Journal of Mathematics and Statistics Studies
ISSN: 2709-4200
DOI: /jmss
Journal Homepage: www.al-kindipublisher.com/index.php/jmss

JMSS
AL-KINDI CENTER FOR RESEARCH AND DEVELOPMENT

| RESEARCH ARTICLE

A Hybrid CMA-ES–GWO Algorithm for Solving Optimization Problems

Elias Ahmad Ahmadi[1]✉, Besmillah Danish[2], Ahmad Ramin Rahnaward[3]
[1,2,3]Department of General Mathematics, Faculty of Mathematical Sciences, Kabul University, Kabul, Afghanistan
Corresponding Author: Elias Ahmad Ahmadi, **E-mail**: eliasahmad102266@gmail.com

| ABSTRACT
Metaheuristic algorithms often suffer from an imbalance between exploration and exploitation, which can lead to premature convergence. In this paper, a hybrid optimization algorithm is proposed by integrating the Grey Wolf Optimizer (GWO) and the Covariance Matrix Adaptation Evolution Strategy (CMA-ES), termed the Hybrid CMA-ES–GWO Algorithm (HCG). The proposed hybridization is motivated by the strong exploration capability of GWO and the efficient exploitation of CMA-ES. The HCG algorithm adopts a multi-population framework in which GWO is first employed to explore the search space. Subsequently, CMA-ES is applied to exploit these regions through covariance-based sampling, enabling accurate convergence toward optimal solutions. This cooperative mechanism allows HCG to effectively balance global search and local refinement while maintaining population diversity. To evaluate the performance of the algorithm, experiments were conducted on 19 benchmark functions, including unimodal, multimodal, and fixed-dimension multimodal problems. The performance of HCG was compared with the original CMA-ES and GWO algorithms. The experimental results demonstrate that HCG consistently achieves superior optimization accuracy compared to the standalone algorithms. These findings confirm that the proposed hybrid strategy effectively overcomes the individual limitations of CMA-ES and GWO, making HCG a reliable method for solving complex continuous optimization problems.

| KEYWORDS

Covariance Matrix Adaptation Evolution Strategy (CMA-ES); Exploitation; Exploration; Grey Wolf Optimizer (GWO); Hybrid CMA-ES–GWO Algorithm (HCG)

| ARTICLE INFORMATION

ACCEPTED: 08 August 2026 **PUBLISHED:** 23 September 2026 **DOI:** 10.32996/jmss.2026.7.6.5

1. Introduction

The increasing demand for fast, reliable, and efficient decision-making systems has made optimization a fundamental research area across science and engineering. Optimization algorithms aim to identify the best feasible solution among a large set of candidate solutions by iteratively improving decision variables with respect to a predefined objective function. However, many real-world optimization problems involve high-dimensional decision spaces, nonlinear objective functions, complex interdependencies among variables, and expensive function evaluations. Under such conditions, obtaining an exact global optimum becomes computationally impractical, and researchers often seek near-optimal solutions that offer a favorable trade-off between solution quality and computational cost (Mirjalili et al., 2014).
According to the No Free Lunch theorem, no single optimization algorithm is universally optimal for all problem types. An algorithm that performs well for a particular problem structure or variable configuration may perform poorly when the problem characteristics, constraints, or variable domains change. This fundamental limitation has encouraged the development of a wide variety of optimization techniques tailored to different problem settings. Optimization problems may involve a single objective or multiple conflicting objectives and may be defined over continuous or discrete decision variables. In particular, continuous optimization plays a crucial role, as many discrete problems are transformed into sequences of continuous sub-problems during the solution process (Faris et al., 2018).

In practical applications, optimization models may include unconstrained variables or be subject to equality and inequality constraints that describe complex relationships among decision variables. These problems can further be classified as linear or nonlinear, convex or non-convex, and differentiable or non-differentiable. Moreover, in many real-world scenarios, the exact values of parameters and variables are not known with certainty, giving rise to stochastic optimization problems (Almufti et al., 2021). In such cases, deterministic analytical methods or gradient-based techniques are often inadequate, especially when the objective function is noisy, discontinuous, or lacks explicit derivatives. To address these challenges, meta-heuristic algorithms have been widely adopted as robust and flexible optimization tools. Meta-heuristics are iterative, stochastic, and derivative-free methods that explore the search space by manipulating candidate solutions represented through sets of decision variables. By balancing exploration of the global search space and exploitation around promising regions, these algorithms are capable of handling complex optimization landscapes with minimal assumptions about problem structure (Nadimi-Shahraki et al., 2021)

Meta-heuristic algorithms are generally classified into physics-based methods, swarm intelligence algorithms, and evolutionary algorithms. Physics-based algorithms simulate natural laws to guide the movement of solution vectors in the search space. Swarm intelligence algorithms model the collective behavior of social agents, where multiple candidate solutions interact and exchange information to update their positions. Evolutionary algorithms, inspired by biological evolution, iteratively improve populations of candidate solutions through operators such as mutation, recombination, and selection, gradually refining the values of decision variables toward optimal regions (Beyer & Sendhoff, 2017).

Among evolutionary algorithms, CMA-ES has proven highly effective for continuous, non-convex, and black-box optimization problems by adaptively learning variable dependencies through covariance matrix updates (Igel et al., 2007). Conversely, GWO, introduced by (Mirjalili et al., 2014), is a swarm intelligence algorithm inspired by the leadership hierarchy and hunting behavior of grey wolves, has gained popularity due to its simplicity, few control parameters, strong exploration capability, and competitive performance. Despite their strengths, both algorithms exhibit limitations: GWO may experience stagnation and premature convergence due to the lack of an explicit operator for exploitation, which results in slow convergence and reduced solution accuracy, while CMA-ES relies on the single adaptive Gaussian sampling distribution, which causes a weak exploration on search space (Karmakar et al., 2023).

The research problem addressed in this study arises from the inability of individual meta-heuristic algorithms to consistently balance global exploration and local exploitation across diverse optimization problems with interacting decision variables. To overcome this problem, researchers have increasingly focused on hybridization strategies that combine complementary features of multiple algorithms. Existing hybrid meta-heuristic algorithms have achieved notable successes in applications such as feature selection in high-dimensional data, intrusion detection for IoT networks, numerical global optimization, and classification tasks. For example, the Hybrid Butterfly-Grey Wolf Optimization (HB-GWO) synergizes the global exploration capabilities of the Butterfly Optimization Algorithm (BOA) with the local exploitation strengths of the GWO (Aly & Alotaibi, 2025). The algorithm incorporates an adaptive switching mechanism that dynamically adjusts the contribution of BOA and GWO throughout the optimization process. The hybridization (GWO-PSO) leverages PSO's exploration ability and GWO's exploitation capacity through adaptive mechanisms to address limitations such as premature convergence and local optima stagnation (Sharma et al., 2025). Even though the original hybrid approach gives better performance, it is appropriate only for problems with a continuous search space. A hybridization algorithm from two inspired algorithms, GWO and differential evolution (DE), is named GWO-DE (Tawhid & Ibrahim, 2020). A new improving encircling prey and a new crossover technique are used for updating the new agents of GWO-DE based on the generated agents of DE and GWO. Since GWO-DE has the advantages over GWO and DE, it subdues the inability of GWO and DE for solving unconstrained optimization problems and systems of nonlinear equations. Other examples include hybridizations of GWO with the Bat Algorithm for gene selection in cancer classification. The BA-GWO is introduced to determine a limited number of biomarker genes that significantly enhance the classification performance. In this approach, the *k*-nearest neighbor algorithm was employed for the classification task. The proposed BA-GWO is mainly introduced to improve the BA search agents' performance in searching for the best candidate gene subset that carries the biomarkers for cancer classification (Tbaishat et al., 2025). Furthermore, the BA-GWO is designed to enhance both exploitation and exploration capabilities while ensuring a balanced approach and preventing stagnation in local optima. The primary function of this proposed approach is to enhance the solutions acquired through the BA by utilizing them as the initial population for the GWO. These studies have shown that combining GWO with complementary meta-heuristics can enhance convergence speed and solution quality by leveraging the strengths of multiple algorithms.

Similarly, Researchers have explored hybridization strategies involving CMA-ES to improve its global search and local refinement capabilities. Examples include cooperative differential evolution assisted by CMA-ES (Song et al., 2024). In the proposed algorithm, jSO, as a variant of Differential Evolution (DE), is used as a global search operator to explore the entire solution space. When the population falls into stagnation, a relatively reliable initial solution for the local search operator is generated by the CMA-ES, which is activated to perturb the optimal candidates in the solution space. The LBFGS utilized as the local search strategy is embedded in CMA-ES to obtain the potential local optimal solutions. A cooperative co-evolutionary dynamic system is formed by jSO and CMA-ES with a local search operator. Another example is the hybridization with Genetic Algorithms for chemical compound classification (Guo et al., 2024). In this paper, we introduce a novel hybrid optimization framework named GA-CMA-ES, which integrates Genetic Algorithms (GA) and the Covariance Matrix Adaptation Evolution Strategy (CMA-ES) to train Recurrent Neural Networks (RNNs) for compound classification. Leveraging the global exploration capabilities of GAs and the local exploration abilities of the CMA-ES, the proposed method achieves notable performance, attaining an 83% classification accuracy on a benchmark dataset, surpassing the baseline method. coupling with PSO for enhanced global optimization performance. In this paper, we first propose a time-window PSO (TW-PSO) as an improvement of PSO, which could enhance the exploration ability of the algorithm. Second, we design a hybrid algorithm of TW-PSO, PSO, and CMA-ES, i.e., HTPC, which combines the advantages of TW-PSO, PSO, and CMA-ES (Xu et al., 2019).

Building on the insights from previous studies, this research proposes a hybrid algorithm, referred to as HCG, which integrates the adaptive covariance learning mechanism of CMA-ES with the social hierarchy and cooperative search behavior of GWO. The HCG algorithm is designed to address the limitations of existing meta-heuristics by enhancing convergence performance, improving solution quality, and mitigating stagnation. By combining the complementary strengths of CMA-ES and GWO, HCG provides a robust framework capable of effectively navigating complex optimization landscapes while maintaining an appropriate balance between exploration and exploitation.

2. The Original Algorithms

2.1 Covariance Matrix Adaptation Evolution Strategy Algorithm

The Covariance Matrix Adaptation Evolution Strategy (CMA-ES) is a powerful evolutionary optimization algorithm designed for solving complex non-convex, nonlinear, and multimodal optimization problems in continuous domains. Originally introduced by Ostermeier and Hansen in 2001 as a completely derandomized self-adaptation mechanism, CMA-ES has become a state-of-the-art method in evolutionary computation due to its fast convergence, robustness, and reduced time complexity. A key advantage of CMA-ES is that it does not require gradient information, making it suitable for noisy, non-smooth, discontinuous, and black-box objective functions. Moreover, its internal parameters are automatically adapted during the search process, which minimizes user intervention and enhances algorithmic reliability. Owing to its second-order adaptation strategy, CMA-ES is particularly effective for high-dimensional, ill-conditioned, large-scale, and non-separable optimization problems(Ikram et al., 2022).

CMA-ES is a population-based stochastic search algorithm that explores the search space by sampling candidate solutions from a multivariate normal distribution. At generation g, a population of λ offspring is generated around the current mean vector $m^{(g)} \in \mathbb{R}^n$ according to

$$x_k^{(g+1)} \sim N\left(m^{(g)}, \left(\sigma^{(g)}\right)^2 \cdot C^{(g)}\right) \text{ for } k = 1,2,\dots,\lambda \tag{1}$$

where $C^{(g)} \in \mathbb{R}^{n\times n}$ is the covariance matrix capturing the dependencies among decision variables, $\sigma^{(g)} \in \mathbb{R}_+, \lambda \geq 2$ is the global step size controlling the mutation scale, and n denotes the problem dimension. Each candidate solution is represented as

$$X_k = [x_1, x_2, \dots, x_D] \text{ for } k = 1,2,\dots,\mu \tag{2}$$

After sampling, all offspring are evaluated using the objective function $f: \mathbb{R}^n \to \mathbb{R}$ and sorted based on their fitness values such that

$$f\left(x_{1:\lambda}^{(g+1)}\right) \leq f\left(x_{2:\lambda}^{(g+1)}\right) \leq \cdots \leq f\left(x_{\lambda:\lambda}^{(g+1)}\right)$$

where only the ranking information is used as feedback, ensuring invariance to monotonic transformations of the objective function.

The update of the search distribution is achieved through the weighted recombination of the best $\mu \leq \lambda$ offspring. The mean vector for the next generation is computed as

$$m^{(g+1)} = \sum_{i=1}^{\mu} \omega_i x_{i:\lambda}^{(g+1)} \tag{3}$$

where the recombination weights satisfy

$$\sum_{i=1}^{\mu} \omega_i = 1, \omega_i > 0 \text{ for } i = 1,2,\dots,\mu \tag{4}$$

This weighting scheme assigns greater influence to better-performing individuals. The variance-effective selection mass is defined by

$$\mu_{eff} = \left(\sum_{i=1}^{\mu} \omega_i^2\right)^{-1} \tag{5}$$

which reflects the effective number of contributing parents and plays a central role in the adaptation of the strategy parameters.

To capture correlations between decision variables and guide the search toward promising regions, CMA-ES adapts the covariance matrix using information accumulated in an evolution path. The evolution path for covariance adaptation is updated as

$$P_\sigma^{(g+1)} = (1 - c_\sigma)P_\sigma^{(g)} + \sqrt{c_\sigma(2 - c_\sigma)\mu_{eff}}\, {C^{(g)}}^{-\frac{1}{2}} \frac{m^{(g+1)} - m^{(g)}}{\sigma^{(g)}} \tag{6}$$

where c_c is a learning rate controlling the memory of past successful steps. The covariance matrix is then updated using a combination of rank-one and rank- μ updates according to

$$C^{(g+1)} = (1 - c_{\text{cov}})C^{(g)} + \frac{c_{\text{cov}}}{\mu_{\text{cov}}} P_c^{(g+1)} P_c^{(g+1)^T} + c_{\text{cov}} \left(1 - \frac{1}{\mu_{\text{cov}}}\right) \times \sum_{i=1}^{\mu} W_i \left(\frac{x_{1:\lambda}^{(g+1)} - m^{(g)}}{\sigma^{(g)}}\right) \left(\frac{x_{1:\lambda}^{(g+1)} - m^{(g)}}{\sigma^{(g)}}\right)^T \quad (7)$$

This mechanism allows the algorithm to learn the shape, orientation, and scaling of the objective function landscape, thereby enabling efficient search in non-separable and ill-conditioned problems(Igel et al., 2007).

In addition to covariance matrix adaptation, CMA-ES employs cumulative step-size adaptation to control the overall scale of the search. The evolution path for step-size control is updated as

$$P_\sigma^{(g+1)} = (1 - c_\sigma) P_\sigma^{(g)} + \sqrt{c_\sigma (2 - c_\sigma) \mu_{eff}}\, C^{(g)^{-\frac{1}{2}}} \frac{m^{(g+1)} - m^{(g)}}{\sigma^{(g)}} \quad (8)$$

where c_σ is the step-size learning rate. The step size is then adapted as

$$\sigma^{(g+1)} = \sigma^{(g)} \exp\left(\frac{c_\sigma}{d_\sigma}\left(\frac{\left\|P_\sigma^{(g+1)}\right\|}{E\|N(0,1)\|} - 1\right)\right) \quad (9)$$

with d_σ being a damping parameter. When the evolution path is longer than expected under random selection, the step size is increased to enhance exploration; otherwise, it is decreased to promote exploitation and prevent premature convergence.

The population size λ represents a trade-off between convergence speed and global search capability. Smaller values of λ favor faster convergence and local exploitation, whereas larger values improve robustness against local optima and enhance exploration of large search spaces. Overall, CMA-ES is particularly effective for high-dimensional, non-separable optimization problems where derivative information is unavailable, and the search space is large. However, for low-dimensional, partly separable problems or problems solvable with a limited number of function evaluations, alternative strategies may achieve better performance.

2.2 Grey Wolf Optimizer

GWO is a nature-inspired metaheuristic algorithm that models the leadership hierarchy and hunting behavior of grey wolves (Canis lupus). Grey wolves are social predators that live in packs of five to twelve members, exhibiting a strict dominance hierarchy with roles assigned to alpha (α), beta (β), delta (δ), and omega (ω) wolves. Alpha wolves act as leaders, making decisions for hunting, migration, and feeding, while beta wolves serve as advisers and enforcers of the alpha's commands. Delta wolves, including subordinates such as caretakers, elders, scouts, hunters, and sentinels, follow the commands of alphas and betas while performing specialized tasks. Omega wolves have the lowest rank and are subordinate to all others; although they appear least important, their presence maintains the internal social balance of the pack. Inspired by this hierarchical social structure, GWO simulates the hunting and leadership dynamics of grey wolves to guide the search for global optima in optimization problems(Mirjalili et al., 2014).

GWO is a population-based algorithm in which candidate solutions correspond to wolves in a pack. At the start, the position of the prey (optimal solution) is unknown, and three top-ranked wolves are identified as alpha (best solution), beta (second-best), and delta (third-best), while the remaining wolves are classified as omega. The hunting process is guided by the positions of α, β, and δ wolves, with ω wolves updating their positions according to the guidance provided by these three leaders. The algorithm models three main behaviors of grey wolves during hunting: tracking, encircling, and attacking the prey. Encircling behavior is mathematically represented as

$$\vec{X}(t+1) = \vec{X}_p(t) - \vec{A} \cdot \vec{D} \quad (10)$$

where

$$\vec{D} = \left|\vec{C} \cdot \vec{X}_p(t) - \vec{X}(t)\right| \quad (11)$$

$$\vec{C} = 2 \cdot \vec{r}_2 \quad (12)$$

$$\vec{A} = 2\vec{a} \cdot \vec{r}_1 - \vec{a} \quad (13)$$

Here, $\vec{X}(t)$ is the position vector of a wolf, $\vec{X}_p(t)$ represents the prey's position, $\vec{r}_1$ and $\vec{r}_2$ are random vectors in [0,1], and $\vec{a}$ is a linearly decreasing parameter from 2 to 0 over the iterations:

$$\vec{a} = 2 - t \cdot \left(\frac{2}{T}\right) \quad (14)$$

where t is the current iteration, and T is the maximum number of iterations. This formulation allows the wolves to explore any point within a hypersphere surrounding the prey, ensuring a stochastic and global search capability.

The positions of each wolf are updated relative to the top three solutions (α, β, δ) to model cooperative hunting and to approximate the prey location. For a wolf, the position update equations are

$$\vec{D}_\alpha = |\vec{C}_1 \cdot \vec{X}_\alpha - \vec{X}|, \vec{X}_1 = \vec{X}_\alpha - \vec{A}_1 \cdot \vec{D}_\alpha \quad (15)$$
$$\vec{D}_\beta = |\vec{C}_2 \cdot \vec{X}_\beta - \vec{X}|, \vec{X}_2 = \vec{X}_\beta - \vec{A}_2 \cdot \vec{D}_\beta \quad (16)$$
$$\vec{D}_\delta = |\vec{C}_3 \cdot \vec{X}_\delta - \vec{X}|, \vec{X}_3 = \vec{X}_\delta - \vec{A}_3 \cdot \vec{D}_\delta \quad (17)$$
$$\vec{X}(t+1) = \vec{X}_1/3 + \vec{X}_2/3 + \vec{X}_3/3 \quad (18)$$

The coefficient vectors $\vec{A}$ and $\vec{C}$ regulate exploration and exploitation. During the search phase ($|\vec{A}| > 1$), wolves diverge from the prey to explore new regions, enhancing global search and avoiding local optima. During the attack or exploitation phase ($|\vec{A}| < 1$), wolves converge toward the estimated prey position, refining the search near the best solutions. The parameter $\vec{a}$ linearly decreases from 2 to 0, controlling the transition from exploration to exploitation. The random vector $\vec{C}$ introduces stochasticity in the wolves' positions to maintain diversity and promote the exploration of unvisited regions (Faris et al., 2018).

The GWO algorithm begins with randomly initialized solutions representing wolves in the search space. At each iteration, the positions of α, β, and δ wolves are updated based on their fitness, and the ω wolves adjust their positions according to the guidance of the top 3 wolves using the position update equations. This iterative process of encircling, hunting, and attacking continues until a stopping criterion is met, typically a maximum number of iterations or a convergence threshold. The best solution identified by the alpha wolf throughout the iterations is returned as the optimal solution. The simplicity, flexibility, and ability of GWO to balance exploration and exploitation make it suitable for large-scale, high-dimensional, and complex optimization problems.

3. Hybrid CMA-ES–GWO Algorithm (HCG)

CMA-ES and GWO update their populations using fundamentally different sources of information, which makes them highly complementary. In CMA-ES, new candidate solutions are generated by sampling from a multivariate normal distribution centered around the current mean of the population. This update mechanism relies heavily on the statistical properties of successful solutions accumulated in recent generations. As a result, CMA-ES performs an efficient local search by exploiting the neighborhood of the current mean and progressively refining promising regions of the search space using second-order information encoded in the covariance matrix. Consequently, CMA-ES exhibits strong exploitation capability and fast convergence when the search is already close to an optimal region, but its performance may degrade if the initial population is far from the global optimum or if the algorithm prematurely converges to local optima due to limited global exploration.

In contrast, GWO updates each candidate solution based on the positions of the three best individuals (α, β, and δ) in the population. This collective guidance mechanism enables search agents to dynamically diverge and converge during the optimization process. When exploration is required, wolves spread out and explore different regions of the search space, whereas during exploitation, they converge toward the best solutions to refine the search. Since the update of each agent depends on multiple elite solutions rather than a single central tendency, GWO effectively utilizes global population information and maintains diversity throughout the search. This property allows GWO to explore the search landscape extensively and avoid premature convergence, although its convergence speed and local refinement ability are generally weaker than those of CMA-ES.

Motivated by these complementary characteristics, a hybrid CMA-ES–GWO algorithm (HCG) is proposed to simultaneously enhance global exploration and local exploitation. The core idea of HCG is to exploit the strong global search and diversification ability of GWO to identify promising regions of the search space, while employing the powerful local refinement and second-order adaptation mechanism of CMA-ES to accurately exploit these regions and accelerate convergence toward the global optimum. By integrating these two algorithms in a structured multi-population framework, HCG overcomes the individual limitations of each method and achieves a more balanced exploration–exploitation trade-off.

In HCG, a minimization problem is considered without loss of generality. The algorithm adopts a multi-population strategy, where N independent populations are initialized and evolved in parallel. A data structure G is used to store all populations along with their corresponding statistical and fitness information. In the first stage, each population is evolved using the GWO algorithm, allowing the search agents to widely explore the landscape and locate multiple promising regions. After a predefined number of GWO iterations, all populations are stored in G.

Next, the population containing the best solution is selected from $G$and denoted as P_b. To preserve population diversity and prevent the search from becoming overly exploitative at an early stage, the second-best population P_s is also selected and combined with P_b to form a temporary population P. The population P_s is then removed from G, ensuring that redundant or less informative populations do not dominate the search process. This controlled population merging strategy allows HCG to retain useful global information while maintaining diversity.

The combined population $P$is subsequently refined using the CMA-ES algorithm. By updating P through covariance matrix adaptation and step-size control, CMA-ES efficiently exploits the promising regions previously identified by GWO. After the CMA-ES update, the best Sindividuals from $P$are selected to form an updated elite population, which replaces P_b. This process ensures that high-quality solutions benefit from both global exploration and local exploitation mechanisms.

The algorithm iteratively repeats this procedure: if more than one population remains in G, the process continues by selecting new candidate populations for hybridization; otherwise, the algorithm returns to the initial exploration stage. The pseudocode of the proposed HCG algorithm is presented as follows.

Pseudocode

N: The number of populations
S: The size of each population
G: The set of all X_k

Best-Solution = $+\infty$;
While termination condition is not satisfied, do
  for $k = 1$ to N
    Initialize the k^{th} population X_k;
    Update population X_k by **GWO**;
    Put updated X_k into G;
    Update Best-Solution;
  end
  Group-Number = $|G|$;
  Select the best population in G, which contain the Best-Solution and set it as P_b;
  While Group-Number > 1
    Get the second-best population P_s from G;
    Set population $P = P_s \cup P_b$;
    Remove P_s from G;
    Update population P using **CMA-ES**;
    Update Best-Solution;
    P_b = the best S individuals in P;
    Groups-Number = Groups-Number $- 1$;
  end
end
Return Best-Solution

4. Experimental Results and Discussion

In this work, the proposed hybrid algorithm of CMA-ES and GWO, referred to as HCG, is evaluated on 19 CEC 2019 benchmark functions that are widely used in the optimization literature. These classical benchmark functions have been selected to enable a fair and direct comparison between HCG and the original CMA-ES and GWO metaheuristic algorithms. The benchmark set includes unimodal, multimodal, and fixed-dimension multimodal functions, thereby providing a comprehensive assessment of algorithmic performance under different landscape characteristics. All benchmark problems considered in this study are minimization problems, and the details of these test functions are summarized in Table 1.

For each benchmark function, 200 independent runs are performed using a population size of 30 and a maximum number of 1000 generations. All simulations are implemented in MATLAB R2024a to ensure numerical consistency. For each function, the best, worst, mean, and standard deviation (SD) of the obtained fitness values are recorded and reported in Table 2. These statistical indicators provide an initial quantitative comparison of convergence quality, robustness, and solution stability among the competing algorithms.

In addition to numerical statistics, a graphical comparison based on box plots is employed to further analyze and visualize the distribution of errors obtained by CMA-ES, GWO, and the proposed HCG algorithm in Figure 1. Each box plot summarizes the log-scale final errors obtained over the 19 benchmark functions for a given algorithm. Since the benchmark functions often yield objective values spanning several orders of magnitude, especially near the global optimum, the box plots are constructed using the logarithmic scale of the absolute error. Specifically, the plotted value is defined as

$$\log_{10}(|\, f(x) - f^* \,|)$$

where $f(x)$ is the obtained solution value and f^*denotes the known global optimum. The use of a log-scale error representation is essential when comparing algorithms that achieve very small error values, as linear-scale plots tend to compress the differences and may obscure meaningful performance distinctions. In the log-scale box plots, lower values correspond to smaller optimization errors and thus higher solution accuracy. An algorithm is considered superior when its box plot is positioned closer to the bottom of the plot, with fewer outliers. This indicates not only higher accuracy but also greater robustness and consistency.

Table 1: Benchmark functions

OBJECTIVE FUNCTION	DIMENSION	LANDSCAPE	OPTIMAL VALUE
$f_1(x)=\sum_{i=1}^{n} x_i^2$	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
$f_4(x)=\max_i\left\{\lvert x_i\rvert, 1\le i\le n\right\}$	30	[-100,100]	0
$f_5(x)=\sum_{i=1}^{n}\left(\left[x_i+0.5\right]\right)^2$	30	[-100,100]	0
$f_6(x)=\sum_{i=1}^{n}-x_i\sin\left(\sqrt{\lvert x_i\rvert}\right)$	30	[-500,500]	-418.9829*30
$f_7(x)=\sum_{i=1}^{n}\left[x_i^2-10\cos\left(2\pi x_i\right)+10\right]$	30	[-5.12,5.12]	0
$f_8(x)=-20\exp\left(-0.2\sqrt{\frac{1}{n}\sum_{i=1}^{n}x_i^2}\right)-\exp\left(\frac{1}{n}\sum_{i=1}^{n}\cos\left(2\pi x_i\right)\right)+20+e$	30	[-32,32]	0
$f_9(x)=\frac{1}{4000}\sum_{i=1}^{n}x_i^2-\prod_{i=1}^{n}\cos\left(\frac{x_i}{\sqrt{i}}\right)+1$	30	[-600,600]	0
$f_{10}(x)=\frac{\pi}{n}\left\{10\sin(\pi y_i)+\sum_{i=1}^{n-1}(y_i-1)^2[1+10\sin^2(\pi y_{i+1})]+(y_n-1)^2\right\}$ $+\sum_{i=1}^{n}u(x_i,10,100,4)$ $y_i=1+\frac{x_i+1}{4}$, $u(x_i,a,k,m)=\begin{cases} k(x_i-a)^m & x_i>a \\ 0 & -a<x_i<a \\ k(-x_i-a)^m & x_i<-a \end{cases}$	30	[-50,50]	0
$f_{11}(x)=0.1\left\{\sin^2(3\pi x_i)+\sum_{i=1}^{n-1}(x_i-1)^2[1+\sin^2(3\pi x_i+1)]+(x_n-1)^2[1+\sin^2(2\pi x_n)]\right\}$ $+\sum_{i=1}^{n}u(x_i,5,100,4)$	30	[-50,50]	0
$f_{12}(x)=\left(\frac{1}{500}+\sum_{j=1}^{25}\frac{1}{j+\sum_{i=1}^{2}(x_i-a_{ij})^6}\right)^{-1}$	2	[-65,65]	1
$f_{13}(x)=\sum_{i=1}^{11}\left[a_i-\frac{x_1(b_i^2+b_i x_2)}{b_i^2+b_i x_3+x_4}\right]^2$	4	[-5,5]	0.00030
$f_{14}(x)=4x_1^2-2.1x_1^4+\frac{1}{3}x_1^6+x_1x_2-4x_2^2+4x_2^4)$	2	[-5,5]	-1.0316
$f_{15}(x)=(x_2-\frac{5.1}{4\pi^2}x_1^2+\frac{5}{\pi}x_1-6)^2+10(1-\frac{1}{8\pi})\cos x_1+10$	2	[-5,5]	0.398
$f_{16}(x)=[1+(x_1+x_2+1)^2(19-14x_1+3x_1^2-14x_2+6x_1x_2+3x_2^2)]$ $\times[30+(2x_1-3x_2)^2\times(18-32x_1+12x_1^2+48x_2-36x_1x_2+27x_2^2)]$	2	[-2,2]	3
$f_{17}(x)=-\sum_{i=1}^{4}c_i\exp\left(-\sum_{i=1}^{3}a_{ij}\left(x_j-p_{ij}\right)^2\right)$	3	[1,3]	-3.86
$f_{18}(x)=-\sum_{i=1}^{4}c_i\exp\left(-\sum_{i=1}^{6}a_{ij}\left(x_j-p_{ij}\right)^2\right)$	6	[0,1]	-3.32
$f_{19}(x)=-\sum_{i=1}^{10}\left[\left(X-a_i\right)\left(X-a_i\right)^T+c_i\right]^{-1}$	4	[0,10]	-10.5363

Table 2: The comparative result

	Algorithms	*MEAN*	*BEST*	*WORST*	*MEDIAN*	*SD*	*ERROR*
f_1	GWO	9.42E-35	2.99E-38	1.06E-32	9.18E-36	7.53E-34	2.99171E-38
	CMA-ES	4.54E-41	3.12E-43	9.64E-40	1.49E-41	9.93E-41	3.12E-43
	HCG	3.37061E-58	1.62414E-60	4.62408E-57	1.16317E-58	6.60013E-58	1.62414E-60
f_2	GWO	3.75E-21	1.14E-22	4.72E-20	2.23E-21	5.07E-21	1.13645E-22
	CMA-ES	1.03E-18	1.04E-19	6.94E-18	7.02E-19	9.54E-19	1.04E-19
	HCG	3.07E-27	3.66168E-28	1.81147E-26	2.3656E-27	2.41683E-27	3.66168E-28
f_3	GWO	1.77E-05	1.87E-12	0.002671221	1.01E-07	0.000189469	1.86753E-12
	CMA-ES	1.90E-20	1.16E-22	3.48E-19	6.23E-21	4.36E-20	1.16E-22
	HCG	1.31371E-49	2.65925E-52	2.48377E-48	4.56229E-50	2.61659E-49	2.65925E-52
f_4	GWO	2.11E-08	7.20E-11	2.78E-07	6.30E-09	3.92E-08	7.20116E-11
	CMA-ES	1.79E-13	8.09E-15	1.63E-12	1.18E-13	2.03E-13	8.09E-15
	HCG	4.77776E-22	4.73308E-23	2.59372E-21	3.87248E-22	3.80884E-22	4.73308E-23
f_5	GWO	1.784775145	0.432110696	2.990176149	1.756678565	0.514744479	0.432110696
	CMA-ES	7.87E-30	3.76E-30	1.47E-29	7.49E-30	2.04E-30	3.76E-30
	HCG	1.72631E-30	8.38165E-31	2.96747E-30	1.67479E-30	4.31756E-31	8.38165E-31
f_6	GWO	-5563.519187	-7717.929778	-2911.566299	-5621.98794	858.2011002	4851.557222
	CMA-ES	-118.3590488	-118.3590488	-118.3590488	-118.3590488	4.93E-14	1.25E+04
	HCG	-5668.616254	-8299.688725	-2798.182946	-5741.509958	1056.493344	4269.798275
f_7	GWO	1.857859883	0	45.26071169	1.14E-13	4.321796689	0
	CMA-ES	24.57547238	9.949590571	52.73269868	22.88404824	8.276463475	9.95E+00
	HCG	5.911907148	0	12.93446774	5.969754343	2.925539502	0
f_8	GWO	3.92E-14	2.89E-14	5.73E-14	3.95E-14	3.99E-15	2.88658E-14
	CMA-ES	7.46E-15	4.00E-15	7.55E-15	7.55E-15	5.56E-16	4.00E-15
	HCG	4.01E-15	4.00E-15	7.55E-15	4.00E-15	2.51E-16	3.9968E-15
f_9	GWO	0.00537415	0	0.066167305	0	0.010985278	0
	CMA-ES	0	0	0	0	0	0.00E+00
	HCG	0.000308049	0	0.012320989	0	0.001797421	0
f_{10}	GWO	0.121360409	0.026346215	0.647788228	0.100287076	0.099989573	0.026346215
	CMA-ES	0.002591726	3.66392E-31	0.10366902	8.53859E-31	0.016225936	3.66E-31
	HCG	0.000518345	4.60E-32	0.10366902	1.04E-31	0.007330507	4.59579E-32
f_{11}	GWO	1.256998704	0.528466342	1.941255636	1.269312022	0.277593706	0.528466342
	CMA-ES	0.000714179	4.19E-30	0.010987366	8.15E-30	0.002715467	4.19E-30
	HCG	1.61E-30	7.23E-31	3.21E-30	1.55E-30	4.73E-31	7.23473E-31
f_{12}	GWO	5.803010484	0.998003838	15.5038168	2.982105157	4.543295399	-0.001996162
	CMA-ES	12.67050581	12.67050581	12.67050581	12.67050581	1.52E-13	1.17E+01
	HCG	2.263897587	0.998003838	10.76318067	1.9920309	1.72507658	-0.001996162
f_{13}	GWO	0.006127689	0.000307492	0.056627951	0.00051279	0.010544073	0.000007492
	CMA-ES	0.082245953	0.000307486	8.193594814	0.000307486	0.81726698	7.49E-06
	HCG	0.000271063	0.000307486	-0.000817961	0.000307486	0.000594784	7.486E-06
f_{14}	GWO	-1.031628426	-1.031628453	-1.031628123	-1.031628439	4.10E-08	-2.8453E-05
	CMA-ES	-1.031628453	-1.031628453	-1.031628453	-1.031628453	4.69E-16	-2.85E-05
	HCG	-1.031627362	-1.031628453	-1.03143967	-1.031628363	1.33E-05	-2.8453E-05
f_{15}	GWO	0.397888237	0.397887358	0.397891165	0.397887962	8.53E-07	-0.000112642
	CMA-ES	0.529011634	0.397887358	26.62274256	0.397887358	1.854377295	-1.13E-04
	HCG	0.397887358	0.397887358	0.397887358	0.397887358	3.65E-16	-0.000112642
f_{16}	GWO	9.480095771	3.000000041	84.00016333	3.000063657	22.02987575	4.1E-08
	CMA-ES	125.27	3	24376	3	1723.412093	0.00E+00
	HCG	3	3	3	3	6.10E-15	0

f_{17}	*GWO*	-3.861287898	-3.862781968	-3.854898833	-3.862627433	0.002482409	-0.002781968
	CMA-ES	-2.837240782	-0.091332415	-3.862782148	-3.862782148	1.608866796	3.77E+00
	HCG	-3.861285317	-3.86278109	-3.85487034	-3.862409384	0.002302554	-0.00278109
f_{18}	*GWO*	-3.255792736	-3.321994294	-2.867555262	-3.32196678	0.0864322	-0.001994294
	CMA-ES	-1.414325288	-3.321995172	-0.027251594	-0.165720596	1.561967767	-2.00E-03
	HCG	-3.26221631	-3.32198683	-3.0379255	-3.321858998	0.078223354	-0.00198683
f_{19}	*GWO*	-10.26485673	-10.53635869	-2.421692272	-10.53551917	1.426548353	-5.869E-05
	CMA-ES	-5.128480787	-5.128480787	-5.128480787	-5.128480787	4.07E-15	5.41E+00
	HCG	-10.19388364	-10.53594119	-2.421039159	-10.52631662	1.546382918	0.00035881

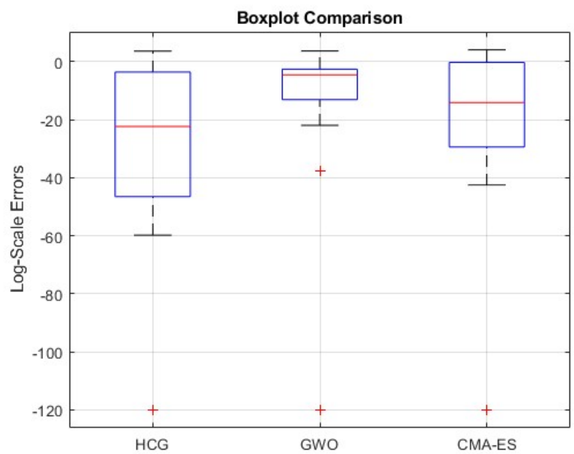

Figure 1: Boxplot Comparison

As observed in Figure 1, the proposed HCG algorithm achieves the lowest median log-scale error among all compared algorithms, demonstrating higher overall accuracy across the benchmark suite.

Overall, both the numerical statistics reported in Table 2 and the log-scale box plot visualizations confirm that the proposed HCG algorithm outperforms the original CMA-ES and GWO algorithms in terms of accuracy, convergence reliability, and robustness. This behavior confirms that the hybridization effectively balances global exploration, inherited from GWO's leader-based population guidance, and local exploitation, achieved through CMA-ES's covariance-based sampling mechanism.

5. Conclusion

This paper proposed a novel hybrid optimization algorithm, termed HCG, by effectively integrating the complementary strengths of CMA-ES and GWO. The motivation behind this hybridization was to address the inherent limitations of the individual algorithms, where CMA-ES exhibits strong local exploitation but limited global exploration, while GWO demonstrates powerful exploration capabilities with comparatively weaker exploitation near optimal regions.

The proposed HCG framework employs a multi-population strategy in which GWO is first utilized to explore the search space and identify promising regions using global population information. Subsequently, CMA-ES is applied to intensify the search within these regions through covariance-guided sampling, enabling precise exploitation. This cooperative mechanism allows HCG to maintain population diversity while simultaneously accelerating convergence toward high-quality solutions.

Extensive experimental evaluations were conducted on 19 well-known benchmark functions, including unimodal, multimodal, and fixed-dimension multimodal problems. The results demonstrate that HCG consistently outperforms the original CMA-ES and GWO algorithms in terms of convergence accuracy, robustness, and stability. Statistical comparisons and log-scale box plot analyses confirmed that HCG achieves lower median errors with reduced variability across different problem landscapes, indicating superior generalization and reduced sensitivity to problem characteristics.

Overall, the findings validate that the proposed hybridization strategy successfully balances exploration and exploitation, leading to a more reliable and efficient optimization algorithm. The strong and consistent performance of HCG across diverse benchmark functions suggests its suitability for solving complex real-world optimization problems. Future work may focus on adaptive control mechanisms, scalability to large-scale optimization tasks, and applications to constrained and multi-objective optimization problems.

Funding: This research received no external funding.
Conflicts of Interest: The authors declare no conflict of interest.
Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers.

References

Almufti, S. M., Ahmad, H. B., Marqas, R. B., & Asaad, R. R. (2021). Grey wolf optimizer: Overview, modifications and applications. *International Research Journal of Science, Technology, Education, and Management*, *1*(1), 1–1.

Aly, M., & Alotaibi, A. S. (2025). Hybrid Butterfly-Grey Wolf Optimization (HB-GWO): A Novel Metaheuristic Approach for Feature Selection in High-Dimensional Data. *Statistics, Optimization & Information Computing*, *13*(6), 2575–2600.

Beyer, H.-G., & Sendhoff, B. (2017). Simplify your covariance matrix adaptation evolution strategy. *IEEE Transactions on Evolutionary Computation*, *21*(5), 746–759.

Faris, H., Aljarah, I., Al-Betar, M. A., & Mirjalili, S. (2018). Grey wolf optimizer: A review of recent variants and applications. *Neural Computing and Applications*, *30*(2), 413–435.

Guo, Z., Hou, D., & He, Q. (2024). Hybrid Genetic Algorithm and CMA-ES Optimization for RNN-Based Chemical Compound Classification. *Mathematics*, *12*(11), 1684.

Igel, C., Hansen, N., & Roth, S. (2007). Covariance matrix adaptation for multi-objective optimization. *Evolutionary Computation*, *15*(1), 1–28.

Ikram, R. M. A., Goliatt, L., Kisi, O., Trajkovic, S., & Shahid, S. (2022). Covariance matrix adaptation evolution strategy for improving machine learning approaches in streamflow prediction. *Mathematics*, *10*(16), 2971.

Karmakar, B., Kumar, A., Mallipeddi, R., & Lee, D.-G. (2023). CMA-ES with exponential based multiplicative covariance matrix adaptation for global optimization. *Swarm and Evolutionary Computation*, *79*, 101296.

Mirjalili, S., Mirjalili, S. M., & Lewis, A. (2014). Grey wolf optimizer. *Advances in Engineering Software*, *69*, 46–61.

Nadimi-Shahraki, M. H., Taghian, S., & Mirjalili, S. (2021). An improved grey wolf optimizer for solving engineering problems. *Expert Systems with Applications*, *166*, 113917.

Sharma, A., Garg, A., & Bala, S. (2025). *Adaptive Hybridization of Particle Swarm Optimization and Grey Wolf Optimization*. *1*, 1–8.

Song, Z., Ren, C., & Meng, Z. (2024). Differential evolution with perturbation mechanism and covariance matrix based stagnation indicator for numerical optimization. *Swarm and Evolutionary Computation*, *84*, 101447.

Tawhid, M. A., & Ibrahim, A. M. (2020). A hybridization of grey wolf optimizer and differential evolution for solving nonlinear systems. *Evolving Systems*, *11*(1), 65–87.

Tbaishat, D., Tubishat, M., Makhadmeh, S. N., & Alomari, O. A. (2025). A hybrid bat and grey wolf optimizer for gene selection in cancer classification. *Knowledge and Information Systems*, *67*(1), 455–495.

Xu, P., Luo, W., Lin, X., Qiao, Y., & Zhu, T. (2019). *Hybrid of PSO and CMA-ES for global optimization*. 27–33.