A Modified Mean Gray Wolf Optimization Approach for Benchmark and Biomedical Problems

Narinder Singh and SB Singh

Department of Mathematics, Punjabi University, Patiala, India.

Evolutionary Bioinformatics
Volume 13: 1–28

DOI: 10.1177/1176934317729413

ABSTRACT: A modified variant of gray wolf optimization algorithm, namely, mean gray wolf optimization algorithm has been developed by modifying the position update (encircling behavior) equations of gray wolf optimization algorithm. The proposed variant has been tested on 23 standard benchmark well-known test functions (unimodal, multimodal, and fixed-dimension multimodal), and the performance of modified variant has been compared with particle swarm optimization and gray wolf optimization. Proposed algorithm has also been applied to the classification of 5 data sets to check feasibility of the modified variant. The results obtained are compared with many other meta-heuristic approaches, ie, gray wolf optimization, particle swarm optimization, population-based incremental learning, ant colony optimization, etc. The results show that the performance of modified variant is able to find best solutions in terms of high level of accuracy in classification and improved local optima avoidance.

KEYWORDS: Gray wolf optimization (GWO), optimization techniques, meta-heuristics

RECEIVED: May 5, 2017. **ACCEPTED:** August 1, 2017.

PEER REVIEW: Two peer reviewers contributed to the peer review report. Reviewers' reports totaled 553 words, excluding any confidential comments to the academic editor.

TYPE: Original Research

FUNDING: The author(s) received no financial support for the research, authorship, and/or publication of this article.

DECLARATION OF CONFLICTING INTERESTS: The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

CORRESPONDING AUTHOR: Narinder Singh, Department of Mathematics, Punjabi University, Patiala 147002, Punjab, India. Email: narindersinghgoria@ymail.com

Literature Review

Many researchers have proposed various nature-inspired techniques to solve the different types of real-life problems to improve the quality of the solutions. The most popular meta-heuristics algorithms are discussed in this section.

Evolution strategies (ES) are evolutionary algorithms that date back to the 1960s and are most commonly applied to black-box global optimization functions in continuous search spaces. Evolution strategy was proposed by Rechenberg.[1] This approach is population-based on ideas of evolution and adaptation. In this use, mutation, recombination, and selection are applied to a crowd of individuals containing member of the population solutions to evolve iteratively better and better optimization problem solutions.

The particle swarm optimization (PSO) algorithm was first introduced by RC Eberhart (Electrical Engineer) and James Kennedy (Social Psychologist).[2] Its fundamental judgment was primarily inspired by the simulation of the social behavior of animals such as bird flocking and fish schooling. While searching for food, the birds either are scattered or go together before they settle in the position where they can find the food. While the birds are searching for food from one position to another, there is always a bird that can smell the food very well, that is, the bird is observable of the position where the food can be found, having the correct food resource message. Because they are transmitting the message, particularly the useful message at any period while searching the food from one position to another, the birds will finally flock to the position where food can be found.

Genetic algorithm (GA) was proposed by Holland.[3] This approach is inspired by Darwin's theory of evolution "survival of the fittest." In this approach, each new population is created by mutation and combination of the individuals in the previous generation. Because the best individuals have a higher probability of participating in generating the new position of the candidate, the new position is likely to be better than the previous position of the candidate.

Ant colony optimization (ACO) approach was proposed by Marco Dorigo et al.[4] This approach is based on the behavior of ants seeking a path between their colony and source of food. The basic idea has since diversified to solve a wider class of numerical problems and improved the quality of the solutions.

Population-based incremental learning (PBIL) was introduced by Shumeet.[5] It is a global optimization approach and an estimation of distribution algorithm. Population-based incremental learning approach is an extension to the Evolutionary GA achieved through the re-examination of the performance of the evolutionary GA in terms of competitive learning. It is easier than a GA and in a number of cases leads to better and good qualities of solutions than a standard GA.

Recently, some of most popular variants are gravitational search algorithm (GSA),[6] gravitational local search,[7] big bang-big crunch,[8] central force optimization,[9] artificial chemical reaction optimization algorithm,[10] charged system search (CSS),[11] ray optimization,[12] galaxy-based search algorithm,[13] black hole,[14] curved space optimization,[15] and small-world optimization algorithm,[16] and many others. All these approaches are different from evolutionary algorithms in the sense that a random set of search agents communicate surrounding the search area according to the physical rules.

Gray wolf optimization (GWO) algorithm was first proposed by Mirjalili et al[17] It is a nature-inspired optimizer approach and mimics the leadership hierarchy and hunting mechanism of gray wolves in nature. Four types of gray wolves alpha (α), beta (β), delta (δ), and omega (ω) are worked for simulating the leadership hierarchy. A wolf very near to a target is assigned by α, second level of near to a target is assigned as β, third level of near to a target is assigned as δ, and remaining wolves are assigned as ω. The main 3 stages of hunting, searching for target, encircling target, and attacking target, have been implemented. The performance of this approach was tested on several benchmark functions and real-life problems. On the basis of results obtained, it was concluded that the present approach is superior to and better than other existing nature-inspired approaches such as PSO, differential evolution, GSA, ES, and evolutionary programming.

The GWO for training multilayer perceptron was first proposed by Mirjalili.[18] On the basis of this existing variant, the author solved 3 function approximation data sets and 8 standard data sets including 5 classifications. The performance of the proposed variant was compared with a number of existing nature-inspired algorithms such as PSO, GA, ACO, ES, and PBIL. The results obtained showed that the proposed variant provides competitive solutions in the forms of improved local optima avoidance and also demonstrates high level of accuracy in approximation and classification of the proposed trainer.

Some of the recent population-based nature-inspired training algorithms are social spider optimization,[19] invasive weed optimization,[20] chemical reaction optimization,[21] teaching-learning–based optimization,[22] biogeography-based optimization,[23] and CSS.[24] Several researchers have used the above variants to solve the real-life medical problems and presented their high performance in terms of approximating the global optimum. In this article, we have also solved these real medical problems using the newly proposed mean gray wolf optimization (MGWO) algorithm. We have also reported that quality of solution of these problems using MGWO algorithm is better than other existing algorithms.

Two novel binary versions of the GWO (bGWO) algorithm were also proposed by Emary et al[25] for feature selection in wrapper mode. These algorithms were applied and used for feature selection in machine learning domain using different initialization methods. The bGWO approaches are hired in the feature selection domain for evaluation, and the results are compared against 2 of the well-known feature selection algorithms—PSO and GA.

Mittal et al[26] developed a modified variant of the GWO called modified GWO. An exponential decay function is used to improve the exploitation and exploration in the search space over the course of generations. On the basis of obtained results, authors proved that the modified variant benefits from high exploration in comparison with the standard GWO, and the performance of the variant is verified on a number of standard benchmarks and real-life NP-hard problems.

Sodeifian et al[27] used the response surface methodology to study the efficiency of supercritical fluid extraction from *Cleome coluteoides*. Chemical compositions extracted by hydrodistillation and SC-CO2 methods were identified by gas chromatography (GC)/mass spectrometry and determined by GC/flame ionization detector. Comparing the 2 techniques, the obtained solutions showed higher total extraction yield with SC-CO2 method.

The rest of the article is organized as follows. The newly proposed algorithm MGWO algorithm is presented in section "MGWO Algorithm." The proposed mathematical model and algorithm have also been discussed in section "MGWO Algorithm." The tested benchmark functions and numerical experiments are presented in sections "Testing Functions" and "Numerical Experiments." Parameter setting, results, discussion of standard benchmark functions, and real-life problems are represented in sections "Parameter Setting," "Analysis and Discussion on the Results," and "Real-Life Data Set Problems." Finally, the conclusion of the work is summarized at the end of the article.

Gray Wolf Optimization

Mirjalili et al[17] proposed a new swarm-based meta-heuristic approach. This variant mimics the hunting behavior and social leadership of gray wolves in nature. In this variant, the crowd is divided into 4 different groups (Figure 1).

The first 3 wolves in the best position (fittest) are indicated as α, β, and δ which guide the other wolves (ω) of the group toward promising areas of the search space. The position of each wolf of the group is updated using the following mathematical equations:

$$\vec{D} = \left|\vec{C} \cdot \vec{X}_p(t) - \vec{X}(t)\right| \tag{1}$$

$$\vec{X}(t+1) = \vec{X}_{p(t)} - \vec{A} \cdot \vec{D} \tag{2}$$

where $\vec{X}_p$ is the position vector of the prey, t indicates the current iteration, and $\vec{X}$ indicates the position vector of a gray wolf.

The vectors $\vec{A}$ and $\vec{C}$ are mathematically calculated as follows:

$$\vec{A} = 2a \cdot \vec{r}_1 - a \tag{3}$$

$$\vec{C} = 2 \cdot \vec{r}_2 \tag{4}$$

where components of a are linearly decreased from 2 to 0 and $\vec{r}_1, \vec{r}_2$ are random numbers lying between [0, 1].

Hunting

To mathematically simulate the hunting behavior of gray wolves, the hunt is usually guided by α, β, and δ which also participate in hunting occasionally. Suppose that α is the best solution of the candidate, β and δ have better knowledge about the potential location of prey.

We save the first 3 best candidate solutions obtained so far and oblige the other search agents to update their positions according to the position of the best search agents. The following mathematical equations are developed for this simulation:

$$\vec{D}_\alpha = \left|\vec{C}_1 \cdot \vec{X}_\alpha - \vec{X}\right|, \vec{D}_\beta = \left|\vec{C}_1 \cdot \vec{X}_\beta - \vec{X}\right|, \vec{D}_\delta = \left|\vec{C}_1 \cdot \vec{X}_\delta - \vec{X}\right| \quad (5)$$

$$\begin{aligned}\vec{X}_1 &= \vec{X}_\alpha - \vec{A}_1 \cdot \left(\vec{D}_\alpha\right), \vec{X}_2 = \vec{X}_\beta - \vec{A}_2 \cdot \left(\vec{D}_\beta\right), \\ \vec{X}_3 &= \vec{X}_\delta - \vec{A}_3 \cdot \left(\vec{D}_\delta\right)\end{aligned} \quad (6)$$

$$\frac{\vec{X}_1 + \vec{X}_2 + \vec{X}_3}{3} \quad (7)$$

The wolves update their positions randomly around the prey as represented in Figure 2.[17]

MGWO Algorithm

In this article, a modified variant MGWO is proposed for the purpose of improving the accuracy, convergence speed, and time performance of the GWO algorithm. In the proposed variant, mathematical equations of encircling and hunting have been modified. Remaining equations/procedure is same as that in GWO.[17]

The main purpose of this variant is to improve the movement or optimal path of each wolf in the searching space.

The MGWO approach is outlined in the following sections.

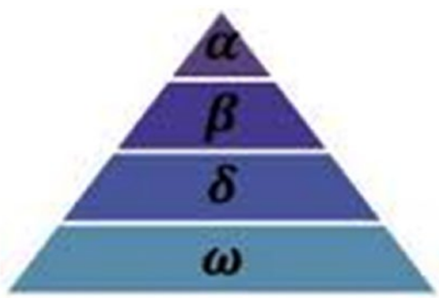

Figure 1. Hierarchy of gray wolf (dominance decreases from top to down). Adapted from Mirjalili et al.[17]

Encircling prey

Gray wolves encircle the prey during the hunt which can be modified using the following mathematical equation:

$$\vec{D} = \left|\vec{C}.\vec{X}_p(t) - \mu(\vec{X}(t))\right| \quad (8)$$

$$\vec{X}(t+1) = \vec{X}_{p(t)} - \vec{A}.\vec{D} \quad (9)$$

where μ is the mean, $\vec{X}_p$ is the position vector of the prey, t indicates the current iteration, and $\vec{X}(t)$ indicates the position vector of a gray wolf.

The vectors $\vec{A}$ and $\vec{C}$ are expressed as follows:

$$\vec{A} = 2a \cdot \vec{r}_1 - a \quad (10)$$

$$\vec{C} = 2 \cdot \vec{r}_2 \quad (11)$$

where components of a are linearly decreased from 2 to 0, and $\vec{r}_1, \vec{r}_2$ are random numbers lying between [0, 1].

Hunting

The hunting of prey is usually guided by α, β, and δ groups which participate occasionally. First 3 best candidate solutions

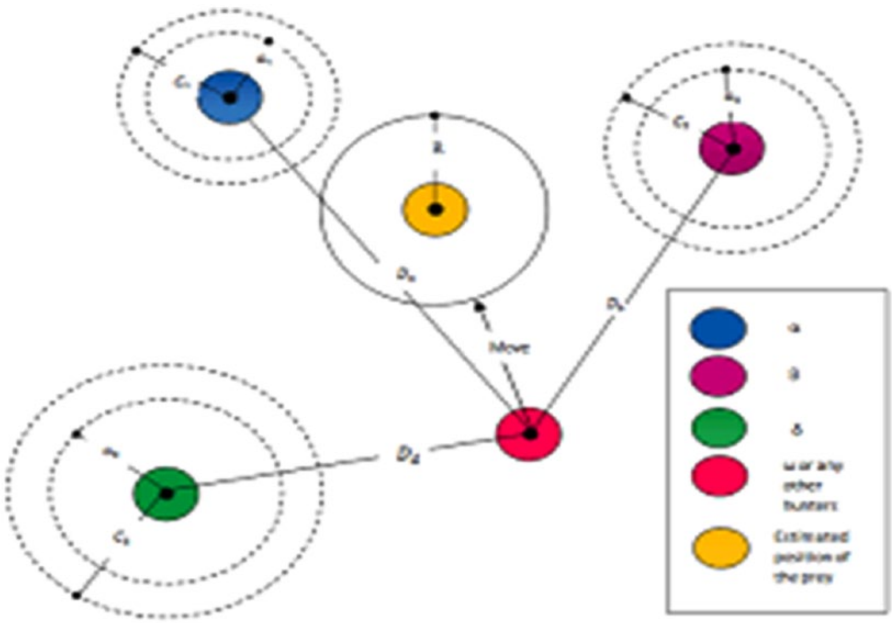

Figure 2. Positions updated by the wolves in gray wolf optimization. Adapted from Mirjalili et al.[17]

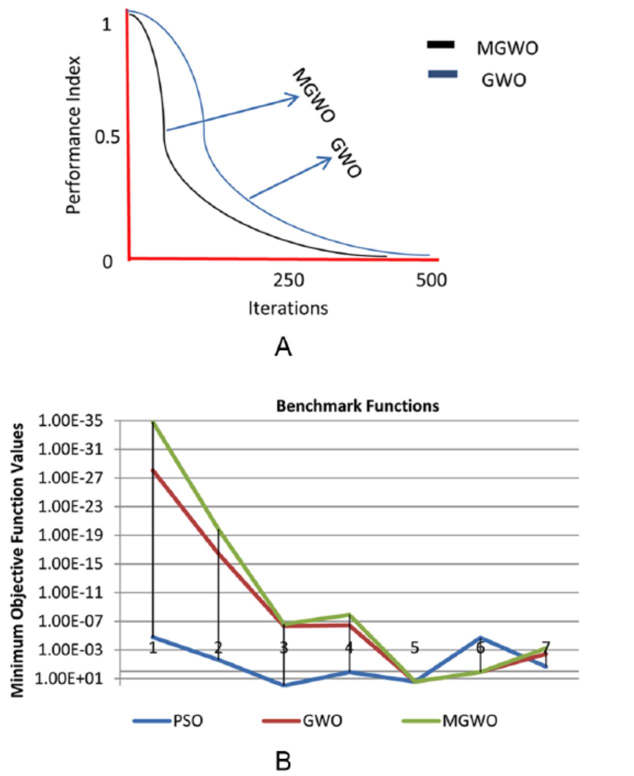

Figure 3. (a) Performance index graph and (b) performance graph of PSO, GWO, and MGWO. GWO indicates gray wolf optimization; MGWO, mean gray wolf optimization; PSO, particle swarm optimization.

are referred by α, β, and δ and the remaining candidate solutions are denoted by ω. The position of each wolf has been modified in the search space area by taking the mean of the positions. The following modified mathematical equations are proposed in this regard (Figure 3):

$$\vec{D}_\alpha = \left|\vec{C}_1 \cdot \vec{X}_\alpha - \mu(\vec{X}(t))\right| \tag{12}$$

$$\vec{D}_\beta = \left|\vec{C}_2 \cdot \vec{X}_\beta - \mu(\vec{X}(t))\right| \tag{13}$$

$$\vec{D}_\delta = \left|\vec{C}_3 \cdot \vec{X}_\delta - \mu(\vec{X}(t))\right| \tag{14}$$

$$\vec{X}_1 = \vec{X}_\alpha - \vec{A}_1 \cdot (\vec{D}_\alpha), \vec{X}_2 = \vec{X}_\beta - \vec{A}_2 \cdot (\vec{D}_\beta), \\ \vec{X}_3 = \vec{X}_\delta - \vec{A}_3 \cdot (\vec{D}_\delta) \tag{15}$$

$$\frac{\vec{X}_1(t)+\vec{X}_2(t)+\vec{X}_3(t)}{3} \tag{16}$$

In GWO and MGWO algorithms, the wolves update positions randomly around the prey which can symbolically be represented as shown in Figure 4.

The pseudocode of the MGWO algorithm:

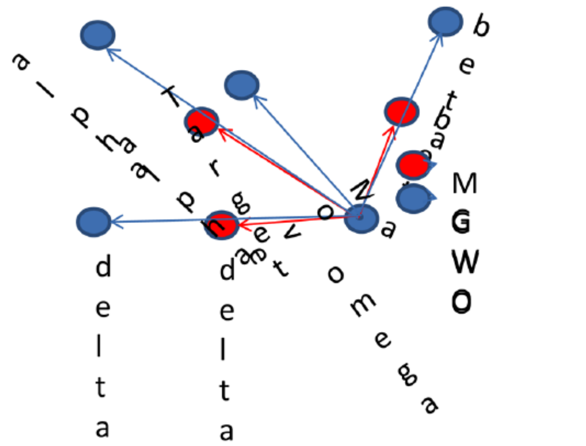

Figure 4. Positions updated in GWO and MGWO. GWO indicates gray wolf optimization; MGWO, mean gray wolf optimization.

Initialize the population $\vec{X}_i \quad (i = 1,2,3,\ldots,n)$

Initialize A, a and C

Calculate the fitness of each search candidate (agent) of the population in the search space

    $\vec{X}_\alpha$ is the first best search candidate (agent)

    $\vec{X}_\beta$ is the second best search candidate (agent)

    $\vec{X}_\delta$ is the third best search candidate (agent)

While ($t <$ max. number of iterations)

    For each search candidate (agent)

        Update the position of the current search (candidate) agent using equation (16)

    End for

    Update $\vec{A}$ and $\vec{C}$ using equations (10) and (11)

    Find the fitness of all search candidate (agent)

    Update $\vec{X}_\alpha$, $\vec{X}_\beta$, and $\vec{X}_\delta$ using equations (12) to (14)

End while

Return $\vec{X}_\alpha$

Testing Functions

The convergence and time-consuming performance of proposed variant have been tested on several types of standard functions, and the results obtained are compared with those obtained using other recent meta-heuristics. These classical functions have divided into 4 different parts, ie, unimodal, multimodal, fixed-dimension multimodal, and composite functions and are listed in Appendix 1 (Tables A to C) where $f_{\min}$ is the minimum objective function value, dim is the dimension, and *range* is the boundary of the standard function's search area. All these classical functions have been used by many scientists in their research (Holland,[3] Eberhart et al,[2] Dorigo et al,[4] Shumeet,[5] and many others).

Numerical Experiments

The MGWO, GWO, PSO, PBIL, and ACO algorithms are coded in MATLAB R2013a and implemented on Intel HD Graphics, 15.6″ 16.9 HD LCD, Pentium-Intel Core, i5 Processor 430 M, 320 GB HDD, and 3 GB Memory.

Parameter Setting

In MGWO, GWO, PSO, PBIL, and ACO algorithms, we have set the following parameters:

1. Number of search agents (candidate) = 30;
2. Maximum number of iterations (generations) = 500;
3. $\bar{a} \in [2,0]$.

Analysis and Discussion on the Results

In this section, effectiveness of using MGWO algorithm has been checked. Usually, it is done by solving a set of benchmark problems. We have used 23 such classical functions for the purpose of comparing the performance of the modified variants with other recent meta-heuristics. These classical functions are divided into 3 types:

1. Unimodal $(F_1\text{-}F_7)$—these functions are suitable for exploitation of the variants because they have one global optimum and no local optima. These functions are given in Appendix 1, Table A.
2. Multimodal $(F_8\text{-}F_{13})$—these functions have a large number of local optima and are helpful to examine local optima avoidance and exploration of the variants. These functions are given in Appendix 1, Table B.
3. Fixed-dimension multimodal $(F_{14}\text{-}F_{23})$—the dimension of these functions is fixed. The mathematical equation of these functions is given in Appendix 1, Table C.

The MGWO and GWO variants were run 30 times on each benchmark function. The numerical results (best solutions, minimum objective function value, maximum objective function value, standard deviation, mean and time performance) are reported in Tables 1 to 18. The modified variants, GWO and PSO algorithms, have to be run at least more than 10 times to find the best statistical results. It is again a common technique that a variant is run on a function many times and best solutions, mean, standard deviation, time-consuming performance, and minimum and maximum objective functions of the superior are obtained in the last generation.

To verify the convergence and time-consuming performance of MGWO variant, PSO and GWO variants are chosen. Here, we use 500 generations and 30 search agents for each of the variants. The convergence performance for unimodal, multimodal, and fixed-dimensional multimodal standard classical functions for the PSO, GWO, and MGWO is given in Figures 5 to 27 and results are presented in Tables 1 to 9. Simulated results in Tables 1 to 9 and Figures 5 to 27 show that the proposed variant is superior to PSO and GWO in terms of rate of convergence and best optimal solution. Hence, all experimental results reveal that the MGWO is relatively better as compared with PSO and GWO.

The experimental statistical results of the MGWO, PSO, and GWO variants on unimodal benchmark functions are shown in Tables 10 and 13. On the basis of results obtained in these tables, we are comparing the performance of modified variant with GWO and PSO variants in terms of minimum and maximum objective value of cost functions, mean, and standard deviation. After analysis, it may be seen that modified variant gives highly competitive solutions as compared with PSO and GWO on unimodal benchmark functions. As previously discussed, the unimodal benchmark problems are competent for benchmarking exploitation of the variants. Hence, all obtained solutions evidence high rate of exploitation capability of the MGWO variant.

Furthermore, the experimental numerical solutions of the proposed variant on multimodal test function are shown in Tables 11 and 14. We observe that modified variant performs better to other meta-heuristics on $F_8, F_9, F_{10}, F_{11}, F_{12}, \text{and}, F_{13}$. The results obtained in Tables 11 and 14 strongly prove that high exploration of MGWO variant is able to explore the search area extensively and give promising regions of the search area.

Furthermore, the statistical results of the modified variant on fixed-dimension multimodal functions are presented in Tables 12 and 15. For these functions, we have checked the rate of convergence performance of the modified variants, PSO and GWO, in terms of minimum and maximum objective functions, mean, and standard deviation values. The solutions are consistent with those of the benchmark test problems. Modified variant gives highly competitive solutions compared with other meta-heuristics, for these problems.

Finally, the performance of the newly proposed algorithm has been verified using starting and end time of the CPU (TIC and TOC), CPUTIME, and CLOCK. These results are provided in Tables 16 and 17, respectively. It may be seen that the modified variant solved most of the benchmark functions in least time as compared with other variants.

To sum up, all simulation results assert that the modified approach is very helpful in improving the efficiency of the GWO in terms of result quality as well as computational efforts.

Real-Life Data Set Problems

In this section, the following 5 data set problems are employed: (1) XOR, (2) Balloon, (3) Breast Cancer, (4) Iris,

Table 1. Best solution obtained by GWO and MGWO on 500 generations.

ITERATIONS	PROBLEM NAME					
	$f_1(x)$		$f_2(x)$		$f_3(x)$	
	GWO	MGWO	GWO	MGWO	GWO	MGWO
10	−1.7002e−15	6.8765e−19	−1.3769e−18	−3.8434e−22	0.00015614	−0.00012367
20	−1.874e−15	7.1702e−19	−1.4543e−18	−3.7102e−22	−2.7089e−05	0.00025874
30	1.6404e−15	6.179e−19	−1.1437e−18	4.1373e−22	−0.00010487	−0.00013463
40	−1.381e−15	7.9917e−19	8.8476e−19	5.411e−22	−7.8006e−05	−0.00011876
50	1.7321e−15	−8.077e−19	−1.2949e−18	−3.931e−22	−0.00014099	0.00019745
60	−1.8335e−15	5.9303e−19	−1.1492e−18	3.7938e−22	0.00037311	−0.00018313
70	1.3993e−15	−8.4851e−19	1.7189e−18	−4.5698e−22	−0.00013835	8.841e−05
60	−1.5119e−15	6.3197e−19	1.2181e−18	5.4712e−22	−0.00010291	0.00016454
80	−1.6507e−15	7.8559e−19	1.4136e−18	5.7181e−22	−8.6107e−05	−0.000233
100	1.8193e−15	−7.5823e−19	−8.3248e−19	−6.2308e−22	0.00033237	0.00017331
120	1.8944e−15	−6.099e−19	1.4121e−18	4.4111e−22	−0.00016013	−0.00020051
140	1.7885e−15	7.9654e−19	1.2573e−18	4.2805e−22	−6.0739e−05	0.00021778
160	1.6329e−15	6.844e−19	−9.7035e−19	−4.765e−22	−0.00014521	−0.00025557
180	1.8544e−15	−7.1333e−19	1.0734e−18	−4.3413e−22	0.00034517	7.789e−05
200	1.5281e−15	8.7761e−19	−1.1473e−18	−6.9381e−22	−0.00021773	6.441e−05
220	−1.8403e−15	8.8858e−19	−1.2264e−18	−4.3322e−22	−0.00010868	5.7033e−05
240	1.3295e−15	7.5088e−19	1.0645e−18	−6.6125e−22	0.00015798	0.00012506
260	−1.4762e−15	6.8189e−19	1.0632e−18	4.212e−22	0.00019816	−0.00030953
280	−1.6359e−15	−6.1216e−19	−1.3139e−18	−5.0504e−22	−0.00029779	0.00010249
300	1.6576e−15	6.6286e−19	1.2966e−18	4.6399e−22	0.00020223	1.5144e−08
320	−1.7016e−15	−7.0218e−19	1.9433e−18	4.4932e−22	−0.00019964	4.7633e−05
340	−1.8177e−15	−6.9387e−19	−9.3109e−19	−4.7001e−22	0.00029937	4.6168e−05
360	−1.6465e−15	−7.3474e−19	1.3297e−18	−4.9018e−22	−0.00013395	6.1066e−05
380	−1.6094e−15	−7.3569e−19	1.088e−18	6.4093e−22	−5.6145e−05	−6.9071e−05
400	1.8205e−15	−7.0419e−19	1.3262e−18	5.3603e−22	−5.6473e−05	−8.3593e−05
420	−1.703e−15	−6.8039e−19	−1.3781e−18	3.7424e−22	−7.5682e−05	6.4413e−06
440	1.7092e−15	−8.0956e−19	1.6653e−18	−4.9875e−22	−1.522e−07	−0.00011816
460	−1.3593e−15	7.4904e−19	1.3845e−18	5.3811e−22	0.00031088	0.00011756
480	−1.7123e−15	6.5486e−19	1.1146e−18	−4.8691e−22	−0.00015196	6.5644e−06
500	1.6999e−15	6.8149e−19	1.2265e−18	−4.8087e−22	−0.00017144	0.00011237

Abbreviations: GWO, gray wolf optimization; MGWO, mean gray wolf optimization.

and (5) Heart. These problems have been solved using modified variant, and results obtained have been compared with several meta-heuristics. Different types of parameters have been used for running code of several meta-heuristics.

Table 2. Best solution obtained by GWO and MGWO on 500 generations.

ITERATIONS	PROBLEM NAME					
	$f_4(x)$		$f_5(x)$		$f_6(x)$	
	GWO	MGWO	GWO	MGWO	GWO	MGWO
10	−3.5971e−07	−1.2152e−08	0.67847	0.82644	−0.49897	−0.49874
20	3.6824e−07	1.1563e−08	0.458	0.68259	−0.50149	−0.50111
30	3.6923e−07	−1.1926e−08	0.20109	0.46306	−0.49817	−0.50052
40	−3.5931e−07	−1.2112e−08	−0.00027583	0.20532	−0.50053	−0.50137
50	−3.667e−07	−1.2199e−08	0.007573	−0.00072299	−0.49961	−0.49994
60	−3.5238e−07	1.1977e−08	−0.00050232	−0.00016668	−0.49997	−0.50213
70	−3.6787e−07	−2.3984e−09	0.0071152	0.011135	0.0032181	−0.49861
60	−3.6908e−07	−1.1464e−08	−1.4391e−05	−5.46e−05	0.0048451	−0.5023
80	−3.6939e−07	1.1874e−08	0.0071846	0.00019418	−0.50109	6.4864e−05
100	3.6842e−07	−1.2112e−08	0.0003128	−8.8968e−05	−0.0032703	0.00011448
120	3.6789e−07	−1.2021e−08	0.0043621	−0.0026679	−0.00052754	−0.50201
140	3.6872e−07	1.2213e−08	−0.00013115	0.0010269	−0.50025	−0.50006
160	−3.6916e−07	1.0023e−08	−0.00047276	0.0011014	−0.50169	−0.49955
180	−3.6928e−07	−1.2188e−08	−0.0018889	−7.6101e−05	−0.50035	−0.49981
200	3.4979e−07	1.213e−08	−0.00071533	−2.7018e−06	−0.49867	−0.50089
220	3.6801e−07	−1.1688e−08	0.0046027	−0.00092194	−0.49925	−0.50265
240	2.1383e−07	−1.2064e−08	0.0093635	0.0062233	0.0040758	4.9444e−06
260	−3.6483e−07	−1.2205e−08	0.01084	0.0030486	−0.50119	0.00053093
280	3.5202e−07	1.219e−08	0.0039531	−2.2467e−05	−0.49935	−0.4998
300	3.6114e−07	1.2123e−08	−0.00016604	−3.0738e−05	−0.5001	−0.50144
320	3.6799e−07	−1.2124e−08	0.0098741	7.4604e−05	−0.50029	−0.49928
340	−3.1058e−07	7.6152e−09	0.00021983	7.0245e−07	−0.49867	−0.50309
360	−3.6749e−07	1.2182e−08	8.5684e−05	−4.2109e−05	−0.50054	−0.49815
380	3.6227e−07	1.204e−08	0.00018917	−1.2425e−05	−0.49908	−0.50015
400	−3.5667e−07	−1.2101e−08	0.012404	3.7472e−06	−0.49769	−0.50063
420	3.6893e−07	1.195e−08	−0.0014015	9.5091e−05	−0.50358	−0.5018
440	−3.3098e−07	1.2011e−08	0.0018281	0.0040767	−0.50143	−0.49828
460	2.708e−08	−1.1703e−08	0.0010668	0.00013148	−0.49762	−0.5041
480	3.5648e−07	−2.9391e−09	0.0021144	−0.00020263	−0.50137	−0.4993
500	−3.6509e−07	−1.2165e−08	0.00077988	−0.0023799	−0.50122	0.0010644

Abbreviations: GWO, gray wolf optimization; MGWO, mean gray wolf optimization.

These parameters are listed in Appendix 1, Table E. The performance of these algorithms has been compared in terms of average, standard deviation, classification rate, and convergence rate of all the variants. All these data set problems have been discussed step-by-step in the following sections.

Table 3. Best solution obtained by GWO and MGWO on 500 generations.

ITERATIONS	PROBLEM NAME					
	$f_7(x)$		$f_8(x)$		$f_9(x)$	
	GWO	MGWO	GWO	MGWO	GWO	MGWO
10	0.029423	0.077502	418.7634	−298.9024	−4.8178e−09	1.4727e−09
20	0.061542	−0.0034678	7.890677	422.5337	5.2078e−09	−2.0444e−09
30	−0.0052652	−0.010899	200.8017	26.25219	1.1065e−08	3.9829e−10
40	−0.10774	0.025158	−500	−3.174072	1.4494e−08	−7.4365e−10
50	−0.027983	0.021111	−125.3959	421.6416	5.4837e−09	−3.9091e−09
60	0.082137	−0.037085	−7.221072	199.6524	1.0833e−08	3.9096e−09
70	−0.0781	0.0064255	420.9101	420.7892	−9.9361e−09	−8.9296e−10
60	0.055795	0.061683	419.77	−499.9521	9.267e−09	−4.7717e−09
80	−0.039566	−0.024718	−128.3696	−299.3261	−9.953e−09	−4.8101e−09
100	0.00024457	−0.020844	421.3294	−109.8219	−8.6132e−09	2.5863e−09
120	0.0012085	−0.02979	200.3792	33.23177	6.9308e−09	1.2765e−09
140	−0.014384	−0.048311	63.42083	199.2015	−9.9047e−09	−2.2088e−09
160	−0.039246	−0.040789	−300.1527	418.7271	−1.0436e−08	4.3848e−09
180	0.018483	−0.013968	13.31976	−302.8847	−1.2663e−08	5.5603e−09
200	−0.00071862	0.049291	−7.574163	−500	−8.3694e−09	2.6931e−09
220	−0.0017537	−0.055844	−301.9192	420.3039	8.1988e−09	3.9673e−09
240	0.061328	0.0054619	421.2066	28.46334	−7.1084e−09	−4.0706e−09
260	−0.010229	0.012456	5.960788	70.37576	6.7086e−09	1.0293e−08
280	0.098731	0.01818	−127.7605	−69.08079	8.1177e−09	−5.977e−09
300	0.020226	0.0040015	60.80246	207.1159	−1.3378e−08	−2.3381e−09
320	−0.013394	0.0064534	−304.0941	−302.4255	−7.3408e−09	6.2421e−09
340	0.0081719	0.0264	201.5527	423.1574	−6.6327e−09	5.2722e−10
360	0.0061936	−0.002891	−500	422.0409	6.5257e−09	1.1726e−09
380	−0.0065873	0.0041979	205.8273	−305.4574	−1.4001e−08	−4.8026e−09
400	−0.050018	−0.0030437	−65.5463	−16.43798	−7.3359e−09	4.0501e−09
420	−0.04869	−0.0062353	−128.2352	−120.6897	6.4397e−09	5.0764e−10
440	0.0005496	0.0033052	−303.8939	206.3322	−8.5935e−09	7.1361e−09
460	−0.021359	0.0080088	201.04	206.2603	−9.9156e−09	−3.2966e−09
480	0.0056483	−0.0077328	−301.9652	70.21548	8.724e−09	7.3021e−10
500	0.0011928	−0.007108	204.9987	−14.75609	−7.6446e−09	8.1855e−09

Abbreviations: GWO, gray wolf optimization; MGWO, mean gray wolf optimization.

Table 4. Best solution obtained by GWO and MGWO on 500 generations.

ITERATIONS	PROBLEM NAME					
	$f_{10}(x)$		$f_{11}(x)$		$f_{12}(x)$	
	GWO	MGWO	GWO	MGWO	GWO	MGWO
10	−1.8167e−14	8.9423e−15	−0.019039	5.3262e−09	−0.9996	−0.99932
20	−1.2321e−14	1.4626e−14	4.4382	2.3876e−09	0.024447	−0.0057794
30	1.4065e−14	6.8127e−15	−5.434	−1.5246e−08	−1.0013	−0.99781
40	−2.0385e−14	−1.3312e−14	0.025776	−6.5272e−09	0.018335	−0.93331
50	−2.8643e−14	4.7171e−15	0.010692	−1.6958e−08	−1.0025	−0.99452
60	2.6769e−14	6.2785e−15	0.081679	−7.0927e−09	−0.0012782	−0.096163
70	−1.9476e−14	9.5679e−15	0.038692	9.7113e−09	−1.0022	−0.99453
60	2.5826e−14	−7.632e−15	−0.0061243	−2.5782e−08	−0.73552	−0.99316
80	−1.9214e−14	8.0977e−15	−0.0018057	2.5804e−08	−1.0026	0.0017965
100	−2.0926e−14	5.941e−15	−0.070671	−5.0612e−09	−0.2654	−1.0008
120	−1.9016e−14	1.5692e−14	0.0023691	−1.8281e−08	−1.0066	−1.0225
140	−1.3046e−14	−1.1063e−14	−0.0066212	3.2331e−08	−0.098088	−1.0127
160	−2.2857e−14	1.9852e−14	0.063109	1.4704e−08	−1.0026	−1.0014
180	−2.0175e−14	−1.0776e−14	−0.10185	−1.7661e−09	−0.07441	7.4273e−05
200	−2.6876e−14	−9.1138e−15	0.002799	−5.9329e−09	−0.99391	−1.0002
220	−2.3615e−14	8.6438e−15	−0.0075882	−3.5188e−08	−0.94804	0.00012775
240	−2.0723e−14	−1.1247e−14	0.042538	−1.427e−09	−0.9851	−0.99864
260	−2.8553e−14	−4.0693e−15	0.0059001	−3.9786e−08	−1.0715	−0.9958
280	−2.5586e−14	1.2396e−14	−0.097501	2.9202e−09	−1.0032	−0.0021754
300	−1.1253e−14	1.6556e−14	−0.011378	4.7113e−08	0.022496	−0.99742
320	1.2973e−14	1.1214e−14	0.026999	4.388e−08	−1.001	−0.9599
340	−7.5563e−15	−1.0387e−14	−0.0060048	−4.2643e−08	−0.94636	−1.0336
360	−2.9824e−14	9.6746e−15	−0.079693	−1.7512e−08	−0.99739	−1.0102
380	−1.6148e−14	−1.1281e−14	−0.099893	1.8579e−08	−0.99301	−1.009
400	−1.5006e−14	−3.6308e−15	−0.054028	4.2624e−08	−1.0046	−0.92644
420	−2.3638e−14	−5.7501e−15	0.051612	−3.9184e−08	−0.017619	−0.99712
440	−1.9547e−14	6.1612e−15	0.010453	−2.9982e−08	−1.0005	−0.001263
460	1.8011e−14	2.5528e−15	−0.08148	3.6022e−08	−0.9995	−0.99998
480	−2.0054e−14	−7.0224e−15	0.11215	−1.1691e−08	0.0091993	−0.97862
500	−1.7128e−14	−5.6568e−15	−0.057493	3.9177e−08	−0.995	−0.98908

Abbreviations: GWO, gray wolf optimization; MGWO, mean gray wolf optimization.

Table 5. Best solution obtained by GWO and MGWO on 500 generations.

ITERATIONS	PROBLEM NAME	
	$f_{13}(x)$	
	GWO	MGWO
10	1.0017	0.33972
20	0.037012	−0.042207
30	0.99873	0.67138
40	0.00048659	−0.017276
50	0.013208	0.99471
60	−0.011551	1.0415
70	0.9999	0.0019544
60	0.85541	0.998
80	0.013833	−0.00011723
100	0.017527	0.99887
120	0.99978	0.80284
140	1.0191	1.0166
160	0.87957	0.0014587
180	0.01684	0.99995
200	0.66956	0.8789
220	1.0006	1.0346
240	−0.11366	−0.029245
260	0.0082416	0.66927
280	0.019867	−0.038115
300	0.035205	0.0001905
320	0.66962	0.998
340	0.99764	1.0259
360	1.0108	0.76088
380	1.0675	−0.0019594
400	0.98336	0.99662
420	0.051414	0.86818
440	1.0004	0.022489
460	1.1023	1
480	0.95524	0.016364
500	0.93389	0.99916

Abbreviations: GWO, gray wolf optimization; MGWO, mean gray wolf optimization.

XOR data set

This data set has 3 attributes (input), 8 training samples, 8 test samples, 2 classes, and 1 output (Appendix 1, Table D[18]). The experimental numerical results obtained through MGWO, GWO, PSO, GA, ACO, ES, and PBIL for this data set are shown in Table 19, and convergence performance of GWO and MGWO variant is shown in Figure 28.

Table 6. Best solution obtained by GWO and MGWO on 500 generations.

ITERATIONS	PROBLEM NAME	
	$f_{14}(x)$	
	GWO	MGWO
200	−0.00565596	−31.9871
500	−31.9651	−31.9763
	$f_{15}(x)$	
100	−0.389	0.19072
200	−5	0.27247
400	1.2998	0.15591
500	5	0.17047

Abbreviations: GWO, gray wolf optimization; MGWO, mean gray wolf optimization.

It is clear from Table 19 that MGWO, GWO, and GA variants give the better quality of statistical results as compared with other meta-heuristics. The results obtained with MGWO, GWO, and GA variants indicate that it has the highest ability to avoid the local optima and is considerably superior to other variants such as PSO, GA, ACO, ES, and PBIL.

The performance of these variants has also been compared in terms of average, standard deviation classification rate (Table 19), and convergence rate (Figure 28). The low average and standard deviation show the superior local optima avoidance of the variant. On the basis of obtained results, we have concluded that newly modified variant MGWO gives highly competitive results as compared with other existing variants, and convergence graph shows that MGWO gives better solutions rather than GWO variant.

Balloon data set

It is clear from Appendix 1, Table D[18] that this data set has 4 attributes, 16 training samples, 16 test samples, and 2 classes. The statistical numerical and convergence results of the variants on this data set are shown in Table 20 and Figure 29.

Here, we are comparing the accuracy of the algorithms in terms of average, standard deviation, classification rate, and convergence rate of the algorithm. First, we observe that all the variants give similar classification rate. Second, on the basis of statistical and convergence results, we observe that modified variant gives highly competitive solutions as compared with other variants such as GWO, PSO, GA, ACO, ES,

Table 7. Best solution obtained by GWO and MGWO on 500 generations.

ITERATIONS	PROBLEM NAME					
	$f_{16}(x)$		$f_{17}(x)$		$f_{18}(x)$	
	GWO	MGWO	GWO	MGWO	GWO	MGWO
200	−0.089882	0.089746	3.1415	3.1413	0.00016791	2.0193e−05
500	0.7126	−0.71261	2.2746	2.2757	−0.99995	−0.99997

Abbreviations: GWO, gray wolf optimization; MGWO, mean gray wolf optimization.

Table 8. Best solution obtained by GWO and MGWO on 500 generations.

ITERATIONS	PROBLEM NAME	
	GWO	MGWO
	$f_{19}(x)$	
100	0.046099	0.058244
300	0.55508	0.55606
500	0.85294	0.85251
	$f_{20}(x)$	
100	0.20152	0.20171
150	0.14643	0.1467
200	0.47763	0.47735
300	0.27539	0.27526
400	0.31165	0.31187
500	0.65724	0.65705

Abbreviations: GWO, gray wolf optimization; MGWO, mean gray wolf optimization.

Table 9. Best solution obtained by GWO and MGWO on 500 generations.

ITERATIONS	PROBLEM NAME					
	$f_{21}(x)$		$f_{22}(x)$		$f_{23}(x)$	
	GWO	MGWO	GWO	MGWO	GWO	MGWO
100	4.0011	4.0025	4.0011	3.9995	4.0029	3.9996
200	4.0058	3.9959	3.9978	3.9979	4.0026	4.0004
400	3.9982	4.0021	4.0036	3.9987	4.0023	3.9978
500	3.9979	3.9971	4.0009	4.0018	3.998	3.9989

Abbreviations: GWO, gray wolf optimization; MGWO, mean gray wolf optimization.

and PBIL algorithms. These convergence results are plotted in Figure 29.

Breast cancer data set

This data set has 9 attributes, 599 training samples, 100 test samples, and 2 classes (Appendix 1, Table D).[18] All problems have been run 10 times using this data set. The numerical results are shown in Table 21. The convergence performance on this data set is plotted in Figure 30.

We have observed that the modified variant (MGWO) gives 99.11% classification rate and better convergence solutions (Figure 30) that are superior to other meta-heuristics.

Table 10. Results of unimodal benchmark functions (maximum and minimum).

PROBLEM	PSO		GWO		MGWO	
	MINIMUM	MAXIMUM	MINIMUM	MAXIMUM	MINIMUM	MAXIMUM
1	1.6301e−05	5.8559e+04	8.3933e−29	6.7832e+04	1.5833e−35	7.2334e+04
2	0.0244	2.2522e+12	3.7699e−17	4.0253e+12	1.4605e−20	8.5140e+11
3	96.0253	1.1455e+05	4.6658e−07	1.1711e+05	2.599e−07	1.2165e+05
4	1.2636	90.8299	3.6939e−07	90.6367	1.2213e−08	91.5194
5	26.7395	2.5613e+08	27.1234	2.2938e+08	26.2201	2.6366e+08
6	1.9802e−05	6.0740e+04	1.2585	6.9920e+04	1.2518	7.1406e+04
7	0.2130	89.4207	0.003646	105.9925	0.00057612	146.7004

Abbreviations: GWO, gray wolf optimization; MGWO, mean gray wolf optimization; PSO, particle swarm optimization.

Table 11. Results of multimodal benchmark functions (maximum and minimum).

PROBLEM	PSO		GWO		MGWO	
	MINIMUM	MAXIMUM	MINIMUM	MAXIMUM	MINIMUM	MAXIMUM
8	−3.6455e+03	−3.3735e+03	−5703.971	−2.5289e+03	−6023.026	−2.1300e+03
9	53.3912	446.0825	2.8422e−13	469.6513	0.0000	479.1380
10	0.0681	20.8251	8.2601e−14	20.5838	3.9968e−14	20.8448
11	0.0099	578.4494	0.015231	646.0588	0.0000	687.4643
12	1.4583e−06	6.9167e+08	0.05501	6.9928e+08	0.04462	7.4944e+08
13	0.0110	1.2165e+09	1.2322	1.2205e+09	1.206	9.7589e+08

Abbreviations: GWO, gray wolf optimization; MGWO, mean gray wolf optimization; PSO, particle swarm optimization.

Table 12. Results of fixed-dimension multimodal benchmark functions (maximum and minimum).

PROBLEM	PSO		GWO		MGWO	
	MINIMUM	MAXIMUM	MINIMUM	MAXIMUM	MINIMUM	MAXIMUM
14	12.6705	23.3017	2.9821	8.3608	0.9980	30.6623
15	0.0010	0.0431	0.020363	0.0786	0.00031732	0.2304
16	−1.0316	0.2656	−1.0316	0.2148	−1.0316	0.7792
17	0.3979	1.1355	0.39789	1.0081	0.39789	1.5247
18	3	62.4398	3	48.3077	3	534.8252
19	−3.8628	−3.7784	−3.8599	−3.3858	−3.8609	−2.9920
20	−3.3220	−1.3907	−3.3220	−1.3471	−3.3220	−0.9232
21	−10.1532	−0.4051	−10.1490	−0.4626	−10.1495	−0.2926
22	−10.4029	−0.4329	−10.4002	−0.6413	−10.4015	−0.3539
23	−10.5364	−1.2565	−10.5346	−1.3843	−10.5359	−0.6351

Abbreviations: GWO, gray wolf optimization; MGWO, mean gray wolf optimization; PSO, particle swarm optimization.

Table 13. Results of unimodal benchmark functions (mean and SD).

PROBLEM	PSO	GWO	MGWO	PSO	GWO	MGWO
	MEAN	MEAN	MEAN	SD	SD	SD
1	874.3217	710.7351	445.2002	5.6498e+03	5.0831e+03	4.2505e+03
2	4.5045e+09	8.0606e+09	1.7028e+09	1.0072e+11	1.8002e+11	3.8076e+10
3	3.3311e+03	2.5686e+03	2.1484e+03	1.3596e+04	1.0040e+04	9.6138e+03
4	5.1486	3.9130	3.3071	9.8558	14.2933	14.0919
5	1.1509e+06	1.4381e+06	9.9549e+05	1.4586e+07	1.4875e+07	1.4018e+07
6	859.7742	709.4915	421.1404	5.4499e+03	5.1678e+03	4.0498e+03
7	44.9171	0.7454	0.5288	38.6615	6.7275	7.2339

Abbreviations: GWO, gray wolf optimization; MGWO, mean gray wolf optimization; PSO, particle swarm optimization.

Table 14. Results of multimodal benchmark functions (mean and SD).

PROBLEM	PSO	GWO	MGWO	PSO	GWO	MGWO
	MEAN	MEAN	MEAN	SD	SD	SD
8	−3.4243e+03	−3.7809e+03	−3.6613e+03	101.8935	978.9551	1.0165e+03
9	209.5102	29.1202	11.7840	101.8575	78.5096	49.7529
10	4.2659	0.7215	0.4159	3.5620	3.0475	2.2324
11	43.0150	5.6194	3.5615	116.0241	41.2943	37.4242
12	2.8521e+06	3.0499e+06	2.9527e+06	3.7532e+07	3.6338e+07	3.8702e+07
13	5.2381e+06	7.7264e+06	2.7124e+06	6.7511e+06	7.4601e+07	4.4432e+07

Abbreviations: GWO, gray wolf optimization; MGWO, mean gray wolf optimization; PSO, particle swarm optimization.

Table 15. Results of fixed-dimension multimodal benchmark functions (mean and SD).

PROBLEM	PSO	GWO	MGWO	PSO	GWO	MGWO
	MEAN	MEAN	MEAN	SD	SD	SD
14	13.8873	1.8232	3.0669	0.6627	1.4463	1.8076
15	0.0014	0.0206	0.0013	0.0035	0.0030	0.0112
16	−1.0194	−1.0287	−1.0273	0.1064	0.0563	0.0817
17	0.4014	0.4019	0.4048	0.0438	0.0342	0.0561
18	3.1443	3.1612	4.3807	2.6660	2.2937	23.9465
19	−3.8612	−3.8465	−3.8486	0.0076	0.0339	0.0521
20	−3.2481	−3.2074	−3.2436	0.1567	0.1729	0.1890
21	−8.9448	−7.1551	−7.1596	2.4085	2.6671	2.2778
22	−6.8767	−7.8294	−7.8380	3.8997	1.8218	2.3014
23	−9.5800	−7.8107	−7.2332	2.0122	2.0366	2.1071

Table 16. Time-consuming results of unimodal benchmark functions.

PROBLEM	PSO			GWO			MGWO		
	TIC TOC	CPUTIME	CLOCK	TIC TOC	CPUTIME	CLOCK	TIC TOC	CPUTIME	CLOCK
1	1.01346	0.0156001	1.013	1.01214	0.0624004	1.012	1.00432	0.0001	1.004
2	1.0116	0.001	1.014	1.01281	0.0156001	1.014	1.00669	0.001	1.014
3	1.00695	0.001	1.014	1.01441	0.001	1.014	1.0132	0.001	1.014
4	1.00365	0.00001	1.016	1.01375	0.00001	1.016	0.999516	0.00001	1.014
5	1.0091	0.0416001	1.014	1.00401	0.0312002	1.014	1.00296	0.00012	1.014
6	1.00479	0.0156001	1.014	1.01376	0.0156001	1.014	1.00578	0.0146001	1.014
7	1.006	0.0666005	1.014	1.00101	0.0468003	1.016	1.01178	0.0156001	1.014

Abbreviations: GWO, gray wolf optimization; MGWO, mean gray wolf optimization; PSO, particle swarm optimization.

Table 17. Time-consuming results of multimodal benchmark functions.

PROBLEM	PSO			GWO			MGWO		
	TIC TOC	CPUTIME	CLOCK	TIC TOC	CPUTIME	CLOCK	TIC TOC	CPUTIME	CLOCK
8	1.01377	0.0156001	1.014	1.00292	0.0156001	1.003	1.00299	0.0109001	1.003
9	1.01261	0.109201	1.013	1.00716	0.0312002	1.007	1.00982	0.00001	1.009
10	1.01401	0.0816002	1.014	1.01014	0.0936006	1.01	1.00116	0.0312002	1.001
11	1.00841	0.0001	1.008	1.0059	0.0001	1.006	1.00169	0.0001	1.001
12	1.01046	0.0312002	1.004	1.00295	0.0312002	1.003	1.00916	0.0301002	1.003
13	1.00584	0.421203	1.006	1.0052	0.0156001	1.005	1.01407	0.0312002	1.006

Abbreviations: GWO, gray wolf optimization; MGWO, mean gray wolf optimization; PSO, particle swarm optimization.

Table 18. Time-consuming results of fixed-dimension multimodal benchmark functions.

PROBLEM	PSO			GWO			MGWO		
	TIC TOC	CPUTIME	CLOCK	TIC TOC	CPUTIME	CLOCK	TIC TOC	CPUTIME	CLOCK
14	1.01128	**0.0251**	1.011	1.00888	0.0312002	1.009	1.00553	0.0001	1.005
15	1.01445	**0.8953011**	1.014	1.00673	0.0624004	1.007	1.01143	0.0156001	1.011
16	1.01589	**0.00081**	1.016	0.998874	0.0312002	1.009	1.01114	0.00001	1.007
17	1.01227	**0.9158702**	1.014	1.01169	0.0156001	1.014	1.00204	0.0156001	1.014
18	1.01162	**0.02305**	1.014	1.00446	0.0156001	1.014	1.00321	0.00001	1.014
19	1.00426	**0.4167091**	1.014	1.00695	0.0156001	1.014	1.0101	0.0156001	1.014
20	1.01494	**0.39031**	1.015	1.01401	0.0312002	1.014	1.00902	0.00001	1.014
21	1.00002	**0.00001**	1.014	1.00717	0.0156001	1.004	1.01156	0.00001	1.003
22	0.996519	**0.51167**	0.999	1.00192	0.0312002	1.014	1.0097	**0.02489**	1.014
23	1.01091	**0.0436013**	1.014	1.00325	0.0312002	1.014	1.01417	0.0312002	1.014

Abbreviations: GWO, gray wolf optimization; MGWO, mean gray wolf optimization; PSO, particle swarm optimization.
Bold values highlight the results of proposed variant.

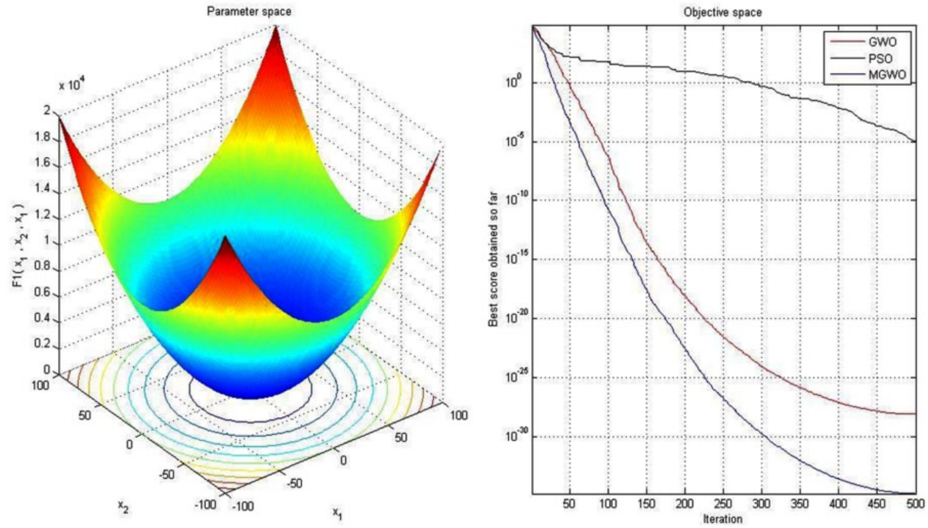

Figure 5. Convergence graph of unimodal benchmark function (F_1). GWO indicates gray wolf optimization; MGWO, mean gray wolf optimization; PSO, particle swarm optimization.

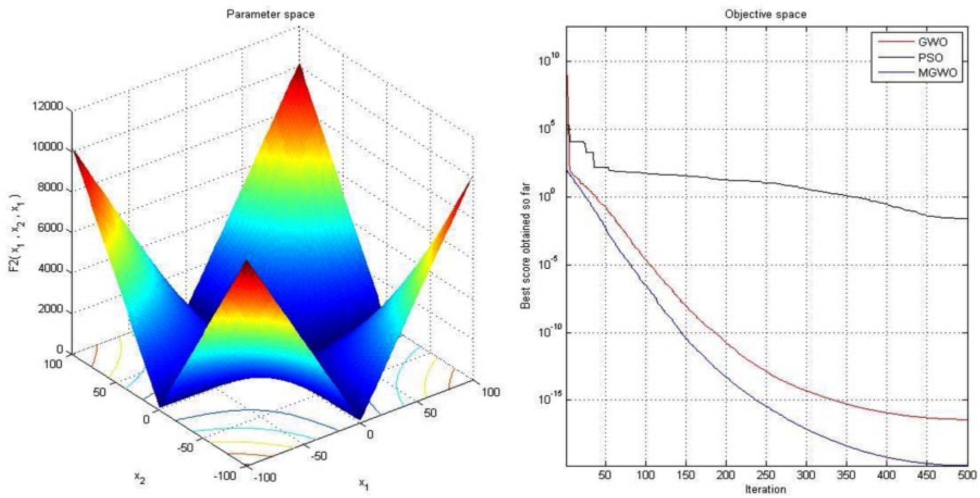

Figure 6. Convergence graph of unimodal benchmark function (F_2). GWO indicates gray wolf optimization; MGWO, mean gray wolf optimization; PSO, particle swarm optimization.

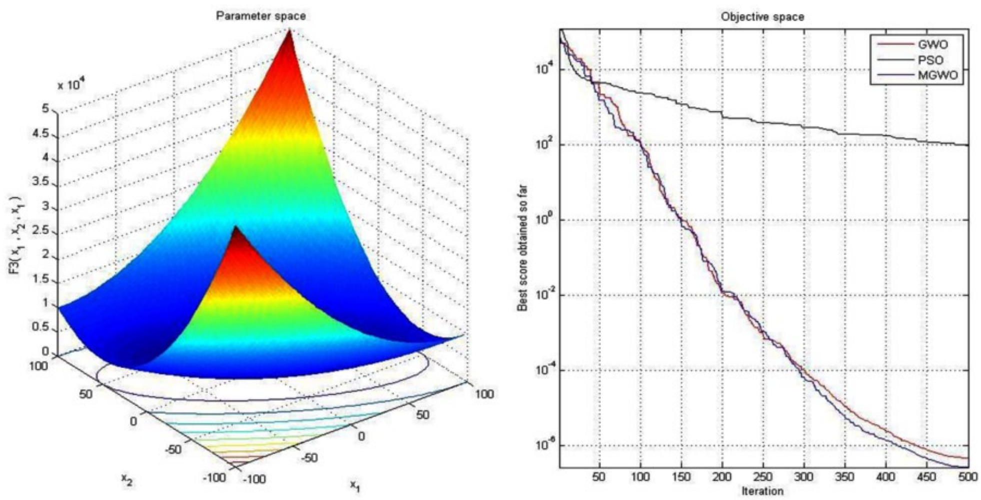

Figure 7. Convergence graph of unimodal benchmark function (F_3). GWO indicates gray wolf optimization; MGWO, mean gray wolf optimization; PSO, particle swarm optimization.

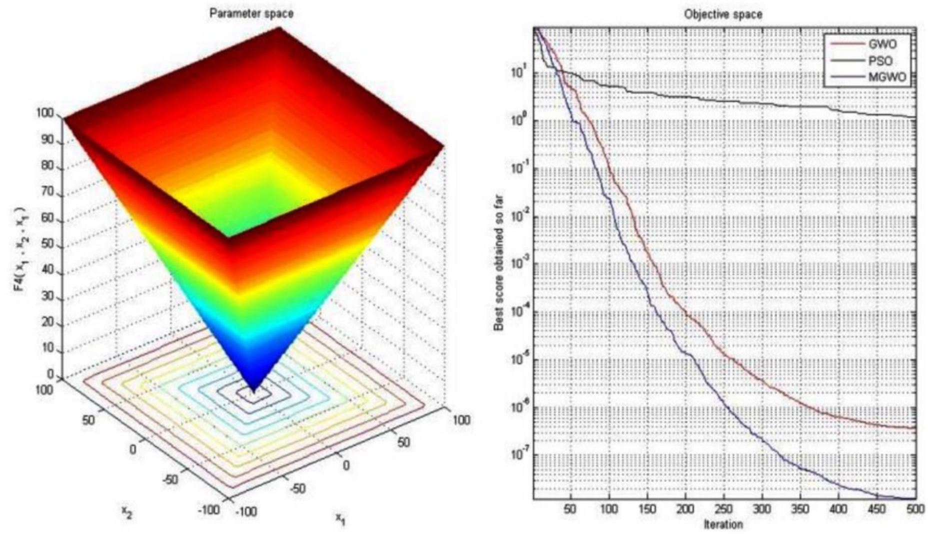

Figure 8. Convergence graph of unimodal benchmark function (F_4). GWO indicates gray wolf optimization; MGWO, mean gray wolf optimization; PSO, particle swarm optimization.

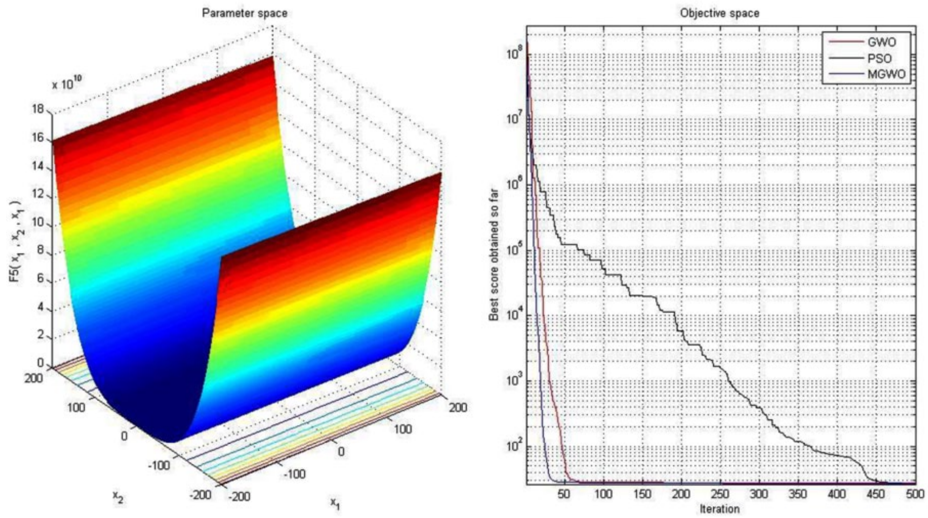

Figure 9. Convergence graph of unimodal benchmark function (F_5). GWO indicates gray wolf optimization; MGWO, mean gray wolf optimization; PSO, particle swarm optimization.

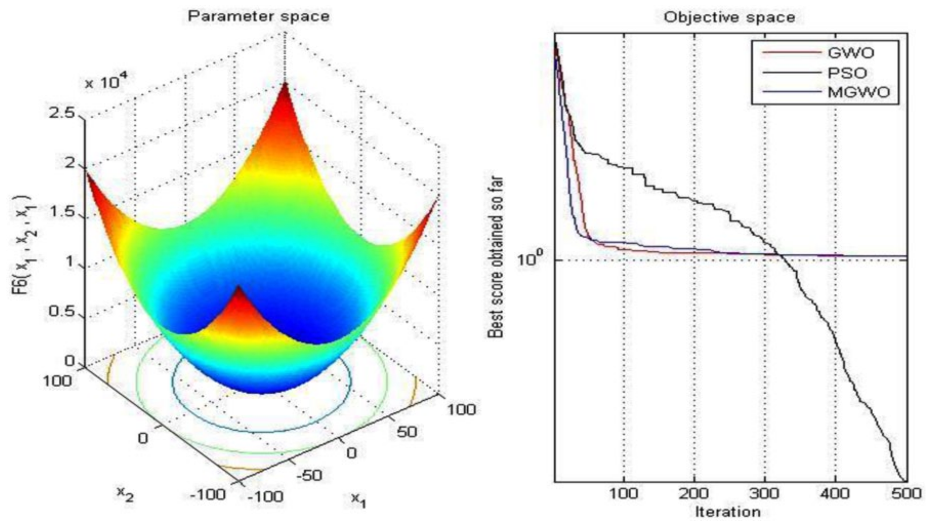

Figure 10. Convergence graph of unimodal benchmark function (F_6). GWO indicates gray wolf optimization; MGWO, mean gray wolf optimization; PSO, particle swarm optimization.

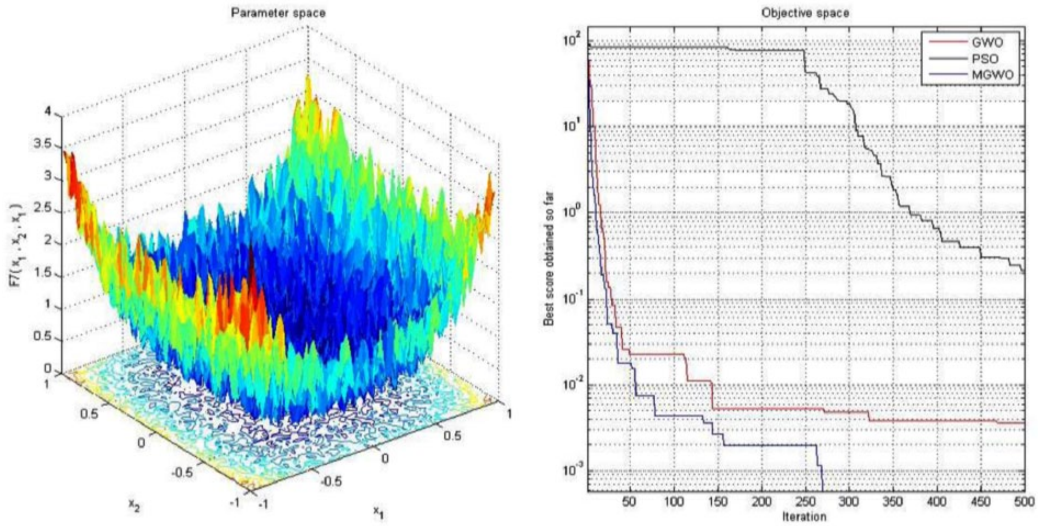

Figure 11. Convergence graph of unimodal benchmark function (F_7). GWO indicates gray wolf optimization; MGWO, mean gray wolf optimization; PSO, particle swarm optimization.

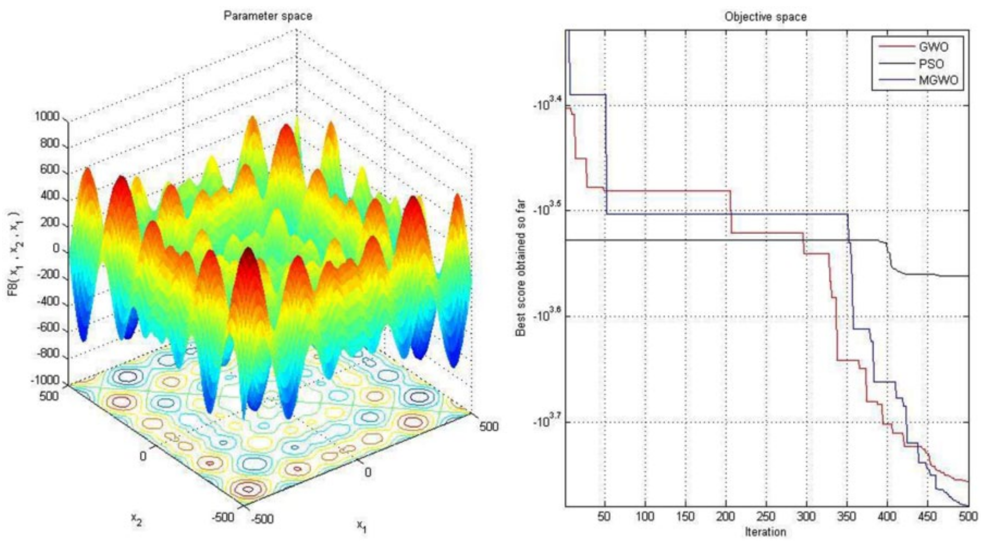

Figure 12. Convergence graph of multimodal benchmark function (F_8). GWO indicates gray wolf optimization; MGWO, mean gray wolf optimization; PSO, particle swarm optimization.

Iris data set

This data set is another well-known testing data set in the text. It consists of 4 attributes, 150 training samples, 150 test samples, and 3 classes as represented in Appendix 1, Table D.[18] The convergence performance of MGWO, GWO, PSO, GA, ACO, ES, and PBIL variants is plotted in Figure 31. The numerical results are shown in Table 22.

We have observed that these variants give the classification rate as MGWO (91.334%), GWO (91.333%), PSO (37.33%), GA (89.33%), ACO (32.66%), ES (46.66%), and PBIL (86.66%), respectively. The modified variant presents the better classification rate as compared with other variants.

The results confirm that MGWO algorithm has better local optima accuracy and avoidance simultaneously.

Heart data set

The heart data set is really one of the most popular data sets in the text. This data set has 22 attributes, 80 training samples, 187 testing samples, and 2 classes, respectively, and these data sets are reported in Appendix 1, Table D.[18] The results of the training these variants are shown in Table 23, and the convergence performance of MGWO and GWO is plotted in Figure 32. The low average and standard deviation show the superior local optima avoidance of the variant.

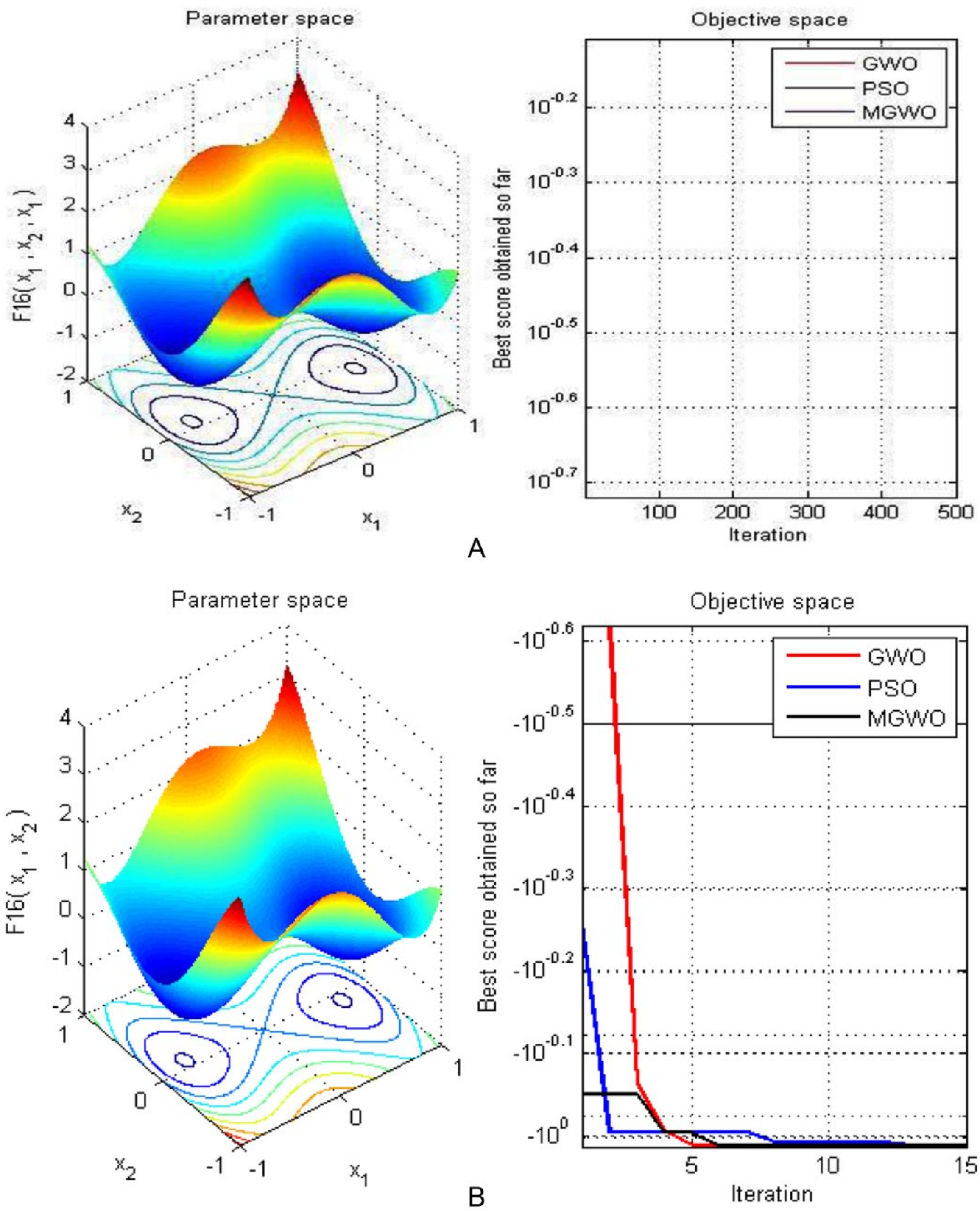

Figure 13. (a) Convergence graph of multimodal benchmark function (F_9) and (b) convergence graph of multimodal benchmark function (F_9) from 0 to 15 iterations. GWO indicates gray wolf optimization; MGWO, mean gray wolf optimization; PSO, particle swarm optimization.

The results of Table 23 reveal that MGWO has the best performance in this data set in terms of improved mean squared error, classification rate, and convergence as compared with other meta-heuristics.

Figure 32 shows that MGWO variant gives better quality of convergence solutions and outperforms GWO variant.

Conclusions

This article proposes a modified variant of GWO, namely, MGWO, inspired by the hunting behavior of gray wolves in nature. A statistical mean is used to balance the exploitation and exploration in the search space over the route of generations. The results reveal that the newly modified variant

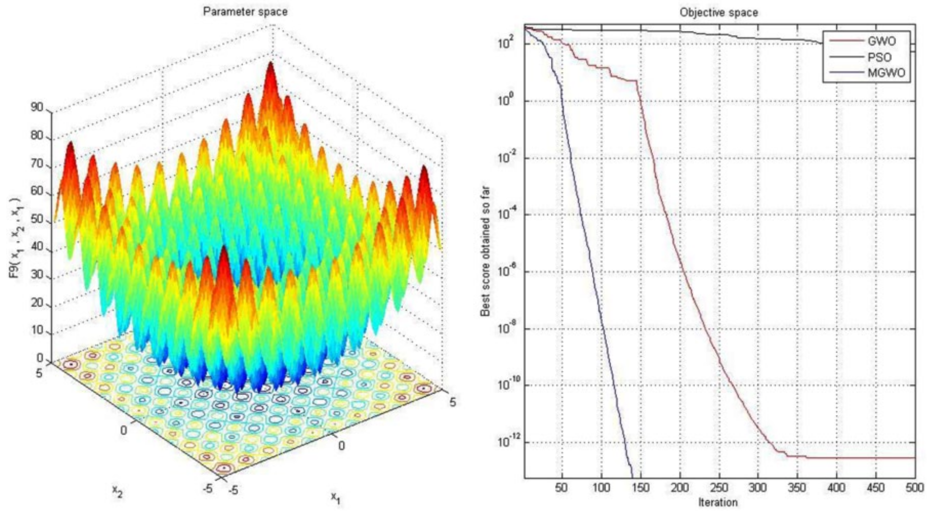

Figure 14. Convergence graph of multimodal benchmark function (F_{10}). GWO indicates gray wolf optimization; MGWO, mean gray wolf optimization; PSO, particle swarm optimization.

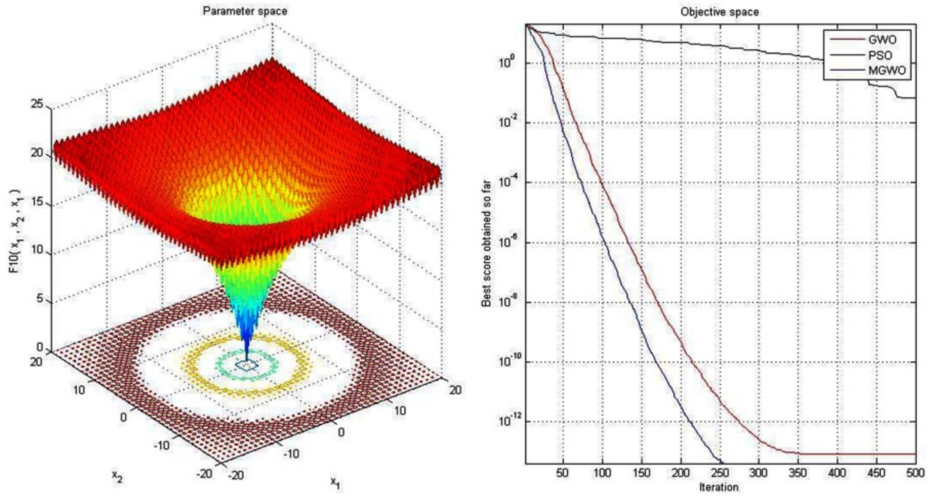

Figure 15. Convergence graph of multimodal benchmark function (F_{11}). GWO indicates gray wolf optimization; MGWO, mean gray wolf optimization; PSO, particle swarm optimization.

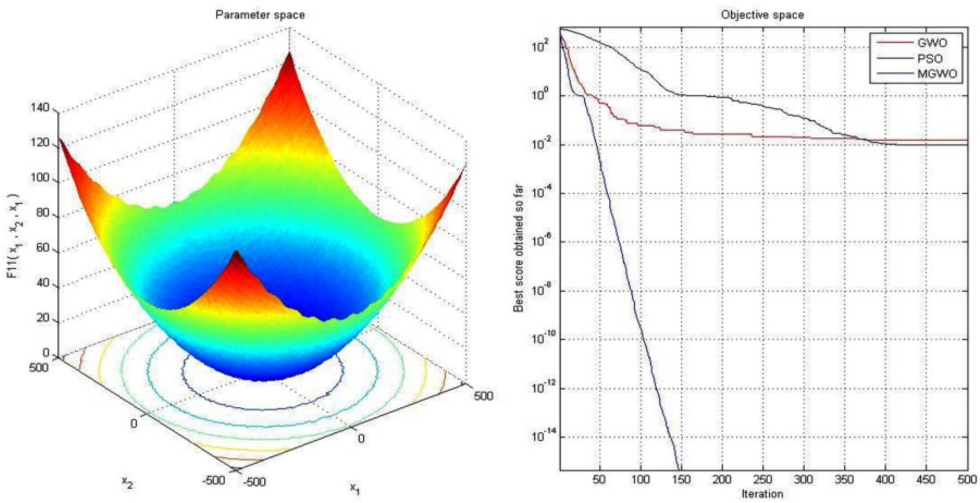

Figure 16. Convergence graph of multimodal benchmark function (F_{12}). GWO indicates gray wolf optimization; MGWO, mean gray wolf optimization; PSO, particle swarm optimization.

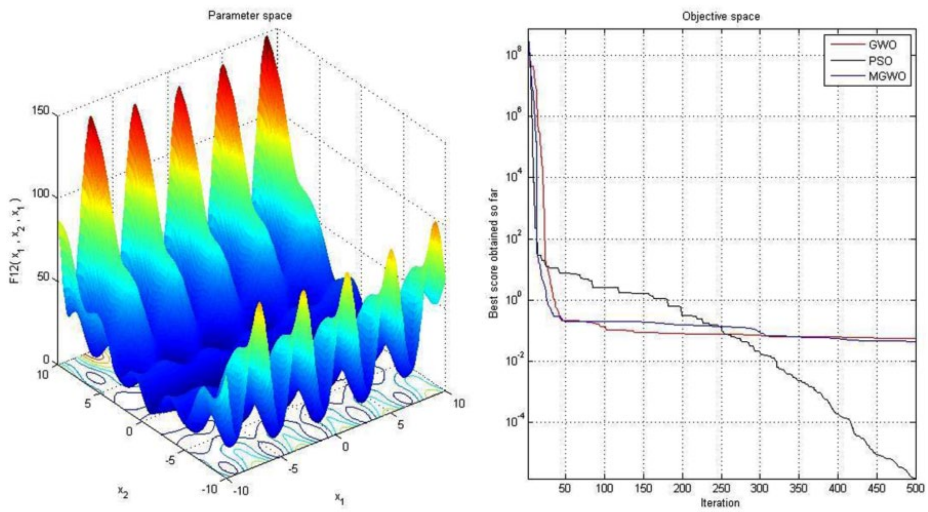

Figure 17. Convergence graph of multimodal benchmark function (F_{13}). GWO indicates gray wolf optimization; MGWO, mean gray wolf optimization; PSO, particle swarm optimization.

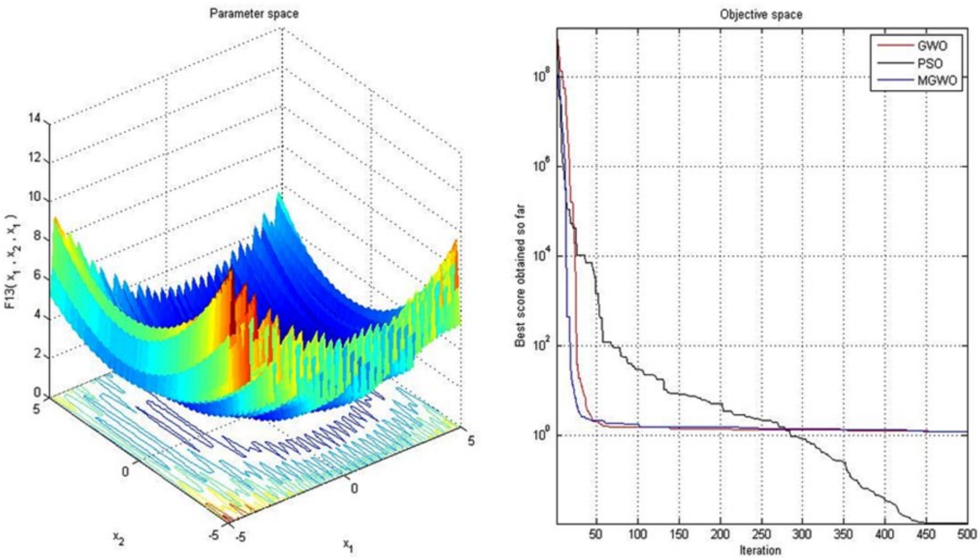

Figure 18. Convergence graph of fixed-dimension multimodal benchmark function (F_{14}). GWO indicates gray wolf optimization; MGWO, mean gray wolf optimization; PSO, particle swarm optimization.

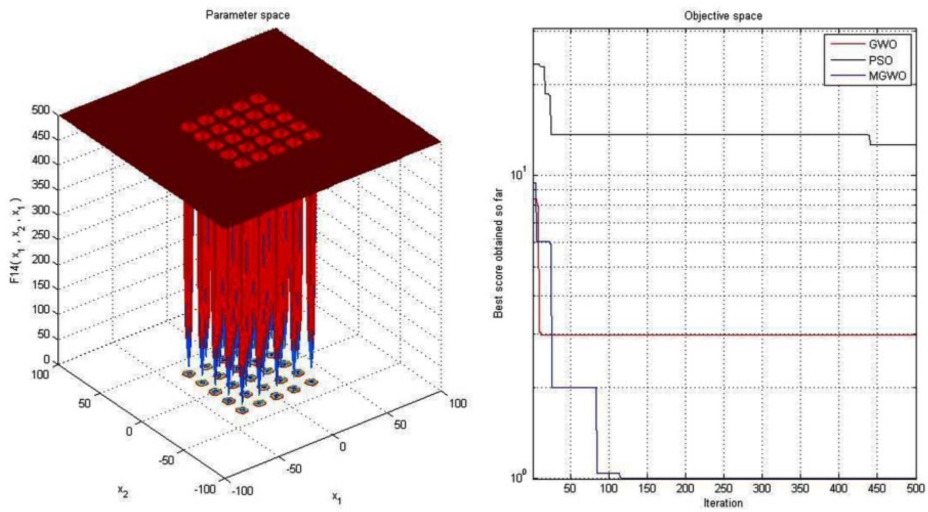

Figure 19. Convergence graph of fixed-dimension multimodal benchmark function (F_{15}). GWO indicates gray wolf optimization; MGWO, mean gray wolf optimization; PSO, particle swarm optimization.

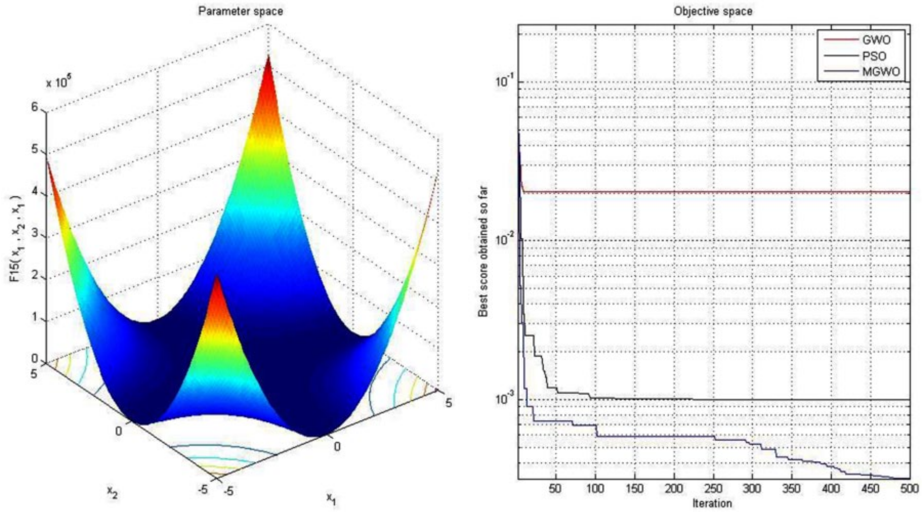

Figure 20. Convergence graph of fixed-dimension multimodal benchmark function (F_{16}). GWO indicates gray wolf optimization; MGWO, mean gray wolf optimization; PSO, particle swarm optimization.

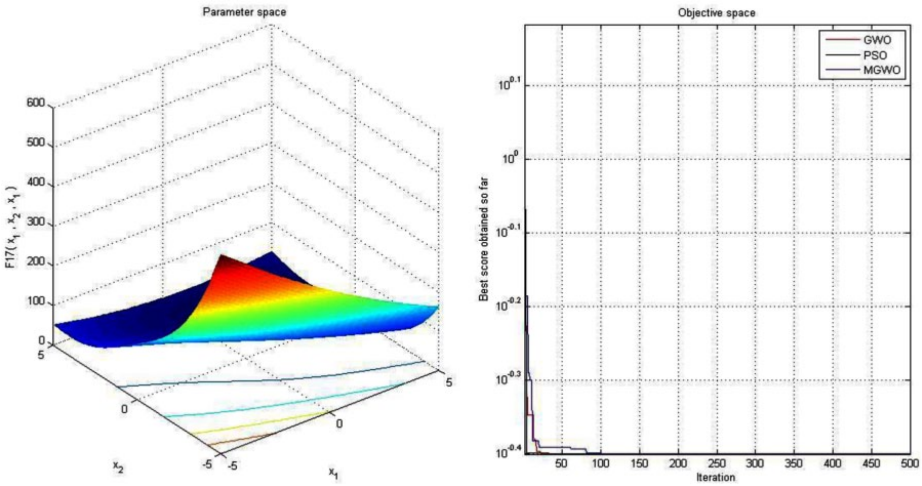

Figure 21. Convergence graph of fixed-dimension multimodal benchmark function (F_{17}). GWO indicates gray wolf optimization; MGWO, mean gray wolf optimization; PSO, particle swarm optimization.

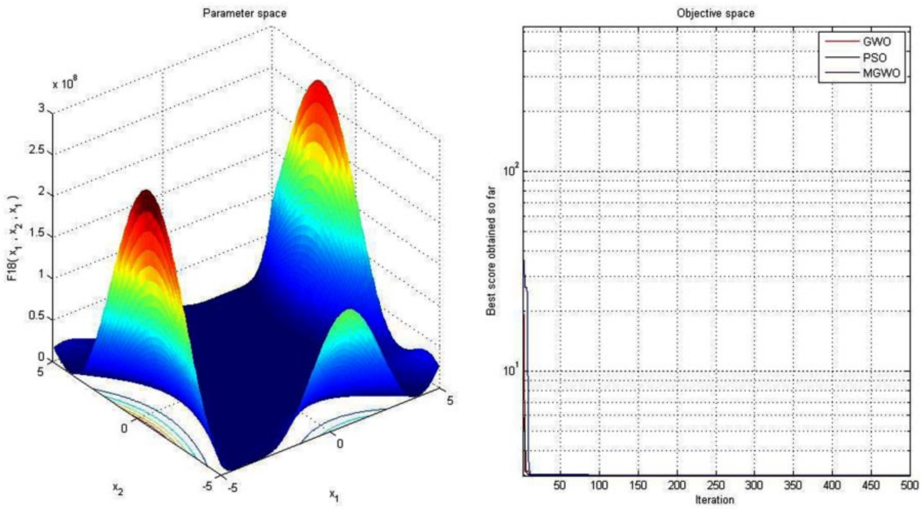

Figure 22. Convergence graph of fixed-dimension multimodal benchmark function (F_{18}). GWO indicates gray wolf optimization; MGWO, mean gray wolf optimization; PSO, particle swarm optimization.

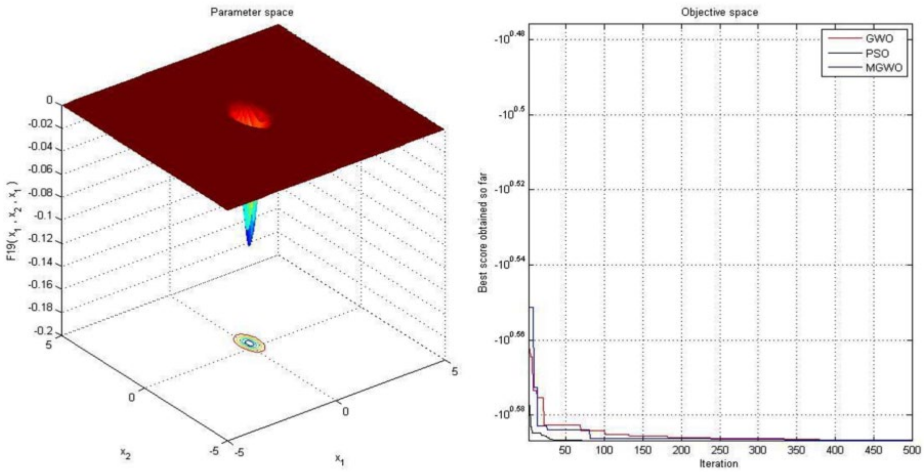

Figure 23. Convergence graph of fixed-dimension multimodal benchmark function (F_{19}). GWO indicates gray wolf optimization; MGWO, mean gray wolf optimization; PSO, particle swarm optimization.

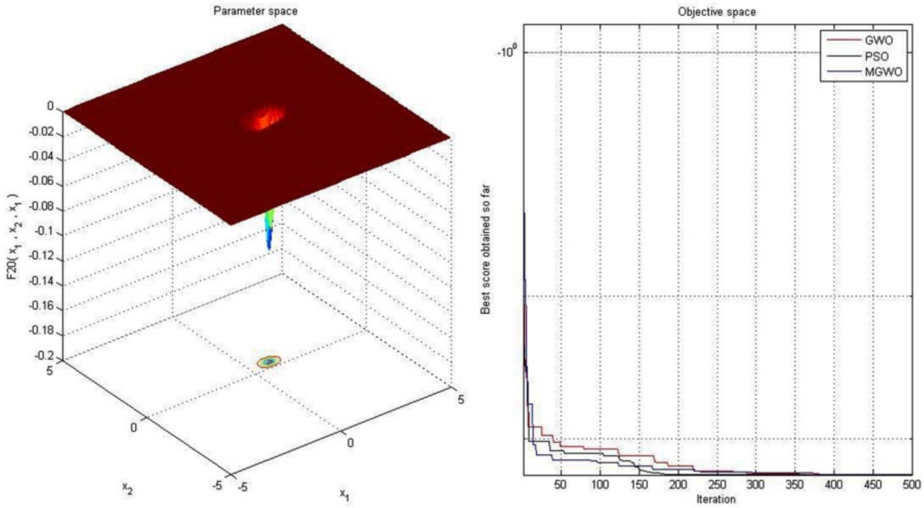

Figure 24. Convergence graph of fixed-dimension multimodal benchmark function (F_{20}). GWO indicates gray wolf optimization; MGWO, mean gray wolf optimization; PSO, particle swarm optimization.

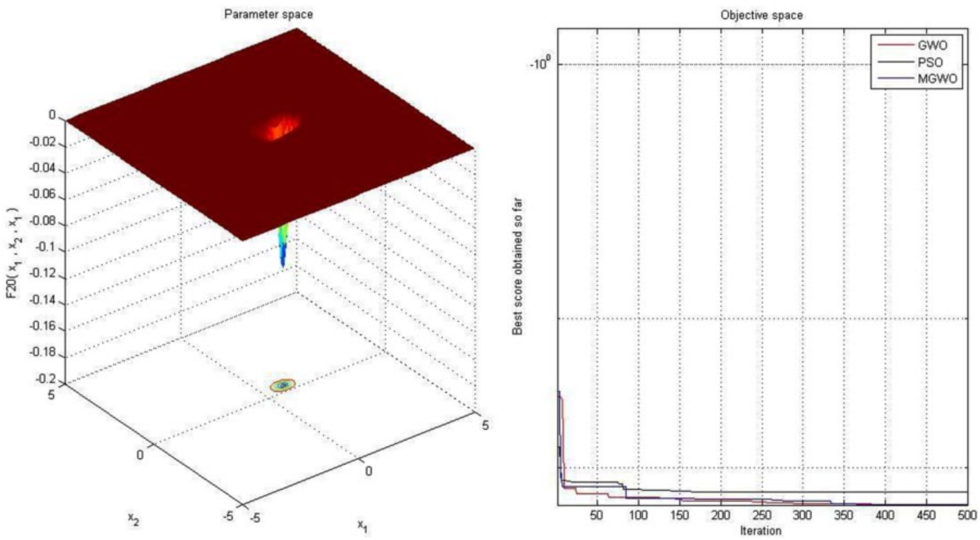

Figure 25. Convergence graph of fixed-dimension multimodal benchmark function (F_{21}). GWO indicates gray wolf optimization; MGWO, mean gray wolf optimization; PSO, particle swarm optimization.

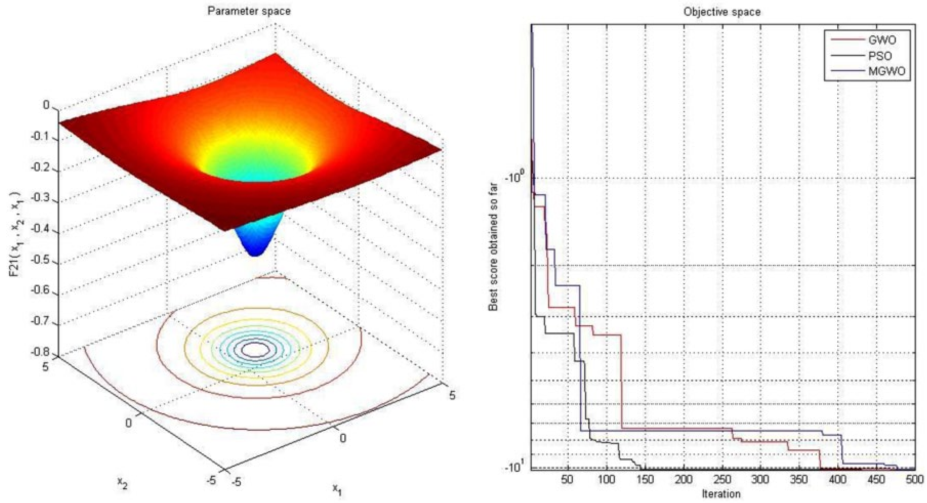

Figure 26. Convergence graph of fixed-dimension multimodal benchmark function (F_{22}). GWO indicates gray wolf optimization; MGWO, mean gray wolf optimization; PSO, particle swarm optimization.

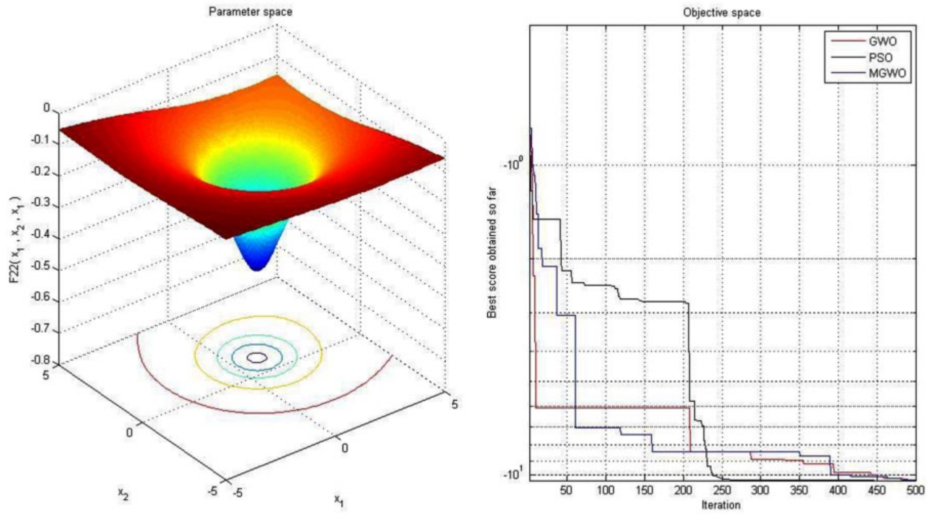

Figure 27. Convergence graph of fixed-dimension multimodal benchmark function (F_{23}). GWO indicates gray wolf optimization; MGWO, mean gray wolf optimization; PSO, particle swarm optimization.

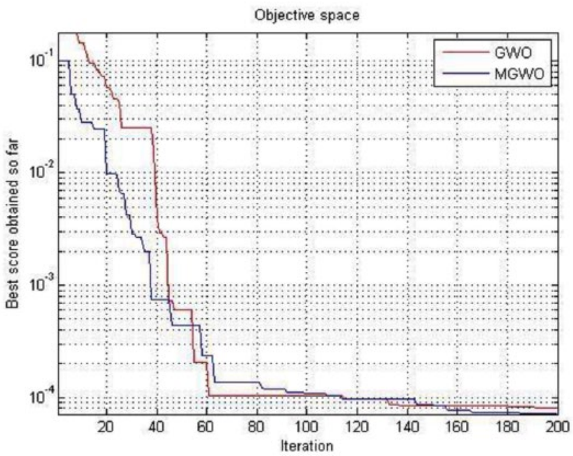

Figure 28. Convergence graph of XOR data set problem. GWO indicates gray wolf optimization; MGWO, mean gray wolf optimization.

benefits from high exploration in comparison with the PSO and GWO algorithms.

Moreover, the performance of the modified variant has also been tested on 5 data set problems, ie, (1) XOR, (2) Balloon, (3) Breast Cancer, (4) Iris, and (5) Heart. For the verification, the statistical results of the MGWO algorithm have been compared with 6 other meta-heuristics trainers: GWO, PSO, GA, ACO, ES, and PBIL. On the basis of results obtained for these data sets, we have discussed and identified the reasons for poor and strong performance of other variants. The experimental statistical results showed that the modified variant gives high competitive solutions in terms of improved local optima avoidance and high level of accuracy in mean, standard deviation, classification, and convergence rate as compared with GWO, PSO, GA, ACO, ES, and PBIL algorithms.

Table 19. Experimental results for the XOR data set.

VARIANT	MSE (AVE.)	MSE (STD)	CLASSIFICATION RATE, %
MGWO	**0.0053**	**0.0173**	**100**
GWO	0.009410	0.029500	100
PSO	0.084050	0.035945	37.50
GA	0.000181	0.000413	100
ACO	0.180328	0.025268	62.50
ES	0.118739	0.011574	62.50
PBIL	0.030228	0.039668	62.50

Abbreviations: ACO, ant colony optimization; ES, evolution strategy; GA, Genetic algorithm; GWO, gray wolf optimization; MGWO, mean gray wolf optimization; MSE, mean squared error; PBIL, population-based incremental learning; PSO, particle swarm optimization.
Bold values highlight the results of proposed variant.

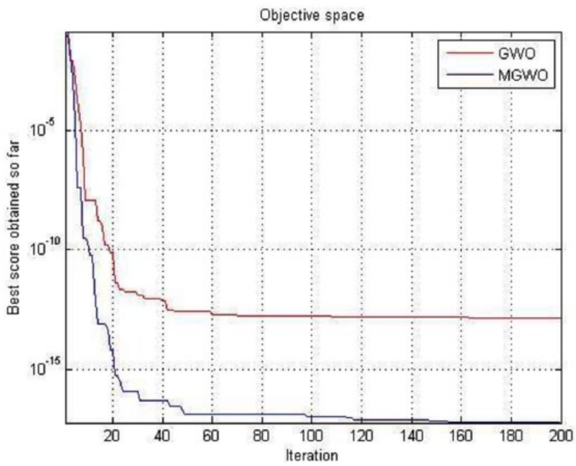

Figure 29. Convergence graph of balloon data set problem. GWO indicates gray wolf optimization; MGWO, mean gray wolf optimization.

Table 20. Experimental results for the balloon data set.

VARIANT	MSE (AVE.)	MSE (STD)	CLASSIFICATION RATE, %
MGWO	**0.0014**	**0.0132**	**100**
GWO	9.38e−15	2.81e−14	100
PSO	0.000585	0.000749	100
GA	5.08e−24	1.06e−23	100
ACO	0.004854	0.007760	100
ES	0.019055	0.170260	100
PBIL	2.49e−05	5.27e−05	100

Abbreviations: ACO, ant colony optimization; ES, evolution strategy; GA, Genetic algorithm; GWO, gray wolf optimization; MGWO, mean gray wolf optimization; MSE, mean squared error; PBIL, population-based incremental learning; PSO, particle swarm optimization.
Bold values highlight the results of proposed variant.

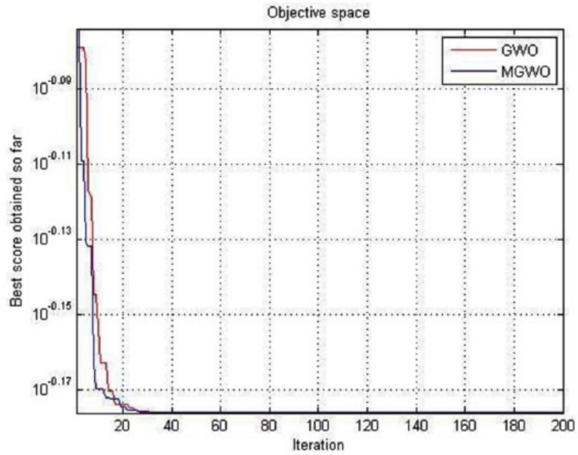

Figure 30. Convergence graph of breast cancer data set problem. GWO indicates gray wolf optimization; MGWO, mean gray wolf optimization.

Table 21. Experimental results for the breast cancer data set.

VARIANT	MSE (AVE.)	MSE (STD)	CLASSIFICATION RATE, %
MGWO	**0.0036**	**0.0063**	**99.11**
GWO	0.0012	7.4498e−05	99
PSO	0.034881	0.002472	11.00
GA	0.003026	0.001500	98.0
ACO	0.013510	0.002137	40.00
ES	0.040320	0.002470	06.00
PBIL	0.032009	0.003065	07.00

Abbreviations: ACO, ant colony optimization; ES, evolution strategy; GA, Genetic algorithm; GWO, gray wolf optimization; MGWO, mean gray wolf optimization; MSE, mean squared error; PBIL, population-based incremental learning; PSO, particle swarm optimization.
Bold values highlight the results of proposed variant.

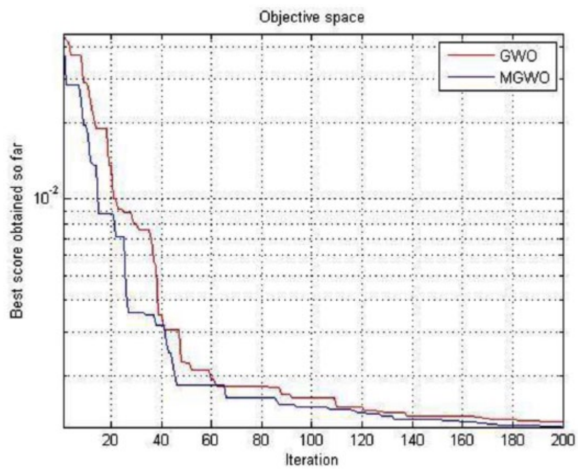

Figure 31. Convergence graph of iris data set problem. GWO indicates gray wolf optimization; MGWO, mean gray wolf optimization.

Table 22. Experimental results for the iris data set.

VARIANT	MSE (AVE.)	MSE (STD)	CLASSIFICATION RATE, %
MGWO	**0.6712**	**0.0024**	**91.334**
GWO	0.0229	0.0032	91.333
PSO	0.228680	0.057235	37.33
GA	0.089912	0.123638	89.33
ACO	0.405979	0.053775	32.66
ES	0.314340	0.052142	46.66
PBIL	0.116067	0.036355	86.66

Abbreviations: ACO, ant colony optimization; ES, evolution strategy; GA, Genetic algorithm; GWO, gray wolf optimization; MGWO, mean gray wolf optimization; MSE, mean squared error; PBIL, population-based incremental learning; PSO, particle swarm optimization.
Bold values highlight the results of proposed variant.

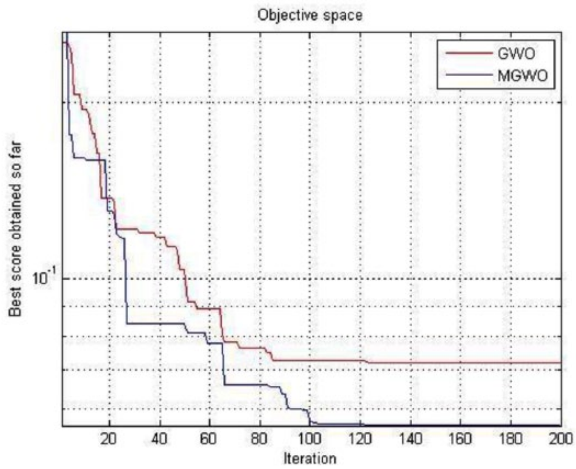

Figure 32. Convergence graph of heart data set problem. GWO indicates gray wolf optimization; MGWO, mean gray wolf optimization.

Table 23. Experimental results for the heart data set.

VARIANT	MSE (AVE.)	MSE (STD)	CLASSIFICATION RATE, %
MGWO	**0.0765**	**0.0376**	**75.14**
GWO	0.122600	0.007700	75.00
PSO	0.188568	0.008939	68.75
GA	0.093047	0.022460	58.75
ACO	0.228430	0.004979	00.00
ES	0.192473	0.015174	71.25
PBIL	0.154096	0.018204	45.00

Abbreviations: ACO, ant colony optimization; ES, evolution strategy; GA, Genetic algorithm; GWO, gray wolf optimization; MGWO, mean gray wolf optimization; MSE, mean squared error; PBIL, population-based incremental learning; PSO, particle swarm optimization.
Bold values highlight the results of proposed variant.

Author Contributions

SBS conceived the idea to develop a new variant of nature inspired technique which can outperform other metaheuristics in terms of solution quality and convergence. NS designed the numerical experiments, developed code and prepared the manuscript. Both the authors revised and finalized the final draft of manuscript.

REFERENCES

1. Rechenberg I. *Evolutions Strategies*. Vol 94. Stuttgart, Germany: Frommann-Holzboog; 1994.
2. Kennedy J, Eberhart R. Particle swarm optimization. In: Proceedings of IEEE International Conference on Neural Networks, Piscataway, NJ; November 27-December 1, 1995:1942–1948.
3. Holland JH. Genetic algorithms. *Sci Am*. 1992;267:66–72.
4. Dorigo M, Birattari M, Stutzle T. Ant colony optimization. *IEEE Comput Intell M*. 2006;1:28–39.
5. Shumeet B. *Population-Based Incremental Learning: A Method for Integrating Genetic Search Based Function Optimization and Competitive Learning*. Technical Report (CMU-CS-94-163). Pittsburgh, PA: Carnegie Mellon University; 1994.
6. Rashedi E, Nezamabadi-Pour H, Saryazdi S. GSA: a gravitational search algorithm. *Inform Sci*. 2009;179:2232–2248.
7. Webster B, Eksin PJ. A new optimization algorithm based on natural principle of gravitation. In: Proceedings of the 2003 International Conference on Information and Knowledge Engineering (IKE'03), Las Vegas, NV; June 23-26, 2003;255–261.
8. Erol OK, Eksin I. A new optimization method: big bang-big crunch. *Adv Eng Softw*. 2006;37:106–111.
9. Fornato RA. Central force optimization: a new metaheuristic with applications in applied electromagnetics. *Prog Electromagn Res*. 2007;77:425–491.
10. Alatas B. ACROA: artificial chemical reaction optimization algorithm for global optimization. *Expert Syst Appl*. 2011;38:13170–13180.
11. Kaveh A, Talatahari S. A novel heuristic optimization method: charged system search. *Acta Mech*. 2010;213:267–289.
12. Kaveh A, Khayatazad M. A new meta-heuristic method: ray optimization. *Comput Struct*. 2012;112:283–294.
13. Shah-Hosseini H. Principal components analysis by the galaxy-based search algorithm: a novel meta-heuristic for continuous optimization. *Int J Comput Sci Eng*. 2011;6:132–140.
14. Hatamlou A. Black hole: a new heuristic optimization approach for data clustering. *Inform Sci*. 2012;222:175–184.
15. Moghaddam FF, Moghaddam RF, Cherict M. Curved space optimization: random search based on general relativity theory. arXivpreprintarXiv:1208.2214, 2012. https://arxiv.org/abs/1208.2214.
16. Du H, Wu X, Zhuang J. Small-world optimization algorithm for function optimization. In Jiao L, Wang L, Gao X, Liu J, Wu F (eds) *Advances in Natural Computation*, Berlin: Springer; 2006;264–273.
17. Mirjalili S, Mirjalili SM, Lewis A. Grey wolf optimization. *Adv Eng Softw*. 2014;69:46–61.
18. Mirjalili S. How effective is the Grey Wolf Optimizer in training multi-layer perceptrons. *Appl Intell*. 2015;43:150–161.
19. Pereira L, Rodrigues D, Ribeiro P, Papa J, Weber SA. Social spider optimization-based artificial neural networks training and its applications for Parkinson's disease identification. In: *IEEE 27th International Symposium on in Computer-Based Medical Systems (CBMS)*; New York, NY; May 27-29, 2014:14–17.
20. Moallem P, Razmjooy N. A multi layer perceptron neural network trained by invasive weed optimization for potato color image segmentation. *Trends Appl Sci Res*. 2012;7:445–455.
21. Yu JJ, Lam AY, Li VO. Evolutionary artificial neural network based on chemical reaction optimization. In: *IEEE Congress on Evolutionary Computation (CEC)*; New Orleans, LA; June 5-8, 2011:2083–2090.
22. Uzlu E, Kankal M, Akpmar A, Dede T. Estimates of energy consumption in Turkey using neural networks with the teaching-learning based optimization algorithm. *Energy*. 2014;75:295–303.
23. Simon D. Biogeography-based optimization. *IEEE T Evolut Comput*. 2008; 12:702–713.
24. Kaveh A, Talatahari S. A novel heuristic optimization method: charged system search. *Acta Mech*. 2010;213:267–289.
25. Emary E, Zawbaa HM, Hassanien AE. Binary grey wolf optimization approaches for feature selection. *Neurocomputing*. 2016;172:371–381.
26. Mittal N, Singh U, Singh Sohi B. Modified grey optimizer for global engineering optimization. *Appl Comput Intel Soft Comput*. 2016;2016:1–16.
27. Sodeifian G, Ardestani NS, Sajadian SA, Ghorbandoost S. Application of supercritical carbon dioxide to extract essential oil from *Cleome coluteoides* Boiss: experimental, response surface and grey wolf optimization methodology. *J Supercrit Fluid*. 2016;114:55–63.

Appendix 1

Table A. Unimodal benchmark functions.

FUNCTION	DIMENSION	RANGE	f_{min}
$F_1(x) = \sum_{i=1}^{n} x_i^2$	30	[−100, 100]	0
$F_2(x) = \sum_{i=1}^{n} \lvert x_i \rvert + \prod_{i=1}^{n} \lvert x_i \rvert$	30	[−10, 10]	0
$F_3(x) = \sum_{i=1}^{n} \left(\sum_{j-1}^{i} x_j \right)^2$	30	[−100, 100]	0
$F_4(x) = \max_i \{\lvert x_i \rvert, 1 \le i \le n\}$	30	[−100, 100]	0
$F_5(x) = \sum_{i=1}^{n-1} \left[100\left(x_{i+1} - x_i^2\right)^2 + \left(x_i - 1\right)^2 \right]$	30	[−30, 30]	0
$F_6(x) = \sum_{i=1}^{n} \left(\left[x_i + 0.5\right]\right)^2$	30	[−100, 100]	0
$F_7(x) = \sum_{i=1}^{n} i x_i^4 + rand[0,1)$	30	[−1.28, 1.28]	0

Table B. Multimodal benchmark functions.

FUNCTION	DIMENSION	RANGE	f_{min}
$F_8(x) = \sum_{i=1}^{n} -x_i \sin\left(\sqrt{\lvert x_i \rvert}\right)$	30	[−500, 500]	−418.9829×5
$F_9(x) = \sum_{i=1}^{n} \left[x_i^2 - 10\cos(2\pi x_i) + 10 \right]$	30	[−5.12, 5.12]	0
$F_{10}(x) = -20\exp\left(-0.2\sqrt{\frac{1}{n}\sum_{i=1}^{n} x_i^2}\right) - \exp\left(\frac{1}{n}\sum_{i=1}^{n}\cos\left(2\pi x_i\right)\right) + 20 + e$	30	[−32, 32]	0
$F_{11}(x) = \frac{1}{4000}\sum_{i=1}^{n} x_i^2 - \prod_{i=1}^{n}\cos\left(\frac{x_i}{\sqrt{i}}\right) + 1$	30	[−600, 600]	0
$F_{12}(x) = \frac{\pi}{n}\left\{10\sin(\pi y_i) + \sum_{i=1}^{n-1}\left(y_i - 1\right)^2\left[1 + 10\sin^2\left(\pi y_{i+1}\right) + \left(y_{n-1}\right)^2\right]\right\} + \sum_{i=1}^{n} u\left(x_i, 10, 100, 4\right)$ $y_i = 1 + \frac{x_i + 1}{4}$ $u\left(x_i, a, k, m\right) = \begin{cases} k\left(x_i - a\right)^m & x_i > a \\ 0 & -a < x_i < a \\ k\left(-x_i - a\right)^m & x_i < -a \end{cases}$	30	[−50, 50]	0
$F_{13}(x) = 0.1\left\{\sin^2(3\pi x_i) + \sum_{i=1}^{n}\left(x_i - 1\right)^2\left[1 + \sin^2\left(3\pi x_i + 1\right)\right] + \left(x_n - 1\right)^2\left[1 + \sin^2(2\pi x_n)\right]\right\}$ $+ \sum_{i=1}^{n} u\left(x_i, 5, 100, 4\right)$	30	[−50, 50]	0

Table C. Fixed-dimension multimodal benchmark functions.

FUNCTION	DIMENSION	RANGE	f_{min}
$F_{14}(x)=\left(\frac{1}{500}+\sum_{j=1}^{25}\frac{1}{j+\sum_{i=1}^{2}\left(x_i-a_{ij}\right)^6}\right)^{-1}$	2	[−65, 65]	1
$F_{15}(x)=\sum_{i=1}^{11}\left[a_i-\frac{x_1\left(b_i^2+b_ix_2\right)}{b_i^2+b_ix_i+x_4}\right]^2$	4	[−5, 5]	0.00030
$F_{16}(x)=4x_1^2-2.1x_1^4+\frac{1}{3}x_1^6+x_1x_2-4x_2^2+4x_2^4$	2	[−5, 5]	−1.0316
$F_{17}(x)=\left(x_2-\frac{5.1}{4\pi^2}x_1^2+\frac{5}{\pi}x_1-6\right)^2+10\left(1-\frac{1}{8\pi}\right)\cos x_1+10$	2	[−5, 5]	0.398
$F_{18}(x)=\left[1+\left(x_1+x_2+1\right)^2\left(19-14x_1+3x_1^2-14x_2+6x_1x_2+3x_2^2\right)\right]\times\left[30+\left(2x_1-3x_2\right)^2\times\left(18-32x_1+12x_1^2+48x_2-36x_1x_2+27x_2^2\right)\right]$	2	[−2, 2]	3
$F_{19}(x)=-\sum_{i=1}^{4}c_i\exp\left(-\sum_{j=1}^{3}a_{ij}\left(x_j-p_{ij}\right)^2\right)$	3	[1, 3]	−3.86
$F_{20}(x)=-\sum_{i=1}^{4}c_i\exp\left(-\sum_{j=1}^{6}a_{ij}\left(x_j-p_{ij}\right)^2\right)$	6	[0, 1]	−3.32
$F_{21}(x)=-\sum_{i=1}^{5}\left[\left(X-a_i\right)\left(X-a_i\right)^T+c_i\right]^{-1}$	4	[0, 10]	−10.1532
$F_{22}(x)=-\sum_{i=1}^{7}\left[\left(X-a_i\right)\left(X-a_i\right)^T+c_i\right]^{-1}$	4	[0, 10]	−10.4028
$F_{23}(x)=-\sum_{i=1}^{10}\left[\left(X-a_i\right)\left(X-a_i\right)^T+c_i\right]^{-1}$	4	[0, 10]	−10.5363

Table D. Classification data sets.

CLASSIFICATION DATA SETS	NUMBER OF ATTRIBUTES	NUMBER OF TRAINING SAMPLES	NUMBER OF TEST SAMPLES	NUMBER OF CLASSES
3-bits XOR	3	8	8 as training samples	2
Balloon	4	16	16 as training samples	2
Iris	4	150	150 as training samples	3
Breast cancer	9	599	100	2
Heart	22	80	187	2

Adapted from Mirjalili.[18]

Table E. The initial parameters of algorithms.

ALGORITHM	PARAMETER	VALUE
MGWO	$\vec{a}$	Linearly decreased from 2 to 0
	Population size	50 for XOR and Balloon, 200 for the rest
	Maximum number of generations	250
GWO	$\vec{a}$	Linearly decreased from 2 to 0
	Population size	50 for XOR and Balloon, 200 for the rest
	Maximum number of generations	250

Abbreviations: GWO, gray wolf optimization; MGWO, mean gray wolf optimization.