Journal of Kyzylorda Scholarly Review

P-ISSN:3006-0265 | E-ISSN:3007-4681. Published by the Open University of Kyzylorda
Copyright © Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC-BY).
https://bulletin.ouk.kz/index.php/bulletin/index



COMPARISON OF THE USE OF BIO-INSPIRED ALGORITHMS AND CONVENTIONAL METHODS FOR ESTIMATING THE POWER FACTOR IN TWEEDY DISTRIBUTIONS

Wakkaa AliI Hadba[1*]
University of Kirkuk
college of management and economic
wakaa2017@uokirkuk.edu.iq

Abstract. This paper will estimate the parameters of Tweedie distribution and make use of both the traditional methods and nature-inspired optimization algorithms to estimate the distribution through realistic rainfall data in Kirkuk, Iraq. These estimation methods were the method of moments (MoM), Bayesian method, maximum likelihood (MLE), and profile likelihood, and evolutionary algorithms, namely, the genetic algorithm (GA), molecular swarm algorithm (PSO), gray wolf algorithm (GWO), and white whale algorithm (WOA) . Different indices of trade-offs (maximum likelihood (MLE) method, PSO and GA algorithms) were used to assess the performance of various models; it was found that these algorithms give better results in terms of goodness of fit (AIC and BIC) and predictive accuracy (RMSE and MAE). Conversely, the moment method showed low performance as compared to other methods though the GWO and WOA algorithm gave acceptable and less stable estimates. In addition, the partial likelihood approach was also found to be sensitive to network settings, which produced unreliable results. The Bayesian model offered a convenient probabilistic method of parameters distributions, yet it needs additional software enhancement to address the Tweedy distribution better. The paper found the maximum likelihood approach (MLE) and the most recent optimization techniques, including PSO and GA, to be the best options to easily and precisely estimate Tweedy distribution parameters. The research suggests that more sophisticated optimization methods should be used and more specific Bayesian applications should be created to assist the investigation of such distribution in the upcoming studies.

Keywords: Tweedy distribution, parameter estimation, heuristic meta-optimization, maximum likelihood, profile likelihood; MoM , MLE ,Bayesian methods.

1. INTRODUCTION

Tweedie distribution is among the most used and wide ranging statistical models because of the massive capabilities of modeling mixed data, both continuous and discrete values. The characteristic of this distribution is its capability to deal with data, which is highly dispersed or skewed and this is the reason it is an appropriate distribution used in the analysis of the data containing zero and positive values. This is Tweedie distribution that is common in many areas, like analysis of insurance claims, renewable energy and weather and rainfall patterns analysis. However, estimating its main parameters the power coefficient (p), the mean coefficient (μ), and the dispersion coefficient (ϕ) remains a challenging task due to the complex nature of the likelihood function, nonlinearity of the variance function and sensitivity of conventional techniques to starting values.This research focuses on coming up with an all-inclusive method to enhance the estimation of the Tweedies distribution parameters by amalgamating conventional and contemporary techniques. Although the literature related to the Tweedy distribution and their application in different areas of life; insurance, energy and the environment, most of the past research has been done using traditional statistics like maximum likelihood estimation (MLE), method of moments (MoM) and the estimation using numerical algorithms like the Newton-Raphson and the Fisher. Simultaneously, nature-inspired metaheuristic algorithms have developed to become potent regression, optimization of nonlinear functions and parameter estimators. Nonetheless, it appears that these algorithms have not been directly applied to enhance the estimation of the parameter p of the Tweedy distribution, so this research gap is the focus of this study where a framework is developed on the basis of metaheuristic algorithms like particle swarm optimization (PSO), whale optimization algorithm (WOA), and others. and gray wolf algorithm (GWO) to estimate the parameter p. The paper is also aimed at comparing the efficiency of these algorithms to traditional statistical algorithms in terms of accuracy of estimation and efficiency. Conventional techniques like the method of moments (MoM) and maximum likelihood estimation (MLE) were covered and the current techniques like Bayesian analysis and nature-inspired algorithms like genetic algorithms (GA), PSO, GWO, and WOA. All these approaches were applied to the similar dataset to guarantee the similarity of the results of the comparison using objective measures of root mean square error (RMSE), mean absolute error (MAE), Akaike information criterion (AIC), and Bayesian information criterion (BIC). The given comparative approach is aimed at discovering the most solid and correct way to estimate the parameters of the Tweedy distribution contributing to the advancement of the statistical models in situations with mixed data. The study focuses on estimating three main parameters: the power parameter (p), the arithmetic mean (μ), and the dispersion parameter (ϕ). The power factor (p) is a key component of the Tweedie distribution due to its role in determining the relationship between the variance and the mean, making it a pivotal element in data modeling. Using the modern estimation techniques, thisdistribution has an opportunity to be greatly improved in the accuracy and efficiency of its statistical models [1]. Numerous researchers have analyzed the strength parameter The index coefficient is usually defined by the actuaries to a default value and the tweedy package in R gives functions to estimate the power of tweedy through maximum likelihood estimation. They examined the existence of implicit estimator of power factor in the Tweedy distribution[2]. Gave a thesis that showed the possible advantages of employing generalized estimating equations to model and explain past rainfall data by examining monthly rainfall regimes in Emerald, Toowoomba and Gatton.[3]. The importance of the power factor in the context of Tweedie models is that this is an indicator which gives some difference between certain significant continuous distributions.[4]. This research will be organized in the following way: Section One: Introduction, research objective, and literature review. Section Two:Materials and Methods, with mathematical definition: probability density, variance function, and parameters (p, μ, ϕ). Estimation methods and typical evaluation metrics.Data Description.Section Three: Discussion of Results. Section five : Conclusion and Recommendations. As in diagram No. (1)

Journal of Kyzylorda Scholarly Review
P-ISSN:3006-0265 | E-ISSN:3007-4681
https://bulletin.ouk.kz/index.php/bulletin/index.php

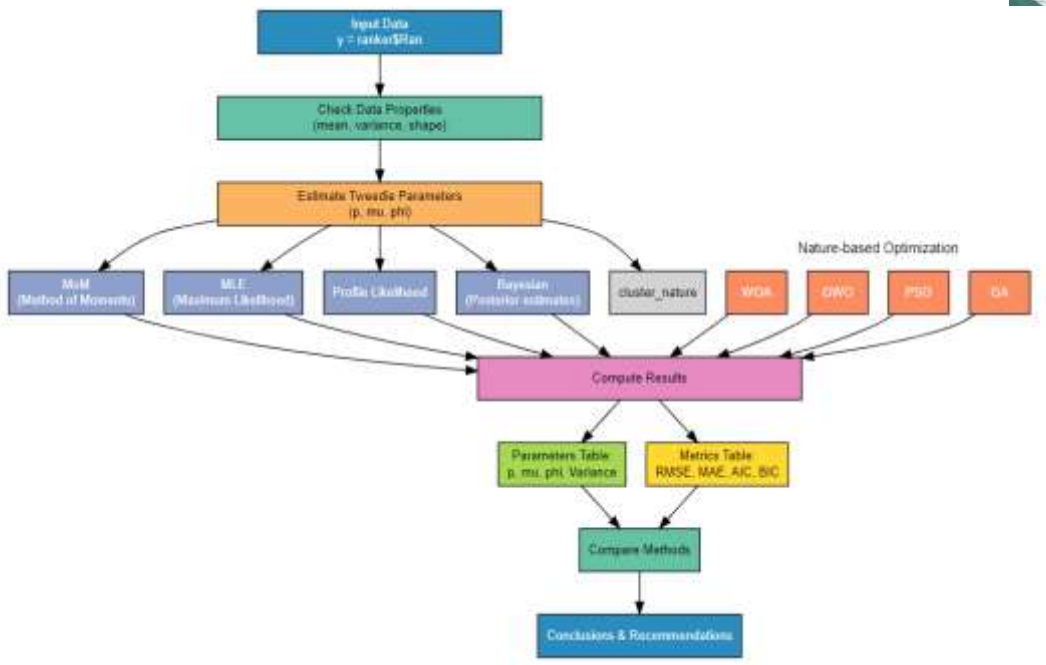


figure 1. flowchart of Estimation and Evaluation for Tweedie Distribution

2. MATERIALS AND METHODS

2.1.Tweedie distribution:

On (2005) Dunn and Smith [5] introduced the models of tweedy distribution applied in a variety of fields including meteorology, actuarial science, analysis of consumption, study of variance, time and cost measurements, and survival analysis. They also find their use in rain forecasting Dunn [6] in 2004 and fisheries research.[7].

whose probability density functions [8]

$$.fZ(z|\theta,\varphi) = a(z,\varphi)exp\left\{\frac{z\theta - \kappa(\theta)}{\varphi}\right\} \ldots\ldots..(1)$$

where $a(\cdot)$ and κ(˙) are given functions, with $\theta \in R$ and $\varphi \in R+$. For Tweedie models, the mean and variance of Z has the property $E(Z) := \mu = \dot{\kappa}(\theta), Var(Z) = \varphi\ddot{\kappa}(\theta)$, where $\dot{\kappa}(\theta)$ and $\ddot{\kappa}(\theta)$ are the first and second derivatives of $\kappa(\theta)$, respectively. The power mean-variance relationship is Var(Z) = φμρ for some index parameter ρ ∈ $(1,2)$, The compound Poisson distribution [9]Jørgensen.[7,8].

2.2.Power Parameter Concept

The underlying index/power parameter characterizing the Tweedy random variable is that of the distribution. This parameter is usually given a default value by the actuaries and the Tweedy package in R offers functions to estimate the power of Tweedy through maximum likelihood estimation. Changing this parameter changes only slightly the regression coefficients and model predictions. Actuaries should apparently apply the default parameter of Tweedy power because of the time limitations, and alternative feature engineering should be taken into account [1].

In distributions like Tweedie:

$Var(Y) = \phi\,\mu^{\rho}$.....(2)

ρ: is called the power parameter, and it determines the type of distribution.

Table 1. shows that Tweedy's models based on the p index exhibit model stability only under certain conditions that depend on the choice of parameters[1,4].

Distribution	Domain	p value
Generated by Extreme Stable Distributions	R	P<0
Normal (Gaussian)	R	P=0
Poisson	N	1
Compound Poisson-gamma R0+	R0+	1<p<2
Gamma	R+	2
Generated by Positive Stable Distribution	R+	2<p<3
Inverse Gaussian	R+	3
Generated by Positive Stable Distributions	R+	p>3

Distributions of variance-form distributions. Any distribution with this variance function will be a member of so-called Tweedy distribution family. This is called by the name of Tweedy, and it was also named by [12]. The Tweedy distribution family includes most of the well-known distributions in the framework of power variance relationships, such as the normal distribution (p = 0), the Poisson distribution (p = 1)[13], the gamma distribution (p = 2), and the inverse Gaussian distribution[14] (p = 3).Furthermore, there are variations of Tweedy distributions for all values of p outside the range (0, 1), but only the four previously mentioned distributions have density functions with clearly defined analytical expressions, as reported by [5]. For values of p > 1, Tweedy distributions have completely positive means, and their more right-skewed nature is evident at values of p > 2, giving them a shape similar to the gamma distribution, but more skewed. For values of p < 0, these distributions are fully continuous on the real axis. Finally, when p < 1 < 2, these distributions are supported on non-negative real numbers and exhibit a distinct mixture of the two distributions. Poisson and gamma, where mass appears at zero according to Dunn and Smyth.[15]. These distributions are called the compound Poisson, compound gamma, and Poisson-gamma distributions because of their unique ability to model both discrete and continuous phenomena. They have distinct applications in modeling both rainfall frequency and rainfall amount. The mean, μ, and the standard parameter, θ, of the Tweedie distribution can be determined by the relationship $\kappa''(\theta) = \mu p$, where the mean is given by the equation $\mu = \kappa'(\theta)$.[16].

2.3.Estimating ρ :

The power factor is estimated by calculating log-likelihoods, as indicated in the sources [17]. The technique adopted by the tweedie.profile function follows specific steps, which are as follows:

1.Input a data vector along with a list of possible power factors as input.
2.Calculate the log-likelihood for each power factor specified in the list.
3.Identify the power factor with the highest log-likelihood value.
4.Optionally, display the log-likelihoods as a graph.
5.Optionally, draw a smooth curve that intersects the points represented on the graph.
6.Finally, if desired, extract the power factor with the maximum log-likelihood value based on the drawn curve.[1]

The idea: We try different values of ρ
(for example, between 1 and 2) and calculate the log-likelihood function at each value
We choose p

R is used to fit the Tweedie model, with the Tweedie library being an essential part of the process (according to Dunn [6]. The best-fitting Tweedie distribution to the data set under consideration is calculated with the help of the tweedie.profile function it applies the maximum likelihood methodology. This test is only effective when the p-value = or above 1 and a maximum probability value of p and φ is obtained and a 95 percent confidence interval of p-

value is also determined. This function will automatically calculate using the chain expansion method, although shear can also re-calculate using interpolation and inversion computational methods. In order to fit a generalized regression model to a Tweedie distribution, the strength of the variance (p is calculated as tweedie.profile) and the correct link function to be used by the model are required [8,9,10].

2.4. Estimation Methods and Mathematical Formulas

2.4.1 Maximum Likelihood Estimation (MLE)

The log-likelihood function for the Tweedie distribution is given by:

$$\ell(\mu, \phi, p) = \sum_{i=1}^{n} \left[y_i \frac{\mu_i^{1-p}}{\phi(1-p)} - \frac{\mu_i^{2-p}}{\phi(2-p)} \right] \text{.....(3)}$$

MLE chooses parameters that maximize (p).

2.4.2 Method of Moments (MoM)

Based on matching the theoretical mean and variance to the sample moments:

$$E(Y) = \mu, \quad Var(Y) = \phi\mu^p$$

Solve for p using:

$$\hat{p} = \frac{\log(Var(Y)/\bar{Y})}{\log(\bar{Y})} \text{.....(4)}$$

2.4.3..Profile Likelihood Estimation :

The power factor is estimated by calculating log-likelihoods, as indicated in the sources[17] . The technique adopted by the tweedie.profile function follows specific steps, which are as follows:

1.Take a data vector and a list of potential power factors as input.
2.Compute the log-likelihood of each of the entries in the list of power factors.
3.The power factor with the maximum log-likelihood value.
4.Plot (optionally) the log-likelihoods.
5.Draw a smooth curve, optionally, which goes through the points depicted on the graph.
6.Finally, if desired, extract the power factor with the maximum log-likelihood value based on the drawn curve.[1]

The idea: We try different values of ρ
(for example, between 1 and 2) and calculate the log-likelihood function at each value

Profile the likelihood over p while estimating other parameters:

$$\hat{p}_{PL} = \arg\max_{p} \ell\left(\hat{\mu}(p), \hat{\phi}(p), p\right) \text{.....(5)}$$

The partial likelihood method (profile likelihood) is used to estimate the power parameter (p) in the Tweedie distribution. The main steps are to fix p to a set of predetermined values (e.g., p.vec = seq(1.1, 1.9, by=0.01)), then analyze the likelihood function for each value in that set. The benefit of this process is that the value of the likelihood function with the highest value is taken as the best estimate of p and this can be abbreviated as profile p.max.The strengths of the process are that it is very accurate since it is based directly on the formula of the likelihood function of the model. It is also not as sensitive to outliers as the method of moments (MoM). To sum up, the given method can be considered as the effective way of estimating the power parameter (p) with the help of the partial likelihood technique.

2.4.4. Bayesian Estimation:[19][20]

Given independent observations y1, y2, …, yn from a Tweedie distribution:

$$L(\mu, \varphi, p \mid y) = \prod f(yi \mid \mu, \varphi, p) \text{.....(6)}$$

Log-likelihood: $\ell(\mu, \varphi, p \mid y) = \Sigma \, log \, f(yi \mid \mu, \varphi, p)$.....(7)Priors can be specified as follows:

$$\mu \sim Gamma(a\mu, b\mu) \; or \; \mu \sim LogNormal(\mu 0, \sigma \text{^} 2)$$
$$\varphi \sim Inverse - Gamma(a\varphi, b\varphi) \; or \; \varphi \sim Half - Cauchy(0, \tau)$$
$$p \sim Uniform(1, 2)$$

Posterior Distribution[21]

$\pi(\mu, \varphi, p \mid y) \propto L(\mu, \varphi, p \mid y) * \pi(\mu) * \pi(\varphi) * \pi(p)$.....(8)

2.4.5. Particle Swarm Optimization (PSO)[22].

Particle swarm optimization (PSO) algorithm is a reference algorithm that was introduced by Kennedy and Eberhart in 1995 and was based on the collective foraging behavior of birds and fish (Bratton & Kennedy, 2007). This algorithm initiates with a population of particles and has been scattered randomly in the search space that each particle defines its current location, its best location (pbest) and the general location (gbest). In the iterations, the particles update their velocities and positions to arrive at the best solution through a fitness functional which is appropriate to the problem at hand[23].

$$v_{id}^{\text{new}} = w \times v_{id}^{\text{old}} + c_1 \times rnd() \times \left(p_{id} - x_{id}^{\text{old}}\right) + c_2 \times Rnd() \times \left(p_{gd} - x_{id}^{\text{old}}\right)$$

$$x_{id}^{\text{new}} = x_{id}^{\text{old}} + v_{id}^{\text{new}} \quad \ldots\ldots (9)$$

In the PSO algorithm, the constants c1 and c2 are positive acceleration factors, often set to 2.0 as proposed by [24] in 2001. w is the inertial weight that governs the influence of the past movement of the particles, pid is the personal best position of the particle i and pgd is the collective best position of all the particles. Moreover, uniformly distributed random numbers in the interval [0,1] are obtained with the help of the rand() and rand() functions. In the search process, the particles keep changing their positions in every iteration until they achieve relative steady state, or they have exhausted the maximum number of iterations.[23].

2.4.6. Grey Wolf Optimization (GWO) [25].

Grey Wolf Optimization (GWO) is one of the most recent meta-optimization algorithms proposed by Mirjalili and colleagues in 2014 [26]. In 2015,[27]. Ghelizadeh developed the GWO algorithm to address the problem of optimizing two-layer networks, taking into account the nonlinear properties of behavior.[25]. The Gray Wolf Optimizer (GWO) algorithm is inspired by the social behavior and hunting approach of gray wolves. These wolves live in packs of typically 5 to 12 individuals, adhering to a clear and deliberate hierarchy:

Alpha (α): The supreme leader who makes important decisions, such as determining where to hunt or when to move.

Beta (β): The Alpha's advisor and assistant, responsible for maintaining internal order within the pack.

Delta (δ): Includes multiple categories such as guards, scouts, elders, hunters, and caretakers.Omega (ω): They are the weakest in the pack, which is very important in relieving internal tension and also keeping the social balance. In a hunt, Alpha, Beta, and Delta wolves give the impetus to the process of chasing and surrounding prey, and the other members of the pack contribute to these actions. This cooperative behavior forms the basis of the GWO algorithm, in that possible solutions are represented as the wolves in a pack, the three best candidate solutions (0, 1, 2) are the leaders who direct the rest of the solutions (the other wolves) to the best possible answer (the location of the prey or the best solution). In this method, hierarchy and spirit of cooperation of the gray wolves inspires the GWO algorithm to locate the best solution of different optimization problems. Besides concentrating on social hierarchy, gray wolves are also dependent on group hunting that takes place in three fundamental phases: first, tracking and approaching to the prey, then encircling it and confusing it, and then attacking it right away. These were behaviors that were manifested by hierarchical structure and hunting strategy which led to the creation of the Gray Wolf Orientation (GWO) algorithm that seeks to recreate natural processes in search of the best solutions.[28]. The wolf α is the best solution and the second- and third-best solutions in the mathematical model of the GWO algorithm are the wolves 2 and 3. The rest of the solutions or wolves are the ones that trail along with this trio when they are seeking the best solution.[25]. As it was already stated above, one of the primary hunting methods of the gray wolves is encircling prey. Studies on this subject reveal that this behavior is effective in terms of surround and search of prey [29].[30] Gray wolves move in circles around their prey during hunting.

Journal of Kyzylorda Scholarly Review
P-ISSN:3006-0265 | E-ISSN:3007-4681
https://bulletin.ouk.kz/index.php/bulletin/index.php

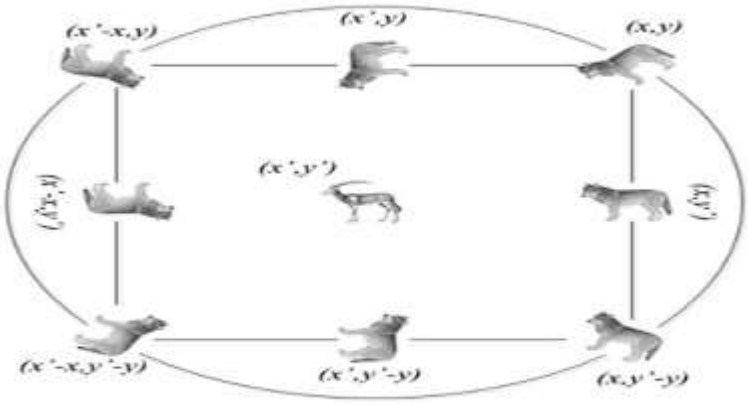


figure 2. Attacking toward prey versus searching for prey[25]

To represent this behavior mathematically, equations () are presented.[29].

$$\vec{D} = \left|\vec{C} \cdot \overrightarrow{X_P}(t) - \vec{X}(t)\right| \text{.....(10)}$$

$$\vec{X}(t+1) = \overrightarrow{X_P}(t) - \vec{A} \cdot \vec{D} \text{ ... . (11)}$$

In this formulation, t represents the number of iterations, while A and C represent the vector coefficients, where $\vec{X}$denotes the prey location vector and X also denotes the gray wolf location vector. The vectors A and C are calculated based on Equations (12-13).

$$\vec{A} = 2\vec{a} \cdot \vec{r}_1 - \vec{a} \text{.....(12)}$$

$$C^{\rightarrow} = 2 \cdot r^{\rightarrow}_2 \quad \text{.....(13)}$$

$$\vec{C} = 2 \cdot \vec{r}_2 \text{.....(14)}$$

Where $\vec{a}$ decreases linearly from 2 to 0 under the path of repetitions,
and r1and r2 are random vectors.

To mathematically simulate hunting behavior, Alpha, Beta, and Delta are assumed to have the best knowledge of the prey location, and the locations of the remaining wolves (Omega) are updated based on these three solutions according to Equations (15–17).

$$\begin{aligned}
\overrightarrow{D_\alpha} &= \left|\overrightarrow{C_1} \cdot \overrightarrow{X_\alpha} - \vec{X}\right|, \\
\overrightarrow{D_\beta} &= \left|\overrightarrow{C_2} \cdot \overrightarrow{X_\beta} - \vec{X}\right|, \text{.....(15)} \\
\overrightarrow{D_\delta} &= \left|\overrightarrow{C_3} \cdot \overrightarrow{X_\delta} - \vec{X}\right|, \\
\vec{X}_1 &= \overrightarrow{X_\alpha} - \overrightarrow{A_1} \cdot \left(\vec{D}_\alpha\right), \\
\vec{X}_2 &= \overrightarrow{X_\beta} - \overrightarrow{A_1} \cdot \left(\vec{D}_\beta\right), \text{.....(16)} \\
\vec{X}_3 &= \overrightarrow{X_\delta} - \overrightarrow{A_1} \cdot \left(\vec{D}_\delta\right), \\
\vec{X}(t+1) &= \frac{\vec{X}_1 + \vec{X}_2 + \vec{X}_3}{3} \text{.....(17)}
\end{aligned}$$

Alpha, beta, and delta positions are considered estimates of the prey's location, with the final location randomly determined within a drawn circle, while other wolves randomly update their positions nearby.[31]

2.4.7. Whale Optimization Algorithm (WOA).

Whale Optimization Algorithm (WOA) is a meta-exploration algorithm based on the foraging processes of the humpback whales. The algorithm has been found to be useful in global optimization in many fields. In comparison to the traditional ways of hunting, WOA employs the use of either the optimal solution presented in the search space or random selection. The nature of this algorithm is inspired by the bubble-net technique of humpback whaling which involves creating spiral rings of bubbles around prey-rich areas near the water surface, and repeated up-down movement on the prey to closely entrap the prey based on a bubble-net model. Next, it will be presented in theory the principles of the algorithm and the formulation of the algorithm mathematically [32].

Journal of Kyzylorda Scholarly Review
P-ISSN:3006-0265 | E-ISSN:3007-4681
https://bulletin.ouk.kz/index.php/bulletin/index.php



Figure 3. shows the diagram illustrating the attack strategy using the spiral bubble network.[33].

The narrowing of the surrounding prey: WOA algorithm takes the most possible solution as the main objective and guides other search agents to reposition themselves in order to reach the solution. The following equation is reflected in the behavioral pattern of this process:[34].

$$\vec{X}(t+1) = \vec{X}^*(t) - A \cdot \vec{D} \text{.....(18)}$$

$$\vec{D} = \left|C \cdot \vec{X}^*(t) - \vec{X}(t)\right| \text{.....(19)}$$

$$A = 2 \cdot a \cdot r - a \text{ ..... (20)}$$

$$C = 2 \cdot r \text{ ..... (21)}$$

To accomplish this goal, the current iteration is denoted by t, $\vec{X}^*$denotes the best previously recorded location, and $\vec{X}$ represents the whale's current location. To gradually decrease the value of a from 2 to 0 in a linear fashion, r is assigned a random number evenly distributed between 0 and 1 at each iteration.

Spiral bubble net feeding maneuver: This method mimics the spiral path taken by humpback whales by relying on a mathematical equation to accurately calculate the locations of the whale and its prey at every moment.

$$\vec{X}(t+1) = e^{bk} \cdot \cos(2\pi k) \cdot \vec{D}' + \vec{X}^*(t) \text{ ..... (22)}$$

$$\vec{D}' = \left|\vec{X}^*(t) - \vec{X}(t)\right| \text{.....(23)}$$

Here, b denotes the aspect that defines the shape of the logarithmic spiral, whereas k is a random variable that is uniformly distributed within the interval of -1 to 1 and this allows the use of a global optimizer. In the event that the value of A is too large as the value of 1 or is too small as the value of -1, the search agent is updated randomly by choosing a new search agent.

$$\vec{X}(t+1) = \vec{X}_{\text{rand}} - A \cdot \vec{D}'' \text{ ..... (24)}$$

$$\vec{D}'' = \left|C \cdot \vec{X}_{\text{rand}} - \vec{X}(t)\right| \text{ ..... (25)}$$

where in the current iteration $\vec{X}_{\text{rand}}$ is a random whale.[34]

2.4.8The genetic algorithm (GA) :[21,22,23]

The genetic evolutionary algorithm (GA) is an effective method based on natural systems that applies a concept of natural selection and genetics in order to optimize solutions. It is best suited to the complex optimization problems because it is based on the gradual improvement of a set of possible solutions. There are a number of fundamental steps in this algorithm.

:1 .Initialization: An initial set of solutions (genes) is randomly generated to represent the range of possible solutions.

2 .Performance evaluation: The performance of each solution is measured using a pre-defined objective function to determine its fitness.

3 .Selection: The best-performing solutions are selected to form a new population aimed at achieving optimal values.

4 .Mating: Genetic characteristics are combined between different solutions to produce new individuals with improved or shared characteristics.

5 .Genetic switching: Random changes are added to the genes to ensure diversity, benefit from searches, and avoid localized solutions.

6 .Replacement: A new generation is formed by replacing the existing population with the solutions generated through mating and switching.

7 .Final Benchmark: The process continues until a specific criterion is achieved, such as a specified number of generations or satisfactory performance.
What distinguishes this algorithm is its ability to adapt to ongoing search dynamics. This makes it an effective tool for tackling multidimensional optimization problems that are difficult to address using traditional methods..[32,33]

$$F(x) = Objective(x) \text{.....(26)}$$

Selection Probability:

$$P_i = f(x_i) / \Sigma f(x_j) \quad for\ j = 1, \ldots, N \text{.....(27)}$$

Crossover Operation:

$$Offspring = \alpha * Parent1 + (1 - \alpha) * Parent2 \quad , \ldots\ldots (28) \quad 0 \leq \alpha \leq 1$$

Mutation Operation:

$$x_i' = x_i + \delta \, , \ldots\ldots (29) \quad where\ \delta \sim N(0, \sigma^2)$$

Termination Criterion:
Stop if max_generations is reached OR if

$$|F(x_best)\text{^}t - F(x_best)\text{^}(t-1)| < \varepsilon$$

2.5.Model Selection Criteria
The Residual Mean Squared Error has been computed as [38][37]

$$\text{RMSE} = \sqrt{\sum_{i=1}^{k} \left(Y_i - \hat{Y}_i\right)^2 / k} \text{.....(30)}$$

mean absolute error (**MAE**) is a measure of errors between paired observations expressing the same phenomenon. Examples of Y versus X include comparisons of predicted versus observed, subsequent time versus initial time, and one technique of measurement versus an alternative technique of measurement. MAE is calculated as the **sum of absolute errors** .

$$\text{MAE} = \frac{\sum_{i=1}^{n} |y_i - x_i|}{n} = \frac{\sum_{i=1}^{n} |e_i|}{n}. \text{ .....(31)}$$

Akaike Information Criterion (AIC):

$$AIC = -2 \cdot LogLik + 2k \text{.....(32)}$$

Bayesian Information Criterion (BIC):

$$BIC = -2 \cdot LogLik + k \cdot log(n) \text{.....(33)}$$

Where k = number of estimated parameters, n = sample size
2.6. Data description:
For this study, data were collected from the Kirkuk Meteorological Department, Iraq, covering the period from 1984 to 2014. The database includes monthly averages for several key climate variables, including rainfall, wind speed, relative humidity, and temperature.The data were analyzed using R software, which provides an integrated environment for applying advanced statistical models, as well as using nature-inspired algorithms to provide innovative optimization solutions. This analytical approach contributed to a deeper understanding of the data characteristics, the estimation of various model parameters, and the selection of the most appropriate method to express the statistical behavior of the climate variables under study.

Table 2. Estimation of the parameters and power parameter of the Tweedie distribution and performance of different estimation methods using model selection metrics, such as the mean absolute error (MAE) and the root mean square error (RMSE).

MAE	RMSE	Variance	phi	mu	p	Method
30.04936	37.26037	1413.1270	9.647373	28.29754	1.491837	Profile Likelihood
30.04936	37.26037	1413.1270	48.296422	28.29754	1.010000	MoM
30.04936	37.26037	1588.5455	10.601548	28.297544	1.498627	MLE
30.04936	37.26037	1413.1270	9.387676	28.297544	1.500000	Bayesian
37.26037	37.26037	1588.5386	10.601567	28.297544	1.498626	PSO
30.05981	37.26038	1597.8728	10.584958	28.328917	1.500350	GWO
27.99965	46.06367	100.7289	14.13127	28.24845	1.443428	WOA

Journal of Kyzylorda Scholarly Review
P-ISSN:3006-0265 | E-ISSN:3007-4681
https://bulletin.ouk.kz/index.php/bulletin/index.php

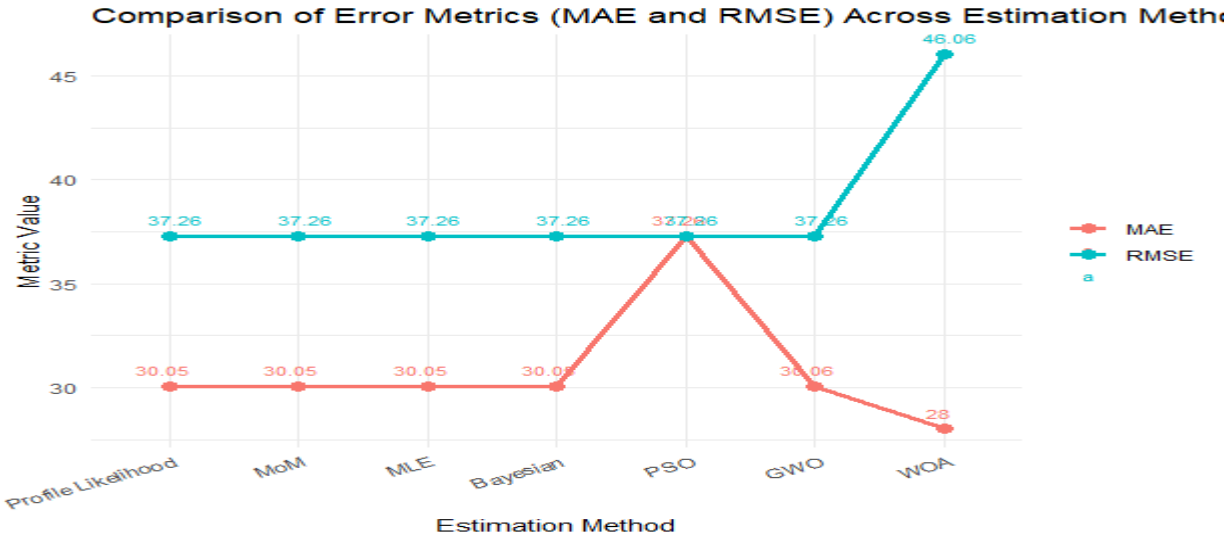


Figure 4. Comparing the performance of different estimation methods using the evaluation metrics RMSE and MAE

Table 3. Performance of different estimation methods using model selection metrics, AIC, BIC.

Method	AIC	BIC		
Profile Likelihood	426.898	433		
MoM	426.3562	432.4854		
MLE	426.3562	432.4854		
Bayesian	534.3343	540.4634		
PSO	426.3570	432.4861		
GWO	426.8980	433.0272		
WOA	427.0461	433.1752		

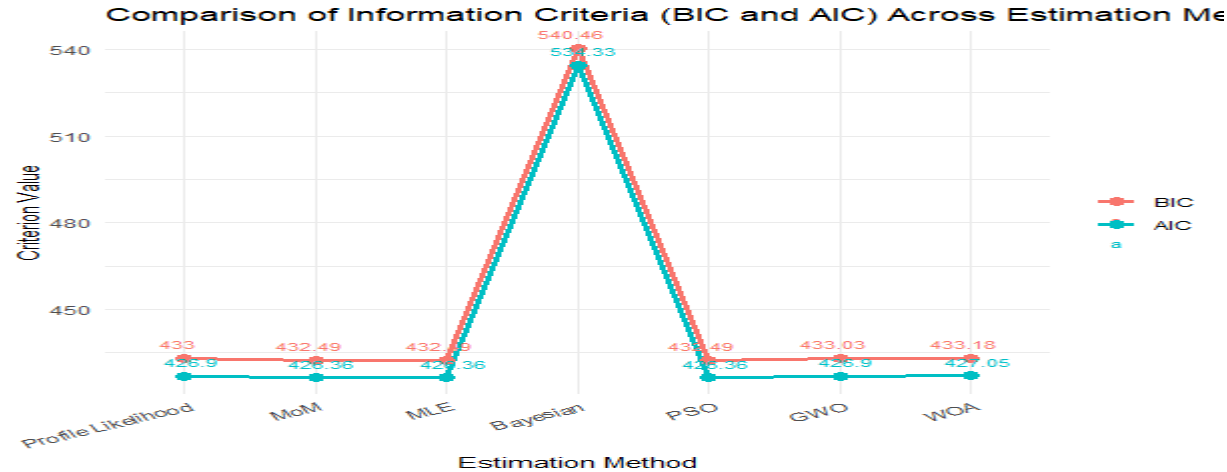


Figure 5. Performance of different estimation methods using model selection metrics, AIC, BIC

Journal of Kyzylorda Scholarly Review
P-ISSN:3006-0265 | E-ISSN:3007-4681
https://bulletin.ouk.kz/index.php/bulletin/index.php

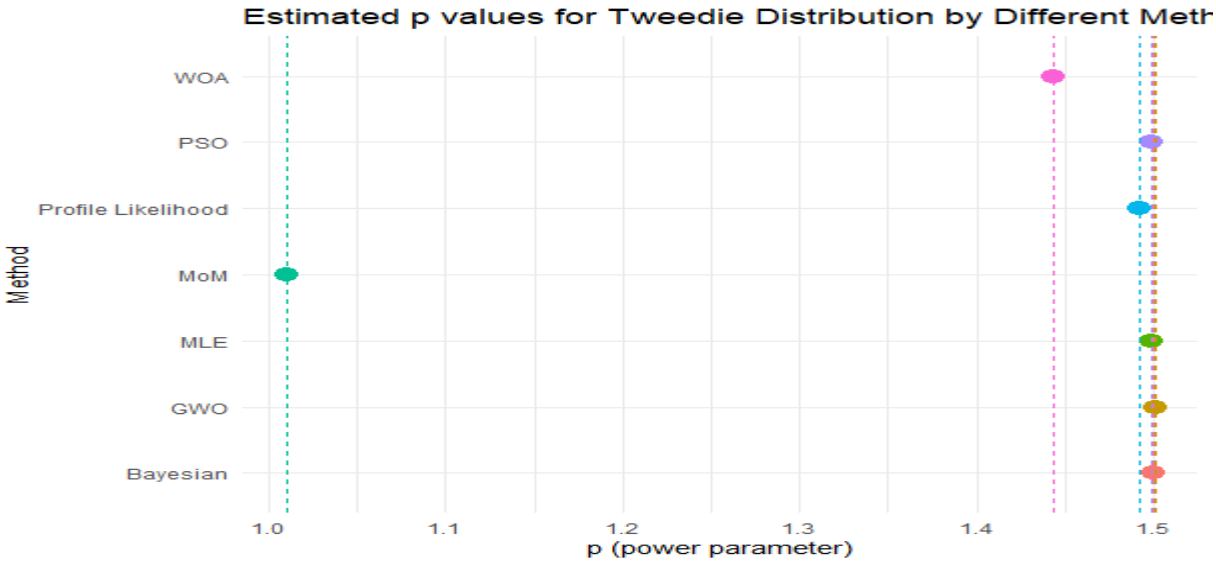


Figure 6. Estimating the value of p in Tweedie distribution using different methods..

3. RESULTS AND DISCUSSION

The primary objective of this study was to accurately estimate the Tweedie power coefficient (p), which is a crucial indicator for determining the variance function of a distribution and its suitability for modeling data using a compound Poisson-gamma fitting. The results showed that the estimation methods provided values for the coefficient p within a relatively narrow range (1.01-1.50), which is consistent with the theory that Tweedie distributions with p between 1 and 2 are particularly suitable for modeling continuous data with zero clustering. Both the likelihood regression method ($p \approx 1.49$) and the maximum likelihood mean error method ($p \approx 1.498$) provided convergent estimates, demonstrating their stability and reliability in estimating coefficients.The Bayesian estimation (p = 1.50) and the Grey Wolf Optimization (GWO, p ≈ 1.5003) method recorded slightly higher values for p, indicating their tendency to deviate towards the upper bound of the power-variance curve. In contrast, the Method of Moments (MoM, p = 1.01) provided the lowest estimates compared to the other methods, indicating its limited accuracy, despite its computational simplicity, on more complex probabilistic structures.Similarly, the Particle Optimization (PSO, p ≈ 1.4986) method performed close to the maximum likelihood (MLE) method, demonstrating that metaheuristics can compete with traditional methods. However, the Whale Optimization (WOA, p ≈ 1.44) method recorded a relatively lower estimate of the parameter, reflecting its tendency to deviate from the local optimum value, as evidenced by a higher mean square error (RMSE).The convergence of most methods around the range $p \approx 1.49–1.50$ significantly supports the reliability of this coefficient in determining the fit of the Tweedie distribution, enhancing its suitability for characterizing data sets where variance exhibits nonlinearity with the mean. The effect of the value of p was also evident in model performance indicators, with methods that yielded stable and high coefficient values (such as MLE, likelihood regression, and GWO) achieving better results for the AIC/BIC indices, compared to the poorer performance of Bayesian algorithms and the WOA method.

4. CONCLUSSION

1. Most estimation methods demonstrate that the power coefficient p ranges between 1.49 and 1.50, enhancing the feasibility of using the Tweedie distribution to model data sets.
2.The convergence of traditional methods such as MLE and likelihood regression with meta-optimization methods such as PSO and GWO indicates the latter's effectiveness in providing accurate estimates.

3.Although the MoM method provides lower estimates of p, this reflects its limitations when dealing with more complex distributions.
4.The WOA method recorded the lowest mean absolute error but was accompanied by a higher variance in the estimate, which affects its reliability and the ease of interpreting its results.
5.The Bayesian method showed mixed results due to overfitting the data and its inability to provide competitive efficiency indicators.
2.9.Recommendations:
1. Future research should focus on estimation methodologies that provide stable and consistent values for p, such as MLE, likelihood regression, and GWO.
2. Meta-optimization algorithms, especially PSO and GWO, show promising results in cases with complex probability surfaces, making them an effective alternative to classical methods.
3.Although WOA is effective in reducing the mean absolute error, improving its performance is necessary to address high variance issues.
4.It is preferable to explore hybrid frameworks that combine the advantages of classical probability-based methods and meta-optimization algorithms to achieve greater flexibility and higher accuracy in power factor estimation.

References

[1] S. D. Temple and C. Sinclair, "The Tweedie Index Parameter and Its Estimator An Introduction with Applications to Actuarial Ratemaking Title: The Tweedie Index Parameter and Its Estimator: An Introduction with Applications to Actuarial Ratemaking Approved," no. June, 2018.
[2] A. Ghribi, A. Hassin, and A. Masmoudi, "on the Existence of the Tweedie Power Parameter Implicit Estimator," Bull. Korean Math. Soc., vol. 59, no. 4, pp. 979–991, 2022, doi: 10.4134/BKMS.b210590.
[3] T. Swan, "Generalized estimating equations when the response variable has a Tweedie distribution: an application for multi-site rainfall modelling," Dissertation, p. 168, 2006, [Online]. Available: http://eprints.usq.edu.au/3388
[4] W. H. Bonat and C. C. Kokonendji, "Flexible Tweedie regression models for continuous data," J. Stat. Comput. Simul., vol. 87, no. 11, pp. 2138–2152, 2017, doi: 10.1080/00949655.2017.1318876.
[5] P. K. Dunn and G. K. Smyth., "Series evaluations of Tweedie exponen- tial dispersion models.," Stat. Comput. Accept. Publ., 2005.
[6] P. Dunn. and Tweedie:, "Tweedie exponential fam- ily models.," R Packag. http//www.sci.usq.edu.au/staff/dunn/twhtml/home.html, vol. version 1., 2004.
[7] S. G. Candy, "Modelling catch and effort data using generalized linear models, the Tweedie distribution, random vessel effects and random stratum-by-year effects.," CCAMLR Sci. Davis, vol. 11, pp. 59–80, 2004.
[8] M. Tweedie, "An index which distinguishes between some important exponential families,"," pp. 579-604 2 wood, 1984.
[9] B. Jørgensen, "The theory of exponential dispersion models and analysis of deviance," 1992.
[10] W. A. Hadba, "Tweedie regularization model and application," Commun. Stat. - Theory Methods, vol. 54, no. 6, pp. 1729–1746, 2025, doi: 10.1080/03610926.2024.2349698.
[11] M. H. Ndegwa, "CLAIM RESERVING USING TWEEDIE DISTRIBUTION School of Mathematics University of Nairobi," no. July, 2013.
[12] B. Jørgensen., "Exponential dispersion models (with discussion).," Jour- nal R. Stat. Soc. Ser. B Stat. Methodol., vol. 49, pp. 127–162, 1987.
[13] babakr Sawen Othman. & Badal Mohammed AbdulMajeed., "Using Multilevel Poisson Regression with Wavelets Filters.," Univ. Kirkuk J. Adm. Econ. Sci., vol. 13, no. 4, pp. 179–192, 2023.
[14] Abduljabbar Luay Adil., "Ridge-Type Estimator for the Poisson-inverse Gaussian

regression model.," Univ. Kirkuk J. Adm. Econ. Sci., vol. 15, no. 2, pp. 405–413, 2025.
[15] P. K. Dunn and G. K. Smyth., "Randomized quantile residuals.," J. Comput. Graph. Stat., vol. 5, pp. 1–10, 1996.
[16] P. K. Dunn and G. K. Smyth, Evaluation of Tweedie exponential dispersion model densities by Fourier inversion, vol. 18, no. 1. 2008. doi: 10.1007/s11222-007-9039-6.
[17] Peter Dunn., "Package 'tweedie'. R Documentation.," Packag. R Program. Dec., 2016.
[18] P. K. Dunn and G. K. Smyth, "Series evaluation of Tweedie exponential dispersion model densities," Stat. Comput., vol. 15, no. 4, pp. 267–280, 2005, doi: 10.1007/s11222-005-4070-y.
[19] D. B. Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D., Vehtari, A., & Rubin, "Bayesian Data Analysis," (3rd ed.). CRC Press.
[20] A. J. Dobson and A. G. Barnett, An introduction to generalized linear models, third edition. 2008. doi: 10.1080/02664760802695900.
[21] Z. Wu, X. Duan, and W. Zhang, "Bayesian Analysis of Tweedie Compound Poisson Partial Linear Mixed Models with Nonignorable Missing Response and Covariates," Entropy, vol. 25, no. 3, 2023, doi: 10.3390/e25030506.
[22] Puguh Jayadi, Khairul Adilah binti Ahmad, Rayhan Zulfitri Dwi Cahyo, and Jofanza Denis Aldida, "Particle Swarm Optimization-based Linear Regression for Software Effort Estimation," J. Inf. Syst. Technol. Eng., vol. 2, no. 2, pp. 261–268, 2024, doi: 10.61487/jiste.v2i2.69.
[23] H.-I. Hsieh, T.-P. Lee, and T.-S. Lee, "A Hybrid Particle Swarm Optimization and Support Vector Regression Model for Financial Time Series Forecasting," Int. J. Bus. Adm., vol. 2, no. 2, pp. 48–56, 2011, doi: 10.5430/ijba.v2n2p48.
[24] Y. (2001b). Eberhart, R. C., & Shi, "Particle Swarm Optimization: Developments, Applications and Resources.," Proc. Congr. Evol. Comput. , IEEE Serv. Center, Seoul, pp. 81–86, 2001.
[25] G. Wolf and O. Gwo, "Chapter 9 Grey Wolf Optimization (GWO) Algorithm," no. July, 2018, doi: 10.1007/978-981-10-5221-7.
[26] A. Mirjalili, S., Mirjalili, S. M., & Lewis, "Grey Wolf Optimizer. Advances in Engineering Software," vol. 69, pp. 46–6.
[27] S. Gholizadeh, "Optimal design of double layer grids considering nonlinear behaviour by sequential grey wolf algorithm.," J. Optim. Civ. Eng., vol. 5, no. 4, pp. 511–523, 2015.
[28] R. Muro, C., Escobedo, R., Spector, L., & Coppinger, "Wolf-pack (Canis Lupus) hunting strategies emerge from simple rules in computational simulations.," Behav. Process., vol. 88, no. 3, pp. 192–197, 2011.
[29] S. Mirjalili, S. M. Mirjalili, and A. Lewis, "Grey wolf optimizer," Adv. Eng. Softw., vol. 69, pp. 46–61, 2014.
[30] D. Perin, A. D. Karaoglan, and K. Yilmaz, "Using grey wolf optimizer to minimize voltage total harmonic distortion of a salient-pole synchronous generator," Sci. Iran., vol. 30, no. 2 D, pp. 592–604, 2023, doi: 10.24200/sci.2021.57657.5349.
[31] A. Goli, H. K. Zare, R. Tavakkoli-Moghaddam, and A. Sadeghieh, "An Improved Artificial Intelligence Based on Gray Wolf Optimization and Cultural Algorithm to Predict Demand for Dairy Products: A Case Study," Int. J. Interact. Multimed. Artif. Intell., vol. 5, no. 6, pp. 15–22, 2019, doi: 10.9781/ijimai.2019.03.003.
[32] R. Sihwail, M. Al Ghamri, and D. Ibrahim, "An Enhanced Model of Whale Optimization Algorithm and K-nearest Neighbors for Malware Detection," Int. J. Intell. Eng. Syst., vol. 17, no. 3, pp. 606–621, 2024, doi: 10.22266/ijies2024.0630.47.
[33] M. H. Nadimi-Shahraki, H. Zamani, Z. Asghari Varzaneh, and S. Mirjalili, "A Systematic Review of the Whale Optimization Algorithm: Theoretical Foundation, Improvements, and Hybridizations," Arch. Comput. Methods Eng., pp. 1–47, 2023.
[34] N. Khodadadi, E. S. M. El-kenawy, S. Faridmarandi, M. S. Ghahfarokhi, A. Ibrahim, and S. Mirjalili, "A novel version of whale optimization algorithm for solving optimization problems," Handb. Whale Optim. Algorithm Var. Hybrids, Improv. Appl., no. November, pp. 591–600, 2023, doi: 10.1016/B978-0-32-395365-8.00048-8.



[35] D. E. (Goldberg, "Genetic Algorithms in Search, Optimization, and Machine Learning.," Addison-Wesley., 1989.
[36] M. A. Iquebal, Prajneshu, and H. Ghosh, "Genetic algorithm optimization technique for linear regression models with heteroscedastic errors," Indian J. Agric. Sci., vol. 82, no. 5, pp. 422–425, 2012, doi: 10.56093/ijas.v82i5.17802.
[37] Z. Y. Algamal and H. M.Thabet, "Using Genetic Algorithm in Outlier Detection for Regression Model," J. Educ. Sci., vol. 27, no. 3, pp. 136–142, 2018, doi: 10.33899/edusj.2018.159314.
[38] S. Paterlini and T. Minerva, "Regression model selection using genetic algorithms," Proc. 11th WSEAS Int. Conf. Neural Networks, NN '10, Proc. 11th WSEAS Int. Conf. Evol. Comput. EC '10, Proc. 11th WSEAS Int. Conf. Fuzzy Syst. FS '10, no. January 2010, pp. 19–27, 2010.