Mathematical Biosciences and Engineering, 16(3): 1258–1279.
DOI: 10.3934/mbe.2019061
Received: 10 December 2018
Accepted: 16 January 2019
Published: 19 February 2019

http://www.aimspress.com/journal/MBE

Research article

Multi-objective cellular particle swarm optimization and RBF for drilling parameters optimization

Jun Zheng[1], Zilong li[2,]*, Bin Dou[1] and Chao Lu[3]

[1] Engineering Faculty, China University of Geosciences, Wuhan, Hubei, P.R. China
[2] Research Institute No. 717 of shipbuilding Industry Corporation-Wuhan National Laboratory for Optoelectronics, China
[3] Computer College, China University of Geosciences, Wuhan, Hubei, P.R. China

* **Correspondence:** Email: lizilong_st@126.com ; Tel: +8602767883507; Fax: +02767883507.

Abstract: Wellbore drilling parameters optimization is one of the most important issue in drilling engineering. Rate of penetration or mechanical specific energy was usually utilized as the optimization objective. The rate of penetration directly relates to the drilling cycle, while mechanical specific energy reflects the drilling efficiency. In this paper, except for rate of penetration and mechanical specific energy, the drilling life of bit is also summarized as a comprehensive assessment indicator in wellbore drilling parameters optimization problem. The drilling life of bit is taken into consideration for the design and manufacturing cost of bit compose a significant part of the drilling cost and the bit drilling life greatly influences the drilling efficiency. However, those objectives are usually related in a highly nonlinear relationship and in conflict with each other. Thus, a multi-objective cellular particle swarm optimization (MOCPSO) is developed to solve the three-objective drilling parameters optimization problem. Moreover, the radius basis function (RBF) method is employed into the formation parameters identification for rate of penetration model. Performance of MOCPSO is investigated by taken a comparison with multi-objective PSO and non-dominated sorting genetic algorithm-II (NSGA-II). Effect of the four commonly used neighborhood function is also investigated by making contrasts with each other. It can be inferred that MOCPSO is statistically superior to both multi-objective PSO, NSGA-II at the 0.05 level of significance on the wellbore drilling parameters optimization problem. And the four commonly used neighborhood templates perform comparable with each other, and are not statistically different for the drilling parameters optimization problem.

Keywords: multi-objective optimization; wellbore drilling parameters optimization; parameter estimation; cellular particle swarm optimization

1. Introduction

Wellbore drilling is one of the most important and costly operations during the oil and gas reservoirs exploration and production. For the high exploration cost and relatively much lower oil and gas prices, most of the worlds' oil and gas companies are urgent to reduce the drilling costs as much as possible. The drilling cost is often denoted with the cost per meterage drilled. The practical drilling engineering is a complex, coupled and time-consuming system which involves the drilling equipment and tools, drilling fluid circulating system, and geological conditions of the construction area, and so on. Therefore, the drilling cost depends on various factors, such as types of the used drilling rigs, geologic feature of the drilling formation and the total drilling depth. In addition, the drilling cost is also closely related to rotating time of drilling bit, connection time, design and manufacturing cost of drilling bits and rigs, and so on. A significant part of the drilling expense is made up with the costs of rig operations, therefore, drilling parameters optimization is a central priority of all operators. One main task of drilling parameters optimization is to research the effects of the various drilling parameters (such as the rotations per minute, weight on bit etc.) on the drilling process and to select the most reasonable drilling parameters to obtain the optimal technical and economic indicators. Through the drilling parameters optimization, a shortened drilling period, a remarkably drilling cost reduction, an improved drilling quality and efficiency could be possibly achieved.

In order to select the best set of drilling operating parameters, various drilling models are developed according to drilling operating parameters and rock properties. Graham and Muench [1] optimized the rotary speed and weight on bit to get the minimum drilling costs with field data. Then, Bourgoyne and Young [2] proposed a comprehensive nonlinear drilling model to estimate the drilling performance, and employed a relatively simple analytical procedure to solve the model and obtained a set of best operating drilling variables with a minimum drilling cost. Reza and Alcocer [3] then proposed a three-equation drilling model with several drilling variables, which consist of rate of penetration, rate of bearing wear and rate of bit dulling. Iqbal [4] employed a stepwise drilling parameters optimization procedure for roller-cutter bit insertion. Amjad et al. [5] minimized drilling cost for the Iranian Khangiran gas field based on Bourgoyne and Young model. T. Eren, M. E. Ozbayoglu [6] implemented a real time optimization of drilling parameter. Zhang et al. [7] employed the biogeography-based optimization to deal with the drilling parameter optimization. Sun et al. [8] developed a real-time surveillance system of mechanical specific energy for drilling parameters optimization. Zhai et al. [9] studied deep water drilling parameters optimization. Cui et al. [10] optimized drilling parameters for a compound drilling with the mechanical specific energy theory.

The above researches utilized the rate of penetration or mechanical specific energy as the optimization objective, which were single objective optimization approaches. The rate of penetration relates to the drilling cycle, while mechanical specific energy reflects the drilling efficiency. Both rate of penetration and mechanical specific energy are important indicators for a practical drilling situation. In addition, the drilling life of bit should also be taken into consideration. For the design and manufacturing cost of bit compose a significant part of the drilling cost and the bit drilling life greatly influences the drilling efficiency. Therefore, a three-objective drilling parameters

optimization model, in which rate of penetration, mechanical specific energy and drilling life of bit are included, is proposed in this paper. However, rate of penetration in the three-objective drilling parameters optimization model depends on many parameters, which should be estimated with the data gained during the drilling. Thus, the radius basis function (RBF) method is employed into the formation parameters identification for rate of penetration model. Those three objectives are in conflict with each other and related in a highly non-linear manner. Take a comprehensive consideration of the above objectives, a multi-objective drilling parameters optimization model is established. Then, a novel multi-objective algorithm based on the cellular particle swarm optimization, MOCSPO, is presented to solve the established optimization model. The performance of MOCPSO is compared with multi-objective PSO [11] and NSGA-II [12] on the multi-objective drilling parameters optimization problem. Moreover, four commonly used neighborhood templates of CPSO are taken into comparison to investigate the most suitable neighborhood for this drilling parameters optimization problem.

The remainder of the paper is organized as follows. Section 2 establishes a multi-objective problem formulation for the drilling parameters optimization with several constraints. Section 3 developed a RBF based inverse parameter estimation method for rate of penetration model. Then, the multi-objective cellular particle swarm optimization is discussed in detail in Section 4. The bit factors and formation parameters identification for an oil well section is present in Section 5. Section 6 describes the optimization results and analysis, including the comparisons with multi-objective PSO and NSGA-II, investigations among the four neighborhood templates. Finally, the paper is closed by recapitulating salient points and concluding remarks in Section 7.

2. Problem formulation for drilling parameters optimization

In the practical drilling applications, biggest rate of penetration, longest drilling life of bit and smallest mechanical specific energy are respected to be achieved simultaneously. Thus, a multi-objective optimization model, including rate of penetration, mechanical specific energy and drilling life of bit, is established in this section.

2.1. Rate of penetration

Rate of penetration is an important indicator to estimate drilling efficiency and technological economy during the drilling engineering. Through investigating the impact of several variables (such as weight on bit, rotary speed, tooth wear and drilling fluid density etc.) on rate of penetration, Florence et al. proposed the following rate of penetration prediction model [13]:

$$v_{pc} = \frac{dH}{dt} = \frac{K C_p C_H (WOB - M) RPM^{\lambda}}{1 + C_2 h} \tag{1}$$

Where, K is the coefficient of formation drillability; C_p is the differential pressure drilling parameter; C_H is the hydraulic efficiency; WOB is the weight on bit (kN); M is the threshold bit weight (kN); RPM is the rotary speed (r/min); λ is the rotary speed exponent; C_1, C_2 are the

wear coefficients of bit teeth; h is the wearing capacity of bit teeth, the teeth is totally wear when $h=1$.

Variation of bit size can also impact the value of rate of penetration. As the diameter of the bit decreases, the pressure applied to the bit per unit area would increase. Therefore, with a certain drilling pressure, the bit with a smaller diameter has a higher rate of penetration, and the bit teeth will wear faster than the bit with a larger diameter.

Thus, the impact of bit diameter is taken into consideration when building the rate of penetration model in this paper. The new rate of penetration is express as following:

$$v_{pc}=\frac{K\,C_p C_H (WOB'-M')RPM^{\lambda}}{1+C_2 h} \tag{2}$$

Where, $WOB'=\frac{WOB}{D}$, $M'=M-\frac{C_e E_h}{D}$, E_h is bit specific hydraulic horsepower (kW/mm^2), C_e is coefficient of bit specific hydraulic horsepower, D is the bit diameter (mm).

2.2. Drill bit tooth wear

In the drilling process, extrusion and friction between the formation and the drill bit is the main factor that cause bit wear and affect the bit life. Here, bit wear consists of tooth wear and bearing wear. The tooth and bearing wear model is established as following:

$$\frac{dh}{dt}=\frac{A_f (S_1\cdot RPM+S_2\cdot RPM^3)}{(M_2-M_1 WOB)(1+C_1 h)} \tag{3}$$

$$\frac{dB}{dt}=\frac{1}{C}WOB^{1.5}\cdot RPM \tag{4}$$

Where, S_1, S_2 are rotary speed influence coefficients, M_1, M_2 are bit weight influence coefficients, A_f is formation abrasiveness factor.

The integral of the above equation can be written:

$$H_f=\frac{1}{C_1}\left(\sqrt{1+C_1\frac{t_f (S_1\cdot RPM+S_2\cdot RPM^3)}{(M_2-M_1\cdot WOB)}}-1\right) \tag{5}$$

$$B_f=\frac{1}{C}WOB^{1.5}\cdot RPM\cdot t_f \tag{6}$$

Eq 5 and Eq 6 represent wear extent of tooth and bearing wear H_f, B_f, after the rig works for t_f hours.

2.3. Mechanical specific energy

Mechanical specific energy refers to the amount of mechanical work required to break up a unit volume of rock. The lower the mechanical specific energy, the higher efficiency of the bit rock-breaking. The initial mechanical specific energy model is present as following [14]:

$$\mathrm{E}_{\mathrm{MSE}} = \frac{WOB}{A_b} + \frac{120\pi \cdot RPM \cdot T}{A_b \cdot v_{pc}} \tag{7}$$

In which, A_b is the cross section area of bit (mm^2), T is the torque applied at the bit.

The mechanical specific energy model expressed in Eq 7 did not take the hydraulic energy into consideration, for the hydraulic energy is hardly utilized in a conventional rotary-drilling. But for a certain formation, hydraulic energy is helpful even indispensable in the actual drilling [15]. In this case, the hydraulic term should be taken into consideration, and mechanical specific energy model turns into the following formula.

$$\mathrm{E}_{\mathrm{MSE}} = \frac{WOB}{A_b} + \frac{120\pi \cdot RPM \cdot T}{A_b \cdot v_{pc}} - \frac{2.8144604 \times 10^6 \times \pi \times E_h}{4D^{1.2} \cdot v_{pc}} \tag{8}$$

2.4. Modeling of drilling parameters optimization

In this paper, the rate of penetration, the drilling life of bit and mechanical specific energy are taken as the optimization objectives. The fastest rate of penetration, longest bit life and biggest mechanical specific energy are expected to be achieved simultaneously. However, these three objectives are often in conflict with each other. A preferable set of drilling parameters are expected to satisfy all these objects to one degree or another and offer a relatively fast rate of penetration, a long bit life and big mechanical specific energy. Thus, the wellbore drilling parameters optimization objective function can be given by:

$$\begin{aligned}
&\min \quad \vec{F}(\vec{X}) = (\vec{f_1}(\vec{X}), \vec{f_2}(\vec{X}), \vec{f_3}(\vec{X}))^T \\
&\vec{f_1}(\vec{X}) = -v_{pc} = -\frac{K\,\mathrm{C}_p \mathrm{C}_H (WOB - M) RPM^{\lambda}}{1 + C_2 h} \\
&\vec{f_2}(\vec{X}) = -t_f = -[(\frac{C_1}{2} + 1) - (\frac{C_1}{2} h^2 + h)] \frac{M_2 - M_1 \cdot WOB}{A_f (S_1 RPM + S_2 RPM^3)} \\
&\vec{f_3}(\vec{X}) = -\mathrm{E}_{\mathrm{MSE}} = -[\frac{WOB}{A_b} + \frac{120\pi \cdot RPM \cdot T}{A_b \cdot v_{pc}} - \frac{2.8144604 \times 10^6 \times \pi \times E_h}{4D^{1.2} \cdot v_{pc}}]
\end{aligned} \tag{9}$$

Several constraints are applied in the wellbore drilling parameters optimization. Weight on bit directly influences the rate of penetration and drilling efficiency. When the weight on bit is too small to break the rock, rate of penetration and drilling efficiency would be decreased sharply. If the weight

on bit is too high, it will cause excessive wear of the tooth bit, so the value range of weight on bit is as follows:

$$Max\{M,0\} < WOB < M_2 / M_1$$

The bit rotation speed cannot be too small in order to made the bit intrude into the formation to break rock. But the bit rotation speed cannot be too large at the same time for the restriction of bit material. Thus the value range of bit rotation speed is as follows:

$$n_{\min} \le n \le n_{\max}$$

As to the practical situation, value of tooth and bearing wear extent should be as following:

$$0 \le H_f \le 1$$
$$0 \le B_f \le 1$$

3. RBF based inverse parameter estimation

3.1. Parameters to estimation in drilling parameters optimization

The drilling parameters optimization problem formulated in Section 2 indicates that there are plenty of parameters needed to be estimated, which can be classified into two categories. One is related to the bit, such as tooth wear influence coefficient C_1, rotary speed influence coefficients S_1, S_2, and bit weight influence coefficients M_1, M_2, which can be determined as the bit is selected. While the other is related to formation, which should be estimated with the data gained during the drilling, such as formation drillability coefficient *K*, threshold bit weight *M*, rotary speed exponent λ, bit tooth wear coefficient C_2 and formation abrasiveness factor A_f.

The formation parameters estimation process is to gain as good agreement as possible between actual experimental responses and simulated responses by models in Section 2, which can be solved with the inverse method [16]. The inverse method treated the parameter estimation as an optimization procedure, aiming to obtain a set of parameters that make the discrepancy between experimental responses and simulated response minimized. Thus, the objective function used to identify those formation parameters is present as following:

$$\min \quad f(\mathbf{p}) = \sum_{i=1}^{N} \left(\frac{T_i - S_i(\mathbf{p})}{T_i}\right)^2 \tag{10}$$

Where $\boldsymbol{p}$ is the set of formation parameters to be identified, N is the number of experiment data, $S_i(\mathbf{p})$ is the simulation responses, T_i is the experimental responses.

3.2. RBF based inverse estimation method

In the above parameter estimation problem, the objective function is not known explicitly.

Therefore, a straightforward optimization on Eq 10 is obviously impractical. Thus, a RBF based inverse estimation method is proposed to identify those formation parameters, in which a RBF model is created to approximate the objective function. Plenty alternative surrogate models can be used as the approximation, such as the artificial neural network [17,18], the response surface method [19], the support vector regression [18,20], and so on. The RBF is utilized for its simple structure, easily implementation, and high accuracy and suitable for high dimensional problem. After the RBF approximation is created, then an optimization algorithm can be conducted on the RBF model.

Procedure flowchart of the RBF based inverse method for parameter estimation in drilling parameters optimization is present in Figure 1. Main process can be concluded as follows.

(1) Determine design space of formation parameters. In the drilling parameters optimization problem, there are five formation parameters to be estimated. Value range of these parameters should be determined first.

(2) Design of experiments. The drilling model in Section 2 is sampled with regular design of experiments (DOE) methods, such as orthogonal arrays [21,22], various Latin hypercube designs [23], D-Optimal Design [24], and uniform designs [25], etc. In this work, Latin hypercube design is employed to locate the sampling points, for its high efficiency, low computation cost for construction.

(3) RBF construction. RBF is employed to approximate the rate of penetration model, for its simple structure, easy implementation and high accuracy [26]. Details to create a RBF approximation is presented in Section 3.3.

(4) Model validation. The RBF approximation model created during the above steps needs to be validated to see its' performance. Various methods can be employed, such as the correlation coefficient, the mean absolute error, and so on [27]. In this paper, the root-mean-square error (RMSE) is computed to accomplish this task. RMSE reflects the global approximation accuracy. The smaller the value of RMSE, the higher the global accuracy of a corresponding approximation model. If the approximation is not acceptable, the process will turn to step 2.

(5) Objective function evaluations. With the response from RBF approximation and the experimental data, objective function to identify the formation parameters is evaluated according to Eq 10.

(6) Optimization. The evaluated objective in step 4 is optimized to figure out the optimal formation parameters. The optimizer is chosen up to the user's preference.

(7) Convergence test. With the optimization result in p_n and $f(p_n)$ step 5, the convergence can be tested. With the optimization result, the convergence can be test. The convergence can be governed by the two inequality functions.

$$\begin{aligned} &\left|f(p_n)-f(p_{n-1})\right|<\varepsilon_f,\\ &\left|p_n-p_{n-1}\right|<\varepsilon_x, \end{aligned} \tag{11}$$

Where ε_f and ε_x are tolerances supplied by the user, and n is the current iteration counter. If any of the two inequalities is true, the algorithm is considered to be converged and the optimum formation parameters are obtained.

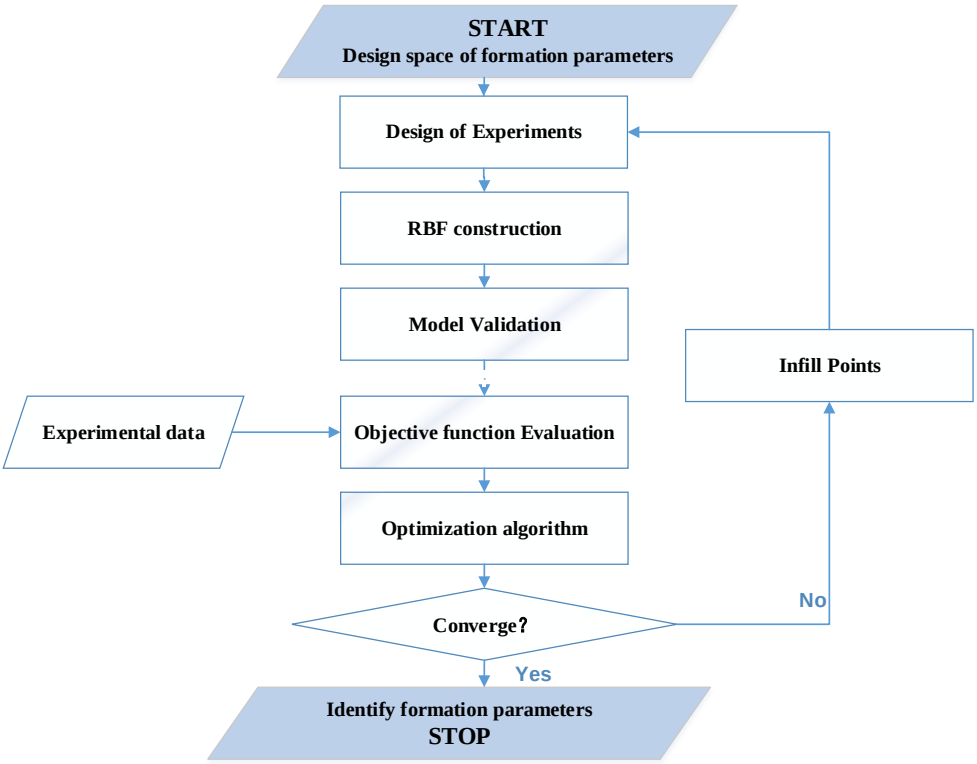

Figure 1. The process of the RBF based inverse method for parameter identification.

3.3. RBF construction

A radial basis function uses a series of basic functions that are symmetric and centered at each sampling point and it has been developed for interpolation of scattered multivariate data [28]. For a given set of inputs $\mathbf{X}^l = \{x_1^l, x_2^l, ..., x_{n_l}^l\}$ and the corresponding outputs $\mathbf{Y}^l = \{y_1^l, y_2^l, ..., y_{n_l}^l\}$, the RBF model can be written as follows:

$$\hat{y}^l(x) = \sum_{i=1}^{n_l} \lambda_i \phi(\left\| x - x_i^l \right\|) \tag{12}$$

Where n_l is the number of sample points, $\left\| x - x_i^l \right\|$ is the euclidean norm between the design variable x and the i_{th} sampling points, λ_i is corresponding weight coefficient of the i_{th} basic function $\phi(\left\| x - x_i^l \right\|)$. With the n_l sampling points, a matrix equation can be denoted the following:

$$\mathbf{A}\boldsymbol{\lambda} = \mathbf{Y} \tag{13}$$

$\mathbf{A}$ is a $n_l \times n_l$ basic function matrix with an element $A_{i,j} = \phi(\left\| x_i^l - x_j^l \right\|)$, and $\boldsymbol{\lambda}$ is the unknown coefficient matrix $\boldsymbol{\lambda} = [\lambda_1, \lambda_2, ..., \lambda_{n_l}]^T$, $\mathbf{Y}$ is the vector of function values at each sampling point

$\mathbf{Y}=[y_1^l, y_2^l, ..., y_{n_l}^l]^T$. The basis function directly affect the performance and complexity of the final RBF model. Various basis functions, such as Gaussian base, multiquadric base, the inverse multiquadric base and the thin-plate spline base, can be utilized in RBF. In our work, the commonly used Gaussian base is used for its good local estimation capacity.

$$\phi(\left\| x_i^l - x_j^l \right\|) = e^{-c(x_i^l - x_j^l)^2}, 0 < c \le 1 \tag{14}$$

In this way, the RBF approximation model is created.

4. Multi-objective cellular particle swarm optimization

4.1. Combination of cellular automata and particle swarm optimization

4.1.1. PSO

Suppose there is a swarm with size *S*, in which each particle can be described by current position X_i , current velocity V_i , best position P_i (*i*=1,2,…,S). One particle flies with the guideline of its own historical experience and other particles'. Suppose P_g^t stands for the global best position thus far and *t* represents the current generation. Aiming to search for the optimum, the velocity and position of a particle is updated as the following equations.

$$\begin{aligned} V_i^{t+1} &= \omega^t V_i^t + c_1 r_1 (P_i^t - X_i^t) + c_2 r_2 (P_g^t - X_i^t) \\ X_i^{t+1} &= X_i^t + V_i^{t+1} \end{aligned} \tag{15}$$

In which, c_1 is a acceleration constant that regulates the relative velocities according to the best global positions and c_2 a acceleration constant that regulates the relative velocities according to the local positions. And $r_j (j=1,2)$ is a random variable drawn from a uniform distribution in the open interval (0, 1). ω^t is a inertia weight that is used to balance the capabilities of global and local search, and can be obtained as following:

$$\omega^t = \omega_{\max} - \frac{\omega_{\max} - \omega_{\min}}{T} t \tag{16}$$

Where $\omega_{\max}$ is the initial weight, $\omega_{\min}$ is the final weight, and T is the maximum number of generations.

4.1.2. Cellular automata

Cellular automata (CA) is firstly invented in late 1940s and then caught a lot of attention and interests of theoretical research because of the famous game of life [29]. It has been successfully applied into different areas, such as bus or traffic route planing system [30,31], generation of digital art works [32], parameters estimation in optimization algorithms [33,34], and fault detection [35] and so on.

A cellular automaton includes a lattice of uniformly arranged finite state automata. Each one takes information from the neighboring automata. And then the neighbor update its next states via a state transition function. Several definitions of CA is described as following.

Definition 1: Cell

Cellular is not only the most basic but also important element in CA, which is also named as cell, which is distributed on the discrete crystal lattice of a cell space. Theoretically, cell space can be arbitrarily dimensional Euclidean space (such as one-dimensional or multi-dimensional). Different dividing ways of the cell space would bring about different cell shapes, but each cell can only has one state at a particular time.

Definition 2: Cell space

The set of all the cells locates in the Euclidean space is the cell space. For the one-dimensional cell space, there is only one dividing way, namely the isometric and piecewise division. While for the two-dimensional cell space, the triangle, hexagonal and square cell shapes is commonly used.

Definition 3: Cell state

Theoretically, each cell at one particular time can only has one cell state. The cell state at a certain time is usually described via multiple variables. Cell state in this work is denoted as $S_i^t = \{P_i^t, P_g^t, V_i^t, X_i^t, ...\}$.

Definition 4: Neighbors

As its name implies, neighbors is the around cells which impact the state of the current cell in the next moment. For a one-dimensional CA, all the cells within the neighborhood radius are treated as neighbors. Definition of neighbors in a two-dimensional CA is more complex. There are four commonly used neighborhood forms, which are illustrated in Figure 2. Detailed description of these neighborhood forms can be refer to literatures [36–38].

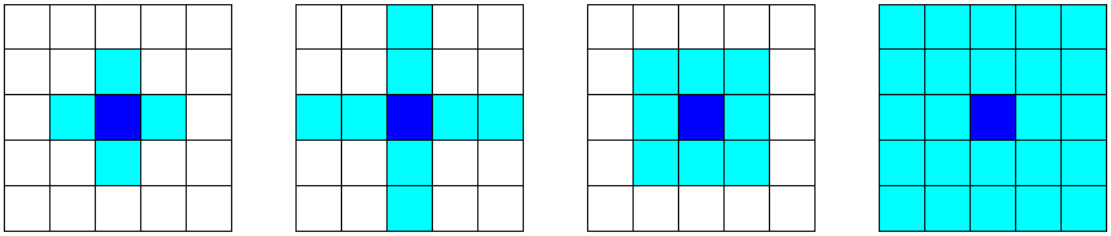

Figure 2. (a) Von Neumann. (b) Extended Von Neumann. (c) Moore. (d) Extended Moore Neighborhood templates.

Except for these four commonly used neighborhood forms, one can propose a proper neighborhood function according to the specific problem and applications.

With the above neighborhood templates, deterministic finite automata (cells) interconnects with each other and work synchronously at discrete time steps. With a collection of such cells, CA could present enormous power to solve problems. In this way, CA could play an important role in modeling complex systems. Thus, the above mentioned CA model was adopted to study the swarm system in our early work [33].

4.1.3. CPSO framework

PSO creates a swarm system in which interaction are involved among different particles. When employ the CA mechanism into the swarm system, each particle in PSO exchanges information not only with other particles but also with its neighbors.

The CA model employed for PSO structure is described as follows:

(1) Cell: Refers to the selected candidate solution;

(2) Cell space: Stands for the set of all cells;

(3) Cell state: Refers to the information of the selected candidate solution at a certain time *t*. The cell state consist of the current position X_i, current velocity V_i etc. State of the *i*th cell can be described by $S_i^t = \{P_i^t, P_g^t, V_i^t, X_i^t, ...\}$;

(4) Neighborhood: Refers to the around particles which could be defined through neighborhood functions.

(5) Transition rule: Refers to the information exchange rules between different cells, which can be denoted as $S_i^{t+1} = f(S_i^t \cup S_{N(i)}^t) = f(S_i^t, S_{i+\delta_1}^t, ..., S_{i+\delta_i}^t)$.

With the above definitions, a Cellular Particle Swarm Optimization (CPSO) is proposed. General framework is present in Figure 3.

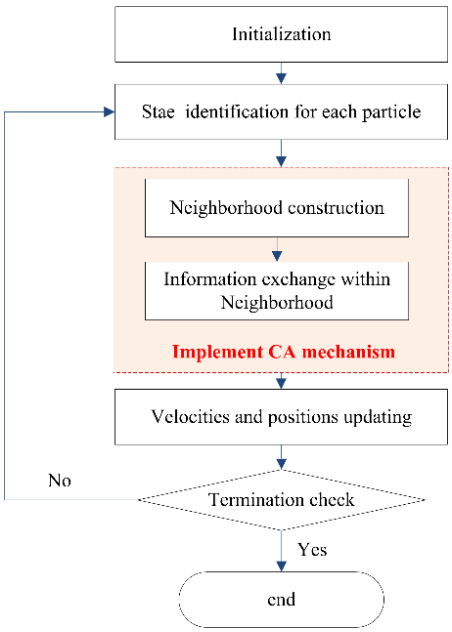

Figure 3. General framework of CPSO.

4.2. Non-dominated ranking

When there are multi-objective functions, individual objective functions $f_i(x), i = 1,...,k$ are expected to be minimized simultaneously, which is unlikely to be achieved in real life. In order to deal with this issue, multi-objective optimization algorithms is developed to distinguish "dominated" and "non-dominated" solutions, and reveal the Pareto optimal solutions that consist of non-dominated solutions.

Suppose the optimization is to minimize $f_i(x), i = 1,...,k$, searching for solutions u and w, if $f_i(u) \le f_i(w), i = 1,...,k$, and $\exists i=1,...,k,\ f_i(u) < f_i(w), i = 1,...,k$, u is a solution that dominates w, and w is a non-dominated solution. Solution that are not dominated by any other solutions is called a non-dominated solution or a Pareto optimal solution. Such a set of non-dominated solutions comprises a Pareto optimal set.

A process of non-dominated ranking is described as following. For a certain solution q, the domination count n_q and dominating set S_q are calculated. The former one is the number of solutions that dominate solution q and the later refers to the set would contain all the individuals that are dominated by solution q.

Solutions that have domination count as zero are in the first non-dominated front. For a solution in the first non-dominated front, visit members in set S_q and reduce its domination count by one. Keeping doing in this way, for a member s, if its domination count becomes zero, it would be put into a separate list Q, which belong to a second non-dominated front. Continue the above procedure with each member of Q and the third front would be identified. The process continues until all fronts are identified.

4.3. Constraint handling

Domination constraint rule is employed to handle the constraints. Suppose there are two assumed feasible solutions, a and b, if a dominates b, then a is deemed to be superior to b. When a is feasible but b is non-feasible, a is deemed to be superior to b. In the case where both a and b are non-feasible, solution that has a smaller overall constraint violation deemed to be superior. In the case where two solutions violates the same number of constraints, the solution with violation of a smaller amount is superior.

4.4. Elitist selection and external archive maintenance

4.4.1. Elitist selection

According to Eq 15, selection of global optimal position and the particle optimal position directly affects the particles' flight trajectory during the optimization. In single-objective PSO, they can be determined simply and uniquely through fitness values. While in multi-objective optimization, there are plenty of potential feasible solutions which cannot be distinguished just according to the fitness values. In this work, the Pareto-based dominance, which adopts non-dominated particle between current position and previous best position is chosen as p_{Best}. If the particles and previous best position do not dominate each other, then p_{Best} is select randomly. While the crowding density (namely, Euclidean distance between current particles and non-inferior solutions in external archive) is employed to choose a particle with minimum Euclidean distance as g_{Best}.

4.4.2. External archive maintenance

An external archive is established to lodge the non-dominated solutions. And non-dominated solutions can be preserved and maintained in a fixed size via an external archive. Process to create and maintain an external archive can be described as following. Firstly, the population particles are compared to the external archive particles to select the non-inferior particles are stored into external archive. When the number of non-inferior solutions exceed the predetermined archive size, particle with a minimum crowded distance is removed. Then sort the remaining particles again according to the crowded distance, and again select a particle with a minimum crowded distance to be removed. Keep these step repeated until meets the predetermined archive size. When the particles number have not reached the predetermined value, all the non-inferior particles are stored in external archive.

Elitist solutions are efficiently preserved in the above external archive maintenance strategy. As the optimization process runs, numbers of non-dominated solutions increase quickly and sharply. External archive size is kept via the above remove mechanism, and several non-dominated solutions would also be lost. Moreover, the evolution and optimizing process that is guided by global optimal position could easily resulting in a precocious convergence.

In order to deal with this issue, a redundant set is proposed to keep the diversity of non-dominated solutions in the external archive, in which the randomly chosen particles in the external archive are varied into particles in the redundant set. Given the redundant set size R_{size}, then $\mathrm{R}_{\text{size}}=\alpha \mathrm{A}_{\text{size}}$, where A_{size} stands for the external archive size. The variation formula is described as follows:

$$\begin{aligned} p_{new}(d) &= p_{new}(d) + \eta(\max(d) - \min(d)) \\ \eta &= \begin{cases} (2r_2)^{\frac{1}{\mu_2+1}} - 1, r_2 < 0.5 \\ 1-(2-2r_2)^{\frac{1}{\mu_2+1}}, 0.5 \le r_2 \end{cases} \\ v_{new} &= e^{-(\sum_{d=1}^{D}(\frac{1}{D}\sum_{i=1}^{N} v_i/N))+0.5} \end{aligned} \tag{17}$$

Where r_2 is a random number $r_2 \in (0,1)$, η is the variation factor and μ_2 is the variation distribution index. And p_{new} is the randomly selected particle in external archive. $\max(d), \min(d)$ are the upper and lower bound of the dth dimensional for particle p_{new}, respectively.

The above described redundant set brings two advantages into the optimization. One is to keep the population from a local optimum. The other is to increase the solutions diversity in the external archive via particle variation mechanism.

5. Bit factors and formation parameters identification

In this section, the proposed method is conducted for drilling parameters optimization of a 2500–2800 m well section in an oil field. An engineering detection test to collect drilling data is implemented in this section. A type 22 milling tooth bit with a diameter of 251 mm is employed in this well section. The relatively wearing capacity of teeth is 0.1. The torque applied at the bit is 15 KN.m. The bit factor can be determined according to the utilized bit type, which is present in Table 1.

Table 1. Bit factors of drilling model.

Bit factor	Value
Teeth wear coefficient, C_1	5
Bit weight influence coefficient, M_1	0.0146
Bit weight influence coefficient, M_2	6.44
Rotary speed influence coefficient, S_1	1.5
Rotary speed influence coefficient, S_1	0.000065

In a practical drilling engineering case, the hydraulic factor can enhance the drilling process. As it all knows, hydraulic energy is helpful to break the rock, and it is expected to be largest in a drilling practice. Therefore, the hydraulic efficiency C_H and differential pressure C_{P} are supposed to be 1 to maximize the hydraulic factor. The real-time drilling data of the 2500–2800 m well section is collected every 5 minutes. Table 2 lists a part of real time drilling data.

Table 2. Part of real time drilling data.

No.	weight on bit (KN)	rotary speed (r/min)	Rate of penetration (m/h)
1	288	60	10.35
2	258	62	10.14
3	286	62	9.87
4	288	63	9.77
5	289	63	9.54
6	292	64	9.50
7	298	65	9.54

Continued on next page

No.	weight on bit (KN)	rotary speed (r/min)	Rate of penetration (m/h)
8	300	64	9.27
9	293	65	8.92
10	294	66	8.85
11	300	69	9.11
12	301	70	9.02
13	295	66	8.28
14	292	67	8.14
15	294	68	8.13
16	286	66	7.60
17	288	65	7.47
18	290	68	7.64
19	293	69	7.67
20	295	70	7.67

With the real time drilling data and the proposed RBF based inverse parameter estimation method, the formation parameters of drilling model is identified as shown in Table 3.

Table 3. Formation parameters of drilling model.

Formation parameters	Values
Coefficient of formation drillability, K	0.00256
Threshold bit weight, M	10.1
Rotary speed exponent, λ	0.653
Bit tooth wear coefficient, C_2	3.569
Formation abrasiveness factor, A_f	0.00298

6. Optimization result and analysis

Performance of the proposed MOCPSO approach is verified through several investigations for drilling parameters optimization problem in this section. The impact of hybridization between CPSO and CA is explored by comparing its pareto front against that of the multi-objective PSO and NSGA-II approach. PSO is an evolutionary algorithm with good global search performance and fast convergence speed. Moreover, PSO has a simple structure, small number of tuning parameters and easy tuning procedures. And PSO has been successfully applied in many industrial field. In addition, the improved versions of PSO give high performance rankings according to the congress on Evolutionary computation (CEC). The NSGA-II is currently one of the most popular multi-objective genetic algorithm with a good convergence, and has become a performance baseline for other multi-objective optimization algorithms. Therefore, the impact of hybridization between cellular automata and particle swarm optimization is analyzed by comparing the obtained Pareto fronts against multi-objective PSO and NSGA-II approaches. Furthermore, effect of different neighborhood function (Von Neumann, Moore, Enhanced Moore and Margolus Neighborhood functions) are also investigated to figure out which one is most suitable for MOCPSO in the drilling parameters optimization problem.

Plenty of metrics including Spread [24], Inverse generational distance (IGD and Hyper volume (HV) can be adopted to estimate the performance of different approaches. IGD is a comprehensive index to measure the euclidean distance between the true Pareto front and the approximate Pareto front that was obtained via a certain optimization algorithms. IGD with a lower value indicates that the obtained approximate Pareto front is more close to the true Pareto front. In other words, the obtained Pareto front has a good diversity and convergence. Thus, IGD is adopted in this paper. Calculation formula for IGD is described as following equations.

$$IGD(P,P^*)=\frac{\sum_{i=1}^{|P|}d\left(P,P^*\right)}{|P|} \tag{18}$$

Where, P is a set of uniform sample on the true Pareto front, while P^* is the Pareto front obtained via a certain optimization algorithms, $d\left(P,P^*\right)$ is the minimum euclidean distance between P and P^*, and $|P|$ is the size of P*.

6.1. Comparison with multi-objective PSO and NSGA-II

In this section, performance of the proposed multi-objective CPSO is investigated through comparison with multi-objective PSO and NSGA-II. Parameters for these approaches are as show in Table 4, which is set according to researches [39,40]. Populations of these three approaches have the same sizes of 150, and the archive size are equal to each other, too. Moreover, they all have the same maximum iteration number, which is 100. Crossover probability in NSGA-II is 0.9, while mutation probability is 0.2 in both multi-objective PSO and NSGA-II. The Von Neumann neighborhood is employed in MOCPSO with the neighbors' number of 6. Each approach conducts 21 runs on the drilling parameters optimization problem, and the IGD of each running is calculated. The boxplot of the 21 runing results (corresponding to IGD) is illustrated in Figure 4. It can be inferred that the proposed MOCPSO method performs better as compared to the other two approaches by having the smallest mean of IGD. In other words, implication of CA into PSO is effective on the multi-objective drilling parameters optimization problem.

Table 4. Parameter settings for different approaches.

Parameters	Multi-CPSO	Multi-PSO	NSGA-II
Population size	150	150	150
Archive size	40	40	40
No. of Max. iteration	100	100	100
Crossover probability	-	-	0.9
Mutation probability	-	0.2	0.2
No. of neighbors	6	-	-

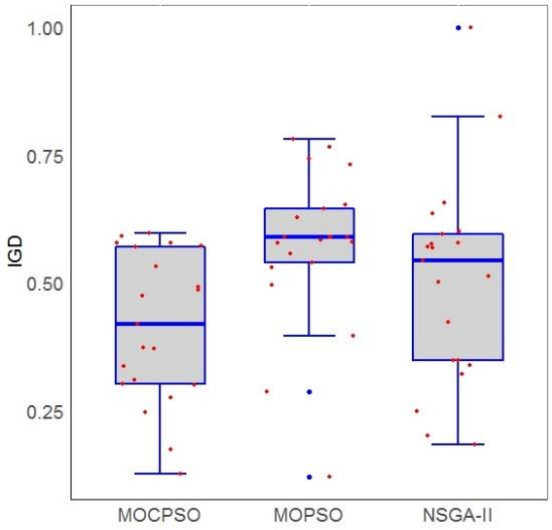

Figure 4. Experimental results box plot for different approaches.

A Wilcoxon signed ranks test [41] is conducted in order to figure out whether there is an algorithm is statistically different with other two approaches on the wellbore drilling parameters optimization problem. Various alternative approaches to find out whether the algorithms are statistically different exist, such as the Analysis of Variance (ANOVA) and Kolmogorov-Smirnov (K-S) test. ANOVA and K-S assume that the tested samples come from a normally distributed population, which is inconformity with the drilling parameters optimization case. Wilcoxon signed ranks tests is a non-parametric statistical test for pairwise comparisons. Therefore, Wilcoxon signed ranks test for pairwise comparison of the proposed multi-objective CPSO method, multi-objective PSO and NSGA-II is conducted. Results are summarized and present in Table 5, in which the *p*-value of the pairwise comparisons are 0.0028 and 0.0420 for NSGA-II v.s. MOCPSO and MOPSO v.s. MOCPSO, respectively. The obtained *p*-values are smaller than 0.05, which means that the calculated IGD samples gained by NSGA-II and MOPSO are statistically different with that via MOCPSO. Besides, Figure 4 shows that MOCPSO method performs better as compared to the other two approaches by having the smallest mean of IGD. Therefore, it can be inferred with *p*-values from Table 5 and the box plot in Figure 4 that the proposed MOCPSO is statistically superior to both multi-objective PSO and NSGA-II at the 0.05 level of significance.

Table 5. Results obtained through the Wilcoxon signed ranks for different approaches.

MOCPSO v.s.	R^+	R^-	*p*-value
NSGA-II	185	25	0.0028
MOPSO	172	38	0.0420

6.2. Different neighborhood functions investigation

In this section, different neighborhood functions including the four commonly used forms described in Section 4.1.2 are investigated. The wellbore drilling parameters optimization is a 3 dimensional problem. Therefore, the neighbor numbers of different neighborhood functions is show

in Table 6.

Table 6. Neighbor numbers of different neighborhood functions.

	Von Neumann	Extended Von Neumann	Moore	Extended Moore
Numbers	6	12	6	12

Similar to the operations in section 6.1, MOCPSO with each neighborhood function is conducted 21runs on the drilling parameters optimization problem. And the IGD of each running is calculated. Boxplot of (corresponding to IGD) according to different neighborhood function results is illustrated in Figure 5, in which "C1", "C2", "C3", "C4" means the Moore, Extended Moore, Extended Von Neumann, Von Neumann neighborhood functions respectively. It can be inferred that the Extended Von Neumann neighborhood function performs better as compared to the other neighborhood functions by having the smallest mean of IGD.

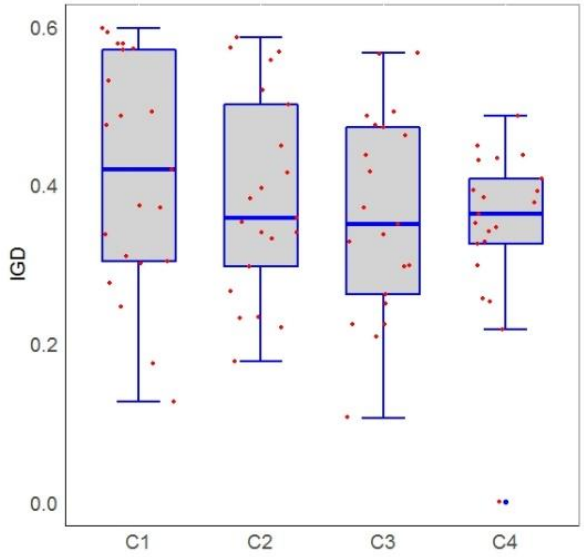

Figure 5. Experimental results box plot for different neighbor functions.

A Wilcoxon signed ranks tests is conducted in order to figure out whether the Extended Von Neumann neighborhood function is statistically different with other neighborhood functions on the wellbore drilling parameters optimization problem. A non-parametric statistical test was applied to the Extended Von Neumann neighborhood function and other three neighborhood functions for pairwise comparisons. The results is summarized and present in Table 7, in which the *p*-value of the pairwise comparisons are 0.7151, 0.2172 and 0.9584 for Extended Moore v.s. Extended Von Neumann, Moore v.s. Extended Von Neumann and Von Neumann v.s. Extended Von Neumann, respectively. The obtained *p*-values are larger than 0.05, which means that the calculated IGD samples gained with Extended Moore, Moore and Von Neumann are not statistically different with that via Extended Von Neumann.

Table 7. Results obtained through the Wilcoxon signed ranks for different neighborhood functions.

Extended Von Neumann v.s.	R^+	R^-	*p*-value
Extended Moore	126	105	0.7151
Moore	150	80	0.2172
Von Neumann	114	117	0.9584

In order to determine which one is the best neighborhood function among the four commonly used neighborhood functions, a Friedman test [41] is conducted to compare among the different neighborhood functions. An alternative approach is Cochran Q test. Both the Friedman and Cochran Q test are non-parametric statistical tests suitable for multiple comparisons. But the Cochran Q test is more suitable to analysis the binary quality type data. Therefore, the Friedman test was applied to the IGD results obtained via different neighborhood functions. The obtained individual ranking and the p-value are summarized and present in Table 8. It infers that the Extended Von Neumann neighborhood function has the highest ranking, which means that it performs better as compared to the other neighborhood functions according to the rank value. However, the *p*-value of the Friedman test in Table 8 is 0.3492, larger than 0.05, which means that the calculated IGD samples gained with the four commonly used neighborhood templates perform comparable with each other, and are not statistically different for the drilling parameters optimization problem. New neighborhood functions may be developed for the drilling parameters optimization to improve the performance of MOCPSO further.

Table 8. Results obtained through the Friedman test for different neighborhood functions.

Neighborhood functions	Rank	*p*-value
Moore	2.9048	0.3492
Extended Moore	2.5238	
Extended Von Neumann	2.2381	
Von Neumann	2.3333	

7. Conclusions

Drilling parameters optimization is an important procedure in drilling engineering, for it concerns quality, efficiency, cost and safety of the practical drilling process. A three-objective drilling parameters optimization model, in which rate of penetration, mechanical specific energy and drilling life of bit are included, is developed in this paper. Besides, the RBF is employed to estimate the formation parameters in rate of penetration model. After that, a multi-objective CPSO is proposed to solve the three-objective drilling parameters optimization problem. Performance of the MOCPSO is compared with multi-objective PSO, NSGA-II and found that MOCPSO is statistically superior. Moreover, performance investigations of different neighborhood functions in MOCPSO are conducted, and indicated that the four commonly used neighborhood templates perform comparable with each other, and are not statistically different for the drilling parameters optimization problem. In the future work, a more efficiency neighborhood function should be developed for this problem. Besides, a new mathematical model formulation for well drilling parameters optimization that

considers geological uncertainty is another interesting research work.

Acknowledgements

This research was supported by the National Natural Science Foundation of China under grant NO.51505439 and 51805495, the Research Fund for the Doctoral Program of Higher Education of China under grant No.2014M562085 and the Fundamental Research Funds for the Central Universities, CUG: Grant no. CUGL150821.

References

1. J. W. Graham and N. L. Muench, Analytical determination of optimum bit weight and rotary speed combinations, in Fall Meeting of the Society of Petroleum Engineers of AIME, Society of Petroleum Engineers: Dallas, Texas, (1959), 35.
2. A. T. Bourgoyne Jr and F. S. Young Jr., A multiple regression approach to optimal drilling and abnormal pressure detection, *Soc. Pet. Eng. J.*, **14** (1974), 371–384.
3. M. R. Reza and C. F. Alcocer, A unique computer simulation model well drilling: Part I—the reza drilling model, in SPE California regional meeting, Society of Petroleum Engineers: Oakland, California. (1986), 10.
4. F. Iqbal, Drilling optimization technique-using real time parameters, in SPE russian oil and gas technical conference and exhibition, Society of Petroleum Engineers: Moscow, Russia, (2008), 6.
5. Q. B. Amjad, S. Waheed and M. S. K. Jadoon, et al., Drilling optimization of kohat/potohar region by mathematical model (using matlab) and comparative method—A case study, in SPE/PAPG pakistan section annual technical conference, Society of Petroleum Engineers: Islamabad, Pakistan, (2015), 12.
6. T. Eren, M. E. Ozbayoglu, Real time optimization of drilling parameters during drilling operations, in SPE oil and gas india conference and exhibition, Society of Petroleum Engineers: Mumbai, India, (2010), 14.
7. L. Zhang, D. Teng and Q. Shen, Parameter optimization of drilling based on biogeography-based optimization, in 2018 2nd International Symposium on Resource Exploration and Environmental Science, REES 2018, April 28, 2018–April 29, 2018. Ordos, China: Institute of Physics Publishing 2018.
8. Y. X. Sun, L. H. Qiao and H. F. Sun, et al., Real-time Surveillance System of Mechanical Specific Energy Applied in Drilling Parameters Optimization, in *Proceedings of the 2nd Annual International Conference on Advanced Material Engineering* (eds. A. K. Bhatnagar, R. Dahai, K. M. Pandey, H. Haeri, R. A. Andrievskiy, M. Ziara, G. Feng), (2016), 771–777.
9. Y. Zhai, Z. Wang and Q. Zhang, et al., The study of deepwater drilling parameters optimization. in 2012 International Conference on Applied Mechanics and Materials, ICAMM 2012, November 24, 2012–November 25, 2012. Sanya, China: Trans Tech Publications 2013.
10. M. Cui, J. Li and G. Ji, et al., Optimize method of drilling parameter of compound drilling based on mechanical specific energy theory, *Pet. Drill. Tech.*, **42** (2014), 66–70.
11. A. Rey, R. Zmeureanu, Micro-time variant multi-objective particle swarm optimization (micro-TVMOPSO) of a solar thermal combisystem, *Swarm. Evolut. Compu.*, **36** (2017), 76–90.

12. L. Vanneschi, R. Henriques and M. Castelli, Multi-objective genetic algorithm with variable neighbourhood search for the electoral redistricting problem, *Swarm. Evolut. Comput.,* **36** (2017), 37–51.
13. F. Florence and F. P. Iversen, Real-time models for drilling process automation: equations and applications, in IADC/SPE drilling conference and exhibition, Society of Petroleum Engineers: New Orleans, Louisiana, USA, (2010), 16.
14. X. Chen, D. Gao and B. Guo, et al., Real-time optimization of drilling parameters based on mechanical specific energy for rotating drilling with positive displacement motor in the hard formation, *J. Nat. Gas. Sci. Eng.,* **35** (2016), 686–694.
15. K. Mohan, F. Adil and R. Samuel, Comprehensive hydromechanical specific energy calculation for drilling efficiency, *J. Energ. Resour-Asme.,* **137** (2015), 8.
16. A. Beck and M. Teboulle, A fast iterative shrinkage-thresholding algorithm for linear inverse problems, *Siam. J. Imaging. Sci.,* **2** (2009), 183–202.
17. C. L. Wu and K. W. Chau, Rainfall-runoff modeling using artificial neural network coupled with singular spectrum analysis, *J. Hydrol.,* **399** (2011), 394–409.
18. P. Hajikhodaverdikhana, M. Nazari and M. Mohsenizadeh, et al., Earthquake prediction with meteorological data by particle filter-based support vector regression, *Eng. Appl. Comput. Fluid. Mech.,* **12** (2018), 679–688.
19. S. F. Ardabili, B. Najafi and S. Shamshirband, et al., Computational intelligence approach for modeling hydrogen production: A review, *Eng. Appl. Comput. Fluid. Mech.,* **12** (2018), 438–458.
20. R. Moazenzadeh, B. Mohammadi and S. Shahaboddin, et al., Coupling a firefly algorithm with support vector regression to predict evaporation in northern Iran, *Eng. Appl. Comput. Fluid. Mech.,* 1**2** (2018), 584–597.
21. R. C. Bose and K. A. Bush, Orthogonal arrays of strength two and three, *Ann. Math. Stat.,* **23** (1952), 508–524.
22. J. Stufken and B. Tang, Complete Enumeration of Two-Level Orthogonal Arrays of Strength d With d+2 Constraints, *Ann. Stat.,* **35** (2007), 793–814.
23. J. Zheng, An output mapping variable fidelity metamodeling approach based on nested Latin hypercube design for complex engineering design optimization, *Appl. Intell.,* **48** (2018), 3591–3611.
24. Y. M. Yu, D-optimal designs via a cocktail algorithm, *Stat. Comput.,* **21** (2011), 475–481.
25. K. T. Fang and D. K. J. Lin, Uniform design in computer and physical experiments, *Grammar. Technol. Dev.,* (2008), 105–125.
26. Q. Zhou, Y. Rong and X. Shao, et al., Optimization of laser brazing onto galvanized steel based on ensemble of metamodels, *J. Intell. Manuf.,* **29** (2018), 1417–1431.
27. B. Najafi, S. F. Ardabili and S. Shamshirband, et al., ANFIS and RSM to estimating and optimizing the parameters that affect the yield and cost of biodiesel production, *Eng. Appl. Comput. Fluid. Mech.,* **12** (2018), 611–624.
28. N. Queipo, R. Haftka and W. Shyy, et al., Surrogate-based analysis and optimization, *Prog. Aerosp. Sci.,* **41** (2005), 1–28.
29. M. Gardner, On Cellular Automata, Self-Reproduction, the Garden of Eden and the Game "Life", *Sci. Am.,* **224** (1971), 112–117.

30. M. E. Larraga and L. Alvarez-Icaza, Towards a realistic description of traffic flow based on cellular automata, in 14th IEEE International Intelligent Transportation Systems Conference, ITSC 2011, October 5, 2011–October 7, 2011. Washington, DC, United states: Institute of Electrical and Electronics Engineers Inc. 2011.
31. Y. J. Luo, B. Jia and X. G. Li, et al., A realistic cellular automata model of bus route system based on open boundary, *Trans. Res. C-Emer.,* **25** (2012), 202–213.
32. M. A. J. Javid, W. Alghamdi and A. Ursyn, et al., Swarmic approach for symmetry detection of cellular automata behaviour, *Soft. Comput.,* **21** (2017), 5585–5599.
33. J. Zheng, X. Y. Shao and L. Gao, et al., A prior-knowledge input LSSVR metamodeling method with tuning based on cellular particle swarm optimization for engineering design, *Expert. Syst. Appl.,* **41** (2014), 2111–2125.
34. B. Placzek, A cellular automata approach for simulation-based evolutionary optimization of self-organizing traffic signal control, *J. Cell. Autom.,* **11** (2016), 475–496.
35. B. K. Mathew, S. K. John and C. Pradeep, New technique for fault detection in quantum cellular automata, in 1st International Conference on Emerging Trends in Engineering and Technology, ICETET 2008, July 16, 2008–July 18, 2008. Nagpur, Maharashtra, India: IEEE Computer Society 2008.
36. L. Gao, J. D. Huang and X. Y. Li, An effective cellular particle swarm optimization for parameters optimization of a multi-pass milling process, *Appl. Soft. Comput.,* **12** (2012), 3490–3499.
37. Y. Shi, H. C. Liu and L. Gao, et al., Cellular particle swarm optimization, *Inform. Sci.,* **181** (2011), 4460–4493.
38. C. Lu, L. Gao and X. Li, et al., A multi-objective approach to welding shop scheduling for makespan, noise pollution and energy consumption, *J. Cleaner. Prod.,* **196** (2018) 773–787.
39. V. Mansouri, R. Khosravanian and D. A. Wood, et al., 3-D well path design using a multi objective genetic algorithm, *J. Nat. Gas. Sci. Eng.,* **27** (2015), 219–235.
40. A. Atashnezhad, D. A. Wood and A. Fereidounpour, et al., Designing and optimizing deviated wellbore trajectories using novel particle swarm algorithms, *J. Nat. Gas. Sci. Eng.,* **21** (2014), 1184–1204.
41. J. Derrac, S. Garcia and D. Molina, et al., A practical tutorial on the use of nonparametric statistical tests as a methodology for comparing evolutionary and swarm intelligence algorithms, *Swarm. Evolut. Comput.,* **1** (2011), 3–18.

AIMS Press