BAREKENG: Journal of Mathematics and Its Applications
December 2025 Volume 19 Issue 4 Page 2803–2814
P-ISSN: 1978-7227 E-ISSN: 2615-3017
https://doi.org/10.30598/barekengvol19no4pp2803-2814

AN IMPROVED HYBRID CONJUGATE GRADIENT METHOD WITH SPECTRAL STRATEGY AND ITS APPLICATIONS IN COVID-19 PREDICTION

Kamilu Kamfa [1], **Rabiu Bashir Yunus** [2*], **Muhammad Auwal Lawan** [3]

[1,2,3] *Department of Mathematics, Faculty of Computing and Mathematical Sciences, Aliko Dangote University of Science and Technology Wudil, Kano, 713101, Nigeria*

*Corresponding author's e-mail: * **rabiubashir@kustwudil.edu.ng***

Article History:

Received: 3rd March 2025
Revised: 12th April 2025
Accepted: 10th June 2025
Available online: 1st September 2025

Keywords:

Algorithm;
Conjugate Gradient;
Covid-19;
Hybrid;
Monotone.

ABSTRACT

This paper introduces a hybrid conjugate gradient (CG) method for unconstrained optimization with a spectral strategy, inspired by key advancements in existing CG techniques. The proposed method guarantees a descent direction at every iteration, regardless of the line search scheme employed. Its global convergence is rigorously established under the Wolfe line search conditions. Numerical experiments on benchmark optimization problems demonstrate that the proposed method outperforms the FR and RMIL methods across multiple performance metrics. Furthermore, its effectiveness is showcased in a neural network framework for predicting chickenpox and COVID-19 infection cases, highlighting its practical applicability in real-world scenarios.

How to cite this article:

K. Kamfa, R. B. Yunus, and M. A. Lawan, "AN IMPROVED HYBRID CONJUGATE GRADIENT METHOD WITH SPECTRAL STRATEGY AND ITS APPLICATIONS IN COVID-19 PREDICTION," *BAREKENG: J. Math. & App.*, vol. 19, iss. 4, pp. 2803-2814, December. 2025.

Journal homepage: *https://ojs3.unpatti.ac.id/index.php/barekeng/*
Journal e-mail: *barekeng.math@yahoo.com; barekeng.journal@mail.unpatti.ac.id*

Research Article · Open Access

1. INTRODUCTION

Optimization, a science of choice [1] aims to identify the best way to solve a certain issue. The issue may be maximizing profit, minimizing expenses, or minimizing energy use. Numerous branches of science, engineering, and economics deal with these problems [2]. When a limitation is placed on the choice variable $x_1, x_2, x_3, \cdots, x_n$ the optimization problem can be confined; otherwise, it can be unconstrained [3]. Numerous techniques, particularly unconstrained optimization, have been employed to tackle optimization issues, including the conjugate gradient method (CG), the Newton method, the BFGS approach, and the steepest descent [4], [5], [6], [7] . Due to its ease of use, low memory requirements, and ability to find the Hessian matrix and invert it, the CG technique has been chosen over the other approaches [3].

The CG method uses an iteration formula to build sequences of iterations. This formula consists of finding the search direction. The search direction is a linear combination of the gradient, the previous search direction, and the conjugate parameter. Fletcher and Reeves (FR) [8], Dai and Yuan [9], Polak–Ribiere–Polyak (PRP) [6], and [10], Liu and Storey (LS) [11], and Hestenes and Stiefel (HS) [7]. There are a few examples of standard CG approaches. To a certain degree, the traditional CG approach can also be supplemented with the work of Rivaie et al. [3] and Wei, Yao, and Liu method [12]. Since they perform similarly to the six conventional CG algorithms [13].

Optimization plays a fundamental role in various scientific and engineering disciplines, particularly in machine learning, data analysis, and predictive modeling. Conjugate gradient (CG) methods are among the most effective iterative techniques for solving unconstrained optimization problems, owing to their simplicity, low memory requirements, and strong convergence properties. Over the decades, numerous CG variants have been developed to enhance convergence speed and robustness. However, the classical CG methods, such as the Fletcher-Reeves (FR) and Polak-Ribière-Polyak (PRP) methods, sometimes suffer from slow convergence and inefficiency when applied to large-scale optimization problems. One promising approach to improving CG methods is incorporating spectral strategies, which dynamically adjust step sizes to accelerate convergence. Spectral CG methods integrate the benefits of spectral gradient methods and classical CG approaches, yielding more efficient optimization techniques. In this research, we propose a new hybrid CG method that leverages a spectral strategy to ensure global convergence and robust performance across various optimization scenarios.

The ongoing COVID-19 pandemic has underscored the importance of accurate and efficient predictive models in epidemiology. Traditional statistical models and machine learning approaches rely heavily on optimization techniques to train neural networks and fit complex predictive functions. Given the necessity for reliable forecasting tools, optimization methods that enhance predictive models' efficiency are paramount. Our proposed hybrid CG method is particularly suited for training neural networks due to its descent property and improved convergence characteristics.

This work aims to develop a hybrid CG method incorporating a spectral strategy to enhance the efficiency and robustness of unconstrained optimization. The proposed method is supported by a theoretical analysis that establishes its descent properties and guarantees global convergence under the Wolfe conditions. To validate its effectiveness, the method is benchmarked against classical CG algorithms, including the Fletcher–Reeves (FR) and RMIL variants, using a set of standard numerical test problems. Additionally, the practical utility of the proposed approach is demonstrated by applying it to improve predictive modeling in epidemiology, particularly for forecasting COVID-19 infection cases.

Please refer to [2] and [14] for most of the research on the CG technique that focused on changing these parameters in terms of convergence, numerical performances, and applications. The FR method is globally convergent under exact line search, according to Zoutendijk [15], based on certain assumptions. However, it performs poorly in numerical computation because the search direction exhibits zigzagging behaviour and continues recycling without moving toward the solution point (refer to Rivaie et al.'s work [3]). Powell produced a counterwork that refuted the Zoutendijk work [16] and [17]. Because of its associated restart mechanism $g_{k+1}^T(g_{k+1} + g_k)$, PRP and its version were demonstrated to be numerically sound but have global convergence issues. For general functions, Powell [19] has demonstrated that PRP is not globally convergent under certain inexact line searches, such as the Strong Wolfe line search. Hager has noted that the PRP parameter's search direction is not descending for the generic function [18].

On a short note, the convergence of the PRP approach and its variations has been examined by a few scholars, including Dai et al. [20], Andrei [13], and Zhang and Hager [18]. Many researchers, including

Touati-Ahmed and Storey [21], Al-Baali [22], Gilbert and Nocedal [23], Rivaie et al. [3], Sulaiman et al. [24], and Kamilu et al. [14], have filled the void left by their contributions by suggesting new PRP-like algorithms using various inexact line searches, with promising results. To accomplish the performance of β_k^{FR}, β_k^{PRP}, and β_k^{RMIL}. In this research, we develop a new conjugate parameter; we employed a hybrid technique [25] to trap the good performance and convergence results [26], [27]. We intend to test the performance of this noble approach on some data sets in neural networks.

The structure of this paper is organized as follows: Section 2 introduces the proposed method and its algorithmic framework, accompanied by a theoretical analysis of its global convergence under certain assumptions. Section 3 presents numerical experiments to assess the performance of the new approach in comparison with established methods and further illustrates its application to a COVID-19 prediction problem. Finally, Section 4 offers a summary of the findings, practical recommendations, and suggestions for future research directions.

2. RESEARCH METHOD

The CG method seeks the minimizer, which we refer to as the solution point x^* is using the iteration formula in **Equation (1)** below, this formula builds a sequence of iterations $\{x_k\}$ which in many situations converges to the solution point [13].

$$x_{k+1} = x_k + \alpha_k d_k \tag{1}$$

where the previous and current iteration points are denoted by t_k and t_{k+1} respectively. α_k is the step length that is required to be moderate, not too wide or too small [4], and it must ensure there is an adequate reduction of the function value with minimum cost [8].

When determining the step size, there are two noteworthy methods to employ. Specifically, the exact and inexact line search. The first method provides an exact step length value, but it is more expensive and time-consuming [4]. To get a cheap step length with minimal computing cost, researchers like Armijo, Goldstein, and Wolfe provide the first inexact line search, which they term the Armijo, Goldstein, and Wolfe line search [29], [30], [31]. The strong Wolfe line search is created by modifying the ordinary Wolfe line search in the manner described below.

$$f(x_k + \alpha_k d_k) \le f(x_k) + \mu \alpha_k \nabla f(x_k)^T d_k \tag{2}$$
$$|\nabla f(x_k + \alpha_k d_k)^T d_k| \le \sigma |\nabla f(x_k)^T d_k|$$

where μ, is between $0 < \mu < \frac{1}{2}$, and ∇, $0 < \nabla < \frac{1}{2}$. Different constant values will lead to differences in the exact line search. Zhong Wan et al. suggested a modification of the Armijo line search [32]. Other researchers like Shuai Huang et al. [19] have proposed a non-monotone line-search-like method, which is an improvement over the non-monotone line-search technique proposed by Zhang and Hager [18]. The α_k is chosen as a trail $\alpha_k = \alpha_k \rho^{h_k} \le \mu$ where h_k is the largest integer,

$$C_{k+1} = \frac{\eta_k Q_k C_k + f(x_k + \alpha_k d_k)}{Qk + 1 + \sigma_k g_k^T d_k} \le C_k$$

For other approaches to finding step length, please see, for example, [28], [33], [34], and [35]. d_k is a search direction. At the initial stage, the CG method employed the steepest descent direction $d_k = -g_k$ where g_k is the gradient at a point k. This is the best direction for the search for the minimum of the objective function f [4]. While in the subsequent iteration point, it uses the linear combination of the gradient, the previous direction, and the conjugate parameter $d_{k+1} = g_k + \beta_k d_k$. Some CG methods consider the three-term linear combination approach, like in [20], [38], and [24] or the fourth-term approach like [39],[40].

Different CG method uses different conjugate parameters β_k. Andrei [4] mentioned the standard CG as the one with either the $\|g_k\|^2$ or $g_{k+1}^T(g_{k+1} - g_k)$ in the numerator and $\|g_k\|^2$ or $d_k^T(g_{k+1} - g_k)$ or $-d_k^T g_k$ in the denominator. This is due to their good numerical results, or they are globally convergent.

A hybrid CG approach was created to make use of the worldwide convergence of the FR method and its variant, as well as the strong performance of the PRP method and its variant [21], [41], [42]. Touti Ahmed et al. proposed the first hybrid CG.

$$\beta_k^{TA} = \begin{cases} \dfrac{g_k^T(g_k - g_{k-1})}{\|g_{k-1}\|^2}, & \text{if } 0 < \dfrac{g_k^T(g_k - g_{k-1})}{\|g_{k-1}\|^2} \leq \dfrac{\|g_k\|^2}{\|g_{k-1}\|^2}, \\ \dfrac{\|g_k\|^2}{\|g_{k-1}\|^2}, & \text{otherwise} \end{cases}$$

The PRP method was incorporated to address the jamming behaviour of the FR method. Hu and Storey [43], further proposed a max-min parameter as follows

$$\beta_k^{HU} = max(0, (min(\frac{g_k^T(g_k - g_{k-1})}{\|g_{k-1}\|^2}, \frac{\|g_k\|^2}{\|g_{k-1}\|^2})))$$

The convex combination of PRP and DY methods was created by Andrei et al. [41] in order to preserve the good numerical results of the PRP approach while utilizing the good convergence property of the DY method.

$$\beta_k^{A} = (1 - \theta_k)\frac{g_k^T(g_k - g_{k-1})}{\|g_{k-1}\|^2} + \theta_k \frac{\|g_k\|^2}{d_{k-1}^T(g_k - g_{k-1})}$$

where $0 < \theta_k \leq 1$ is scalar defined as follows:

$$\theta_k = \frac{(y_k^T g_k)(y_k^T s_k) - (y_k^T g_k)(g_{k-1}^T g_{k-1})}{(y_k^T g_k)(y_k^T s_k) - (g_k^T g_k)(g_{k-1}^T g_{k-1})}$$

Irrespective of any line search used, this parameter is designed to ensure the conjugacy rule is fulfilled [44]. Real-world issues have been resolved using the CG technique [24], [45], and [37].

To include the benefits of CG methods, Jian et al. [44] created a hybrid CG method that incorporates the conjugate parameter in recent years.

$$\beta_k^{N} = \frac{\|g_k\|^2 - max\left\{\frac{\|g_k\|}{\|g_{k-1}\|} g_k^T g_{k-1}, 0\right\}}{max\{\|g_{k-1}\|^2, d_{k-1}^T y_{k-1}\}} \tag{3}$$

The direction of search with Equation (3) is always descending, regardless of the line search. For general functions, global convergence is also achieved using a weak search for the Wolfe line. We suggest a hybrid approach based on the method's exceptional properties, which guarantees that the search direction achieves adequate descent irrespective of the line search employed.

To accomplish this, based on β_k^{FR}, β_k^{PRP}, andβ_k^{RMIL}. To develop a new conjugate parameter, we employed a hybrid technique, which is specified as

$$\beta_k^{IHCG} = \frac{\|g_k\|^2 - max\{0, g_k^T g_{k-1}\}}{max\{\|g_{k-1}\|^2, \|d_{k-1}\|^2\}}. \tag{4}$$

We defined the new composite search direction as follows:

$$d_k = \begin{cases} -g_k + \beta_k^{IHCG} d_{k-1}, & \text{if } |g_k^T g_{k-1}| \leq \|g_{k-1}\|^2, \\ -\lambda_k^* g_k + \beta_k^{IHCG} d_{k-1}, & \text{otherwise} \end{cases} \tag{5}$$

where the control parameter $\lambda_k^* > 1$ serves as a spectral parameter.

Algorithm 1. An Improved Hybrid Conjugate Gradient Algorithm (IHCG)

Step 1. Initialization: $x_0 \in \mathbb{R}^n$

Step 2. If $\|g_k\| \leq \epsilon$, end. If not, proceed to Step 3.

Step 3. Calculate the step length (α_k) by using the Wolfe line search.

Step 4. Create the subsequent iteration by $(x_{k+1} = x_k + \alpha_k d_k)$;

Step 5. Compute $\left(g_k := g(x_k)\right)$ and (β_k^{IHCG}) by **Equation (4)**.

Step 6. Let $d_k = \begin{cases} -g_k + \beta_k^{IHCG} d_{k-1}, & \text{if } |g_k^T g_{k-1}| \le \|g_{k-1}\|^2, \\ -\lambda_k^* g_k + \beta_k^{IHCG} d_{k-1}, & \text{otherwise.} \end{cases}$

Step 7. Set $(k := k + 1)$, return to Step 2.

Below, we demonstrate the algorithm framework's global convergence under the Wolfe line search condition in **Equation (2)**, which means that **Equation (2)** determines the step length. Furthermore, the following lemma asserts that all search directions in the Algorithm 1 framework are descent paths, irrespective of the line search approach used. It is important to make the following assumptions to accomplish this goal.

Assumption1. The objective function $f(x)$ is bounded from below on the level set $\Omega = \{x \in \mathbb{R}^n \mid f(x) \le f(x_0)\}$, where x_0 is the initial point.

Assumption 2. In a neighbourhood N of the level set Ω. The objective function $f(x)$ is continuously differentiable, and its gradient $g(x) = \nabla f(x)$ satisfies the Lipschitz condition. That is, there exists a constant $L > 0$ such that

$$\|g(x) - g(y)\| \le L\|x - y\|, \quad \forall x, y \in N.$$

Lemma 1. *Let Algorithm 1 generate d_k, if the objective function $f(x)$ is continuously differentiable. Then, for any $k \ge 1$, $g_k^T d_k < 0$.*

Proof.

For $k = 1$, it is evident that $g_1^T d_1 = -\|g_1\|^2 < 0$. Assume that $g_{k-1}^T d_{k-1} < 0$ holds for $k = 1$. To establish $g_k^T d_k < 0$ for $k > 1$, we divide the proof into the following four cases. If $\beta_k^{IHCG} = 0$, from **Equation (4)**, it follows that $g_k^T d_k = -\|g_k\|^2 < 0$. Therefore, in the subsequent analysis, we always assume $\beta_k^{IHCG} \ne 0$.

Case I. If $g_k^T g_{k-1} \le 0$ and $\|d_{k-1}\|^2 \ge \|g_{k-1}\|^2$, then from **Equation (4)** we obtain, $\beta_k^{IHCG} = \frac{\|g_k\|^2}{\|d_{k-1}\|^2} = \beta_k^{RMIL}$, noting that $\|g_{k-1}\|^2 > 0$, from Algorithm 1, so $\|d_{k-1}\|^2 > 0$ holds. Therefore, we obtain

$$g_k^T d_k = g_k^T(-g_k + \beta_k^{IHCG} d_{k-1}) = -\|g_k\|^2 + \frac{\|g_k\|^2}{\|d_{k-1}\|^2} g_k^T d_{k-1} \le \frac{\|g_k\|^2 g_k^T d_{k-1}}{\|d_{k-1}\|^2} < 0.$$

Case II. If $g_k^T g_{k-1} \le 0$ and $\|d_{k-1}\|^2 < \|g_{k-1}\|^2$, then, from **Equation (4)**, one has $\beta_k^{IHCG} = \frac{\|g_k\|^2}{\|g_{k-1}\|^2} = \beta_k^{FR}$ noting that, $\|g_{k-1}\|^2 > 0$. Therefore, we obtain

$$g_k^T d_k = g_k^T(-g_k + \beta_k^{IHCG} d_{k-1}) < -\|g_k\|^2 + \frac{\|g_k\|^2}{\|g_{k-1}\|^2} g_k^T d_{k-1} \le \frac{\|g_k\|^2 g_k^T d_{k-1}}{\|g_{k-1}\|^2} < 0.$$

Case III. If $g_k^T g_{k-1} \ge 0$ and $\|d_{k-1}\|^2 \ge \|g_{k-1}\|^2$, then, from **Equation (4)**, one has $\beta_k^{IHCG} = \frac{\|g_k\|^2 - g_k^T g_{k-1}}{\|d_{k-1}\|^2} = \beta_k^* = \frac{\|g_k\|^2 - g_k^T g_{k-1}}{\|d_{k-1}\|^2}$, therefore, we obtain

$$g_k^T d_k = g_k^T(-g_k + \beta_k^{IHCG} d_{k-1}) = -\|g_k\|^2 + \frac{\|g_k\|^2 - g_k^T g_{k-1}}{\|d_{k-1}\|^2} g_k^T d_{k-1} \le \frac{\|g_k\|^2 g_k^T d_{k-1}}{\|d_{k-1}\|^2} < 0$$

Case IV. If $g_k^T g_{k-1} \ge 0$ and $\|d_{k-1}\|^2 \le \|g_{k-1}\|^2$, then $\beta_k^{IHCG} = \frac{\|g_k\|^2 - g_k^T g_{k-1}}{\|d_{k-1}\|^2} = \beta_k^{PRP} = \frac{\|g_k\|^2 - g_k^T g_{k-1}}{\|g_{k-1}\|^2}$, therefore, we obtain

$$g_k^T d_k = g_k^T(-g_k + \beta_k^{IHCG} d_{k-1}) = -\|g_k\|^2 + \frac{\|g_k\|^2 - g_k^T g_{k-1}}{\|g_{k-1}\|^2} g_k^T d_{k-1} \le \frac{\|g_k\|^2 g_k^T d_{k-1}}{\|g_{k-1}\|^2} < 0. \blacksquare$$

The well-known Zoutendijk condition [15] are presented in the following lemma, and a thorough proof can be found in [46] and [15].

Lemma 2. *Assume that assumptions 1 and 2 are true. In the common iteration $x_{k+1} = x_k + \alpha_k d_k$, if the direction d_k satisfies $g_k^T d_k < 0$ and the step length α_k meets the Wolfe line search condition in* **Equation (2)**, *then*

$$\sum_{k=1}^{\infty} \frac{(g_k^T d_k)^2}{\|d_k\|^2} < \infty \tag{6}$$

Corollary 1. *Assume that $\{d_k\}$ is generated by Algorithm 1. Based on Assumption 1, we have*

$$\sum_{k=0}^{\infty} \frac{\|g_k\|^4}{\|d_k\|^2} < \infty \tag{7}$$

Proof. From the Zoutendijk condition in **Equation (6)** and the descent property, we have

$$\sum_{k=0}^{\infty} \frac{\|g_k\|^4}{\|d_k\|^2} \leq \frac{1}{\vartheta} \sum_{k=1}^{\infty} \frac{(g_k^T d_k)^2}{\|d_k\|^2} < \infty. \blacksquare$$

The following theorem shows the suggested method's global convergence.

Theorem 1. *Suppose that Assumptions 1 and 2 hold. Let $\{x_k\}$ be generated by Algorithm 1. Then*

$$\lim_{k \to \infty} \inf \|g_k\| = 0. \tag{8}$$

Proof. From **Equation (4)** and **Equation (5)**, we have

$$\|d_k\| = \left\| -\lambda_k^* g_k + \beta_k^{IHCG} d_{k-1} \right\|$$

$$\leq |\lambda_k^*| \|g_k\| + \left|\beta_k^{IHCG}\right| \|d_{k-1}\|$$

By simplifying the above, we obtain

$$\leq \left(1 + \frac{\|g_k\|}{\|d_{k-1}\|}\right) \|g_k\|$$

Hence, we have

$$\frac{1}{\|d_k\|} \geq \frac{1}{\gamma} \|g_k\| \tag{9}$$

By setting $\gamma := 1 + \frac{Lb}{\varphi}$

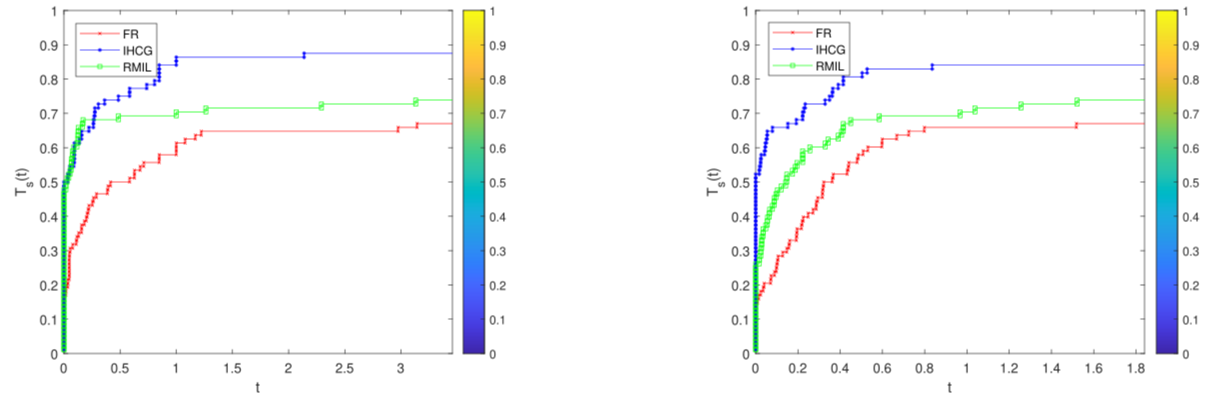
Considering **Equation (7)** and **Equation (9)**, we obtain

$$\sum_{k=0}^{\infty} \|g_k\|^2 \leq \gamma^2 \sum_{k=0}^{\infty} \frac{\|g_k\|^4}{\|d_k\|^2} < \infty$$

This leads to **Equation (8)**. Consequently, the proof is finished. ■

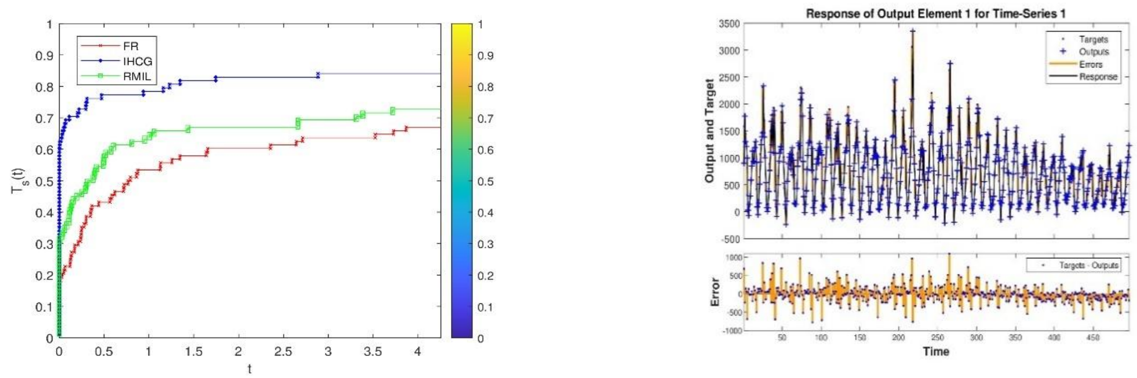
3. RESULTS AND DISCUSSION

The numerical performance of the novel approach in comparison to the traditional methods is presented in this section. Consideration has been given to a significant number of conventional test functions, ranging from minor to major issues. Every issue originates from [4]. Every problem is rerun five times with varying numbers of variables (from 1000 to 10,000). The robustness of the IHCG approach is established using the performance profile by Dolan and More [47]. Numerous researchers have employed this well-known profile [3], [1], [28], and [36]. When there are more than 1000 iterations, failure is indicated. **Table 1**, **Table 2**, and **Table 3** show each method's numerical performance according to the number of iterations, function evaluations, and CPU time.



(a) **(b)**

Figure 1. (a) Performance Overview Based on Function Evaluation and (b) Iteration Count



(a) **(b)**

Figure 2. (a) Performance Overview Based on the CPU time. (b) Performance Results of the First Application

Table 1. Performance of the IHCG in Terms of the Number of Iterations when Compared with FR and RMIL Methods

No.	Description	FR Method	IHCG Method	RMIL
1	Minimum among them	15.000	42.0000	44.0000
2	Percentage of better	0.1705	0.5000	0.4773
3	Number of failures	29.0000	11.0000	23.0000
4	Percentage of failure	0.6705	0.8750	0.7386

Table 2. Performance of the IHCG in Terms of the Function Evaluation when Compared with FR and RMIL Methods

No.	Description	FR Method	IHCG Method	RMIL
1	Minimum among them	13.0000	46.0000	23.0000
2	Percentage of better	0.1477	0.5227	0.2614
3	Number of failures	29.0000	14.0000	23.0000
4	Percentage of failure	0.6705	0.8409	0.7386

Table 3. Performance of the IHCG in Terms of the CPU time when Compared with FR and RMIL Methods

No.	Description	FR Method	IHCG Method	RMIL
1	Minimum among them	13.0000	53.0000	27.0000
2	Percentage of better	0.1477	0.6023	0.3068
3	Number of failures	29.0000	14.0000	24.0000
4	Percentage of failure	0.6705	0.8409	0.7273

All the methods utilized the same line search procedure, with stopping conditions as constant parameters. To carry out a comparison, a MATLAB code was developed and run on a PC Intel Core i5 processor, 16 GB RAM, 64-bit Windows 8 operating system. The results indicate that the IHCG method demonstrates competitive performance compared to the FR and RMIL methods. Notably, the minimum value obtained by IHCG (42.0000) is significantly higher than that of the FR method (15.0000) but slightly lower than RMIL (44.0000). This suggests that IHCG is more effective in achieving better minimum values than the FR method, but slightly underperforms compared to the RMIL method.

In terms of the percentage of better outcomes, IHCG achieves 50.00%, which is considerably higher than FR (17.05%) and slightly better than RMIL (47.73%). This indicates that IHCG frequently outperforms the other methods in terms of optimality. However, when considering the number of failures, IHCG records 11 failures, which is significantly lower than FR (29 failures) and RMIL (23 failures). This highlights the robustness of IHCG, as it is less prone to failure in comparison to the other two methods. Despite this, the percentage of failure for IHCG is 87.50%, which is relatively high compared to RMIL (73.86%) and FR (67.05%). This suggests that while IHCG performs well in finding better solutions, it may still struggle in some cases, leading to a relatively higher failure rate. Overall, the IHCG method strikes a balance between achieving better solutions and minimizing failures, making it a promising alternative to the FR and RMIL methods. However, further analysis may be needed to reduce its failure rate while maintaining or improving its performance.

In this part, we look at forecasting the number of infections using data sets for COVID-19 and chickenpox, respectively, by applying the CG method to a feed-forward neural network. The neural network mimics the brain's workings. The brain is a massive neuronal network. Neural networks are made up of connections between nodes that have weights and biases, just as the brain is made up of connections between neurons. The feed-forward neural network with backpropagation is the most significant neural network **[48]**. Numerous studies have demonstrated the significance of neural networks in prediction **[2]**, **[14]**, **[45]**, and **[48]**.

In the search for the global minimum, the feed-forward neural network used the conjugate gradient method among other methods to update its weight function. The CG method is prepared due to its rapid convergence when compared with Steepest Descent methods **[45]**. Predicting the number of infections can be considered as a univariate time series forecasting problem. The input data of the problem is the past, and the expected output is the future values.

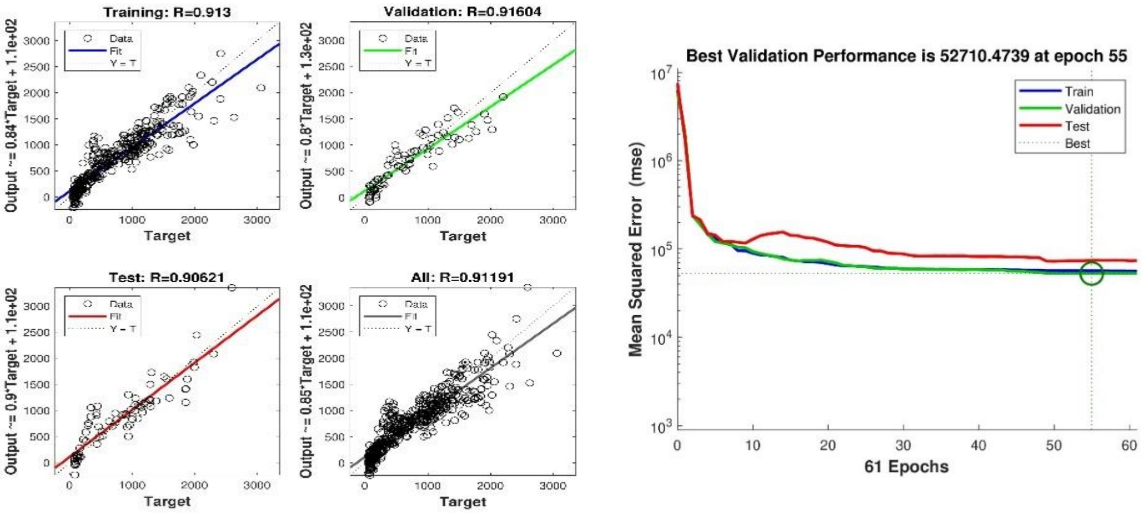
For the first application, we consider the Monthly chickenpox instances dataset. This data set is used to train a neural network to predict monthly cases of chicken pox. Chickenpox data consists of a 1x498 cell array of scalar values representing 498 months of chickenpox cases in New York City.

We design a network with the input layer, some hidden layers, and output layers. In the hidden layers, we have the weight matrix at $k-1$ layers and biases at layers. In the training, we used the backpropagation of error derivatives to update the weights for better results. The CG training process can be seen by minimizing the mean square error function state by

$$MSE = E(e^T e) \tag{10}$$

It is crucial to note that the CG approach is essential to finding the ideal weight, which can then be applied to reduce the error function in **Equation (10)**. Our challenge is a single time-series prediction, which is estimating a time series next value based on its historical values. Using an auto-regression time-series problem with an NAR Neural Network, we show the effectiveness of the model in forecasting the number of cases using our new conjugate gradient.

The data was divided into three categories: 70% for training, 15% for validation, and 15% for testing. The performance function that is selected is the mean square error. The results are displayed in **Figure 3** and **Figure 4**.



(a) **(b)**

Figure 3. (a) Regression Performance and (b) Validation Performance of the First Application

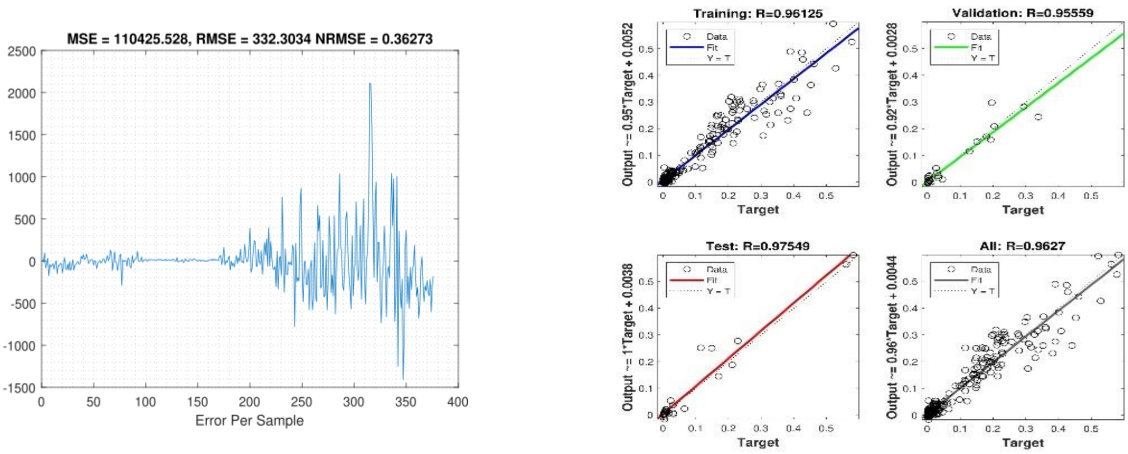
In the second application, we use a feed-forward neural network on a COVID-19 data set in Nigeria to show how well the novel method performs on a single time-series prediction. The Nigerian Centre for Disease Control provides us with the information. For a specific amount of time, it comprises a daily number of infections. The time series data are divided into 80% training data and 20% test data. The designed NN model can be classified into three categories. The first is the initialization of the required parameters; in this category, the data is collected and transformed into a standard form by normalizing it using

$$v = \frac{0.8(x-a)}{b-a} + 0.1$$

In the second category, we set all the parameters for our NN model as follows: Feed Forward Layers = 4, with the number of neurons as follows: first layer 20, the second layer 20, and the third layer 15. 100 was set as the maximum number of epochs. For the training function, we used the proposed CG method. The outcome of the best test is presented in Table 4.

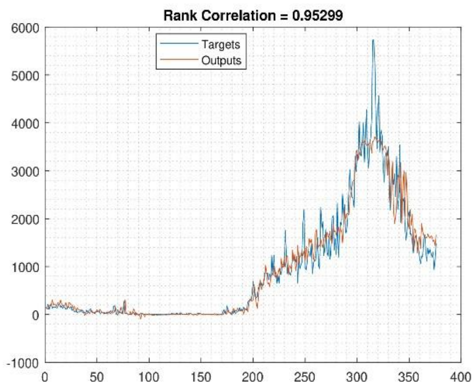
Table 4. Performance of the IHCG on the COVID-19 Data Set

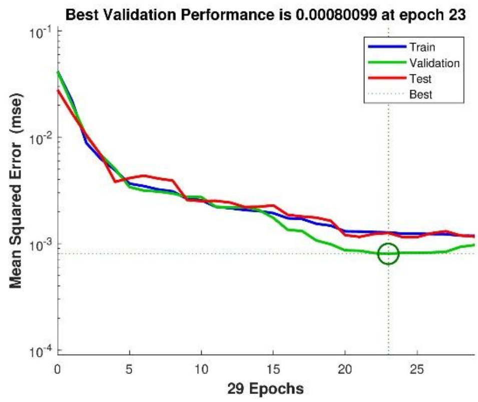
Epoch	Time	Performance	Gradient	Validation Checks	Step Size
0/100	7.075	0.0414/0	0.40268/1e-10	0/6	1/1e-06
25/100	7.58	0.0012409/0	0.0021256/1e-10	2/6	0.002081/1e-06
29/100	7.609	0.0011843/0	0.0040409/1e-10	6/6	0.01332/1e-06



(a) **(b)**

Figure 4: (a) MSE and (b) Regression Performance of the Second Application





(a) **(b)**

Figure 5. (a) Rank Correlation and (b) Validation Performance of the Second Application

4. CONCLUSION

A hybrid CG approach to unconstrained optimization issues was introduced in this study. The suggested approach outperforms the FR and RMIL methods in terms of effectiveness and efficiency, according to comparative analysis. The convergence study verified that the approach reaches global convergence under an imprecise line search and satisfies the necessary descent criterion. Its usefulness goes beyond optimization as well, since it has been shown to be a dependable technique for time series prediction, including the prediction of COVID-19 and chickenpox cases. These results demonstrate the method's resilience and adaptability in real-world predictive modeling and mathematical optimization. We observed that as the number of parameters or decision variables increases, the search space expands exponentially, thereby intensifying the difficulty of finding optimal solutions. This phenomenon, often referred to as the curse of dimensionality, poses a significant challenge, especially for the proposed method. Future research could benefit from extending conjugate gradient (CG) techniques to broader application domains, including optimal control, parameter estimation, data assimilation, and imaging problems.

AUTHOR CONTRIBUTIONS

Kamilu Kamfa: Conceptualization, Funding Acquisition, Investigation, Methodology, Writing - Original Draft. Rabiu Bashir Yunus: Formal Analysis, Methodology, Software, Visualization, Writing - Review and Editing. Muhammad Auwal Lawan: Project Administration, Supervision, Resources, Validation, Writing - Original Draft. All authors discussed the results and contributed to the final manuscript.

FUNDING STATEMENT

This study was funded by the Government of Nigeria through the Institute-Based Research (IBR) TetFund Research Grant for the project titled ‘Predicting and Pattern Analysis of a Novel Conjugate Gradient Method Utilizing Exact and Inexact Line Search: Applications in COVID-19 Modelling.

ACKNOWLEDGMENT

The authors would like to acknowledge the financial support provided by the Government of Nigeria through the Institute-Based Research (IBR) TetFund Research Grant, and the staff of the Directorate of Research and Innovation (DRIP) Aliko Dangote University of Science and Technology, Wudil.

CONFLICT OF INTEREST

The authors declare no conflicts of interest to report study.

REFERENCES

[1] K. U. Kamfa, A. A. Mustafa Mamat, M. Rivaie, P. L. B. Ghazali, and Z. Salleh, "ANOTHER MODIFIED CONJUGATE GRADIENT COEFFICIENT WITH GLOBAL CONVERGENCE PROPERTIES," *Applied Mathematical Sciences*, vol. 9, no. 37, pp. 1833–1844, 2015, doi: https://doi.org/10.12988/ams.2015.411996.

[2] K. K. Ibrahim Mohammed Sulaiman Mustafa Mamat and M. Danlami, "A DESCENT MODIFICATION OF CONJUGATE GRADIENT METHOD FOR OPTIMIZATION MODELS," *Iraqi Journal of Science*, vol. 7, no. 61, 2020, [Online]. Available: https://ijs.uobaghdad.edu.iq/index.php/eijs/article/view/1654

[3] M. Rivaie, M. Mamat, L. W. June, and I. Mohd, "A NEW CLASS OF NONLINEAR CONJUGATE GRADIENT COEFFICIENTS WITH GLOBAL CONVERGENCE PROPERTIES," *Appl Math Comput*, vol. 218, no. 22, pp. 11323–11332, 2012, doi: https://doi.org/10.1016/j.amc.2012.05.030 .

[4] N. Andrei, "NONLINEAR CONJUGATE GRADIENT METHODS FOR UNCONSTRAINED OPTIMIZATION.," *Springer, Berlin*, 2020.doi: https://doi.org/10.1007/978-3-030-42950-8

[5] Z. Dai, "COMMENTS ON A NEW CLASS OF NONLINEAR CONJUGATE GRADIENT COEFFICIENTS WITH GLOBAL CONVERGENCE PROPERTIES," *Appl Math Comput*, vol. 276, pp. 297–300, 2016, doi: https://doi.org/10.1016/j.amc.2015.11.085 .

[6] E. Polak and G. Ribiere, "NOTE SUR LA CONVERGENCE DE MÉTHODES DE DIRECTIONS CONJUGUÉES.," *ESAIM: Math. Model. Numer. Anal.* , vol. R1, no. 3, pp. 35–43, 1969.doi: https://doi.org/10.1051/m2an/196903R100351

[7] E. Stiefel, "METHODS OF CONJUGATE GRADIENTS FOR SOLVING LINEAR SYSTEMS.," *J. Res. Natl. Bur. Stand.*, vol. 49, pp. 409–435, 1952.doi: https://doi.org/10.6028/jres.049.044

[8] M. J. D. Fletcher R Powell, "A RAPIDLY CONVERGENT DESCENT METHOD FOR MINIMIZATION.," *Comput. J.*, vol. 2, no. 49, pp. 163–168, 1963.doi: https://doi.org/10.1093/comjnl/6.2.163

[9] Y. H. Dai and Y. Yuan, "A NONLINEAR CONJUGATE GRADIENT METHOD WITH A STRONG GLOBAL CONVERGENCE PROPERTY," *SIAM Journal on Optimization*, vol. 10, no. 1, pp. 177–182, 1999, doi: https://doi.org/10.1137/S1052623497318992.

[10] E. Polak and G. Ribiere, "THE CONJUGATE GRADIENT METHOD IN EXTREMAL PROBLEMS.," *USSR Comput. Math. Math. Phys.*, vol. 4, no. 9, pp. 94–112, 1969.doi: https://doi.org/10.1016/0041-5553(69)90035-4

[11] Y. Liu and C. Storey, "EFFICIENT GENERALIZED CONJUGATE GRADIENT ALGORITHMS," *part 1: theory. J. Optim. Theory Appl.* , vol. 1, no. 69, pp. 129–137, 1991.doi: https://doi.org/10.1007/BF00940464

[12] Z. Wei, S. Yao, and L. Liu, "THE CONVERGENCE PROPERTIES OF SOME NEW CONJUGATE GRADIENT METHODS," *Appl Math Comput*, vol. 183, no. 2, pp. 1341–1350, 2006, doi: https://doi.org/10.1016/j.amc.2006.05.150 .

[13] N. Andrei, "OPEN PROBLEMS IN NONLINEAR CONJUGATE GRADIENT ALGORITHMS FOR UNCONSTRAINED OPTIMIZATION.," *Bulletin of the Malaysian Mathematical Sciences Society. Second Series*, vol. 34, no. 2, pp. 319–330, 2011.

[14] K. Kamilu, M. Sulaiman, A. Muhammad, A. Mohamad, and M. Mamat, "PERFORMANCE EVALUATION OF A NOVEL CONJUGATE GRADIENT METHOD FOR TRAINING FEED FORWARD NEURAL NETWORK," *Math. Model. Comp*, vol. 10, pp. 326–337, 2023.doi: https://doi.org/10.23939/mmc2023.02.326

[15] G. Zoutendijk, "COMPUTATIONAL METHODS IN NONLINEAR PROGRAMMING," *Studies in optimization*, vol. 1, p. 125, 1970.

[16] M. J. D. Powell, "RESTART PROCEDURES FOR THE CONJUGATE GRADIENT METHOD," *Math Program*, vol. 12, pp. 241–254, 1977.doi: https://doi.org/10.1007/BF01593790

[17] M. J. D. Powell, "SOME CONVERGENCE PROPERTIES OF THE CONJUGATE GRADIENT METHOD," *Math Program*, vol. 11, pp. 42–49, 1976.doi: https://doi.org/10.1007/BF01580369

[18] H. Zhang and W. W. Hager, "A NONMONOTONE LINE SEARCH TECHNIQUE AND ITS APPLICATION TO UNCONSTRAINED OPTIMIZATION," *SIAM Journal on Optimization*, vol. 14, no. 4, pp. 1043–1056, 2004, doi: https://doi.org/10.1137/S1052623403428208.

[19] W. Z. & C. X. Huang S., "A NEW NONMONOTONE LINE SEARCH TECHNIQUE FOR UNCONSTRAINED OPTIMIZATION," *Numer Algor*, vol. 68, pp. 671–689, 2015, doi: https://doi.org/10.1007/s11075-014-9866-4 .

[20] Y. Dai, J. Han, G. Liu, D. Sun, H. Yin, and Y. X. Yuan, "CONVERGENCE PROPERTIES OF NONLINEAR CONJUGATE GRADIENT METHODS.," *SIAM J. Optim.* , vol. 2, no. 10, pp. 345–358, 2000.doi: https://doi.org/10.1137/S1052623494268443

[21] T.-A. D. S. C, "EFFICIENT HYBRID CONJUGATE GRADIENT TECHNIQUES.," *J. Optim. Theory Appl.*, vol. 2, no. 64, pp. 379–397, 1990.doi: https://doi.org/10.1007/BF00939455

[22] M. Al-Baali, " DESCENT PROPERTY AND GLOBAL CONVERGENCE OF THE FLETCHER–REEVES METHOD WITH INEXACT LINE SEARCH," *IMA J. Numer. Anal.*, vol. 5, no. 1, pp. 121–124, 1985.doi: https://doi.org/10.1093/imanum/5.1.121

[23] J. C. N. j Gilbert, "GLOBAL CONVERGENCE PROPERTIES OF CONJUGATE GRADIENT METHODS FOR OPTIMIZATION," *SIAM J. Optim.*, vol. 2, no. 1, pp. 21–42, 1992.doi: https://doi.org/10.1137/0802003

[24] M. M. Sulaiman I M, "A NEW CONJUGATE GRADIENT METHOD WITH DESCENT PROPERTIES AND ITS APPLICATION TO REGRESSION ANALYSIS.," *J. Numer. Anal. Ind. Appl. Math.*, vol. 1–2, no. 14, pp. 25–39, 2020.

[25] E. Ali and S. Mahdi, "A FAMILY OF DEVELOPED HYBRID FOUR-TERM CONJUGATE GRADIENT ALGORITHMS FOR UNCONSTRAINED OPTIMIZATION WITH APPLICATIONS IN IMAGE RESTORATION," *Symmetry (Basel)*, vol. 15, no. 6, 2023, doi: https://doi.org/10.3390/sym15061203.

[26] W. W. Hager and H. Zhang, "A SURVEY OF NONLINEAR CONJUGATE GRADIENT METHODS,"," *Pacific Journal of Optimization*, vol. 2, pp. 35–58, 2006.
[27] Y. F. Hu and C. Storey, "GLOBAL CONVERGENCE RESULT FOR CONJUGATE GRADIENT METHODS," *J Optim Theory Appl*, vol. 71, no. 2, pp. 399–405, 1991.doi: https://doi.org/10.1007/BF00939927
[28] M. Malik, M. Mamat, S. S. Abas, I. M. Sulaiman, and F. Sukono, "A NEW COEFFICIENT OF THE CONJUGATE GRADIENT METHOD WITH THE SUFFICIENT DESCENT CONDITION AND GLOBAL CONVERGENCE PROPERTIES," *Engineering Letters*, vol. 28, no. 3, pp. 62–72, 2020.
[29] L. : Armijo, "MINIMIZATION OF FUNCTIONS HAVING LIPSCHITZ CONTINUOUS FIRST PARTIAL DERIVATIVES," *Pac. J. Math*, vol. 16, pp. 1–3, 1966.doi: https://doi.org/10.2140/pjm.1966.16.1
[30] A. A. : Goldstein, "GOLDSTEIN ON STEEPEST DESCEN.," *Pac. J. Math*, vol. 3, pp. 147–151, 1965.doi: https://doi.org/10.1137/0303013
[31] P. Wolfe, "CONVERGENCE CONDITIONS FOR ASCENT METHODS.," *SIAM Review*, vol. 11, pp. 226–235, 1969.doi: https://doi.org/10.1137/1011036
[32] X. S. Zhong Wan Kok Lay Teo and C. Hu, "NEW BFGS METHOD FOR UNCONSTRAINED OPTIMIZATION PROBLEM BASED ON MODIFIED ARMIJO LINE SEARCH," *Optimization*, vol. 63, no. 2, pp. 285–304, 2014, doi: https://doi.org/10.1080/02331934.2011.644284.
[33] Z. S. Zhou w Chalabi, "MODIFICATIONS OF THE WOLFE LINE SEARCH RULES TO SATISFY SECOND-ORDER OPTIMALLY CONDITIONS IN UNCONSTRAINED OPTIMIZATION.," *SIAM Review*, vol. 11, pp. 226–235, 1969.
[34] G. M. Al-Naemi, "A NEW MODIFIED HS ALGORITHM WITH STRONG POWELL-WOLFE LINE SEARCH FOR UNCONSTRAINED OPTIMIZATION," *Eastern-European Journal of Enterprise Technologies*, vol. 4, no. 2, pp. 14–21, 2022.doi: https://doi.org/10.15587/1729-4061.2022.254017
[35] W. W. Z. H. Hager, "A SURVEY OF NONLINEAR CONJUGATE GRADIENT METHODS.," *Pac. J. Optim.*, vol. 1, no. 2, pp. 35–58, 2006.
[36] I. M. Sulaiman, M. Malik, A. M. Awwal, P. Kumam, M. Mamat, and S. Al-Ahmad, "ON THREE-TERM CONJUGATE GRADIENT METHOD FOR OPTIMIZATION PROBLEMS WITH APPLICATIONS ON COVID-19 MODEL AND ROBOTIC MOTION CONTROL," *Advances in Continuous and Discrete Models*, vol. 2022, no. 1, p. 1, 2022.doi: https://doi.org/10.1186/s13662-021-03638-9
[37] M. Malik, I. M. Sulaiman, A. B. Abubakar, G. Ardaneswari, and Sukono, "A NEW FAMILY OF HYBRID THREE-TERM CONJUGATE GRADIENT METHOD FOR UNCONSTRAINED OPTIMIZATION WITH APPLICATION TO IMAGE RESTORATION AND PORTFOLIO SELECTION," *AIMS Mathematics*, vol. 8, no. 1, pp. 1–28, 2023, doi: https://doi.org/10.3934/math.2023001.
[38] Y. Yamakawa, H. Sato, and K. Aihara, "MODIFIED ARMIJO LINE-SEARCH IN RIEMANNIAN OPTIMIZATION WITH REDUCED COMPUTATIONAL COST," 2023. [Online]. Available: https://arxiv.org/abs/2304.02197
[39] K. Amini and P. Faramarzi, "GLOBAL CONVERGENCE OF A MODIFIED SPECTRAL THREE-TERM CG ALGORITHM FOR NONCONVEX UNCONSTRAINED OPTIMIZATION PROBLEMS," *J Comput Appl Math*, vol. 417, p. 114630, 2023, doi: https://doi.org/10.1016/j.cam.2022.114630 .
[40] A. Alhawarat, G. Alhamzi, I. Masmali, and Z. Salleh, "A DESCENT FOUR-TERM CONJUGATE GRADIENT METHOD WITH GLOBAL CONVERGENCE PROPERTIES FOR LARGE-SCALE UNCONSTRAINED OPTIMISATION PROBLEMS," *Math Probl Eng*, vol. 2021, no. 1, p. 6219062, 2021.doi: https://doi.org/10.1155/2021/6219062
[41] N. Andrei, "HYBRID CONJUGATE GRADIENT ALGORITHM FOR UNCONSTRAINED OPTIMIZATION," *J Optim Theory Appl*, vol. 141, pp. 249–264, 2009.doi: https://doi.org/10.1007/s10957-008-9505-0
[42] M. Lotfi and S. M. Hosseini, "AN EFFICIENT HYBRID CONJUGATE GRADIENT METHOD WITH SUFFICIENT DESCENT PROPERTY FOR UNCONSTRAINED OPTIMIZATION," *Optim Methods Softw*, vol. 37, no. 5, pp. 1725–1739, 2022.doi: https://doi.org/10.1080/10556788.2021.1977808
[43] S. C. Hu Y.F., "GLOBAL CONVERGENCE RESULT FOR CONJUGATE GRADIENT METHODS.," *J. Optim. Theory Appl.*, vol. 2, no. 71, pp. 399–405, 1991.doi: https://doi.org/10.1007/BF00939927
[44] J. Jian, L. Han, and X. Jiang, "A HYBRID CONJUGATE GRADIENT METHOD WITH DESCENT PROPERTY FOR UNCONSTRAINED OPTIMIZATION," *Appl Math Model*, vol. 39, no. 3, pp. 1281–1290, 2015.doi: https://doi.org/10.1016/j.apm.2014.08.008
[45] Yunus, R. B. and K. Kamilu, and M. S. I. and M. Mustafa, "A NOVEL THREE TERM CONJUGATE GRADIENT METHOD FOR UNCONSTRAINED OPTIMIZATION VIA SHIFTED VARIABLE METRIC APPROACH WITH APPLICATION," Springer International Publishing, 2022, pp. 581–596. doi: https://doi.org/10.1007/978-3-031-04028-3_37.
[46] X. Z. Jiang, G. D. Ma, and J. B. Jian, "A NEW GLOBAL CONVERGENT CONJUGATE GRADIENT METHOD WITH WOLFE LINE SEARCH," *Chinese Journal of Engineering Mathematics*, vol. 28, no. 6, pp. 779–786, 2011.
[47] E. D. Dolan and J. J. Moré, "BENCHMARKING OPTIMIZATION SOFTWARE WITH PERFORMANCE PROFILES," *Math Program*, vol. 91, pp. 201–213, 2002.doi: https://doi.org/10.1007/s101070100263
[48] F. Shen, J. Chao, and J. Zhao, "FORECASTING EXCHANGE RATE USING DEEP BELIEF NETWORKS AND CONJUGATE GRADIENT METHOD," *Neurocomputing*, vol. 167, pp. 243–253, 2015.doi: https://doi.org/10.1016/j.neucom.2015.04.071