Enhancing linear regression forecasting with an improved conjugate gradient method: Application to economic data

Sijun Tao[1,2,#], Nur Hamizah Abdul Ghani[1], Mohd Rivaie[3]
1 College of Computing, Informatics and Mathematics, Universiti Teknologi MARA, 40450 Shah Alam, Selangor, MALAYSIA.
2 College of Mathematics and Computer, Xinyu University, Jiangxi, P.R. China, Xinyu University, 338004 Xinyu Jiangxi, P.R. CHINA.
3 Department of Computer Science and Mathematics, Universiti Teknologi MARA (UITM) Terengganu, 21080 Kuala Terengganu, MALAYSIA.
#Corresponding author. E-Mail: taosijunjx001@gmail.coml; Tel: +8613507908050

ABSTRACT This study proposes an improved conjugate gradient (CG) method: the Tao-Rivaie-Hamizah (TRH) method and applies it to regression analysis of China's per capita disposable income over the past fifteen years. Unlike previous CG variants, TRH combines the numerator of the Hybrid Polak–Ribière–Polyak (HPRP) method with the denominator of the Hamoda–Rivaie–Mamat (HRM) method, thereby achieving faster convergence and enhanced robustness. Numerical experiments demonstrate that the TRH method generally outperforms several representative CG methods in terms of convergence speed and robustness. In economic applications, the TRH method yields smaller total relative errors and superior forecast values compared to Least Squares and Trend Line methods. These results validate the innovation and efficacy of the TRH method, providing a stable and effective solution for linear regression forecasting while further enriching and advancing the CG method.

KEYWORDS: Conjugate gradient method, Least Squares method, TRH method, Trend Line method, linear regression forecasting
Received 3 September 2025 **Revised** 25 September 2025 **Accepted** 28 September 2025 **In press** 2 October 2025 **Online** 4 October 2025

Original Article

INTRODUCTION

Regression analysis is an essential statistical tool widely used in economics, meteorology, and medicine (Sulaiman & Mamat, 2020; Majumder *et al.*, 2023). It is commonly utilised to forecast the values of the dependent variable or to understand the interplay among explanatory variables. Among various regression models, linear regression is the most widely used. The linear regression model is typically represented as follows:

$$y = a_0 + a_1 x_1 + a_2 x_2 + \cdots + a_p x_p + \varepsilon, \quad (1)$$

where a_0 is the intercept, $a_1, a_2, \ldots, a_n$ are the regression coefficient; ε is the error term, which represents the deviation between the actual value and the value predicted by the model. When there is only one independent variable, the model is:

$$y = a_0 + a_1 x_1 + \varepsilon. \quad (2)$$

This model is known as a one-dimensional linear regression model. The main task of regression analysis is to use mathematical methods to fit the actual data of the independent and dependent variables and to estimate each parameter in the regression model to obtain the regression equation. The Least Squares method is the most commonly used for solving regression equations. Least Squares is a mathematical optimization method used to fit data points in the presence of error, aiming to find a function that minimizes the sum of the squares of the deviations between the data points and that function (Armstrong & Fildes, 1995). Using the method of Least Squares, the multiple linear regression model can be transformed into solving for the minimum value of the function E(a), which is mathematically modelled as:

$$\min E\left(a_0,a_1,\cdots,a_n\right)=\sum_{i=1}^{n}\left(y_i-\left(a_0+a_1x_{i1}+a_2x_{i2}\cdots+a_px_{ip}\right)\right)^2 \quad (3)$$

where y_i is the actual value, $(x_{i1},x_{i2},\cdots,x_{ip})$ is the value of the independent variable at point $i(i=1,2,\cdots,n)$. Since the above problem is an unconstrained optimization problem, it can be solved using the CG method. Usually, unconstrained optimization problems have the following general form (Mohamed *et al.*, 2016).

$$\min f\left(x\right), \ \forall x\in R^n \quad (4)$$

where the function $f\left(x\right):R^n\rightarrow R$ is a real-valued function, which is continuous and differentiable.
The iterative formula for the CG method is:

$$x_{k+1}=x_k+\alpha_kd_k. \quad (5)$$

where x_k represents the k-th iterate point, and $\alpha_k>0$ denotes the step size, α_k determined by a certain type of line search. The Strong Wolfe Powell (SWP) line search is used to solve for step size α_k, so that it satisfies the following.

$$f\left(x_k+\alpha_kd_k\right)\le f\left(x_k\right)+\delta\alpha_kg_k{}^Td_k, \quad (6)$$

$$\left|g\left(x_k+\alpha_kd_k\right)^Td_k\right|\le\sigma\left|g_k{}^Td_k\right|, \quad (7)$$

where $g_k=\nabla f\left(x_k\right)$, δ,σ are constant and $0<\delta<\sigma<1$, d_k is the search direction. In CG methods, d_k is defined as

$$d_k=\begin{cases}-g_k & for \quad k=0\\ -g_k+\beta_kd_{k-1} & for \quad k\ge1\end{cases}, \quad (8)$$

where β_k is the conjugate coefficient of the CG methods. In this paper, according to paper (Hamoda *et al.*, 2016, Ghani *et al.*, 2017), β_k is defined as

$$\beta_k^{TRH}=\begin{cases}\dfrac{\|g_k\|^2-\dfrac{\|g_k\|}{\|g_k-g_{k-1}\|}\left|g_k^Tg_{k-1}\right|}{\mu\|g_{k-1}\|^2+(1-\mu)\|d_{k-1}\|^2}, if\ \|g_k\|^2>\dfrac{\|g_k\|}{\|g_k-g_{k-1}\|}\left|g_k^Tg_{k-1}\right|\\ 0, \quad otherwise\end{cases} \quad (9)$$

where μ is a positive parameter, $0<\mu<1$.

In this paper, a new formulation was used which is Tao-Rivaie-Hamizah (TRH), the new CG variant method. Distinct from previous CG methods, TRH method is not merely a reparameterization but a structured integration of two existing methods: it adopts the numerator term from Ghani's Hybrid Polak–Ribière–Polyak (HPRP) method, which guarantees sufficient descent and avoids non-descent directions when gradient variations are small, and the denominator structure from Hamoda's Hamoda–Rivaie–Mamat (HRM) method, which improves the robustness of the algorithm.

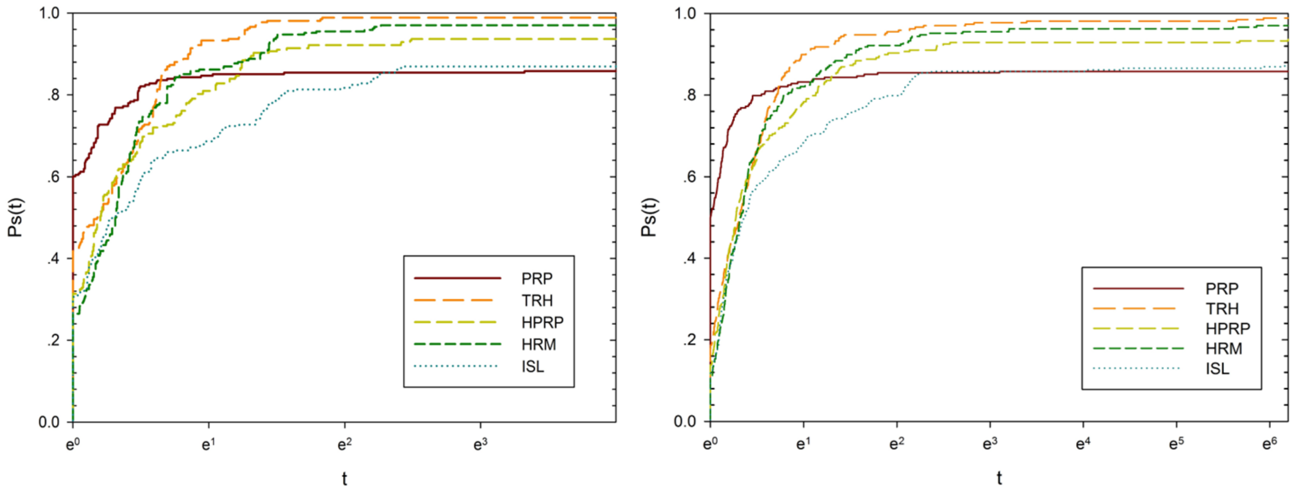
NUNERICAL EXPERIMENTS

Numerical experiments were conducted to evaluate the performance of the TRH method. A total of 20 different test functions were selected from (Andrei, 2008), with four different initial points for each test functions. The dimensionality of the tests ranged from a minimum 2 dimensions to a maximum of 10000 dimensions, based on the characteristics of the test functions. Iteration was terminated either when the criterion was satisfied or the maximum of 10,000 iteration was reached. The test functions and their characteristics are summarized in Table 1.

Table 1. Functions and Related Information.

No.	Test Functions	Test Dimensions	Initial Points
1	Zettl	2	1,10,15,25
2	Booth	2	1,5,10,20
3	Three Hump Camel	2	3,10,20,44
4	Treccani	2	1,4,8,15
5	Six Hump Camel	2	2,5,10,15
6	Powell	4	-5, -1,1,5
7	Extended Wood	4	-1,5,10,20
8	Extended Powell	4,100,500	1,6,16,20
9	Hager	2, 4, 10, 100	2,5,10,20
10	Quadratic QF1	2,500,1000,5000	1,5,10,20
11	Extended Tridiagonal 1	2, 10, 100, 500, 1000	5,10,15,50
12	Fletchcr	2, 10, 100, 500, 1000	12,22,32,62
13	Sphere function	2,500,5000,10000	5,15,40,60
14	Extended Beale	2, 500, 1000, 5000, 10000	-1,0.5,1,2
15	Extended DENSCHNB	2, 500, 1000, 5000, 10000	1,5,10,15
16	Diagonal 4	2, 500, 1000, 5000, 10000	3,10,50,80
17	Extended Himmelblau	2, 500, 1000, 5000, 10000	1,5,20,25
18	Extended White & Holst	2, 500, 2000, 5000, 10000	-2,4,10,12
19	Extended Rosenbrock	2, 500, 1000, 5000, 10000	2,5,10,20
20	Shallow	2, 500, 1000, 5000, 10000	2,5,25,50

To assess the effectiveness of the TRH method, a comparison was made with four other methods, which are the HRM, PRP, HPRP, and ISL method. As the PRP method is regarded as one of the classical CG methods with the best numerical performance (Zhang *et al.*, 2006), and the ISL method shares the same denominator as the TRH method, so these two methods were included for comparative analysis. The comparison focused on two area, which are the number of iterations and CPU times of these five methods under the same conditions. The Dolan (2002) performance profile were applied to analyse the numerical performance, as shown in Figure 1.



a. PRP method b. TRH method

Figure 1. Performance Profile for the Number of Iterations and CPU times.

Figure 1 presents the performance profiles of the five methods, with the number of iterations shown on the left and CPU time on the right. The left panel indicates that the PRP method exhibits the fastest convergence, followed by the TRH method. In contrast, the right panel shows that the TRH curve attains the highest level, demonstrating its ability to solve the largest number of test problems and thus its superior robustness. In contrast, the PRP curve lies at the lowest level, reflecting the weakest robustness. Taken together, these results suggest that the TRH method achieves a more balanced and overall superior performance compared with the other methods.

APPLICATION TO ECONOMIC

This section applies the TRH method to regression analysis to solve practical problems and compares it with the Least Squares method and the Trend Line method. In China, residents' per capita disposable income has been increasing yearly with the continuous improvement of people's living standards. In 2008, Chinese residents' per capita disposable income was 9,957 yuan; by 2014, it reached 20,000 yuan, and in 2019, it exceeded 30,000 yuan. Specific annual per capita disposable income and growth rates are shown in Table 2 below (China, 2025).

Table 2. 2008-2023 China Disposable Income per Person.

No.	Year	Disposable Income per Person (Yuan)	Annual Growth Percentage (%)
1	2008	9,957	9.50
2	2009	10,977	11.00
3	2010	12,520	10.40
4	2011	14,551	10.30
5	2012	16,510	10.60
6	2013	18,311	8.10
7	2014	20,167	8.00
8	2015	21,966	7.40
9	2016	23,821	6.30
10	2017	25,974	7.30
11	2018	28,228	6.50
12	2019	30,733	5.80
13	2020	32,189	2.10
14	2021	35,128	8.10
15	2022	36,883	2.90
16	2023	39,218	6.10

Based on the data in Table 2, it is evident that although the growth rate of per capita disposable income in China varies from year to year, statistically, there is a certain correlation between the year and the growth rate. This paper aims to establish a regression model describing the relationship between the growth rate of income and the year and to find a regression equation that can reflect the growth rate trend. According to the formula of linear correlation coefficient of paired data:

$$r = \frac{\sum_{i=1}^{n}\left(x_i - \bar{x}\right)\left(y_i - \bar{y}\right)}{\sqrt{\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2}\sqrt{\sum_{i=1}^{n}\left(y_i - \bar{y}\right)^2}}, \tag{10}$$

where $\left(x_i, y_i\right), i = 1, \cdots, 16$, x_i denotes the serial number, which takes the value of $x_i = 1, 2, \cdots, 16$, and y_i denotes the annual growth rate for the corresponding year. $\bar{x}$, $\bar{y}$ represents the average value of x_i, y_i respectively. By substituting x and y into the calculation, the correlation coefficient was found to be $r = 0.81$. representing that the data points are highly correlated with the linear regression equation. Therefore, this paper uses the Least Squares method, the TRH method, and the Trend Line method to solve the linear regression of the data in the table and predict the changes in the growth rate of express delivery in the coming years.

Least Squares Method

The Least Squares method is a mathematical optimisation method for fitting data points in the presence of error. It aims to find a function that minimizes the sum of the squares of the deviations between the data points and that function. The method is widely used in regression analysis, especially when exploring linear or non-linear relationships between variables. This section will explore the specific application of the Least Squares method in a univariate linear regression model using the data from Table 2.

Given an arbitrary set of data points $(x_i, y_i), i = 1, 2, \cdots, n$, it is required to find an optimal straight-line

$$y = a_0 + a_1 x, \tag{11}$$

to make the sum of the squares of the deviations between the straight-line and each data point reach the minimum value. From the question, define the Error $E = y_i - (a_0 + a_1 x)$ at point i. The objective function is defined as

$$\min E(a_0, a_1) = \sum_{i=1}^{n} \left(y_i - (a_0 + a_1 x_i) \right)^2. \tag{12}$$

In Table 2, set the serial number in No. as x_i, annual growth rate as y_i, $i = 1, 2, \cdots, 15$. Using the Least Squares method, obtain the univariate linear regression equation:

$$y = 11.5828571429\text{-}0.4953571429x, \tag{13}$$

TRH Method

In regression analysis, the CG method is an efficient and commonly used method, which is particularly suitable for large-scale data processing and solving unconstrained minimization problems (Florens, 2015, Qasim *et al.*, 2021, Sulaiman *et al.*, 2022). For the linear regression equations presented earlier, our goal is to find the optimal regression parameter in the regression equation by minimizing the sum of squares of the errors, thus minimizing the deviation between the fitted function and the actual data. Specifically, transform the linear regression problem into an unconstrained minimization problem by using Equation (11) and Equation (12), yield

$$\min f(a) = \sum_{i=1}^{n} \left[y_i - (a_0 + a_1 x_i) \right]^2, \tag{14}$$

where (x_i, y_i), $i = 1, \cdots, n$ is a pre-given data point.

Similarly to the preceding, set x_i denotes the serial number, y_i denotes the annual growth rate for the corresponding year, $i = 1, \cdots, 15$. To calculate the relative error between the predicted and actual values, this paper excluded the 2023 data from the calculation. These data are substituted into Equation (14), and yield

$$f(a_0, a_1) = 15a_0^2 + 240a_0 a_1 - 228.6a_0 + 1240a_1^2 - 1551.4a_1 + 969.53 \tag{15}$$

Equation (15) was solved as a new test function using the TRH method; four different initial points were arbitrarily selected and solved using the TRH method under SWP line search. The solution results are presented in Table 3.

Table 3. The TRH method for solving linear regression models.

Regression Model	Initial Point	Iterations	CPU Time	Value of Variable
Linear	(-8, -8)	4	0.0652	[11.5828571429, -0.4953571429]
Linear	(-3, 3)	2	0.0067	[11.5828571429, -0.4953571429]
Linear	(6, 6)	3	0.0012	[11.5828571429, -0.4953571429]
Linear	(26, 26)	2	0.00028	[11.5828571429, -0.4953571426]

The linear regression function under the TRH method is derived by averaging the solution results across all initial points in Table 3. The resulting function is as follows.

$$y = 11.5828571429 - 0.4953571428x. \tag{16}$$

Trend Line Method

The Trend Line method is a solution method for regression analysis using Microsoft Excel software (Hanachi *et al.*, 2023). Microsoft Excel 2021 software generated the x and y scatter plots of the linear trend lines for the data presented in Table 2. The years of data used were from 2008 to 2022. Using the Trend Line method, a linear trend line was plotted through Microsoft Excel software, and the specific graph is shown in Figure 2 below.

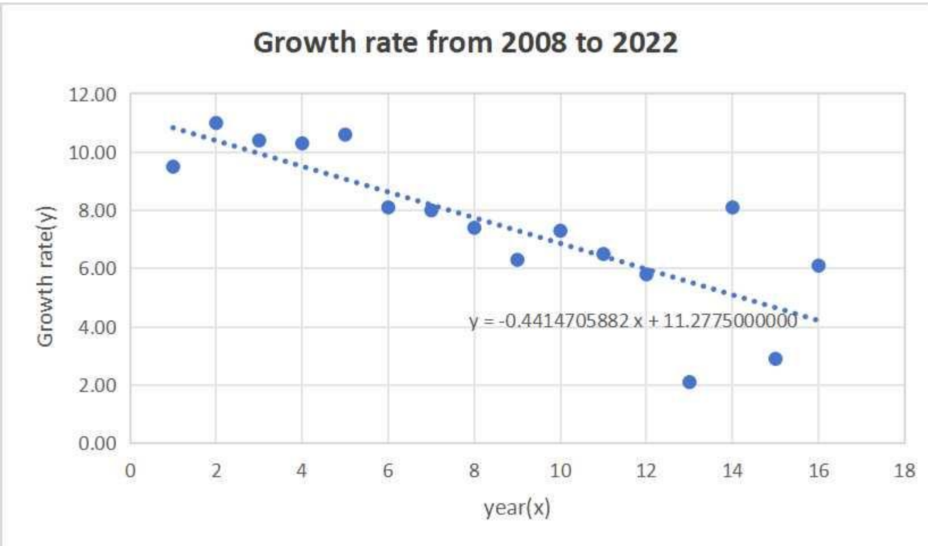


Figure 2. Growth rate of disposable income per capita in China over the past 15 years.

From the figure, the linear regression equation for the Trend Line method is given as

$$y = 11.2775000000 - 0.4414705882x. \tag{17}$$

ERROR ANALYSIS

This error analysis is performed to compare the accuracy of the three regression models: the Least Squares, the TRH, and the Trend Line methods. Firstly, we need to predict the per capita income in 2023 according to the three regression functions and then compare it with the actual per capita income in 2023 to calculate the relative error of the three regression models. The formula for calculating the relative error is written as

$$\text{relative error} = \frac{|\text{exact value - predictive value}|}{|\text{exact value}|} \quad (18)$$

Secondly, the total error of the three methods over the last 15 years was calculated to find the fitting method with the smallest total error, i.e., the one closest to the optimal solution. Table 4 shows the predicted values and relative errors from the three methods each year.

Table 4. Predicted values and relative errors of the three methods.

Year	Annual growth rate (yi)	Least Square method prediction	Trend line method prediction	TRH method prediction	Least Square method relative error	Trend line method relative error	TRH method relative error
2008	9.50	11.0875	10.83602941	11.0875	0.167105263	0.140634675	0.167105263
2009	11.00	10.59214286	10.39455882	10.59214286	0.037077922	0.055040107	0.037077922
2010	10.40	10.09678571	9.953088235	10.09678571	0.02915522	0.042972285	0.02915522
2011	10.30	9.601428571	9.511617647	9.601428572	0.067822469	0.076541976	0.067822469
2012	10.60	9.106071428	9.070147059	9.106071429	0.140936658	0.144325749	0.140936658
2013	8.10	8.610714286	8.628676471	8.610714286	0.063051146	0.0652687	0.063051146
2014	8.00	8.115357143	8.187205883	8.115357143	0.014419643	0.023400735	0.014419643
2015	7.40	7.62	7.745735294	7.620000001	0.02972973	0.046720986	0.02972973
2016	6.30	7.124642857	7.304264706	7.124642858	0.130895692	0.159407096	0.130895692
2017	7.30	6.629285714	6.862794118	6.629285715	0.091878669	0.059891217	0.091878669
2018	6.50	6.133928571	6.42132353	6.133928572	0.056318681	0.012104072	0.056318681
2019	5.80	5.638571428	5.979852942	5.638571429	0.027832512	0.031009128	0.027832512
2020	2.10	5.143214285	5.538382353	5.143214287	1.44914966	1.63732493	1.44914966
2021	8.10	4.647857142	5.096911765	4.647857144	0.426190476	0.370751634	0.426190476
2022	2.90	4.152499999	4.655441177	4.152500001	0.431896552	0.605324544	0.431896552
2023	6.10	3.657142857	4.213970589	3.657142858	0.400468384	0.309185149	0.400468384
Minimum total error					**3.563928677**	**3.779902984**	**3.563928677**

According to the data in Table 4, it can be found that the TRH method performs well in the regression analysis. The fifteen-year total error of the TRH method is the same as that of the Least Squares method, which is 3.563928677, which is significantly better than that of the Trend Line method, which is 3.779902984, reflecting the higher overall prediction accuracy of the TRH method. For the prediction of 2023, the predicted value of the TRH method is 3.657142858, which is slightly better than that of the least squares method of 3.657142857; and the relative error with the actual value is 0.400468384, which is the same as that of the Least Squares method, and slightly higher than that of the Trend Line method.

DISCUSSION

This study verified the numerical performance of the improved TRH method through numerical experiments and applied it to regression analysis for solving economic problems. The experimental results demonstrate that the TRH method consistently achieves lower or comparable errors compared to traditional Least Squares and Trend Line methods, which corroborate some studies that highlighted the efficiency of improved CG methods in regression analysis (Florens, 2015). By integrating the numerator formulation of HPRP with the denominator modification of HRM, TRH balances the strengths and limitations of these approaches, thereby improving overall efficiency. From an economic perspective, the ability to provide more stable forecasts of income growth is valuable for policymakers, especially during periods of economic fluctuation. While this study is limited to univariate regression, the results highlight TRH's potential for broader applications in econometrics and large-scale data analysis.

CONCLUSION

In this paper, using China's per capita disposable income data over the past fifteen years, this study applied the TRH, Least Squares, and Trend Line methods to linear regression forecasting. The numerical experiments show that TRH not only achieves lower overall forecasting errors than the Least Squares and Trend Line methods but also demonstrates faster convergence and greater robustness compared with several classical CG methods. These findings indicate that TRH provides a more effective and stable approach for practical regression forecasting, while further extending the applicability of CG methods in applied econometrics. In future, researchers will address the formal proofs of sufficient descent and global convergence of the new method.

ACKNOWLEDGEMENTS

The authors are grateful to acknowledge the support of Universiti Teknologi Mara (UiTM), College of Computing, Informatics, and Mathematics, Shah Alam, Selangor, Malaysia, for providing the facilities support on this research. This work is also supported by the research project of Xinyu University: An Improved Conjugate Gradient Method for Unconstrained Optimization Problems and Its Applications (Project No. XJ2406).

REFERENCES

[1] Andrei, N. 2008. An unconstrained optimization test functions collection. *Advanced Modeling and Optimization,* 10(1)**,** 147-161.

[2] Armstrong, J. S. & Fildes, R. 1995. Correspondence on the selection of error measures for comparisons among forecasting methods. *Journal of Forecasting,* 14(1)**,** 67-71.

[3] China, N. B. O. S. O. 2025. *China Statistical Data Query Platform*. (https://data.stats.gov.cn/easyquery.htm?cn=C01). Last access on 20 September 2025.

[4] Dolan, E. D. & Moré, J. J. 2002. Benchmarking optimization software with performance profile. *Mathematical programming,* 91(2), 201-213.

[5] Florens, J. P. & Van Bellegem, S. 2015. Instrumental variable estimation in functional linear models. *Journal of Econometrics,* 186(2)**,** 465-476.

[6] Ghani, N. H. A., Mamat, M. & Rivaie, M. 2017. A new family of Polak-Ribiere-Polyak conjugate gradient method with the strong-Wolfe line search. *AIP Conference Proceedings (SKSM24).* 27–29 September 2016. Terengganu, Malaysia. Vol. 1870, No. 1, p. 040060.

[7] Hamoda, M., Mamat, M., Rivaie, M. & Salleh, Z. 2016. A conjugate gradient method with strong Wolfe-Powell line search for unconstrained optimization. *Applied Mathematical Sciences,* 10(721-734.

[8] Hanachi, B., Sabrina, Sellami, B. & Belloufi, M. 2023. A new family of hybrid conjugate gradient method for unconstrained optimization and its application to regression analysis. *RAIRO - Operations Research,* 58(1), 613-627.

[9] Majumder, S. C., Voumik, L. C., Rahman, M. H., Rahman, M. M. & Hossain, M. N. 2023. A quantile regression analysis of the impact of electricity production sources on CO2 emission in South Asian Countries. *Strategic Planning for Energy and the Environment,* 42(02), 307-330.

[10] Mohamed, N. S., Mamat, M., Mohamad, F. S. & Rivaie, M. 2016. A new coefficient of conjugate gradient methods for nonlinear unconstrained optimization. *Jurnal Teknologi (Sciences & Engineering),* 78(6-4)**.**

[11] Qasim, A. M., Salih, Z. F. & Hassan, B. A. 2021. A new conjugate gradient algorithms using conjugacy condition for solving unconstrained optimization. *Indonesian Journal of Electrical Engineering and Computer Science,* 24(3), 1654-1660.

[12] Sulaiman, I. M., Malik, M., Awwal, A. M., Kumam, P., Mamat, M. & Al-Ahmad, S. 2022. 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,* 2022(1), 1.

[13] Sulaiman, I. M. & Mamat, M. 2020. A new conjugate gradient method with descent properties and its application to regression analysis. *Journal of Numerical Analysis, Industrial and Applied Mathematics,* 14(1-2), 25-39.

[14] Zhang, L., Zhou, W. & Li, D.-H. 2006. A descent modified Polak–Ribière–Polyak conjugate gradient method and its global convergence. *IMA Journal of Numerical Analysis,* 26(4), 629-640.

E-ISSN 2289-8786. http://tost.unise.org/