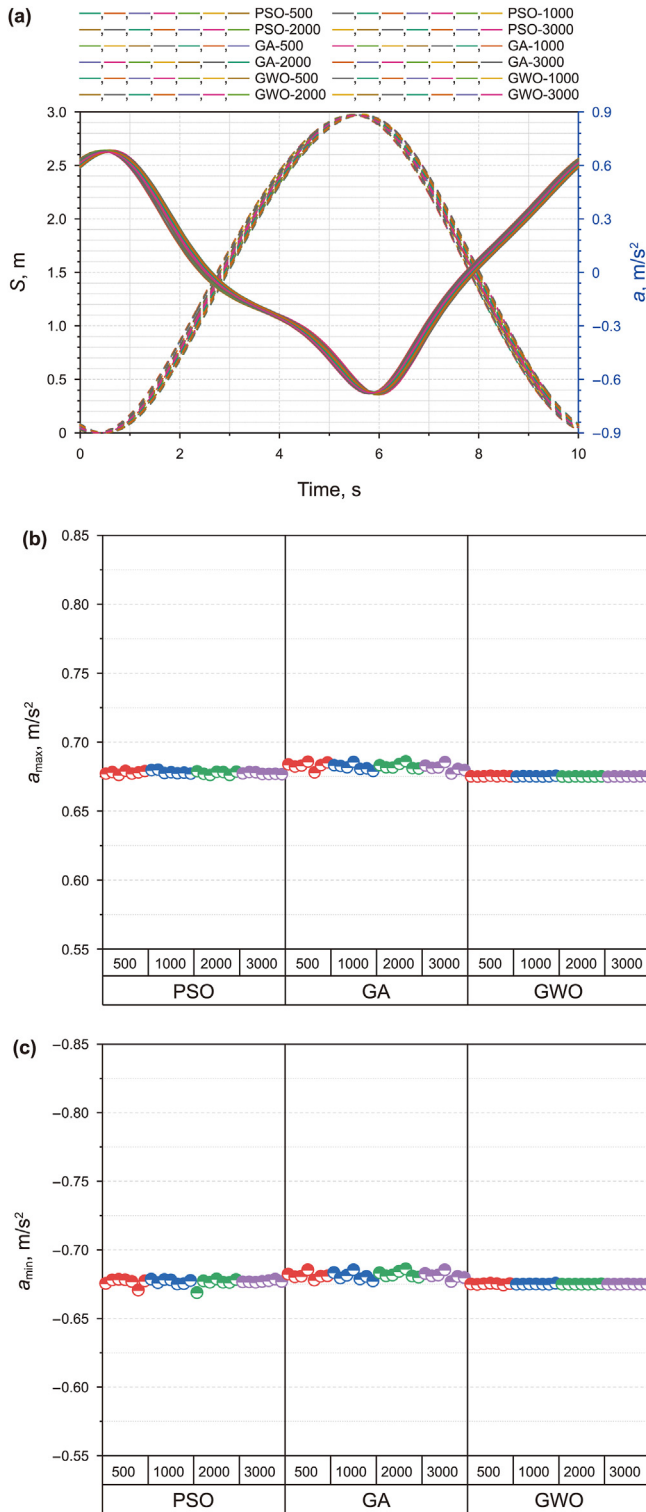


Fig. 13. (a) The displacement of suspension points. (b)–(c) The acceleration of suspension points.

The maximum strokes per minute are noted as 6 min^{-1} . The proposed optimization method is then applied to determine the basic structural parameters of the BPU. The range of design variables was determined based on the structural parameters of these six types of BPUs. These parameters provide a data basis for determining the intervals of design variables. The detailed intervals for design

variables are shown in the last column of Table 1. This method to determining the intervals of design variables can ensure that the overall shape of optimized PBU changes minimally, and also helps reduce computational costs.

In the first stage, the optimal design variables are primarily determined based on the acceleration of the suspension point. Fig. 11 illustrates the variation pattern of polished rod acceleration for the provided BPUs during one stroke. It is evident that there are significant differences in the polished rod acceleration among various types of BPUs, particularly concerning the minimum acceleration when the polished rod reaches the lower dead point. Additionally, there is a notable disparity in the extreme acceleration of the polished rod during the up and down strokes, as observed in types C and D.

Due to the use of three optimization algorithms in the parameter optimization scheme proposed in this article, 7 sets of calculations were conducted for comparative analysis to verify the convergence of the optimization scheme. An initial sample size of 100,000 was generated via the LHS method. The number of qualified samples after screening totaled 110, 118, 119, 121, 123, 124, and 125, respectively. Each optimization algorithm was tested in four sets of trials, with iterations set at 500, 1000, 2000 and 3000, respectively. The specified number of iterations was achieved by setting the threshold of error to 0. These 7 sets of random sampling were primarily used to test the robustness of each algorithm against initial sample variations. The four sets of iterations are primarily used to assess the convergence of the algorithm. This trial arrangement produced 84 sets of design variable values. The values of five key design variables, namely C , H , T , P and R , are depicted in Fig. 12(a)–(e). Fig. 12(f) shows the actual stroke corresponding to each parameter combination. Fig. 13(a) illustrates the variation of stroke, polished rod acceleration, and displacement in a single stroke. The specific values of the maximum and minimum polished rod accelerations are detailed in Fig. 13(b) and (c).

The accelerations obtained by the three algorithms, as depicted in Fig. 13(b) and (c), demonstrate good consistency. Notably, the results obtained via the GWO algorithm appear more robust. Particularly after exceeding 2000 iterations, the maximum and minimum acceleration values stabilize at approximately 0.675 and -0.675 respectively. Despite this consistency, Fig. 13 reveals that the optimization results for the design variables exhibit significant variability, suggesting that optimizing the key structural dimensions of the BPU represents a non-convex problem. It is also evident that none of the three optimization methods achieved convergence for variables H and I . To address this issue and potentially simplify the problem to a convex one, the number of indeterminate variables was reduced. It was assumed that variable H is constant, with values set at 3482 mm, 3567 mm, and 3663 mm respectively. These values represent the mean values obtained from the three optimization algorithms, as illustrated in Fig. 12(b). The range of values for the remaining variables, as listed in Table 1, was kept unchanged, with H fixed. Similarly, the number of iterations was set to 500, 1000, 2000, and 3000, respectively. The optimization outcomes for each variable are presented in Fig. 14.

Upon a thorough examination of the data distribution characteristics, it is clear that the GWO-based method consistently delivers stable results for variable values, as particularly evident in Fig. 14. When juxtaposed with the outcomes depicted in Fig. 12, it becomes apparent that the results derived from the GA-based method exhibit the greatest variability, with those from the PSO-based method following closely behind. This observation necessitates a thorough analysis of the intrinsic characteristics of each algorithm and the methods of their refinements. In the case of the PSO-based method, the implementation of mutation to update particle positions is designed to bolster the algorithm's capacity for

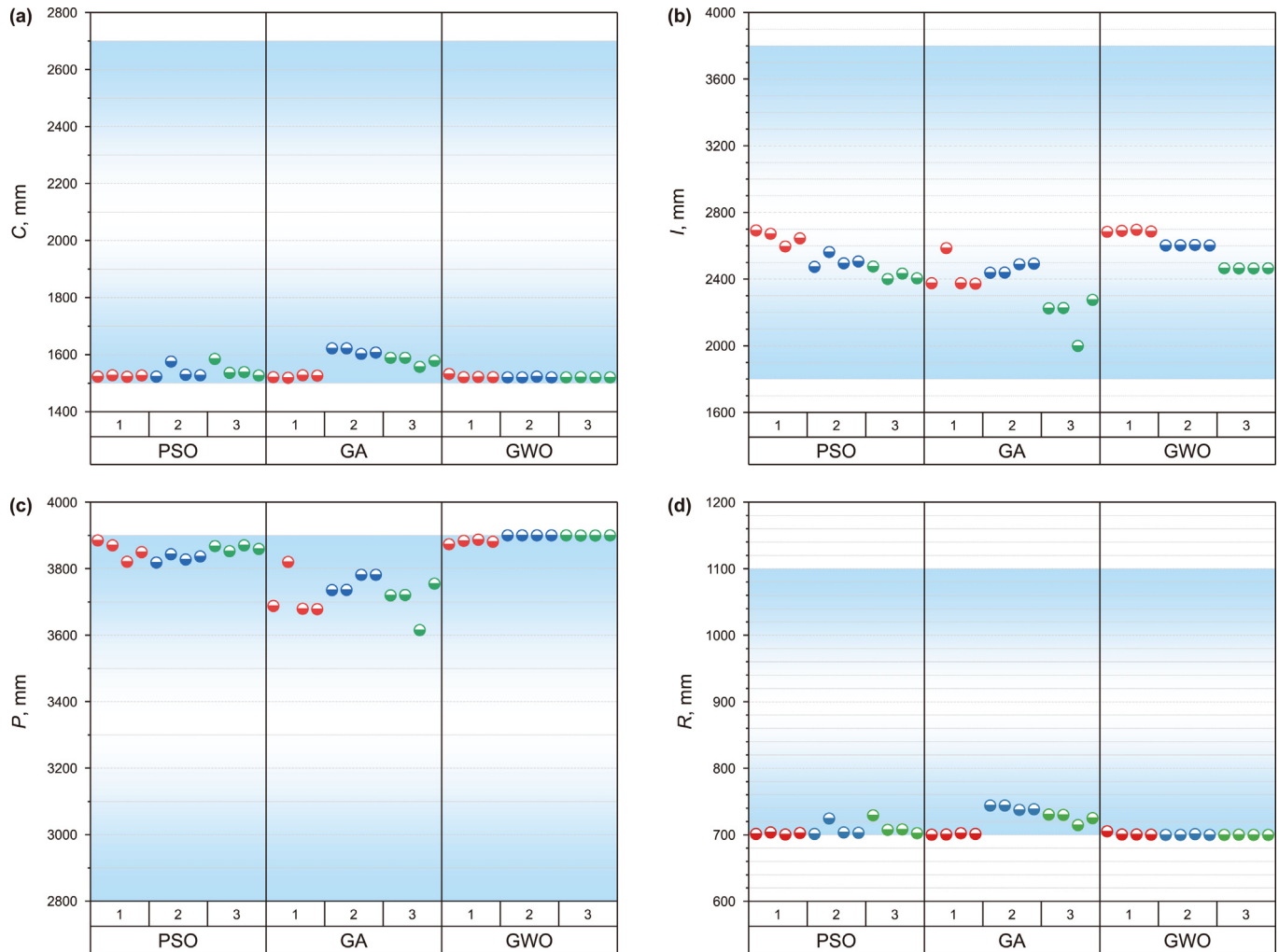


Fig. 14. (a)–(d) Optimization results for variables C , I , P , R .

global optimization. However, achieving consistent outcomes is challenging due to the close proximity of objective function values within the search domain. With respect to the GA-based method, the direct cancellation of mutation and crossover operations for individuals that do not satisfy the constraint conditions somewhat diminishes the algorithm's global and local optimization capabilities. Consequently, the modified GA algorithm performs the least effectively among the three in terms of computational performance. The GWO algorithm's primary advantage stems from the pivotal role of the convergence factor " a ", which effectively contracts the search domain, thereby ensuring that the solutions obtained exhibit a high degree of consistency. This feature is instrumental in the GWO's superior performance, particularly in the context of the optimization problems addressed in this study.

Fig. 14(b) displays the three values of variable I , which are 2685, 2602, and 2464 mm, respectively. This consistency shows that the current optimization problem is convex, implying that there exists an optimal solution. Additionally, Fig. 16(a) and (b) also show that the maximum and minimum acceleration obtained by the GWO algorithm stabilize at constant values of 0.675 and -0.675 m/s^2 , respectively. In contrast, there is still a slight fluctuation in the results obtained through PSO and GA algorithms. These findings further underscore the robustness of the GWO algorithm. Interestingly, even though the variable values

obtained by the PSO and GA algorithms differ, the error in acceleration does not surpass 1.5%. Fig. 15 presents the convergence curves of the three algorithms, with Fig. 15(a)–(c) corresponding to different H values. Each algorithm provides two sets of comparison data. The graphs clearly show that the GWO algorithm achieves consistency in the objective function region after 2000 iterations, with the PSO algorithm following in terms of convergence. While the objective function shows a tendency to converge with the PSO algorithm, it is not entirely consistent. The GA algorithm, which exhibits the worst convergence, is significantly influenced by the initial population, impacting the objective function. Despite the GWO algorithm demonstrating superior convergence, Fig. 15(d) indicates that it has the longest computation time, averaging 677 s for 2000 iterations. However, this computation time is still acceptable in practical applications. Based on this analysis, the structure parameter optimization scheme utilizing the GWO algorithm is more recommendable for its reliability and consistency. To reduce the production costs of the experimental prototype, the structure of the CYJW7-3-23HF BPU was adjusted align with the parameters of sample 3, as these models have quite similar structural parameters, as listed in Table 2. The horsehead, main beam, and downward beam can still use the original components. The relevant parameters are listed in Table 3.

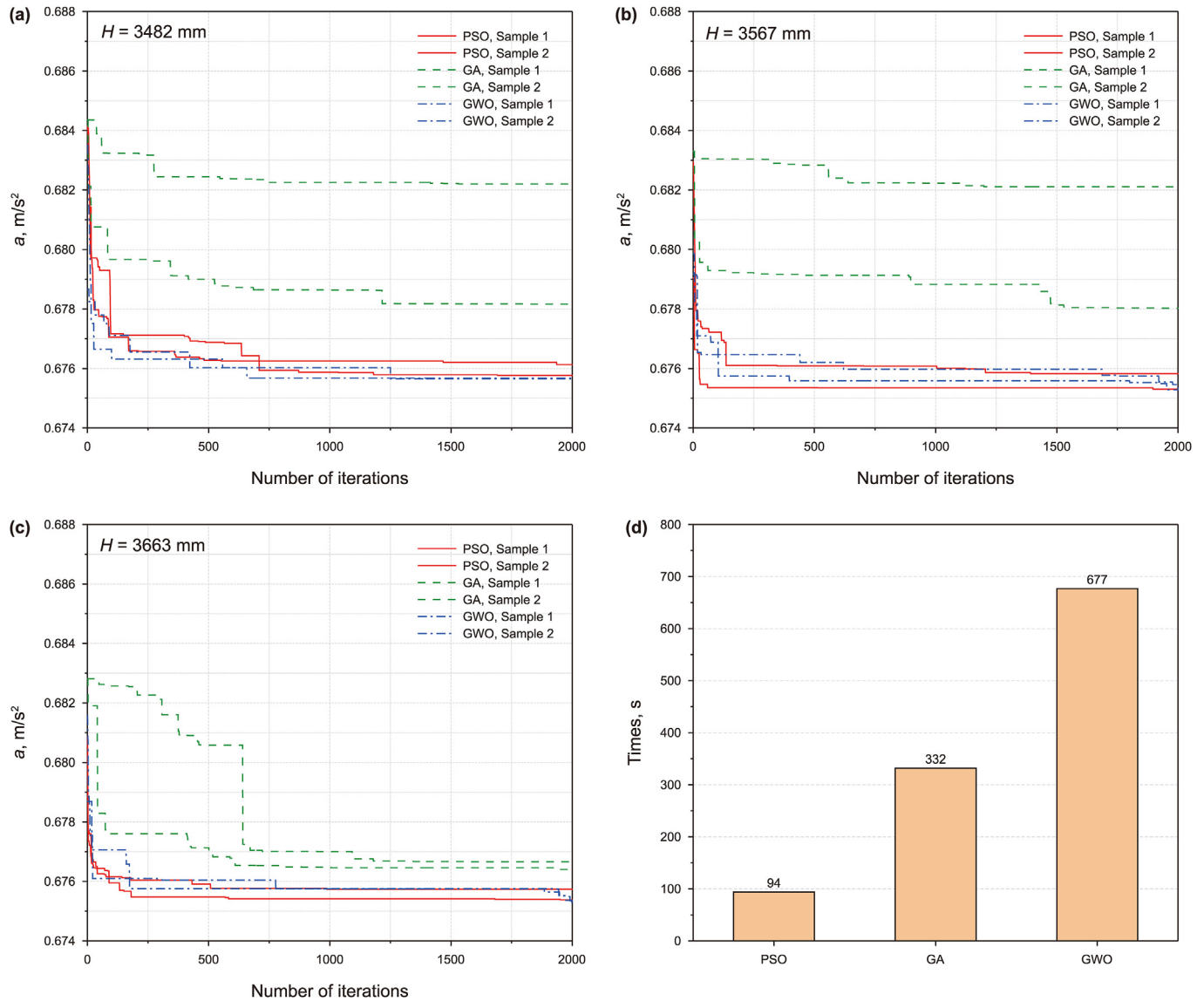


Fig. 15. (a)–(c) Optimization convergence curves of the PSO, the GA, and the GWO. (d) Comparison of time for three optimization algorithms.

5.2. Step 2: Balance optimization

The balance rate of a BPU is a crucial indicator in oilfield production management. During operation, the pumping unit's hanging point is subjected to asymmetric pulsating loads—large during the upstroke and smaller during the downstroke. The imbalance can cause several issues:

- Motor Load and Efficiency:** During the upward stroke, the motor bears an extremely large load, leading to significant electrical energy wastage as the unit drives the electric motor to work during the downstroke. This not only reduces the efficiency but also the service life of the motor.
- Equipment Durability:** The highly uneven load causes the BPU to vibrate violently, which in turn reduces the service life of the equipment.
- Operational Uniformity:** The imbalance disrupts the uniformity of the crank's rotational speed, causing the horsehead to swing unevenly. This affects the normal operation of the sucker rod and pump, thereby impacting the oil well's production and

increasing the pump repair rate. To mitigate these issues, the methods of adjusting the balance must be used to ensure that the single well balance rate is above 85% during normal operation.

In summary, there are two primary goals for adjusting the balance of the BPU: ensuring the safe operation of the unit and saving energy.

From a safety perspective, achieving balance aims to minimize the output torque of the reducer. The CLF defined as the ratio of the RMS torque to the mean torque, reflects the actual load changes. According to Eq. (43), the closer this value is to 1, the smaller the fluctuation in the BPU's operation, resulting in a more stable performance.

$$\text{CLF} = \frac{M_e}{\bar{M}} \quad (43)$$

in which,

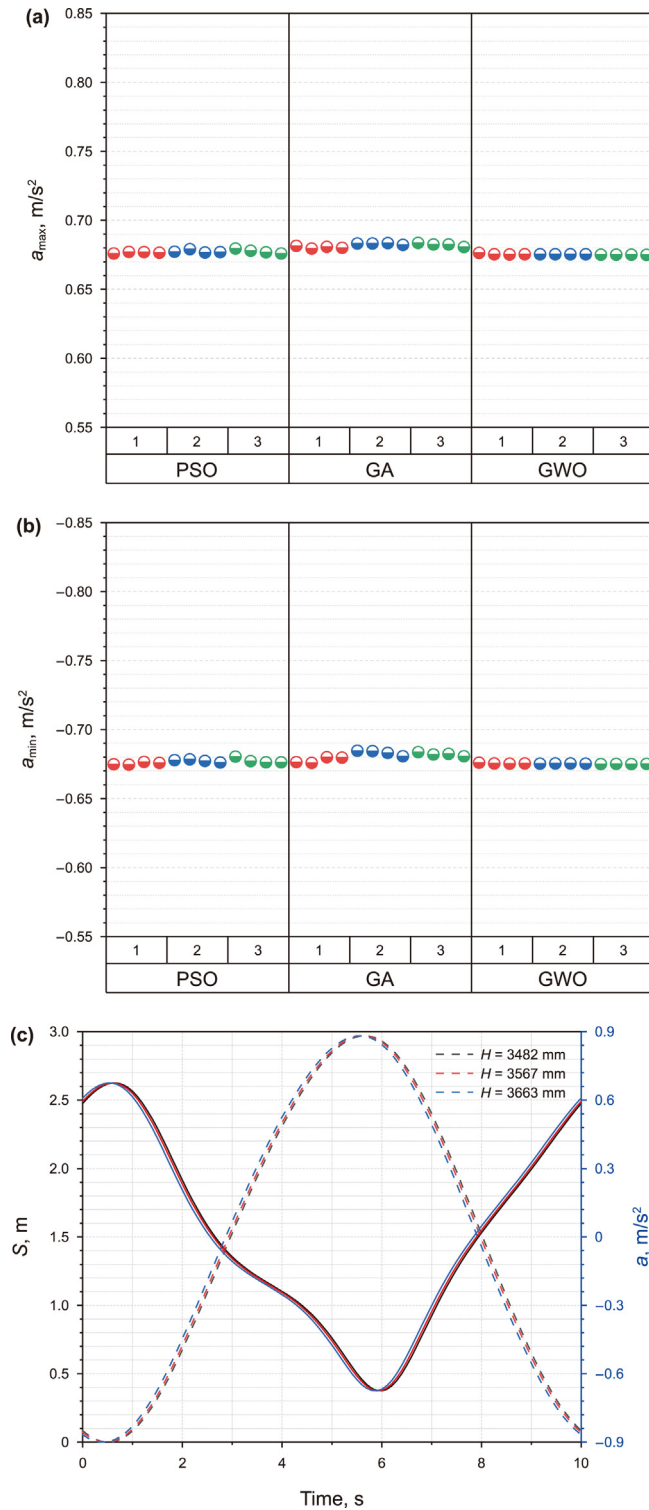


Fig. 16. (a)–(b) The acceleration of suspension points. (c) The displacement of suspension points.

$$\bar{M} = \frac{1}{2\pi} \int_0^{2\pi} M d\theta = \frac{\sum (M_i \Delta\theta_i)}{\sum \Delta\theta_i} \quad (44)$$

From an energy-saving perspective, balance should target the minimization of the RMS value of the electric current. Given that fixed losses during the motor's operation and the mechanical

transmission losses during the BPU's operation are relatively constant, reducing the energy consumption of the BPU can only be accomplished by minimizing the motor's variable losses. This involves reducing the RMS value of the motor's current. Since the torque of the motor's load determines the current's magnitude, minimizing the RMS value of the motor's load torque will in turn minimize the RMS value of the motor's current. The RMS torque and CLF are utilized as indices to evaluate the balance of BPUs.

Next, an optimization of the balance of the BPU for the two production wells will be detailed. These wells are located in the Laohekou block of the Shengli Oilfield. The detailed parameters and undetermined variables of the BPU are listed in Table 3. The range of values for the mass of the beam and crank counterweight is primarily based on the CYJW7-3-23HF, PCYJ8-3-26HF and several in-service BPUs. The allowable bearing capacity of the component must be considered. The lag angle τ , which is related to the installation status of the crankshaft counterweight block, does not have a clear reference value, thus a wider range of values is provided.

Three established optimization methods based on PSO, GA and GWO are utilized to verify their effectiveness in balance optimization for BPU. Three sets of initial samples are generated using the LHS method, each set containing 500 samples. Since there are no constraints applied in this scenario, all initial samples are considered qualified without the need for screening. The number of iterations for each method is uniformly set at 2000.

5.2.1. Well number: YUGDD7-20

The maximum strokes per minute for the BPU used in this production well is 2.5 min^{-1} . According to the polished rod evaluation method described in Section 3, the dynamometer card of the BPU is depicted in Fig. 17. The shape characteristics of the measured dynamometer card seem to indicate high production performance as well as vibration of the polished rod (Fakher et al., 2021). Fig. 18(a), (b) and (c) display the optimization results of variables τ , m_T and m_R , respectively. In these figures, "Load 1" and "Load 2" correspond to two types of polished rod loads estimated by theoretical models, while "Load 3" corresponds to the measured polished rod load. From the distribution of these results, the outcomes obtained by the PSO based and GWO based methods are completely consistent. However, the results obtained by the GA based method are highly dispersed. Another interesting observation is that the optimization results for variables under "Load 1" and "Load 3" are almost identical, suggesting a strong correlation between theoretical models and actual measurements in these scenarios.

The variable results obtained from three types of polished rod loads are listed in Table 4. The serial numbers of the parameter groups correspond to the three types of loads depicted in Fig. 18. The primary purpose of the theoretical estimation of the polished rod load is to provide convenient guidance for the balance optimization of BPUs in actual production settings. To compare the balance performance of BPUs under these three parameter groups, the torque of the BPU under the measured load is evaluated, as shown in Fig. 19. Additionally, the RMS torque and CLF are calculated, as detailed in Table 4. Parameter group 3, which was derived based on the measured load, exhibits relatively smaller RMS torque and CLF values. However, overall, the differences among the groups are not substantial. Therefore, for the current measured well condition, the balance performance optimization based on the two theoretical polished rods described in this paper is also deemed acceptable.

5.2.2. Well number: YUGDD10P428

The maximum strokes per minute for the BPU used in this production well is 1.48 min^{-1} . The theoretically evaluated and

Table 2
Basic structure parameters of three BPUs.

Structure parameter	A^* , mm	C , mm	H^* , mm	l , mm	P , mm	R , mm	S_{\max} , m
Sample 1	3000	1520	3482	2685	3880	700	2.970
Sample 2	3000	1520	3567	2602	3900	700	2.970
Sample 3	3000	1520	3663	2464	3900	700	2.970
CYJW7-3-23HF	3000	1500	3700	2000	3820	708	2.980

Table 3
Parameters of improved BPU.

Structure parameter				Mass parameter	
Nomenclature	Value	Nomenclature	Value	Nomenclature	Value
γ_1	40, °	A	3000, mm	m_A	625.6, kg
γ_4	5.2, °	C	1520, mm	m_C	521.6, kg
l_1	2677, mm	P	3900, mm	m_E	450.2, kg
l_2	2238, mm	R	700, mm	m_H	445.5, kg
l_3	3825, mm	H	3567, mm	m_N	1134.7, kg
l_4	450, mm	l	2602, mm	m_T	[0, 3000], kg
l_5	0, mm	τ_c	8, °	m_R	[0, 4500], kg
l_6	700, mm	τ	[0, 40], °		
l_7	1030, mm				

measured dynamometer card for this BPU is displayed in Fig. 20. The measured dynamometer card suggests potential issues with gas interference or insufficient liquid supply, indicated by the delayed closing of the travelling valve. Fig. 21(a)–(c) show the optimization results of variables τ , m_T and m_R , respectively. The distribution characteristics of these results reveal that the outcomes obtained via PSO based algorithm and GWO based algorithm are consistent, further demonstrating the robustness of these two methods. It is noted that the values of m_R corresponding to Load 2 and Load 3 have reached the upper limit, while the optimization results for the other two variables show significant differences.

The variable results obtained from three types of polished rod loads are listed in Table 5. The impact of these three parameter groups on the balance of BPUs is still evaluated based on the net torque of gearbox and the CLF, with their values displayed in Fig. 22 and Table 5. The RMS torque and CLF corresponding to parameter group 3 are the smallest, with a minimum CLF of 1.749, indicating that the torque fluctuation is still significant. The CLF values for the other two parameter groups have increased by 12% and 17%,

respectively. This result indicates that the two theoretical models of polished rod evaluation previously mentioned do not suit the current well conditions.

Based on the analysis results from the two wells discussed, it is clear that balance optimization of the BPU, particularly when based on the measured polished rod load, significantly reduces power consumption and enhances operational stability. Furthermore, the balance optimization methods utilizing the PSO algorithm and GWO algorithm have demonstrated their ability to achieve global optimal solutions. Figs. 18(d) and 21(d) both demonstrate good convergence of these two algorithms. The calculations further reveal that after 1000 iterations, the PSO and GWO algorithms take relatively less time, 197 s and 182 s respectively, while the GA algorithm takes considerably longer than them, possibly due to its sensitivity to population size in terms of computation time. Among the three optimization algorithms examined in this study, the GWO-based algorithm exhibits superior adaptability for both structural and balance optimization of BPUs.

6. Conclusion

This paper addresses the optimization design of compound balanced BPUs, focusing on structure and balance optimizations. A detailed optimization scheme leveraging three algorithms, such as PSO, GA and GWO, was established. This scheme offers effective methods and technical support for the design and operation of BPUs. The main findings attained, and the recommendations given can be summed up as follows.

- 1) A dynamic model of compound balanced BPU was developed based on the D'Alembert's principle. This dynamic model vividly illustrates the influence of the polished rod and the dynamic characteristics of the moving components of the BPU on the gearbox torque. It serves as a theoretical foundation for both structural and balance optimizations.
- 2) Constraint conditions derived from the fundamental principles of the 4-bar linkage and practical design experience, ensure the rationality of the structural dimensions obtained through the proposed optimization method. An optimal solution exists for variables C , l , P , and R , assuming a fixed height H . Among the

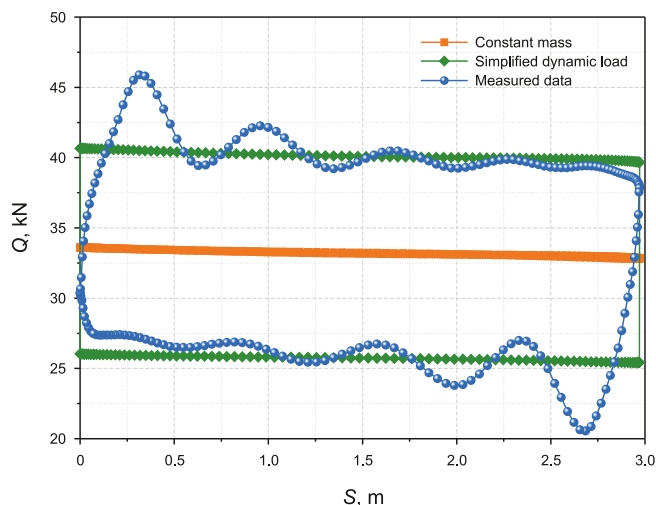


Fig. 17. The theoretical evaluation and measured dynamometer card.

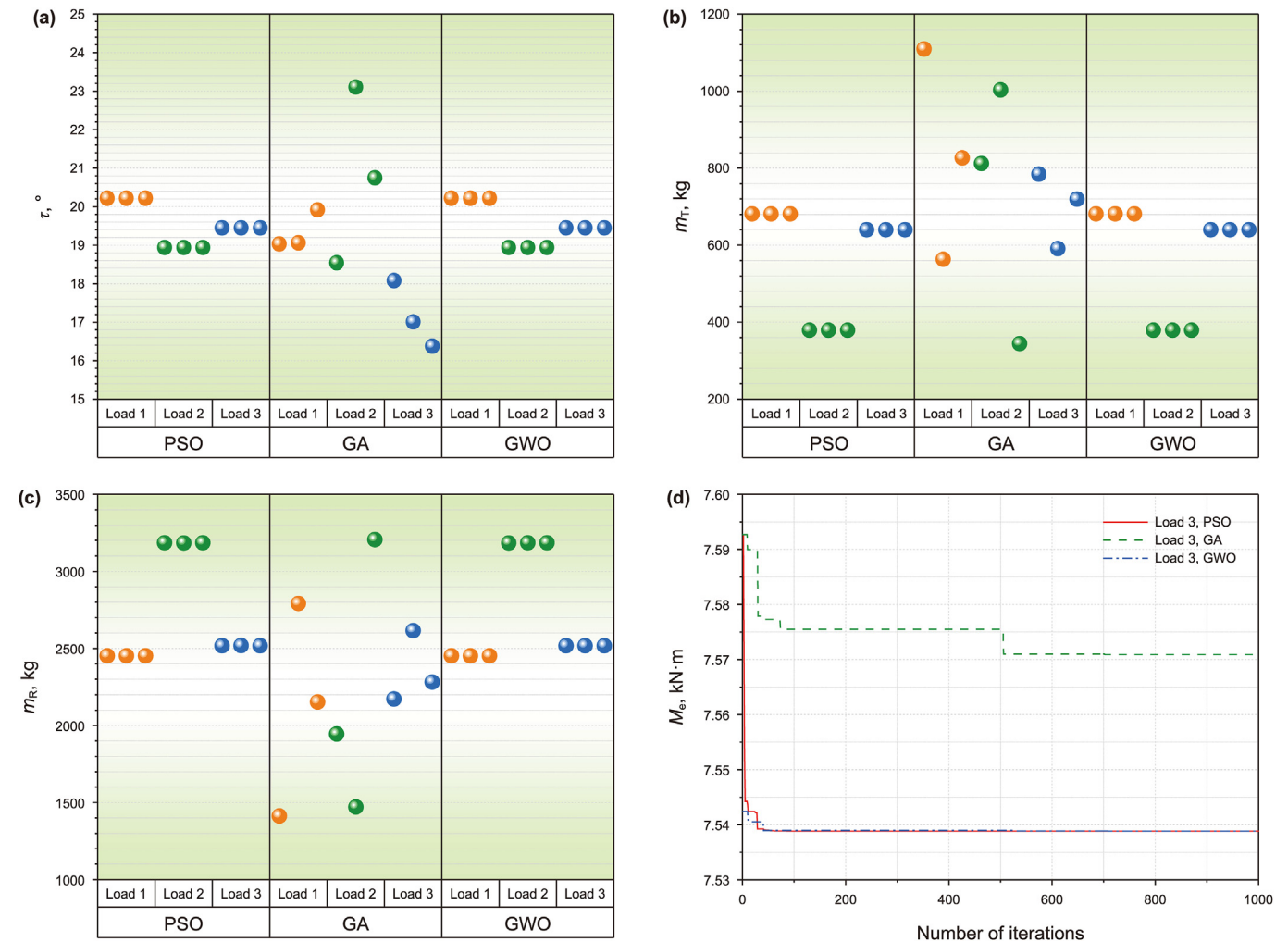


Fig. 18. Optimization results for undetermined variables.

Table 4
Optimization results and balance performance indicators.

Parameter group	τ_r , °	m_T , kg	m_R , kg	RMS torque M_e , kN·m	CLF
1	20.22	681.36	2452.57	7.548	1.113
2	18.94	379.34	3185.15	7.566	1.117
3	19.45	640.15	2518.19	7.539	1.112

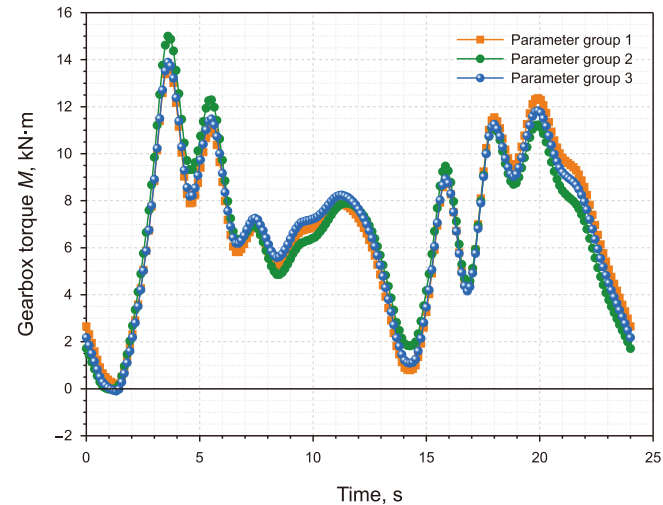


Fig. 19. Net torque curve of gearbox.

algorithms tested, the GWO-based method demonstrated superior robustness in structural optimization.

3) A strategy for optimizing the balance of BPUs based on the measured dynamometer card was proposed, using the PSO, GA and GWO based optimization algorithms. Example calculations

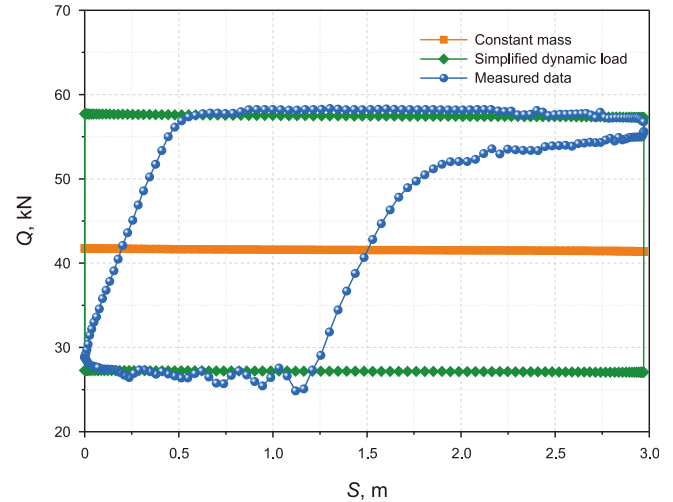


Fig. 20. Theoretical evaluation and measured dynamometer card.

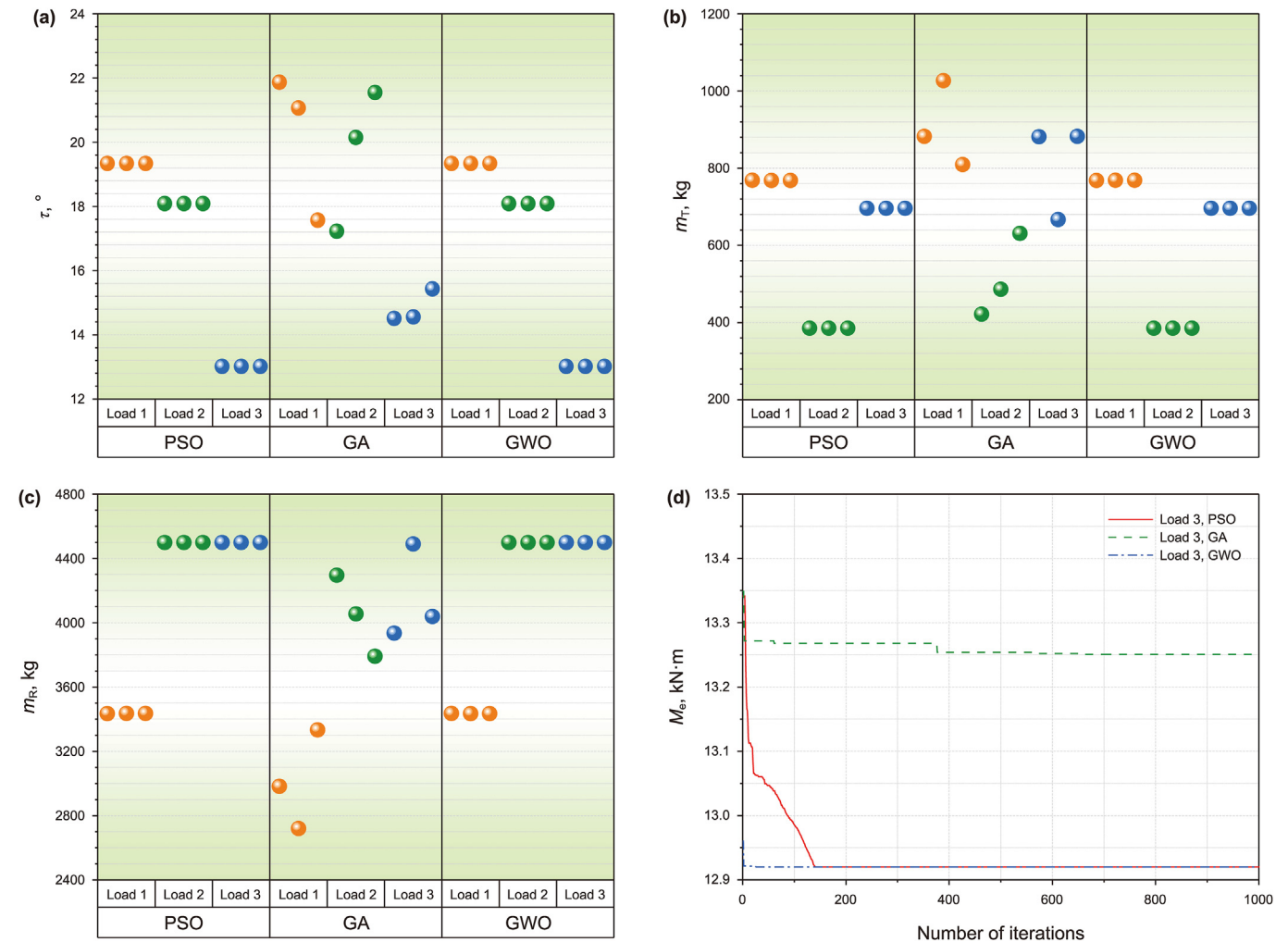


Fig. 21. Optimization results of undetermined variables.

Table 5
Optimization results and balance performance indicators.

Parameter group	$\tau_1, ^\circ$	m_T, kg	m_R, kg	RMS torque $M_e, \text{kN}\cdot\text{m}$	CLF
1	19.34	768.58	3435.48	15.095	2.041
2	18.09	385.46	4500.00	14.455	1.956
3	13.02	696.22	4500.00	12.921	1.749

reveal that both PSO based and GWO based optimization methods exhibit robust performance in balance optimization. Adjusting the counterweight according to these optimization results can significantly reduce the gearbox torque. This strategy is critically important for enhancing energy efficiency and operational stability of BPUs.

CRedit authorship contribution statement

Jie Wang: Writing – original draft, Methodology. **Quan-Ying Guo:** Software. **Cheng-Long Fu:** Visualization. **Gang Dai:** Validation, Data curation. **Cheng-Yu Xia:** Writing – review & editing. **Li-Qin Qian:** Supervision.

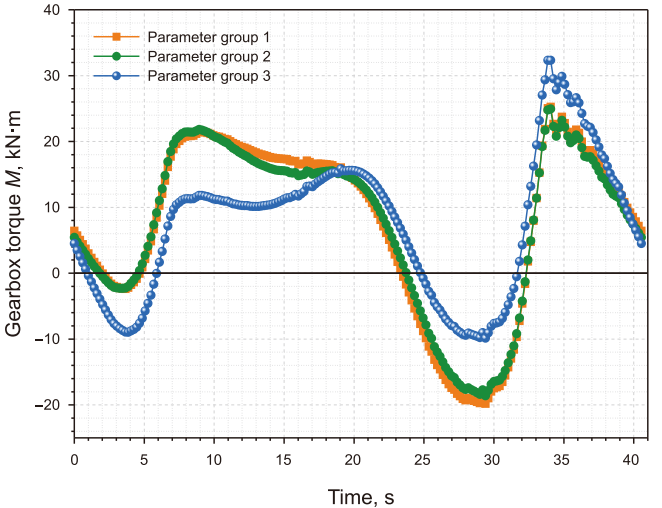


Fig. 22. Net torque curve of gearbox.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by the Key Laboratory of Petroleum and Natural Gas Equipment, Ministry of Education (No. OGE202303-08) and Engineering Technology Research Center for Industrial Internet of Things and Intelligent Sensing, Hubei Province (No. KXZ 202203).

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.petsci.2025.01.007>.

References

- An, Z.B., Zhou, K., Hou, J., et al., 2022. Accelerating reservoir production optimization by combining reservoir engineering method with particle swarm optimization algorithm. *J. Petrol. Sci. Eng.* 208, 109692. <https://doi.org/10.1016/j.petrol.2021.109692>.
- API Spec 11E, 2013. Specification for Pumping Units, nineteenth ed. American Petroleum Institute, Washington, D.C.
- Catalina, U.N., Erik, G.M., Camilo, A.G.M., 2022. Evaluation of the carbon footprint produced by conventional artificial lift systems in a Colombian field. *J. Petrol. Sci. Eng.* 208, 108865. <https://doi.org/10.1016/j.petrol.2021.108865>.
- Chen, S.Y., Shui, X.F., Huang, H., 2017. Improved genetic algorithm with two-level approximation using shape sensitivities for truss layout optimization. *Struct. Multidiscip. Optim.* 55, 1365–1382. <https://doi.org/10.1007/s00158-016-1583-2>.
- Fakher, S., Khlaifat, A., Hossain, M.E., et al., 2021. A comprehensive review of sucker rod pumps' components, diagnostics, mathematical models, and common failures and mitigations. *J. Pet. Explor. Prod. Technol.* 11 (10), 3815–3839. <https://doi.org/10.1007/s13202-021-01270-7>.
- Feng, Z.M., Guo, C.H., Zhang, D.S., et al., 2020. Variable speed drive optimization method and analysis of comprehensive performance of beam pumping unit. *J. Petrol. Sci. Eng.* 191, 107155. <https://doi.org/10.1016/j.petrol.2020.107155>.
- Ghosal, A., 2010. The freudenstein equation: design of four-link mechanisms. *Reson* 15 (8), 699–710. <https://doi.org/10.1007/s12045-010-0079-4>.
- He, Y.P., Zang, C.Z., Zeng, P., et al., 2023. Few-shot working condition recognition of a sucker-rod pumping system based on a 4-dimensional time-frequency signature and metalearning convolutional shrinkage neural network. *Petrol. Sci.* 20, 1142–1154. <https://doi.org/10.1016/j.petsci.2023.02.017>.
- Hu, P., He, Z.G., Zhang, X.J., et al., 2006. Optimum design of an energy-saving pumping unit for combination beam and crank offset balance. *Oil Field Equipment* 6, 39–41. <https://doi.org/10.3969/j.issn.1001-3482.2006.06.012> (in Chinese).
- Kaveh, A., 2017. Applications of Metaheuristic Optimization Algorithms in Civil Engineering. Springer, Switzerland. <https://doi.org/10.1007/978-3-319-48012-1>.
- Kaveh, A., 2021. Advances in Metaheuristic Algorithms for Optimal Design of Structures, third ed. Springer, Switzerland. <https://doi.org/10.1007/978-3-319-46173-1>.
- Kaveh, A., Eslamlou, A.D., 2020. Metaheuristic Optimization Algorithms in Civil Engineering: New Applications. Springer, Switzerland. <https://doi.org/10.1007/978-3-319-48012-1>.
- Kaveh, A., Zaerreza, A., 2023. Structural Optimization Using Shuffled Shepherd Meta-Heuristic Algorithm: Extensions and Applications. Springer, Switzerland. <https://doi.org/10.1007/978-3-031-25573-1>.
- Kaveh, A., Hamedani, K.B., 2022. Advanced Metaheuristic Algorithms and Their Applications in Structural Optimization. Springer, Switzerland. <https://doi.org/10.1007/978-3-031-13429-6>.
- Kaveh, A., Ghazaan, M.I., 2018. Meta-Heuristic Algorithms for Optimal Design of Real-Size Structures. Springer, Switzerland. <https://doi.org/10.1007/978-3-319-78780-0>.
- Kaveh, A., Bakhshpoori, T., 2019. Metaheuristics: Outlines, MATLAB Codes and Examples. Springer, Switzerland. <https://doi.org/10.1007/978-3-030-04067-3>.
- Kaveh, A., Mahdavi, V.R., 2015. Colliding Bodies Optimization: Extensions and Applications. Springer, Switzerland. <https://doi.org/10.1007/978-3-319-19659-6>.
- Kumar, S., Natee, P., Ghanshyam, G.T., et al., 2023. A two-archive multi-objective multi-verse optimizer for truss design. *Knowl. Base Syst.* 270, 110529. <https://doi.org/10.1016/j.knsys.2023.110529>.
- Li, K., Gao, X.W., Zhou, H.B., et al., 2015. Fault diagnosis for down-hole conditions of sucker rod pumping systems based on the FBH-SC method. *Petrol. Sci.* 12, 135–147. <https://doi.org/10.6052/10.1007/s12182-014-0006-5>.
- Li, K., Han, Y., 2016. Modelling for motor load torque with dynamic load changes of beam pumping units based on a serial hybrid model. *Trans. Inst. Meas. Control* 40 (3), 1–15. <https://doi.org/10.1177/0142331216670454>.
- Li, W.C., Dong, S.M., Sun, X.R., 2018. An improved sucker rod pumping system model and swabbing parameters optimized design. *Math. Probl. Eng.*, 4746210. <https://doi.org/10.1155/2018/4746210>.
- Liu, Z., Zhu, P., Chen, W., et al., 2015. Improved particle swarm optimization algorithm using design of experiment and data mining techniques. *Struct. Multidiscip. Optim.* 52, 813–826. <https://doi.org/10.1007/s00158-015-1271-7>.
- Liu, Z.J., 1985. A new method for geometric design of sucker rod pumping units. *China Petrol. Machines.* 13 (5), 39–46 (in Chinese).
- Lv, H.Q., Liu, J., Han, J.Q., et al., 2016. An energy saving system for a beam pumping unit. *Sensors* 16 (5), 685. <https://doi.org/10.3390/s16050685>.
- Ma, L.B., Cheng, S., Shi, Y.H., 2020. Enhancing learning efficiency of brain storm optimization via orthogonal learning design. *IEEE T. Syst. Man Cy-S.* 99, 1–20. <https://doi.org/10.1109/TSMC.2020.2963943>.
- Meng, Z., Yildiz, B.S., Li, G., et al., 2023. Application of state-of-the-art multiobjective metaheuristic algorithms in reliability-based design optimization: a comparative study. *Struct. Multidiscip. Optim.* 66 (191), 1–27. <https://doi.org/10.1007/s00158-023-03639-0>.
- Mirjalili, S., 2015. How effective is the Grey Wolf optimizer in training multi-layer perceptrons. *Appl. Intell.* 43, 150–161. <https://doi.org/10.1007/s10489-014-0645-7>.
- Mirjalili, S., Jangir, P., Seyedeh, Z.M., et al., 2017. Optimization of problems with multiple objectives using the multi-verse optimization algorithm. *Knowl.-Based Syst.* 134, 50–71. <https://doi.org/10.1016/j.knsys.2017.07.018>.
- Ongkunaruk, P., Wahab, M.I.M., Chen, Y., 2016. A genetical algorithm for a joint replenishment problem with resource and shipment constraints and defective items. *Int. J. Prod. Econ.* 175, 142–152. <https://doi.org/10.1016/j.ijpe.2016.02.012>.
- Pineda, L.R., Serpa, A.L., Biazussi, J.L., 2023. Online trained controller for electrical submersible pumps in liquid-gas flow. *Geoen. Sci. Eng.* 225, 211713. <https://doi.org/10.1016/j.geoen.2023.211713>.
- Rao, R.V., Savsani, V.J., Vakharia, D.P., 2011. Teaching-learning-based optimization: a novel method for constrained mechanical design optimization problems. *Comput. Aided Des.* 43, 303–315. <https://doi.org/10.1016/j.cad.2010.12.015>.
- Sun, J.J., Huang, Z.Q., Zhu, Y., et al., 2022. Real-time kinematic analysis of beam pumping unit: a deep learning approach. *Neural Comput. Appl.* 34, 7157–7171. <https://doi.org/10.1007/s00521-021-06783-0>.
- Sun, X.R., Dong, S.M., Wang, H.B., et al., 2018. Comparison of multistage simulation models of entire sucker rod with spatial buckling in tubing. *J. Jilin Univ. Eng. Technol. Ed.* 48 (4), 1124–1132. <https://doi.org/10.13229/j.cnki.jdxbgxb20170512> (in Chinese).
- Takacs, G., Kis, L., 2021. A new model to find optimum counterbalancing of sucker-rod pumping units including a rigorous procedure for gearbox torque calculations. *J. Petrol. Sci. Eng.* 205, 108792. <https://doi.org/10.1016/j.petrol.2021.108792>.
- Tejani, G.G., Savsani, V.J., Patel, V.K., et al., 2017a. Topology, shape, and size optimization of truss structures using modified teaching-learning based optimization. *Adv. Comput. Des.* 2 (4), 313. <https://doi.org/10.12989/acd.2017.2.4.313>.
- Tejani, G.G., Savsani, V.J., Patel, V.K., 2017b. Modified sub-population based heat transfer search algorithm for structural optimization. *Int. J. Appl. Metaheur.* 8 (3), 1–22. <https://doi.org/10.4018/IJAMC.2017070101>.
- Tian, H., Deng, S., Wang, C.B., 2021. A novel method for prediction of paraffin deposit in sucker rod pumping system based on CNN indicator diagram feature deep learning. *J. Petrol. Sci. Eng.* 206, 108986. <https://doi.org/10.1016/j.petrol.2021.108986>.
- Wang, F., Zhang, H., Zhou, A.M., 2021. A particle swarm optimization algorithm for mixed-variable optimization problems. *Swarm Evol. Comput.* 60 (2), 100808. <https://doi.org/10.1016/j.swevo.2020.100808>.
- Wang, H.B., Dong, S.M., 2020. A Model for the transverse vibration simulation of sucker rod strings with axial reciprocating motion in curved wellbores. *Eng. Mech.* 37 (10), 228–237. <https://doi.org/10.6052/j.issn.1000-4750.2019.12.0731> (in Chinese).
- Xing, M.M., Dong, S.M., 2015. A new simulation model for a beam-pumping system applied in energy saving and resource-consumption reduction. *SPE Prod. Oper.* 30 (2), 130–140. <https://doi.org/10.2118/173190-PA>.
- Xu, J.C., Li, W., Meng, S.Y., 2022. Kinematic and dynamic simulation analysis of modified conventional beam pumping unit. *Energies* 15, 5496. <https://doi.org/10.3390/en15155496>.
- Yang, M.Z., Wang, X.W., Huang, M., et al., 2022. A novel evolutionary algorithm based on Judgment-Rule evolution strategy for structural balance in signed social networks. *IEEE Trans. Emerging Top. Comput. Intell.* 6 (4), 805–817. <https://doi.org/10.1109/TETCI.2021.3089714>.
- Zakian, P., Kaveh, A., 2024. Multi-objective seismic design optimization of structures: a review. *Arch. Comput. Methods Eng.* 31, 579–594. <https://doi.org/10.1007/s11831-023-09992-z>.
- Zakian, P., Ordoubadi, B., Alavi, E., 2021a. Optimal design of steel pipe rack structures using PSO, GWO, and IGWO algorithms. *Adv. Struct. Eng.* 24 (11), 2529–2541. <https://doi.org/10.1177/13694332211004116>.
- Zakian, P., Nadi, M., Tohidi, M., 2021b. Finite cell method for detection of flaws in plate structures using dynamic responses. *Structures* 34, 327–338. <https://doi.org/10.1016/j.istruc.2021.07.070>.
- Zamfirache, I.A., Precup, R.E., Roman, R.C., et al., 2022. Policy iteration reinforcement learning-based control using a grey wolf optimizer algorithm. *Inf. Sci.* 585, 162–175. <https://doi.org/10.1016/j.ins.2021.11.051>.
- Zhang, J.J., Li, X.Q., Shi, H.N., 2005. Design on Beam Pumping Unit. Petroleum Industry Press, Beijing (in Chinese).