https://doi.org/10.31272/jae.i152.1566
https://admics.uomustansiriyah.edu.iq
P-ISSN: 1813–6729 E-ISSN: 2707-1359

Estimating Gamma Regression Model Parameters Using Fire Hawk Optimiser with an Application

Basheer Jameel Khaleel
Ministry of Education, Baghdad, Iraq.
Email: bashjameel73@gmail.com, ORCID: https://orcid.org/0009-0006-3377-4055
Ali Mohammed Ali
Dept of Statistics, College of Administration & Economics, Mustansiriyah University, Baghdad, Iraq.
Email: alialfredawy@uomustansiriyah.edu.iq, ORCID: https://orcid.org/0009-0001-8322-2121

Article Information

Article History:
Received: 15 / 02 / 2026
Revised: 03 / 05 / 2026
Accepted : 13 / 05 / 2026
Available Online: 01 / 06 / 2026
Pages no : 123 – 135

Keywords:
Gamma Regression, Fire Hawk Optimizer, Generalized Linear Models, Parameter Estimation, Mean Squared Error.

Correspondence:
Researcher name:
Basheer Jameel Khaleel

Email: bashjameel73@gmail.com

Abstract

This manuscript focuses on one useful regression model for positive data: Gamma Regression. This is a GLM (Generalised Linear Model) used when we know the dependent variable follows a Gamma distribution. To assess the need for determining the Fire Hawk Optimiser algorithm, which may provide an efficient and acceptable alternative to IWLS for estimating Gamma Regression model parameters as a function of variable numbers, sample sizes, and shape-parameter numbers, served as the main problem of this research. In this application, Fiore, Hawkes, and Oksanen used the Fire Hawk Optimiser (FHO) algorithm to maximise the log-likelihood function, as others would do using the iterative reweighted least squares (IWLS) method, which has been used in the literature of Generalised Linear Models.

We compared the two methods in a series of simulation experiments which varied the sample size, the number of variables, and the shape parameter (λ) under different scenarios. The results demonstrated that the proposed FHO algorithm provides competitive, comparable performance to the IWLS algorithm, with a notable reduction in MSE as the sample size increases, particularly for studies with smaller sample sizes or complex parameters.

In the practical aspect, 46 observations were processed and analysed to assess the impact of several variables on air quality, and the data were sourced from the Iraqi Meteorological Organisation. The results of the goodness-of-fit test confirmed that the dependent variable follows a Gamma distribution. In addition, all estimates are statistically significant, indicating a well-fitted model with the Fire Hawk Optimiser (FHO) algorithm, which can find the true values and provide good estimates with high reliability.

1. Introduction

As one of the most meaningful areas of statistics, regression deals with the mathematical connection between a response and explanatory variables. This relationship is modeled as a regression equation, which is a linear combination of variables. The regression line is simply linear to or non-linear to provide an infinite number of quadratic, exponential, and logarithmic models.

Classical linear regression assumes that the response variable follows a normal distribution. However, this assumption is often not met, and the data may follow other distributions, such as the Gamma distribution. The resulting regression model in this case is called Gamma Regression.

Gamma regression is one of the most significant statistical methods used when the response variable does not follow a Normal distribution, or more precisely, when the response variable follows a Gamma distribution and its values are greater than 0 (meaning no negative values). The concept of

Gamma regression was introduced within the general framework of Generalized Linear Models (GLMs) by [1].

The Maximum Likelihood Estimation (MLE) method is one of the most commonly used parameter estimation methods for estimating the parameters of the Gamma regression model. The maximum likelihood estimates of GLMs are generally based on numerical iterative procedures like the Newton-Raphson and Iteratively Reweighted least squares (IRLS or IWLS) algorithm. Standard IRLS/IWLS is the traditional GLM fitting procedure; however, it may exhibit some computational limitation (e.g., convergence problems, sensitive to initial values, numerical instability), leading to some challenges even for standard non-linear GLM models, especially as sample-size gets smaller, number of explanatory variables increases and parameter structure becomes increasingly complex.

Accordingly, the focus of the Fire Hawk Optimizer (FHO) applied in this work is not to supersede the MLE principle but to enhance the numerical optimization approach employed to extract the maximum likelihood estimates by adapting the log-likelihood function more coherently. Thus, for the FHO algorithm to estimate the parameter vector of Gamma regression model in a metaheuristic optimization way, we further compare the present performance of the FHO-based MLE procedure with the well-known IRLS/IWLS algorithm in the sense of estimated accuracy and efficiency using the MSE criterion.

Gamma regression has been studied in many recent research papers, including those by [4][3][2] [8][7][6][5] and others. The rest of this research is organized as follows: Section 2 explains Generalized Linear Models. The Gamma regression model, Likelihood and Log-likelihood functions are introduced in Section 3. In Section 4 a new algorithm, Fire Hawk Optimizer (FHO) is presented. Section 5 focuses on simulation experiments, and Section 6 on real data analysis. Finally, Section 7 included the conclusions.

2. Importance of the Research

1. Gamma Regression: Data Like and Not Linear or Additive Data modeling Among the many forms that data can take -that we teach students during their degree programs or to a small extent- decoupling the Gaussian norm, the gamma regression model is an effective modeling of the response to situations where the dependent variable is not of the normal.
2. Modern Cutting-edge Optimization Techniques: It uses the Fire Hawk Optimizer (FHO) as a modern meta-heuristic optimization method to estimate the parameters of the Gamma regression model by determining the values that maximize the log-likelihood function (in a maximum likelihood estimation framework).
3. Performance in Small Samples: The paper shows that the algorithm could still achieve comparably competitive performance with classic methods in small samples or with non-measurable hyperparameters or simple sample sizes and hyperparameters.
4. However, the research also has a practical aspect, as these statistical models will be applied to establish specific air quality data and their effects on environmental pollutants, [CO2, particulate matter], depending on the real data provided by the Iraqi Meteorological Organization.

3. Objectives of the Research

1. Examine how well the FHO-based MLE procedure is able to provide accurate and stable estimates of the parameters of the Gamma regression model out of sample.
2. Explore the behavior of the FHO algorithm for various sample sizes, number of explanatory variables, and values of the shape parameter (λ).
3. Show that the estimation error can be competitive to the conventional iterative reweighted least squares (IWLS) algorithm in small-sample cases and in more complex parameters setup.
4. Investigating the practical robustness of the FHO-based Gamma regression model for real positive environmental data on air quality.

4. Theoretical Concepts

4.1 Generalized Linear Models

The Generalized Linear Model (GLM), proposed by [1], is an extension of the General Linear Model (GLM). This extension includes all probability distributions belonging to the exponential family. Most common probability distributions belong to the exponential family, such as Bernoulli, binomial, normal, gamma, beta, Poisson, and other continuous and discrete distributions [9]. The Generalized Linear Model (GLM) contains the following components [10] [11]:

1. The random component, which represents the response variable with an exponential probability distribution.
2. The linear systematic component, which contains the observations of the explanatory variables (X) and the parameter vector (β) contained in the general linear model equation. The systematic component is $\eta_i = x_i^T \beta$, which is called the linear predictor. It is defined by the following formula:

$$\eta_i = x_i^T \beta = \beta_0 + \beta_1 x_{i1} + \ldots + \beta_{(p-1)} x_{i(p-1)} \tag{1}$$

3. The differentiable link function g(.), which links the random part and the systematic part, as the inverse of this function at the systematic part is equal to the expectation of the response variable conditional on both x_i^T and β, as follows:

$$\mu_i = E\left(y_i | x_i^T, \beta\right) = g^{-1}(\eta_i) \tag{2}$$

4.2. Gamma Regression Model

The Gamma distribution is one of the most popular continuous distributions. It is an excellent alternative to the normal distribution for modeling positively skewed data. The probability density function of the Gamma distribution is shown by the following equation [12]:

$$f(y; \alpha, \beta) = \frac{1}{\Gamma(\alpha)\beta^{\alpha}} y^{\alpha-1} e^{-\frac{y}{\beta}} \quad y, \alpha, \beta > 0 \tag{3}$$

where α represents the shape parameter and β represents the scaling parameter. The gamma distribution can be expressed descriptively as follows: $Y \sim G(\alpha, \beta)$

The expectation and variance of the gamma distribution are:

$E[Y] = \alpha\beta = \mu, Var[Y] = \alpha\beta^2$

When using the gamma distribution as the response variable in regression models, the parameterization process is performed after replacing β with μ/α, and its probability density function is as follows: [13]

$$f(y; \alpha, \mu) = \frac{1}{\Gamma(\alpha)} \left(\frac{\alpha}{\mu}\right)^{\alpha} y^{\alpha-1} e^{-\frac{\alpha y}{\mu}} \quad x, \alpha, \mu > 0 \tag{4}$$

This formula for the gamma distribution can be expressed descriptively as follows:

$Y \sim G(\mu, \alpha)$

The expectation and variance of the gamma distribution are: They are:

$E[Y] = \mu, Var[Y] = \frac{\mu^2}{\alpha}$

The Gamma Regression Model (GRM) is used to model positively skewed continuous data. The probability density function for this distribution is stated in Equation (4), and its likelihood function can be formulated as follows [13] [3]:

$$f(y_i; \alpha_i, \mu_i) = \frac{1}{\Gamma(\alpha_i)} \left(\frac{\alpha_i}{\mu_i}\right)^{\alpha_i} y_i^{\alpha_i - 1} e^{-\frac{\alpha_i y_i}{\mu_i}} \tag{5}$$

$Let\ \lambda = \alpha_i$

$$L(\mu, \lambda; y) = exp\left[\sum_{i=1}^{n}\left(\frac{y_i/\mu_i + ln(\mu_i)}{-1/\lambda} + \lambda\{ln(\lambda) + ln(y_i)\} - ln\left(\Gamma(\lambda)\right) - ln\{y_i\}\right)\right] \tag{6}$$

$$L(\beta,\lambda;y)=exp\left[\sum_{i=1}^{n}\left(\frac{y_i\left(x_i^T\beta\right)-ln\left(x_i^T\beta\right)}{-1/\lambda}+\lambda\{ln(\lambda)+ln(y_i)\}-ln\big(\Gamma(\lambda)\big)-ln\{y_i\}\right)\right] \tag{7}$$

The natural logarithm of the possibility function can also be defined as It comes out as:

$$\ell(\beta,\lambda;y)=\sum_{i=1}^{n}\left(\frac{y_i(x_i^T\beta)-ln(x_i^T\beta)}{-1/\lambda}+\lambda\{ln(\lambda)+ln(y_i)\}-ln\big(\Gamma(\lambda)\big)-ln\{y_i\}\right) \tag{8}$$

The maximum likelihood estimators are obtained by maximizing the natural logarithm of the likelihood function defined in Equation (6).

It is worth noting that the link function in the case of gamma regression is an inverse function, and in this case it is called a reciprocal function. However, it is common to use other link functions, such as the logarithmic function.

4.3. Fire Hawk Optimizer Algorithm (FHO)

The Fire Hawk Optimizer (FHO) algorithm, proposed by [14], is a metaheuristic optimization algorithm inspired by the behavior of fire hawks, which deliberately set fires to flush prey from their hiding places. The algorithm is based on dividing individuals into hawks and prey. Hawks seek to direct "fire" toward prey to encourage them to flee toward areas likely to be more favorable in terms of the objective function. The location of both hawks and prey evolves over iterations based on exploration and exploitation mechanisms inspired by nature. The algorithm can be summarized in the following steps:

First: Fire Hawk Position Update

Each fire hawk updates its position based on the main fire (the current best solution) as well as the position of another fire hawk in the solution space, according to the equation:

$$X_{firehawk}^{t+1}=X_{firehawk}^{t}+r_1(X_{best}^{t})-r_2\left(X_{near\_firehawk}^{t}\right) \tag{9}$$

Where: $X_{firehawk}^{t}$: The current position of the firehawk. X_{best}^{t}: The best position (representing the main fire). $X_{near\_firehawk}^{t}$: The position of another nearby firehawk. r_1, r_2: Two random numbers following the uniform distribution U(0,1).

Second: Prey Position Update

The prey reacts to the fire with one of two movements: toward the hawk or toward a safe area within the territory:

$$X_{prey}^{t+1}=X_{prey}^{t}+r_3\left(X_{firehawk}^{t}\right)-r_4\left(X_{safe}^{t}\right) \tag{10}$$

Where: X_{safe}^{t}: A safe location within the hawk's territory, calculated as the average of the prey locations in that area. r_3, r_4: Two random numbers following the uniform distribution U(0,1).

Third: Alternative Prey Update

When the prey escapes its home territory, it may head to another falcon or to a safer area outside the current territory:

$$X_{prey}^{t+1}=X_{prey}^{t}+r_5\left(X_{other\_firehawk}^{t}\right)-r_6\left(X_{safe\_global}^{t}\right) \tag{11}$$

Where: $X_{other\_firehawk}^{t}$: Another firehawk not associated with the prey. $X_{safe\_global}^{t}$: A safe location calculated as the average of the locations of all prey in the search space. r_5, r_6: Two random numbers following the uniform distribution U (0,1).

The Fire Falcon Optimizer (FHO) algorithm relies on bio-inspired mechanisms to guide individuals (falcons and prey) toward the optimal solution. This algorithm can be harnessed in the context of regression models, specifically in maximizing the logarithm of the likelihood function of a gamma regression model, as shown in Equation (6).

In this research, the FHO algorithm was used as an optimization tool to estimate model parameters by finding the optimal values of the parameter vector ψ that achieve the highest possible value of the likelihood function. The objective function within the algorithm was represented by the negative form of the natural logarithm of the likelihood function, i.e.,

$$\min_{\psi \in R^n} \ell(\psi; y)$$

Where: $\psi = \begin{bmatrix} \beta \\ \lambda \end{bmatrix}$

This transforms a maximization problem into a minimization problem, consistent with the nature of optimization algorithms. The MLE principle is employed in the fitness evaluation step of FHO procedure. For all candidate solutions created by the algorithm, whether belonging to a fire hawk or representing prey, they map to a potential combination of Gamma regression parameters. For every candidate of the parameter vector, the log-likelihood function of Gamma regression model on the observed data is calculated. And next, we can define negative log-likelihood value which FHO is going to minimize as the objective function. So, MLE is not a standalone step after the algorithm — it is contained inside the FHO search by definition via the objective function. At each iteration, FHO updates the candidate solutions, which it then selects the parameter vector, to yield the minimum negative log-likelihood, thus maximizing the original log-likelihood function and obtaining the maximum-likelihood estimates.

4.4 Simulation

Simulations were conducted by generating data from the Gamma distribution and then applying the two study algorithms in R. The glm function in the stats package developed by the R Core Team (R Core Team, 2025) was used to implement the IWLS algorithm. The FHO algorithm was programmed in R using the original code written in MATLAB by the researchers [14], published on the MATLAB Central File Exchange at the following link:

https://www.mathworks.com/matlabcentral/fileexchange/114325-fire-hawk-optimizer-fho-a-novel-metaheuristic-algorithm

For the comparison criterion, the Mean Squared Error (MSE) criterion was used, defined as follows:

$$MSE = \frac{1}{R}\sum_{r=1}^{R}\left(\hat{\beta}_r - \beta\right)^T\left(\hat{\beta}_r - \beta\right) \qquad (12)$$

Where: R: Number of simulation iterations. $\hat{\beta}_r$: Vector of estimated parameters. β: Vector of actual parameters.

Several scenarios were used with different numbers of independent variables (1, 2, 3, 4, 5), different values of the λ parameter (0.5, 1.5, 3, 5), and sample sizes (25, 50, 100, 200, 300). The following values represent the values of the β parameters:

1. Predictor: (0.8, 1.2)
2. Predictors: (0.5, -0.5, 0.7)
3. Predictors: (1.5, -0.7, 0.4, 0.6)
4. Predictors: (1.3, 0.8, -0.9, 1.0, -0.5)
5. Predictors: (1.0, -0.6, 0.5, -0.4, 0.7, 0.2)

To ensure the consistency and reliability of the results, each simulation scenario was repeated 1,000 times.

Tables (1-5) represent the mean square error (MSE) values for both algorithms (IWLS, FHO) for different scenarios and sample sizes.

Table (1): MSE values in the case of one independent variable

Sample Size		n=25	n=50	n=100	n=200	n=300
λ=0.5	IWLS	5.69511	1.34602	0.67145	0.28398	0.15578
	FHO	5.79876	1.34896	0.68926	0.28863	0.15981
λ=1.5	IWLS	0.60927	0.41428	0.19882	0.10684	0.05384
	FHO	0.60532	0.41358	0.19671	0.10629	0.05562
λ=3	IWLS	0.42343	0.20895	0.09297	0.04432	0.03166

	FHO	0.41917	0.20799	0.09185	0.04376	0.03262
λ=5	IWLS	0.23065	0.11893	0.06897	0.02802	0.01491
	FHO	0.23021	0.11879	0.06748	0.02845	0.01514

As shown in Table (1), the mean squared errors (MSEs) of IWLS and FHO decrease as the sample size increases, indicating more accurate estimation as the sample size increases. Furthermore, the MSE decreases with increasing sample size, confirming the model's stability. The MSE value decreases as the shape scale of the gamma distribution increases. Differences between the two methods are small in all cases, but MSEs are usually slightly lower for IWLS, especially for tiny values of parameter (λ). This means that both methods are adequate, but that IWLS outperforms the RV pass in some situations.

Table (2): MSE values for two independent variables

Sample Size		n=25	n=50	n=100	n=200	n=300
λ=0.5	IWLS	1.19425	0.37495	0.11994	0.06352	0.03980
	FHO	1.17459	0.36109	0.11685	0.06796	0.04017
λ=1.5	IWLS	0.31190	0.13449	0.05349	0.02763	0.01285
	FHO	0.30645	0.13331	0.05513	0.02722	0.01312
λ=3	IWLS	0.10485	0.05115	0.02445	0.01271	0.00824
	FHO	0.10308	0.04976	0.02475	0.01288	0.00856
λ=5	IWLS	0.09161	0.03568	0.01346	0.00714	0.00458
	FHO	0.09273	0.03596	0.01449	0.00728	0.00480

The MSE values in Table (2) indicate that as the sample size and λ increase, the performance of IWLS and FHO in estimating two independent variables is accurate. We see that the differences between the two methods are still very close, and that FHO sometimes outweighs IWLS, for example, when λ = 0.5 and complex = 1.5, while IWLS sometimes outweighs FHO, for example, when λ = 3. In conclusion, both techniques demonstrate high performance and good consistency, indicating that the FHO algorithm can serve as a viable alternative to the classic IWLS method.

Table (3): MSE values for three independent variables

Sample Size		n=25	n=50	n=100	n=200	n=300
λ=0.5	IWLS	20.23650	5.72306	2.56450	1.20395	0.81476
	FHO	19.90693	5.63924	2.55795	1.15663	0.84446
λ=1.5	IWLS	3.72788	1.62014	0.80620	0.42805	0.23521
	FHO	3.63408	1.59642	0.81520	0.43446	0.26517
λ=3	IWLS	2.09142	0.86778	0.38288	0.20566	0.11184
	FHO	2.08257	0.84523	0.39116	0.21128	0.11917
λ=5	IWLS	1.13605	0.43605	0.24396	0.11951	0.07773
	FHO	1.09654	0.43311	0.24490	0.11229	0.08155

In Table (3), even as the independent variable counts to 3, the performance between IWLS and FHO stays very similar. It should be noted that because the FHO algorithm is flexible in light of model complexity, it can yield slightly lower estimates in certain cases (e.g. (λ = 0.5, 1.5) at small sample sizes). For both methods, the MSE values decrease significantly with an increasing sample size, and the differences between the two methods remain limited. The objective is achieved by obtaining efficient estimates, with a small bias of FHO in some situations, especially in small samples.

Table (4): MSE values for four independent variables

Sample Size		n=25	n=50	n=100	n=200	n=300
λ=0.5	IWLS	17.01836	4.94248	2.32650	0.96753	0.73433
	FHO	16.80803	4.91560	2.36273	0.95921	0.77052
λ=1.5	IWLS	3.80419	1.42533	0.65208	0.34521	0.21888
	FHO	3.78888	1.49163	0.62035	0.34666	0.21691
λ=3	IWLS	2.60575	0.78007	0.34299	0.15966	0.11657
	FHO	2.03533	0.75425	0.34576	0.17153	0.12407
λ=5	IWLS	1.14740	0.45244	0.21901	0.10243	0.06583
	FHO	1.13633	0.43512	0.22758	0.11157	0.07032

Table (4) shows that the estimated mean square errors (MSEs) decrease significantly with increasing sample size, whether using IWLS or FHO, confirming the effectiveness of both methods in improving accuracy with increasing information. The FHO algorithm is noted to provide results similar to IWLS in most cases, and in some instances, the FHO results are even better. However, the two methods still show very small differences, and the table, in general, supports improved performance with increased λ and sample size.

Table (5): MSE values for five independent variables

Sample Size		n=25	n=50	n=100	n=200	n=300
λ=0.5	IWLS	14.76163	5.06454	1.66708	0.90329	0.57961
	FHO	14.55323	4.82933	1.62723	0.88917	0.60649
λ=1.5	IWLS	3.70501	1.35370	0.73597	0.26389	0.18476
	FHO	3.40403	1.36737	0.71267	0.26732	0.19116
λ=3	IWLS	1.79699	0.28966	0.29482	0.13134	0.09597
	FHO	1.74340	0.30505	0.30370	0.14554	0.10255
λ=5	IWLS	1.15617	0.45838	0.15183	0.08705	0.06035
	FHO	1.16336	0.47684	0.15379	0.09149	0.06445

Table (5) continues the previously observed pattern, with MSE values decreasing as the sample size and λ values increase, indicating improved estimation accuracy in models with five independent variables. The FHO algorithm shows very close performance with IWLS and even outperforms it in some cases, such as λ = 0.5 and λ = 1.5, for small samples, reflecting its flexibility in optimal search. However, the difference between the two methods remains limited, and the convergence of results improves with increasing sample size, confirming the effectiveness of both methods in estimation under various complex conditions.

5. Application

The data in Table (6) are air quality data from the Iraqi Meteorological Authority, as reported in the study [7]. The dependent variable (y) in the study is the air quality index. Several independent variables were also selected as follows:

X1: Represents the particulate matter variable less than 10 micrometres.

X2: Represents the carbon dioxide variable.

The following table contains the study data:

Table (6): Air quality data

Y	X1	X2	Y	X1	X2
137	144	26	137	148	25
139	148	26	143	150	32
136	151	25	131	147	26
153	153	31	155	150	32
136	147	27	146	153	31
144	153	31	134	148	27
163	151	37	133	147	25
125	144	23	160	161	33
156	157	35	153	154	32
156	149	33	148	145	28
108	141	19	162	153	27
121	145	20	141	152	26
145	153	28	130	141	17
137	147	27	145	149	24
151	152	29	150	150	29
151	154	33	148	143	23
171	161	39	206	165	42
137	147	28	168	156	34
142	150	28	129	140	20
157	150	33	172	156	32
135	150	26	144	146	25
146	150	30	145	143	21
142	144	28	135	144	25

To ensure that the data for the dependent variable shown in Table (6) follow the Gamma distribution, the goodness-of-fit tests available in the ready-made statistical program EasyFit 5.6 were used, namely the Kolmogorov-Smirnov (KS) test and the Anderson-Darling (AD) test, as follows:

$H_0: Y \sim Gamma$

$H_1: Y \nsim Gamma$

The results of the two tests are shown in the table below:

Table (7): Goodness-of-fit tests

	KS	AD	χ^2
Statistic	0.08717	0.60744	3.714
Critical Value	0.19625	2.5018	11.07

It is clear from Table (7) that the test statistic values for the tests are less than the critical value at the significance level of 0.05, and thus the null hypothesis is accepted. This means that the data for the dependent variable follow the gamma distribution. The following figure shows the gamma distribution curve:

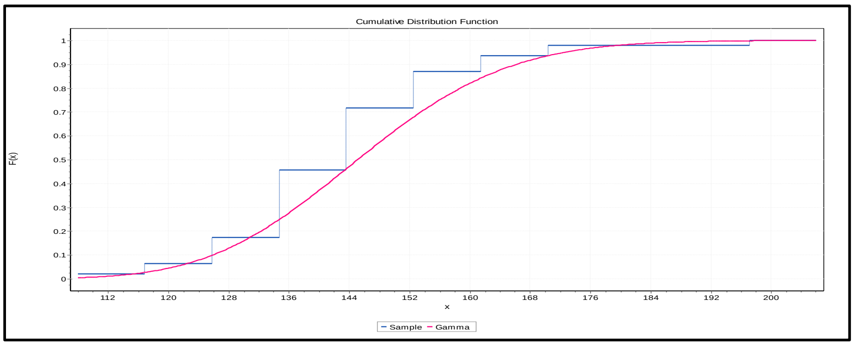

Figure (1): The PDF function for the gamma distribution based on real data.

The following figure represents the cumulative distribution function curve for the gamma distribution:

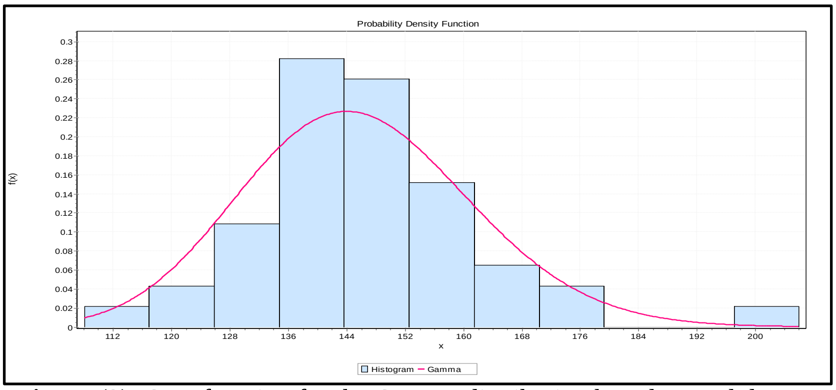

Figure (2): CDF function for the Gamma distribution based on real data.

The Gamma regression parameters were estimated using the inverse link function, and the results were as follows:

Table (8): Parameter estimates using the FHO algorithm

Parameters	Estimate	Std Error	Z statistic	P-Value
β_0	0.011581606	0.001040368	11.132	<0.00001
β_1	-0.000009522	0.000000292	-32.604	<0.00001
β_2	-0.000113658	0.000035064	-3.241	0.00119

Results in Table (8) show that all parameters estimated using the FHO algorithm are statistically significant (P-value < 0.05). The negative coefficients indicate an inverse relationship between the explanatory variables and the linear predictor. But because the Gamma regression model is based on its inverse link function, the interpretation of the effect on the response variable is not as straightforward. Thus, the estimates above confirm the statistical significance of the chosen explanatory variables and support the adequacy of the fitted Gamma regression model.

Because all parameters are significant, these estimates were used to estimate the dependent variable, as follows:

$$\hat{Y} = \frac{1}{\eta_i} = \frac{1}{\beta_o + \beta_1 x_{1i} + \beta_2 x_{2i}} = \frac{1}{0.011581606 - 0.000009522 x_{1i} - 0.000113658 x_{2i}} \quad (13)$$

Based on this equation, the observations of the dependent variable were estimated. Table (9) shows the results of these estimates:

Table (9): Actual and estimated values of the dependent variable

Y	$\hat{Y}$	Y	$\hat{Y}$
137	137.83	137	136.41
139	138.56	143	153.46
136	136.94	131	138.37
153	151.48	155	153.46
136	140.59	146	151.48
144	151.48	134	140.77
163	168.39	133	136.23
125	131.64	160	158.78
156	163.7	153	154.36
156	155.95	148	142.48
108	123.77	162	141.72
121	126.14	141	139.29
145	144.04	130	120.38
137	140.59	145	134.5
151	146.24	150	145.83

151	157.12	148	131.48
171	178.06	206	190.95
137	142.87	168	160.47
142	143.45	129	125.39
157	156.19	172	154.82
135	138.92	144	136.05
146	148.29	145	127.66
142	142.29	135	135.7

These values are drawn as follows:

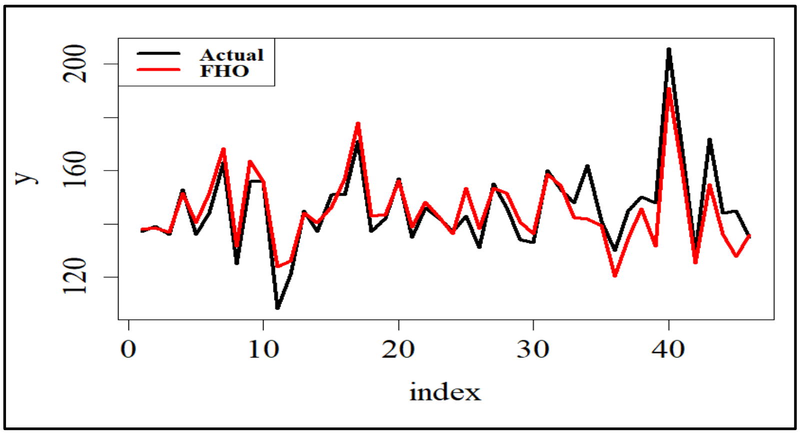

Figure (3) Air Quality Between Actual and Estimated

Figure (3) shows a comparison between actual values and predicted values using the FHO algorithm across 46 observations. We note that the red FHO line closely matches the black FHO line at most points, indicating the FHO algorithm's ability to track the general trend of the actual values efficiently.

6. Conclusions

Results showed that the Fire Hawk Optimiser (FHO) algorithm provides estimates similar to those of the Iteratively Reweighted Least Squares (IWLS) algorithm for the Gamma regression model, with differences between FHO and IWLS estimates varying with sample size, the number of explanatory variables, and the shape-parameter settings. According to the simulation results, the process of estimating the values can be improved by increasing the sample size or the value of (λ) measured by MSE, as this was reduced when the size of the sample or the value of (λ) was increased. In the actual air quality data, all estimated parameters were statistically significant, demonstrating the suitability of the fitted Gamma regression model. Additionally, the estimated values appear to exhibit good consistency with the observed values in general, thus ensuring an acceptable agreement between predicted and actual values. In general, these results indicate that the FHO algorithm is a promising metaheuristic optimisation technique for estimating Gamma regression parameters using the maximum likelihood method.

7. Recommendation

1. Thus, it is essential to observe the alerts when dust and particulate matter are announced and to enforce the strictest regulations. The 10-micrometre particle article shows a statistically significant effect on the AQI. This means that people who are sensitive to respiratory tract effects should preferably avoid outdoor exposure and use protective measures.
2. Reducing Carbon Emissions in Residential Areas: Based on the analysis indicating the significant association between ambient CO2 levels and air quality index values, cutting down on carbon emissions in residential areas can improve the air quality. Thus, cleaner energy alternatives should be promoted, and over-reliance on personal power generators should be reduced, particularly in urban areas.

3. Urban Green Investments: Since the model is efficient in monitoring air trends, expanding "green belts" and planting trees in the vicinity of houses is an important need of this social sector. These serve as a natural filter for the pollutants and particles that the statistical analysis confirmed were harmful to the environment.
4. A Call for Transparent and Sophisticated Monitoring Systems — Given that the FHO research found FHO providing very strong environmental estimates, society should call for incorporating such modern technology at the local scale, thereby enabling local authorities to adopt FHO over traditional monitoring mechanisms. This guarantees the public precise, real-time data about the air they breathe.

8. Supplementary material

(None).

9. Author's Contributions

Ali Mohammad Ali was responsible for preparing the theoretical framework of the research. Basheer Jameel Khaleel conducted the practical and applied part of the study.

10. Funding

(None).

11. Data availability statement

Meteorological Authority

12. Acknowledgements

(None).

13. Conflict of interest

The authors declare no conflict of interest.

14. Declaration of generative AI use

During the preparation of this work, the authors used (Chat GPT)and (Humanizer) for grammar checking and language polishing. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication

References

[1] Nelder, J.A., and Wedderburn, R.W.M. (1972). Generalized Linear Models. Journal of the Royal Statistical Society: Series A, 135(3), 370–384. https://doi.org/10.2307/2344614

[2] Abdaljabbar, L.A. and Nayef, Q.N. (2021). Comparison Between Maximum Likelihood and Bayesian Methods for Estimating the Gamma Regression with Practical Application. Journal of Economics and Administrative Sciences, 27(125), pp. 477–492. https://doi.org/10.33095/jeas.v27i125.2088

[3] Asar, Y., & Algamal, Z. (2022). A new two-parameter estimator for the gamma regression model. Statistics, Optimization & Information Computing, 10(3), 750-761. https://doi.org/10.19139/soic-2310-5070-822

[4] Ali, T., Al-Saffar, A., & Ismael, S. S. (2023). Using Bayes weights to estimate parameters of a Gamma Regression model. Iraqi Journal of Statistical Sciences, 20(1), 43-54. DOI: 10.33899/IQJOSS.2023.178687

[5] Abdullah, M. S., & Abood, S. N. (2024). A Comparison between Methods for Estimating the Restricted Gamma Ridge Regression Model Using the Simulation. Journal of Economics and Administrative Sciences, 30(143), 510-522. https://doi.org/10.33095/jeas.v27i125.2088

[6] Algamal, Z. Y., Salih, A. M., & Khaleel, M. A. (2024). 0544A New Ridge-Type Estimator for the Gamma regression model. Iraqi Journal for Computer Science and Mathematics, 5(1), 22. https://doi.org/10.52866/ijcsm.2024.05.01.006

[7] Abed, A. R., & Nayef Al-Qazaz, Q. N. (2025). Comparison of two approaches robust-ridge estimator in restricted additive partially regression model. In AIP Conference Proceedings (Vol. 3264, No. 1, p. 050023). AIP Publishing LLC. https://doi.org/10.1063/5.0259093

[8] Dawoud, I. (2025). A new improved estimator for the gamma regression model. Communications in Statistics - Simulation and Computation, 1–12. https://doi.org/10.1080/03610918.2025.2450722
[9] Al-Sinjary, A., & Raheem, A. (2022). Gauss-Hermite Cubature Method to estimate parameters of a multivariate GLMM. Journal of Education and Science, 31(2), 29-41. DOI:10.33899/edusj.2022.133041.1216
[10] Salinas Ruíz, J., Montesinos López, O. A., Hernández Ramírez, G., & Crossa Hiriart, J. (2023). Generalized linear mixed models with applications in agriculture and biology. Springer Nature. https://doi.org/10.1007/978-3-031-32800-8
[11] Stroup, W. W., Ptukhina, M., & Garai, J. (2024). Generalized linear mixed models: modern concepts, methods and applications. 2nd Edition. Chapman and Hall/CRC. https://doi.org/10.1080/01621459.2025.2506201
[12] Chattamvelli, R., & Shanmugam, R., (2021). Continuous Distributions in Engineering and the Applied Sciences. Morgan & Claypool. https://doi.org/10.1007/978-3-031-02435-1
[13] Hardin, J. W., & Hilbe, J. M., (2018), "Generalized linear models and extensions", 4th Edition, Stata press. https://doi.org/10.1007/978-981-96-4726-2
[14] Azizi, M., Talatahari, S., & Gandomi, A. H. (2023). Fire Hawk Optimizer: A novel metaheuristic algorithm. Artificial Intelligence Review, 56(1), 287-363. https://doi.org/10.1007/s10462-022-10173-w
[15] Hamza Obeid, R. and J. Sadik, N. (2023) "Selection of variables Affecting Red Blood Cell by Firefly Algorithm", Journal of Economics and Administrative Sciences, 29(136), pp. 90–100. https://doi.org/10.33095/jeas.v29i136.2610
[16] R Core Team. (2025). glm: Fitting generalized linear models. In R: A language and environment for statistical computing. https://stat.ethz.ch/R-manual/R-devel/library/stats/html/glm.html

المصادر

[1] نيلدر، جيه. إيه.، وودربيرن، آر. دبليو. إم .(1972) .النماذج الخطية المعممة .مجلة الجمعية الإحصائية الملكية: السلسلة أ، 135(3)، 370–384 . https://doi.org/10.2307/2344614
[2] عبد الجبار، ل. أ. ونايف، ق. ن .(2021) .مقارنة بين طريقتي الإمكان الأعظم وبيز لتقدير انحدار غاما مع تطبيق عملي .مجلة الإدارة والاقتصاد، 27(125)، ص 477–492 https://doi.org/10.33095/jeas.v27i125.2088 .
[3] عصار، ي.، والجمال، ز .(2022) .مقدار مَعلمي جديد لنموذج انحدار غاما .الإحصاء، الأمثلية وحوسبة المعلومات، 10(3)، 750-761 . https://doi.org/10.19139/soic-2310-5070-822
[4] علي، ت.، الصفار، أ.، وإسماعيل، س. س .(2023) .استخدام أوزان بيز لتقدير مَعالم نموذج انحدار غاما .المجلة العراقية للعلوم الإحصائية، 20(1)، DOI: 10.33899/IQJOSS.2023.17868754-43 .
[5] عبد الله، م. س.، وعبود، س. ن .(2024) .مقارنة بين طرق تقدير نموذج انحدار غاما الحرفي المقيد باستخدام المحاكاة .مجلة الإدارة والاقتصاد، 30(143)، 510-522 https://doi.org/10.33095/jeas.v27i125.2088 .
[6] الجمال، ز. ي.، صالح، أ. م.، وخليل، م. أ 0544 .(2024) .مقدر جديد من نوع "ريدج (Ridge) "النموذج انحدار غاما .المجلة العراقية لعلوم الحاسوب والرياضيات، 5(1)، 22 https://doi.org/10.52866/ijcsm.2024.05.01.006 .
[7] عبد، أ. ر.، ونايف القزاز، ق. ن .(2025) .مقارنة بين نهجين لمقدر "ريدج" المتين في نموذج الانحدار الجزئي الجمعي المقيد .في وقائع مؤتمر AIP)المجلد 3264، رقم 1، ص 050023). منشورات AIP. https://doi.org/10.1063/5.0259093
[8] داود، إ .(2025) .مقدر مُحسّن جديد لنموذج انحدار غاما .اتصالات في الإحصاء - المحاكاة والحوسبة، 1–12 . https://doi.org/10.1080/03610918.2025.2450722
[9] السنجري، أ.، ورحيم، أ .(2022) .طريقة غاوس-هيرميت التكعيبية لتقدير مَعالم نموذج GLMM متعدد المتغيرات .مجلة التربية والعلم، 31(2)، 29-41 DOI:10.33899/edusj.2022.133041.1216 .
[10] سالیناس رويز، جيه.، مونتيسينوس لوبيز، أو. إيه.، هيرنانديز راميريز، جي.، وكروزا هيريارت، جيه .(2023) .النماذج الخطية المختلطة المعممة مع تطبيقات في الزراعة والبيولوجيا .سبرينغر نيتشر https://doi.org/10.1007/978-3-031-32800-8 .
[11] ستروب، دبليو. دبليو.، بتوخينا، م.، وغاراي، جيه .(2024) .النماذج الخطية المختلطة المعممة: المفاهيم الحديثة، الطرق والتطبيقات .الطبعة الثانية. تشابمان وهال/CRC. https://doi.org/10.1080/01621459.2025.2506201
[12] تشاتامفيلي، آر.، وشانموغام، آر .(2021) .التوزيعات المستمرة في الهندسة والعلوم التطبيقية .مورغان وكليبول- https://doi.org/10.1007/978-3-031-02435-1 .
[13] هاردين، جيه. دبليو.، وهيلبي، جيه. إم" .(2018) .النماذج الخطية المعممة وتوسعاتها"، الطبعة الرابعة، مطبعة ستاتا . https://doi.org/10.1007/978-981-96-4726-2
[14] عزيزي، م.، تالاتاهاري، س.، وغاندومي، أ. هـ .(2023) .خوارزمية صقر النار :(Fire Hawk Optimizer) خوارزمية ميتا-هيورستية جديدة . مراجعة الذكاء الاصطناعي، 56(1)، 287-363 https://doi.org/10.1007/s10462-022-10173-w .
[15] حمزة عبيد، ر. و ج. صادق، ن" .(2023) .اختيار المتغيرات المؤثرة على خلايا الدم الحمراء باستخدام خوارزمية اليراع"(Firefly Algorithm) ، مجلة الإدارة والاقتصاد، 29(136)، ص 90–100 https://doi.org/10.33095/jeas.v29i136.2610 .
[16] فريق لغة R الأساسي :glm .(2025) .ملاءمة النماذج الخطية المعممة .في لغة R وبيئة الحوسبة الإحصائية- https://stat.ethz.ch/R-manual/R-devel/library/stats/html/glm.html .

https://doi.org/10.31272/jae.i152.1566
https://admics.uomustansiriyah.edu.iq
P-ISSN: 1813–6729 E-ISSN: 2707-1359

JAE
OPEN ACCESS

تقدير معلمات نموذج انحدار كاما باستخدام محسن صقر النار مع تطبيق عملي

بشير جميل خليل
وزارة التربية، بغداد، العراق.
Email: bashjameel73@gmail.com, ORCID: https://orcid.org/0009-0006-3377-4055

علي محمد علي
قسم الاحصاء، الجامعة المستنصرية، كلية الادارة والاقتصاد، بغداد، العراق.
Email: alialfredawy@uomustansiriyah.edu.iq, ORCID: https://orcid.org/0009-0001-8322-2121

المستخلص

يتناول هذا البحث أحد نماذج الانحدار المهمة المستخدمة في نمذجة البيانات الموجبة، وهو نموذج انحدار كاما. ويُعد هذا النموذج من النماذج الخطية المعممة، ويُستخدم عندما يكون المتغير التابع متبعًا لتوزيع كاما. وتتمثل مشكلة البحث الرئيسة في تقييم مدى قدرة خوارزمية محسن صقر النار Fire Hawk Optimizer (FHO) على تقديم بديل كفوء ومقبول لطريقة المربعات الصغرى المعاد وزنها تكراريًا Iteratively Reweighted Least Squares (IWLS) في تقدير معلمات نموذج انحدار كاما، وذلك في ظل اختلاف عدد المتغيرات، وأحجام العينات، وقيم معلمة الشكل (λ). وفي هذا البحث، تم توظيف خوارزمية FHO لتعظيم لوغاريتم دالة الإمكان ضمن إطار تقدير الإمكان الأعظم، ثم تمت مقارنة أدائها مع طريقة IWLS المستخدمة تقليديًا في النماذج الخطية المعممة.

وقد أجريت المقارنة بين الطريقتين من خلال سلسلة من تجارب المحاكاة التي تضمنت أحجام عينات مختلفة، وأعدادًا مختلفة من المتغيرات، وقيمًا متعددة لمعلمة الشكل (λ) ضمن سيناريوهات متنوعة. وأظهرت النتائج أن خوارزمية FHO تحقق أداءً مقاربًا لأداء خوارزمية IWLS، ولا سيما في السيناريوهات التي تتضمن أحجام عينات صغيرة، مع بقاء الفروق بين الطريقتين محدودة نسبيًا، فضلًا عن الانخفاض الواضح في قيم متوسط مربعات الخطأ MSE مع زيادة حجم العينة.

أما في الجانب التطبيقي، فقد تم تحليل تأثير عدد من المتغيرات في جودة الهواء بالاعتماد على 46 مشاهدة مأخوذة من هيئة الأنواء الجوية العراقية. وأكدت نتائج اختبار حسن المطابقة أن المتغير التابع يتبع توزيع كاما. كما أظهرت نتائج التقدير أن جميع المعلمات كانت معنوية إحصائيًا، مما يدعم ملاءمة نموذج انحدار كاما المستخدم، ويشير إلى قدرة خوارزمية FHO على تقديم تقديرات دقيقة وموثوقة ضمن إطار تقدير الإمكان الأعظم.

معلومات البحث

تواريخ البحث:
التقديم: 15 / 02 / 2026
المراجعة: 03 / 05 / 2026
قبول النشر: 13 / 05 / 2026
نشر الكتروني: 01 / 06 / 2026
تسلسل الصفحات: 123 - 135

الكلمات المفتاحية:
انحدار كاما، محسن صقر النار، النماذج الخطية المعممة، تقدير المعلمات، متوسط مربعات الخطأ.

المراسلة:
أسم الباحث: بشير جميل خليل

Email: bashjameel73@gmail.com