**Original Research Article**

**Estimating Parameters of the Beta Regression Model Using the Dung Beetle Optimizer: Comparative Analysis with BFGS and Application to Thalassemia Data**

**ABSTRACT**

The beta regression model is used for data that is continuous but confined between zero and one, such as percentages. To study these models, it is necessary to estimate their parameters using common estimation methods, primarily the Maximum Likelihood Estimator (MLE). In this research, we relied on a modern metaheuristic algorithm, the Dung Beetle Optimizer (DBO), and compared its behavior with a conventional method, the BFGS algorithm. Simulations were conducted to compare the two methods using the Mean Squared Error (MSE) criterion, the results showed that DBO provided more accurate estimates, particularly in small-sample and high-precision scenarios. Furthermore, the impacts of factors including age, hemoglobin level, and the quantity of blood units on the packed cell volume (PCV) were investigated using the Beta regression model on actual data from thalassemia patients. The application demonstrated how well the model handled proportional data with a 0–1 boundary and demonstrated how important some variables were in describing the event under study.

Keyword / Beta Regression Model, Metaheuristic Algorithms, BFGS Algorithm, Dung Beetle Optimizer.

1. **Introduction**

Analyzing the relationship between explanatory variables and the dependent variable is a fundamental focus of applied statistics, and the multiple linear regression model is widely used for this purpose. However, one of the most important underlying assumptions of this model is that the dependent variable follows a normal distribution. In practical applications, we often deal with relative data between zero and one, such as success rates, financial allocation ratios, or performance indicators, which do not conform to this assumption, necessitating the use of alternative models.

Among these models, the beta regression model stands out as a suitable option for analyzing continuous data confined to the open interval (0,1), due to its flexibility in representing the distribution of relative data. This model relies on the beta distribution, combined with an appropriate link function (often the logit) to link the mathematical prediction of the dependent variable to the independent variables.

Estimating the parameters of this model is often done using the maximum likelihood Estimation (MLE) method, which requires maximizing the likelihood function, a nonlinear numerical task that can be complex in some cases, especially when multiple variables or non-typical data are involved. For this reason, researchers have recently begun employing intelligent metaheuristic optimization algorithms, which have demonstrated high efficiency in this area.

In this research, the Dung Beetle Optimizer (DBO), a recent algorithm proposed in 2022, which is inspired by the behavior of the dung beetle in navigation, searching, and adaptation, is adopted and compared with the classic BFGS algorithm to estimate the parameters of a beta regression model, using the mean square error (MSE) criterion as a trade-off measure using simulation experiments involving many cases and sample sizes.

The rest of the paper will be divided into the following sections: The second section deals with the beta regression model in terms of its definition, properties, and possibility function. Section 3 explains the BFGS algorithm used to estimate the model parameters, explaining its mechanism of action. Section 4 reviews the Dung Beetle Optimizer (DBO) algorithm as a metaheuristic algorithm, outlining its steps. Section 5 addresses the design of the experimental study and the presentation and analysis of simulation results. Section 6 examines the application to real data. Finally, Section 7 presents the most important conclusions drawn from the research.

1. **Materials and methods**

**2.1- Beta Regression Model:**

The Beta distribution is one of the important exponential family distributions, characterized by its boundaries between (0 and 1), which makes it suitable for ratios. Examples of this distribution include the success rate of students in a certain school over several years, the percentage of fat in the human body, the percentage of cells in the body of a sick person, or the percentage of spending on a specific item relative to total spending. The Beta distribution can be expressed as follows [1]:

The expectation and variance of the distribution are:

"Beta regression, first introduced by Ferrari and Cribari-Neto (2004) [2], has since gained considerable attention and has been further extended in recent years in various fields such as biostatistics, econometrics, and machine learning [3–6].

Since we assume [7]:

Thus, the beta distribution can be reformulated as follows:

The following figure shows the plot of the probability density function (PDF) of the beta distribution for different values ​​of the parameters μ and ω:

|  |  |
| --- | --- |
|  |  |

Fig. 1: The probability density function of the Beta distribution

The figures illustrate how the shape of the beta distribution is affected by changes in the parameters μ and ω, where the parameter 𝜇 controls the location of the center (right or left), while the parameter 𝜔 determines the degree of sharpness or dispersion of the distribution. The figures show that the distribution can be symmetric, skewed, or highly centered, reflecting its high flexibility in representing relative data within a range of (0,1).

A beta regression model can be obtained by making:

Where represents the link function, which in the beta distribution is non-linear, but rather represents the logit function, as follows:

Therefore:

The likelihood function will be, after substituting as equal to it, as follows:

The parameters of the beta regression model will be estimated by the Maximum Likelihood Estimation (MLE) method [8]. The maximization of the log-likelihood function (Equation 6) is typically performed using optimization algorithms such as the Conjugate Gradient [9] Method, Artificial Bee Colony Algorithm [10], and Hybrid HS-CG Update Method [11], as well as other techniques [12]. However, in this paper, the parameters are estimated using both the BFGS algorithm and the Dung Beetle Optimizer (DBO) algorithm.

**2.2- BFGS Algorithm**

The BFGS (Broyden, Fletcher, Goldfarb, and Shanno) algorithm is a quasi-Newton method for optimizing multivariable nonlinear functions. It was proposed independently by Broyden, 1970 [13], Fletcher, 1970 [14], Goldfarb, 1970 [15], and Shanno, 1970 [16], hence the name. It is used to find the minimum endpoints of a differentiable objective function. It was originally developed to minimize nonlinear functions, but it can also be used to maximize functions by minimizing the negative of the function to be maximized, i.e.:

This method is distinguished by the fact that it relies only on the first derivative of the function and not the Hessian matrix, which represents the matrix of the second derivatives, because it calculates this matrix iteratively in an approximate manner using the following equation [17, 18]:

so, if

Then the following formula is used for updating:

where represents the vector of parameters, and is determined using linear search to ensure that the function is sufficiently minimized.

**2.3- Dung Beetle Optimizer (DBO)**

The Dung Beetle Optimizer (DBO) algorithm is a metaheuristic algorithm inspired by the natural behavior of dung beetles. It was first proposed in 2022 by Xue & Shen [19] as a novel technique based on collective intelligence, accurately mimicking the life cycle of the dung beetle. The algorithm relies on five key behaviors inspired by nature to guide the search within the solution space, as follows:

* *Rolling Behavior*

This behavior simulates the movement of a dung beetle as it pushes a dung ball toward a target, and includes updating the position using drift coefficients and the effect of the worst-world solution:

where:

f: is a random direction (-1 or +1).

b: is a constant.

: is the global worst position.

* *Dancing Behavior*

When the beetle encounters obstacles, it redirects its path by "dancing" which is represented using the tangent function:

where:

: is the deflection angle.

* *Brood Ball Behavior*

This stage represents the process of laying eggs in dung balls within an incubation area that is dynamically determined during iterations:

*Step 1*: Define Bounds:

*Step 2*: Update Positions:

where:

: is the current local best.

R: dynamic shrinking factor.

and : are random vectors.

* *Foraging Behavior*

After the eggs hatch, the young beetles begin searching for food within an ideal area that has been identified:

where:

and : are bounds of the foraging area.

* *Stealing Behavior*

This behavior mimics beetles that steal dung balls from others:

where:

: is the global best.

: is the local best.

: is a random normal vector.

: is a constant.

After each position update, we ensure that all sites remain within the permitted limits:

By converting the estimation problem into a numerical optimization problem, the Dung Beetle Optimizer (DBO) technique is used to estimate the parameters of statistical models. A fitness function—here specified as the negative logarithm of the likelihood function is used to assess each solution, with each agent in the swarm representing a vector of suggested parameters (such as the regression model parameters). To get the optimal estimate of the parameters, the algorithm aims to minimize this function. DBO uses behaviors like rolling, dancing, brood balling, foraging, and thieving that are modeled after those of dung beetles to improve its capacity to seek globally and avoid local optima. The algorithm is appropriate for difficult or nonlinear issues because it doesn't require knowledge of the function's derivatives. The optimal agent, or optimal collection of parameters, is selected as the ultimate estimation solution at the conclusion of the iterations.

1. **Results and Discussion**

**3.1- Simulation**

The BFGS algorithm and the metaheuristic Dung Beetle Optimizer (DBO) were both used in a simulation study to assess and contrast the performance of the beta regression parameter estimation techniques. All simulations were carried out in the R programming environment, considering multiple scenarios to ensure the robustness of the findings, and several situations were taken into account to guarantee the findings' robustness.

Using the "BFGS" approach and R's built-in optim function, the classical estimation was performed [20]. For the Beta regression model, a bespoke log-likelihood function was designed, in which the logit function links the linear predictor to the mean of the Beta distribution.

The multi-stage nature of the Dung Beetle Optimizer (DBO) led to its complete implementation in R. The log-likelihood function, along with parameter constraints and optimization parameters, was intended to be accepted as input by the DBO function.

The Mean Squared Error (MSE) was employed as the evaluation tool in order to compare the two approaches' performances objectively. The MSE was computed as follows:

Where:

R: is the total number of replications,

: is the vector of estimated parameters in replication r.

: is the true parameter vector.

The simulation was carried out under multiple scenarios by varying both the number of predictors (1, 3, 5, 7) and the precision parameter ω (4 cases: 5, 10, 15, 20). The true parameter vectors, , used in each case were as follows:

1 Predictor: (1.0, 0.5)

3 Predictors: (1.0, 0.5, -0.8, 0.3)

5 Predictors: (1.0, 0.5, -0.8, 0.3, 0.7, -0.4)

7 Predictors: (1.0, 0.5, -0.8, 0.3, 0.7, -0.4, 0.2, -0.6)

To guarantee the consistency and dependability of the outcomes, every simulation scenario was repeated a thousand times.

The Mean Squared Error (MSE) for both the BFGS and DBO methods is displayed in Tables (1-4) under various circumstances, sample sizes (ranging from 25 to 250), and precision parameter ω values (5, 10, 15, 20).

Table 1: MSE for Scenario with 1 Predictor across Different ω Values and Sample Sizes

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | ω=5 | | ω=10 | | ω=15 | | ω=20 | |
|  | **BFGS** | **DBO** | **BFGS** | **DBO** | **BFGS** | **DBO** | **BFGS** | **DBO** |
| 25 | 0.46577 | 0.46588 | 2.05677 | 1.66881 | 3.59185 | 1.74359 | 7.01208 | 3.28977 |
| 50 | 0.14851 | 0.14852 | 0.62050 | 0.61686 | 1.31559 | 0.90188 | 2.44862 | 2.09960 |
| 100 | 0.05365 | 0.05365 | 0.25142 | 0.25128 | 0.59734 | 0.48652 | 1.10476 | 1.05736 |
| 150 | 0.03902 | 0.03902 | 0.14833 | 0.14832 | 0.25904 | 0.26910 | 0.64987 | 0.64998 |
| 200 | 0.03269 | 0.03270 | 0.11695 | 0.11694 | 0.24268 | 0.26066 | 0.45848 | 0.45835 |
| 250 | 0.01853 | 0.01852 | 0.08701 | 0.08699 | 0.13255 | 0.16821 | 0.35567 | 0.35574 |

Improved estimation accuracy with larger sample sizes is confirmed by the results in table 1, which clearly indicate a drop in MSE as sample size grows. Furthermore, DBO generally performs marginally better than BFGS, especially for smaller sample sizes and higher ω values. As the sample size increases, both approaches generally converge identically.

Table 2: MSE for Scenario with 3 Predictors across Different ω Values and Sample Sizes

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | *ω=5* | | *ω=10* | | *ω=15* | | *ω=20* | |
|  | *BFGS* | *DBO* | *BFGS* | *DBO* | *BFGS* | *DBO* | *BFGS* | *DBO* |
| 25 | 0.79773 | 0.79708 | 2.30782 | 2.27401 | 7.63776 | 6.06802 | 15.57228 | 8.33128 |
| 50 | 0.26196 | 0.26189 | 1.21141 | 1.20096 | 1.50071 | 1.48561 | 3.34444 | 3.16977 |
| 100 | 0.10149 | 0.10138 | 0.22599 | 0.22591 | 0.73279 | 0.73789 | 1.19796 | 1.34805 |
| 150 | 0.05645 | 0.05692 | 0.22602 | 0.22562 | 0.40522 | 0.40059 | 0.70298 | 0.94167 |
| 200 | 0.04735 | 0.04711 | 0.11032 | 0.11033 | 0.25693 | 0.25286 | 0.46576 | 0.46502 |
| 250 | 0.03378 | 0.03385 | 0.08240 | 0.08207 | 0.19384 | 0.19165 | 0.38871 | 0.48451 |

Improved parameter estimation is confirmed as sample size increases, since the MSE values steadily decline with bigger sample sizes. In a number of situations, DBO performs marginally better than BFGS, particularly when ω = 20 and sample sizes are modest. As the sample size grows to 250, the performance of both approaches converges.

Table 3: MSE for Scenario with 5 Predictors across Different ω Values and Sample Sizes

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | *ω=5* | | *ω=10* | | *ω=15* | | *ω=20* | |
|  | *BFGS* | *DBO* | *BFGS* | *DBO* | *BFGS* | *DBO* | *BFGS* | *DBO* |
| 25 | 1.45300 | 1.48764 | 5.86146 | 5.03035 | 13.8377 | 6.91296 | 15.2595 | 7.47610 |
| 50 | 0.39591 | 0.43857 | 1.28198 | 1.27899 | 3.33580 | 2.31201 | 5.39101 | 2.84046 |
| 100 | 0.15501 | 0.19212 | 0.34429 | 0.31045 | 0.80120 | 0.82682 | 2.03093 | 1.70156 |
| 150 | 0.08521 | 0.13412 | 0.18095 | 0.19527 | 0.43901 | 0.46846 | 0.77810 | 1.09005 |
| 200 | 0.05439 | 0.07121 | 0.16296 | 0.18399 | 0.27104 | 0.34452 | 0.65225 | 0.95620 |
| 250 | 0.05167 | 0.06419 | 0.09325 | 0.10025 | 0.20176 | 0.28178 | 0.45572 | 0.63728 |

As the sample size grows, the findings in table 3 show a consistent MSE decrease, indicating increased estimation accuracy. In a number of high-ω cases, DBO outperforms BFGS, especially for small samples. But as the sample size increases, the disparities between approaches become less pronounced.

Table 4: MSE for Scenario with 7 Predictors across Different ω Values and Sample Sizes

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | *ω=5* | | *ω=10* | | *ω=15* | | *ω=20* | |
|  | *BFGS* | *DBO* | *BFGS* | *DBO* | *BFGS* | *DBO* | *BFGS* | *DBO* |
| 25 | 2.64108 | 3.34451 | 8.84767 | 4.76584 | 21.25468 | 9.78904 | 39.66581 | 9.39966 |
| 50 | 0.53154 | 1.68773 | 1.85347 | 2.67295 | 2.21837 | 3.80604 | 6.21505 | 8.67853 |
| 100 | 0.15120 | 0.63269 | 0.58819 | 2.16603 | 0.78004 | 2.57998 | 1.47338 | 7.20768 |
| 150 | 0.10829 | 0.31143 | 0.27342 | 1.97620 | 0.37661 | 1.71014 | 0.92925 | 5.62315 |
| 200 | 0.08503 | 0.22524 | 0.16027 | 1.03053 | 0.28340 | 1.49708 | 0.61820 | 4.86785 |
| 250 | 0.05641 | 0.12086 | 0.12683 | 0.89957 | 0.18901 | 1.29930 | 0.41390 | 3.79288 |

Larger sample sizes numbers typically result in lower MSE values, indicating better estimating. It's interesting to note that although DBO performs better than BFGS in certain small-sample, high-ω scenarios, it exhibits instability in others, especially when predictors increase. Although DBO exhibits heterogeneity, both approaches converge as sample numbers increase. This finding highlights the importance of adjusting the optimization approach to the specifics of the data, especially in high-dimensional, small-sample scenarios.

As seen by declining MSE values, the data in Tables 1-4 consistently demonstrate that estimation accuracy is improved with larger sample sizes. Although DBO occasionally exhibits unpredictability as the number of predictors rises, it often performs better than BFGS in small-sample and high-ω circumstances. The efficacy of both approaches in beta regression modeling is confirmed by the tendency for their performance to converge with larger datasets.

**3.2- Application**

The information utilized in this study (Table 5) came from patients with β-thalassemia (Mediterranean anemia) at the Ibn Al-Atheer Teaching Hospital for Maternity and Children in Mosul. Al-Nu’eimy [21] is the source of the dataset. In this study, the Packed Cell Volume (PCV), which stands for the hematocrit level (the percentage of blood volume occupied by red blood cells), is the dependent variable (y). The following list of explanatory variables was chosen:

X1: Real age of the patient, measured in months (Real Age – Month).

X2: Hemoglobin level (Hemoglobin – g/dL).

X3: Number of blood units transfused (Numbers of Blood Units).

Table 5: Dataset of Thalassemia Patients

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| *No* | *Y* | *X1* | *X2* | *X3* | *No* | *Y* | *X1* | *X2* | *X3* |
| 1 | 0.22 | 147 | 6.9667 | 150 | 26 | 0.16 | 18 | 4.9167 | 0 |
| 2 | 0.17 | 153 | 5.6333 | 162 | 27 | 0.23 | 109 | 7.3333 | 102 |
| 3 | 0.23 | 114 | 7.4833 | 106 | 28 | 0.18 | 77 | 6.05 | 66 |
| 4 | 0.18 | 84 | 5.75 | 72 | 29 | 0.22 | 60 | 7.0667 | 23 |
| 5 | 0.2 | 96 | 6.5833 | 81 | 30 | 0.16 | 46 | 5.4167 | 24 |
| 6 | 0.18 | 82 | 5.9 | 57 | 31 | 0.14 | 74 | 4.0833 | 66 |
| 7 | 0.19 | 68 | 6.0667 | 54 | 32 | 0.32 | 113 | 10.5833 | 80 |
| 8 | 0.2 | 24 | 6.3333 | 0 | 33 | 0.25 | 50 | 8.4167 | 18 |
| 9 | 0.17 | 65 | 5.4333 | 56 | 34 | 0.28 | 82 | 9.4167 | 54 |
| 10 | 0.14 | 83 | 5.0333 | 74 | 35 | 0.21 | 90 | 6.6 | 78 |
| 11 | 0.16 | 51 | 5.4167 | 25 | 36 | 0.08 | 50 | 2.8333 | 30 |
| 12 | 0.18 | 145 | 5.85 | 145 | 37 | 0.23 | 42 | 7.4833 | 7 |
| 13 | 0.24 | 89 | 7.85 | 70 | 38 | 0.23 | 54 | 7.5833 | 24 |
| 14 | 0.18 | 95 | 6 | 75 | 39 | 0.25 | 157 | 8.4167 | 149 |
| 15 | 0.18 | 70 | 5.85 | 57 | 40 | 0.24 | 175 | 8.1 | 170 |
| 16 | 0.1 | 44 | 3.1667 | 39 | 41 | 0.1 | 67 | 3.4167 | 60 |
| 17 | 0.2 | 39 | 6.3333 | 20 | 42 | 0.12 | 45 | 4.05 | 30 |
| 18 | 0.2 | 75 | 6.3333 | 60 | 43 | 0.3 | 109 | 10.25 | 89 |
| 19 | 0.24 | 56 | 7.75 | 24 | 44 | 0.26 | 58 | 8.4333 | 18 |
| 20 | 0.19 | 24 | 6.1667 | 0 | 45 | 0.18 | 147 | 5.9 | 142 |
| 21 | 0.21 | 88 | 6.7 | 70 | 46 | 0.27 | 137 | 9.6 | 125 |
| 22 | 0.19 | 97 | 6.0667 | 82 | 47 | 0.21 | 40 | 6.75 | 18 |
| 23 | 0.14 | 20 | 4.8333 | 0 | 48 | 0.21 | 46 | 6.9 | 10 |
| 24 | 0.14 | 109 | 4.3833 | 102 | 49 | 0.16 | 88 | 5.4667 | 84 |
| 25 | 0.15 | 45 | 5 | 21 | 50 | 0.23 | 65 | 7.4833 | 46 |

The statistical program EasyFit 5.6's goodness-of-fit tests were used to make sure the dependent variable data in Table 5 adhered to the Beta distribution. In particular, three tests were carried out: Kolmogorov-Smirnov (KS) Test, Anderson-Darling (AD) Test and Chi-Squared Test. The hypotheses for tests were formulated as follows:

The results of these tests are summarized in Table 6.

Table 6: Goodness-of-Fit Tests for the Dependent Variable

|  |  |  |  |
| --- | --- | --- | --- |
|  | KS | AD | χ2 |
| Statistic | 0.07357 | 0.21072 | 1.2905 |
| Critical Value | 0.18841 | 2.5018 | 11.07 |

As can be seen in Table 6, the null hypothesis cannot be rejected since the computed test statistics are less than the critical values at the 0.05 significance level. This demonstrates that the data from the dependent variable comply with the beta distribution. Furthermore, the fitted Beta distribution curve is depicted in the following figure:

****

Fig. 2: The PDF of the Beta Distribution Based on the Real Data

As explained in the theoretical section, the Dung Beetle Optimizer (DBO), a metaheuristic optimization technique, was used to estimate the Beta regression parameters. A specially written function that maximizes the Beta regression log-likelihood function was used to carry out the implementation in the R environment. The following is how the estimation results are displayed:

Table 7: Estimated Parameters Using the Dung Beetle Optimizer (DBO)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| *Parameters* | *Estimate* | *Std. Error* | *z value* | *Pr(>|z|)* |
|  | -1.9953 | 0.2902 | -6.8767 | <0.001 |
|  | -0.0331 | 0.0158 | -2.0908 | 0.0365 |
|  | 0.2511 | 0.0641 | 3.9191 | <0.001 |
|  | 0.0264 | 6.8440 | 2.0651 | 0.0389 |

The estimated parameters of the beta regression model using the Dung Beetle Optimizer (DBO) are displayed in Table 7. Standard errors, z-values, and matching p-values are included for every parameter. Interestingly, β₁ and β₃ are significant at the 5% level, although β₀ and β₂ are very significant (p < 0.001). Because this is a beta regression model with a logit link, the coefficients show how the predictors affect the dependent variable's log-odds of mean rather than the dependent variable directly. Therefore, rather than concentrating on a direct change in the response variable, interpretation should concentrate on the strength and direction of the link. The Beta regression model, which represents the connection between the dependent variable and the explanatory variables as follows, was developed using the estimations that were obtained:

The plot of the estimated values against the actual data in the following image shows how closely the fitted Beta regression model matches the actual data.

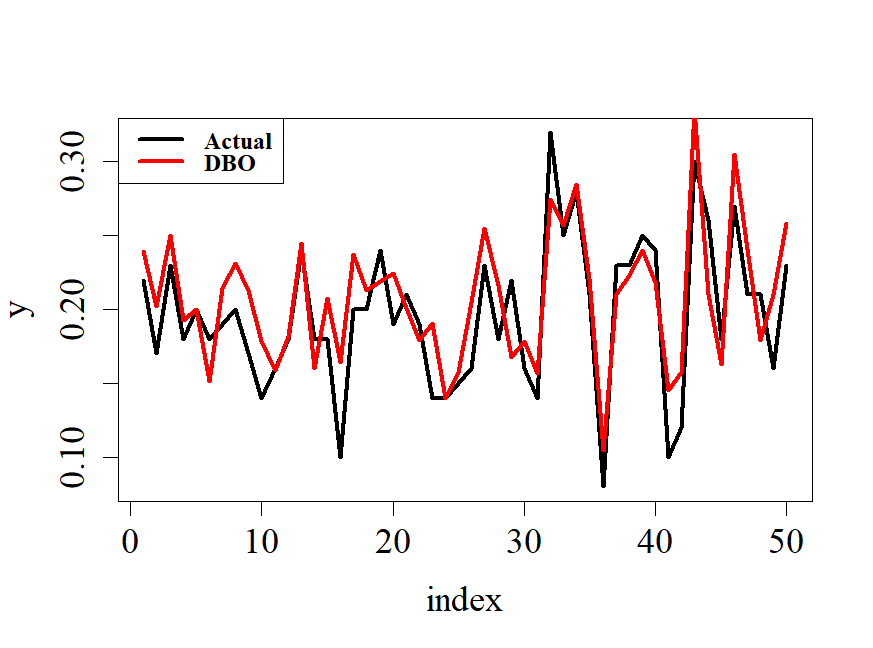


Fig. 3: Actual and Fitted Packed Cell Volume (PCV) Values

The study's sample size and the medical dataset's specificity are its limitations, despite its encouraging results. It is advised to conduct additional validation with bigger datasets and other optimizers. However, the combination of statistical modeling with bio-inspired techniques offers a useful tool for reliable parameter estimate.

1. **Conclusion**

The following conclusions can be made in light of the simulation and real-data results:

1. According to the simulation results, both the BFGS and DBO approaches exhibit better estimation accuracy, with the Mean Squared Error (MSE) constantly decreasing as the sample size grows.
2. The Dung Beetle Optimizer (DBO) demonstrated its strength in difficult optimization tasks by outperforming the standard BFGS algorithm in small-sample and high-precision (ω) circumstances.
3. The beta regression model's appropriateness for the actual dataset under analysis was confirmed by the goodness-of-fit tests, which showed that the dependent variable data had a beta distribution.
4. As demonstrated by the close fit between observed and predicted values, the beta regression model, which was estimated using DBO, effectively captured the relationship between the variables. The application results also indicated that a number of predictors had a statistically significant impact on the dependent variable.

COMPETING INTERESTS DISCLAIMER:

Authors have declared that they have no known competing financial interests OR non-financial interests OR personal relationships that could have appeared to influence the work reported in this paper.

Option 2:

Author(s) hereby declare that generative AI technologies such as Large Language Models, etc. have been used during the writing or editing of manuscripts. This explanation will include the name, version, model, and source of the generative AI technology and as well as all input prompts provided to the generative AI technology

Details of the AI usage are given below:

1. Technology used: ChatGPT (OpenAI GPT-4)

2. Purpose: Discussion and clarification of statistical modeling concepts

3. No manuscript content was generated or modified directly by AI tools.

**References**

1. Shanmugam RCaR. Continuous Distributions in Engineering and the Applied Sciences: Springer Nature; 2021.<https://doi.org/10.1007/978-3-031-02430-6>

2. Ferrari S, Cribari-Neto F. Beta Regression for Modelling Rates and Proportions. Journal of Applied Statistics. 2004;31(7):799–815.<https://doi.org/10.1080/0266476042000214501>

3. GEISSINGER, Emilie A., et al. A case for beta regression in the natural sciences. Ecosphere, 2022, 13.2: 339-353.‏ <https://doi.org/10.1002/ecs2.3940>

4. CRIBARI-NETO, Francisco. A beta regression analysis of COVID-19 mortality in Brazil. Infectious Disease Modelling, 2023, 8.2: 309-317.‏ <https://doi.org/10.1016/j.idm.2023.02.005>

5. CAPELLETTI, Marco; RAIMONDO, Davide M.; DE NICOLAO, Giuseppe. Wind power curve modeling: A probabilistic Beta regression approach. Renewable Energy, 2024, 223: 119970.‏ <https://doi.org/10.1016/j.renene.2024.119970>

6. OLALUWOYE, Olalekan T., et al. Robust estimation methods for addressing multicollinearity and outliers in beta regression models. Scientific Reports, 2025, 15.1: 11649.‏ <https://doi.org/10.1038/s41598-025-85553-7>

7. W. W. Stroup MP, & J. Garai. Generalized linear mixed models: modern concepts, methods and applications: Chapman and Hall/CRC; 2024.<https://doi.org/10.1201/9780429092060>

8. E. A. Hussain ANA-S, & H. A. Saied. Using maximum likelihood method to estimate parameters of the linear regression t truncated model. NTU Journal of Pure Sciences. 2022;1(4):26–34.<https://doi.org/10.56286/ntujps.v1i4.313>

9. Saeed IO, Mahmood HA. A new formula for conjugate-gradient in unconstrained optimization. Tikrit J Pure Sci. 2022;26(3):75–83. <doi:10.25130/tjps.26.3.2022.07>

10. Salih AM, Hussein MH. Projects selection in knapsack problem by using artificial bee colony algorithm. Tikrit J Pure Sci. 2023;27(2):137–142. <doi:10.25130/tjps.27.2.2023.17>

11. Abdullah SM, Latif SA. An effective HS-CG update: avoiding oscillations in nonlinear optimization. Tikrit J Pure Sci. 2025;29(4):95–106. <doi:10.25130/tjps.29.4.2025.14>

12. Ayyed HH, Al-Sinjary AM. Estimating parameters of the Beta regression model using dung beetle optimizer. Tikrit J Pure Sci. 2025;29(2):99–108. [doi: 10.25130/tjps.29.2.2025.10](doi:%2010.25130/tjps.29.2.2025.10)

13. Broyden CG. The convergence of a class of double-rank minimization algorithms. IMA Journal of Applied Mathematics. 1970;6(1):76–90.<https://doi.org/10.1093/imamat/6.1.76>

14. Fletcher R. A new approach to variable metric algorithms. The computer journal. 1970;13(1):317–22.<https://doi.org/10.1093/comjnl/13.3.317>

15. Goldfarb D. A family of variable-metric methods derived by variational means. Mathematics of computation. 1970;24(109):23–6.<https://doi.org/10.2307/2004873>

16. Shanno DF. Conditioning of quasi-Newton methods for function minimization. Mathematics of computation. 1970;24(111):647–56.<https://doi.org/10.2307/2004840>

17. Corriou JP. Numerical Methods of Optimization Springer International Publishing; 2022.<https://doi.org/10.1007/978-3-030-89366-8>

18. Al-Sinjary, A M. Comparison of the Performance of Some Estimation Methods for the Parameters of the McDonald Fréchet Distribution with an Empirical Application. 2025; 17(57): 7681-7661.

19. Xue J, Shen B. Dung beetle optimizer: a new meta-heuristic algorithm for global optimization. The Journal of Supercomputing. 2022;79(7):7305–36.<https://doi.org/10.1007/s11227-022-04959-6>

20. Team RC. optim: General-purpose optimization. In R: A language and environment for statistical computing. 2025.<https://stat.ethz.ch/R-manual/R-devel/library/stats/html/optim.html>

21. Al-Nu’eimy AMTR. Variables Selection in Ridge Regression: University of Mosul; Master Thesis. 2005.