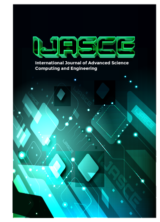
International Journal of
Advanced Science Computing and Engineering

Journal Homepage www.ijasce.org/index.php/ijasce



Improving the Performance of Rank Regression Using Fast Minimum Covariance Determinant in Estimating Weibull Distribution Parameters

Mohammed Abduljabar Ibrahim Hasawy [a,*], Taha Hussein Ali [a], Bekhal Samad Sedeeq [a]

[a] *Department of Statistics and Informatics, College of Administration and Economics, Salahaddin University- Erbil, Iraq*
Corresponding author: [*]mn.hasawy@gmail.com

***Abstract*—Outliers hurt the accuracy of life distribution parameters, including the Weibull distribution. Therefore, researchers have suggested employing the fast minimum covariance determinant method in rank regression estimators (which are robust but not efficient) to obtain robust estimators for the shape and scale parameters of the Weibull distribution. The proposed method is based on the robust means vector and the robust covariance matrix obtained from the fast minimum covariance determinant method, and it employs the rank regression estimation method, which depends on the ordinary least squares estimators of the simple linear regression model. The estimated parameters of the Weibull distribution obtained using the proposed technique have been compared with those derived from conventional maximum likelihood estimation and rank regression, using mean square error as the comparison metric, via both simulation and real data. The study's findings demonstrated the efficacy of the proposed strategy in addressing outliers and yielding highly effective estimators for the shape and scale parameters of the Weibull distribution.**

***Keywords*—Rank regression; Weibull distribution; maximum likelihood estimation; robust estimation; outliers.**

Manuscript received 19 Nov. 2024; revised 17 Mar. 2025; accepted 20 Aug. 2025. Date of publication 30 Dec. 2025.

I. Introduction

A statistical distribution that describes the expected time until a significant event occurs, such as the death of an organism, the failure of a mechanical component, or the time required to complete a task is a lifetime distribution [1]. These distributions are widely used in reliability engineering (Scheduling maintenance, estimating the lifespan of systems, and improving product designs), risk management, and survival analysis (time-to-event data in the medical study, as the time until recurrence of the disease or death). Known lifetime distributions included Weibull, exponential, Normal, lognormal, Gamma, and Gompertz distributions [2]. Weibull distribution is a flexible distribution that can model decreasing, constant, or increasing mean time to failure and includes two parameters, shape and scale [3]. The nature of the failure average determines the shape parameter (for example, the increasing failure average for a shape parameter greater than one).

On the other hand, outliers are a significant challenge that affects the precision of parameter estimation for life distributions, including the Weibull distribution. Maximum likelihood estimators are ineffective in estimating the parameters of a mathematical distribution when outliers are present, but the rank regression technique, which relies on the linear function of the Weibull distribution parameters, exhibits a degree of robustness to outliers. (Boudt et al. [4] proposed three robust Weibull distribution parameter estimators: the quantile regression, the median/Qn estimator, and the repeated median estimator. Derived their breakdown point, asymptotic variance, and influence function.

The methods are illustrated on real Weibull distribution data affected by outliers. Coria et al. [5] proposed estimating both parameters straightforwardly for the Weibull distribution. they analyzed right-censored lifetime data sets with varying sample sizes and censoring percentages to evaluate the effectiveness of the proposed estimator. demonstrate that the parameter estimator yields a high level of accuracy. Sinha, [6] proposed the robust method for fitting the accelerated failure time model to Weibull distribution data by determining the influence of outliers in both the response variable and associated covariates.

The finite-sample properties of the estimators are investigated based on simulated results and the real data from breast cancer patients. Roohanizadeh et al. [7] delineated various estimation methodologies for the two-parameter Weibull distribution utilizing intuitionistic fuzzy lifetime data, encompassing maximum likelihood estimation (employing

Newton-Raphson and Expectation-Maximization techniques) and Bayesian estimation (utilizing Tierney and Kadane's approximation for the shape and scale parameters, with Gamma and inverse-Gamma priors, respectively). The simulation is conducted to identify the most efficient estimator inside the intuitionistic fuzzy methodology.

Gómez et al. [8] delineated five parameters of the Weibull distribution, each derived from distinct moments of the mean, quantile, and mode. Furthermore, it explores the meaning of regression results when integrating linear regression models into these parameters. They provided insights into its implications for modelling failure periods and its prospective contributions to disciplines that require reliability and survival analysis. In this paper, the proposed method is based on the robust means vector and the robust covariance matrix obtained from the fast minimum covariance determinant method, and it employs the rank regression estimation method that depends on the ordinary least squares estimators of the simple linear regression model.

II. Material and Method

A. Fast Minimum Covariance Determinant

The Fast Minimum Covariance Determinant (FMCD) method, as presented by Rousseeuw and Leroy [9], is a modification of the minimum covariance determinant (MCD). This method selects l data points from a total of N, where $N/2 < l \leq N$, to achieve the least possible determinant from the classical variance-covariance matrix. The estimate comprises the mean vector and covariance matrix of the defined l points, adjusted by a consistency value to ensure adherence to the multivariate normal distribution (MND), along with a correction value to address bias in small sample sizes. Based on this, the estimates of the robust mean vector (MR) and the robust variance-covariance matrix (SR) are calculated to be resilient against outliers. These estimators can be configured to achieve a theoretically maximal breakdown point of 50%. This configuration enables the detection of outliers, even when their quantity approaches nearly half of the sample size [10].

The purpose of MCD is to identify h observations (out of N) that minimize the determinant of the classical covariance. The MCD estimate of location is the mean of these h points, whereas the MCD estimate of dispersion is represented by their covariance matrix. The resultant breakdown value is equivalent to that of the minimal-volume ellipsoid (MVE); however, the minimum covariance determinant (MCD) offers significant benefits over the MVE, as it exhibits superior statistical efficiency due to its asymptotic normality. [11].

A novel method for the MCD, termed FMCD, has been developed to address such issues. The fundamental concept involves an inequality involving order statistics and determinants, using methods known as "selective iteration" or "nested extensions". FMCD generally identifies the precise MCD for small datasets, but for larger datasets it yields more accurate results than current methods and runs much more quickly. FMCD can identify a precise hyperplane that encompasses h or more observations. The algorithm makes the MCD technique a standard instrument for analyzing multivariate data [12].

B. Weibull Distribution

The Weibull distribution is a probability distribution commonly used in reliability engineering and survival analysis to model the time to a specific event. It is named after Walloddi Weibull, who introduced it in the mid-20th century [13]. The distribution is characterised by two parameters: the shape parameter (often denoted as "β") and the scale parameter (often denoted as "η"). The probability density function (PDF) of the Weibull distribution is expressed as [14]:

$$f(x) = \frac{\beta}{\eta}\left(\frac{x}{\eta}\right)^{\beta-1} e^{-\left(\frac{x}{\eta}\right)^{\beta}} \quad (1)$$

Where:

$$f(x) \geq 0, \quad x \geq 0, \ \beta \geq 0, \quad \eta \geq 0$$

C. The Estimation

Determining the values of the parameters that best fit the observed data to a probability distribution is essential for estimation [15]. The selection of the estimation technique depends on the distribution you are addressing. This section will focus on the techniques used to estimate the parameters of the two-parameter Weibull distribution, as employed in this study.

1) Maximum Likelihood Estimation: Maximum likelihood estimation (MLE) is a prevalent and robust technique for parameter estimation over several distributions. Maximum Probability Estimation (MLE) seeks to identify parameter values that optimize the probability of the observed data [16], [17]. The equations for the partial derivatives of the log-likelihood function are formulated and shown below:

$$\frac{\partial\Lambda}{\partial\beta} = \frac{n}{\beta} + \sum_{i=1}^{n} ln\left(\frac{x_i}{\eta}\right) - \sum_{i=1}^{n}\left(\frac{x_i}{\eta}\right)^{\beta} ln\left(\frac{x_i}{\eta}\right) = 0$$

And:

$$\frac{\partial\Lambda}{\partial\eta} = \frac{-\beta}{\eta}.n + \frac{\beta}{\eta}\sum_{i=1}^{n}\left(\frac{x_i}{\eta}\right)^{\beta} = 0$$

Solving the above equations simultaneously, we obtain the estimated parameter values.

2) Rank Regression on Y: Conducting rank regression (RRY) involves mathematically fitting a straight line to a collection of data points in a manner that minimizes the sum of the squares of the vertical deviations from the points to the line [18]. This methodology mirrors the probability plotting method; however, it employs the principle of least squares to ascertain the line through the points rather than relying on visual estimation. The initial step involves transforming our function into a linear format [19]. For the two-parameter Weibull distribution, the cumulative density function is:

$$F(x) = 1 - e^{-\left(\frac{x}{\eta}\right)^{\beta}} \quad (2)$$

Applying the natural logarithm to both sides of the equation results [20]:

$$ln[1 - F(x)] = -\left(\frac{x}{\eta}\right)^{\beta}$$

$$ln\{-ln[1-F(x)]\} = \beta ln\left(\frac{x}{\eta}\right) = -\beta ln(\eta) + \beta ln(x)$$

Now let:

$$y = ln\{-ln[1-F(x)]\} \tag{3}$$

And:

$$b = \beta \tag{4}$$

which results in the linear equation of:

$$y = a + bx$$

The technique of least squares parameter estimation, which is sometimes referred to as regression analysis, was covered in the article on parameter estimation [13], [21], and the equations for regression on Y were constructed as follows:

$$\hat{a} = \bar{y} - \hat{b}\,\bar{x} \tag{5}$$

And:

$$\hat{b} = \hat{\rho}\sqrt{\frac{S_y}{S_x}}$$

$$\hat{\rho} = \frac{S_{xy}}{\sqrt{S_x \times S_y}} \quad ; \quad S = \begin{pmatrix} S_x & S_{xy} \\ S_{yx} & S_y \end{pmatrix} \tag{6}$$

The sample variances are located on the main diagonal of the matrix S, whereas the sample covariances are situated on the off-diagonal of the matrix S. In this case, the equations for y_i and x_i are:

$$y_i = ln\{-ln[1-F(x_i)]\}$$

And:

$$x_i = ln(x_i)$$

The F(Xi) values are derived from the median ranks (MR). The Median Ranks approach is used to quantify the unreliability associated with each failure. The median rank represents the genuine probability of failure, Q(Xj), at the jth failure among a sample of N units, corresponding to the 50% confidence level [22]. An alternative and simpler approach to determining median ranks involves applying two transformations to the cumulative binomial equation, namely the Beta and F Distribution Approach [23], first to the beta distribution and then to the F distribution.

$$MR = \frac{1}{1+((N-j+1)/j)F_{0.50;\,m;\,n}} \tag{7}$$

Where $m = 2(N\text{-}j+1)$ and $n = 2j$. $F_{0.50;\,m;\,n}$ is the F distribution for 0.50 point (m and n degrees of freedom), for failure j from N units.

The estimated parameters are:

$$\hat{\beta} = \hat{b} \tag{8}$$

$$\hat{\eta} = e^{-\hat{a}/\hat{b}} \tag{9}$$

3) Proposed Method: To deal with the outlier problem, the proposed method employs the robust mean vector, and the robust covariance matrix calculated by the Fast Minimum Covariance Determinant (FMCD) method in the rank regression method to get the Robust Rank Regression of dependent Y (RRRY) estimate Weibull distribution parameters through the following.

- Let x = Log(T), where T represents the failure time, is ranked in ascending order, and x is (n × 1) vector.
- Compute the cumulative density function $F(x_i)$ from the median ranks (MR) in equation (7).
- $y = Log\left(-Log\left(1-F(x_i)\right)\right)$ and y is (n × 1) vector.
- D represents (n × 2) matrix for x and y.
- The robust mean vector and covariance matrix (MR and SR, respectively) are calculated by the FMCD method (from paragraph 2.1):

$$MR = [\bar{x}R \quad \bar{y}R], \text{ and } SR = \begin{bmatrix} SxR & SxyR \\ SyxR & SyR \end{bmatrix}$$

- The discussion centered around the least squares parameter estimation method applied to the estimation of Weibull distribution parameters, leading to the formulation of the following linear equation:

$$y = aR + bRx$$

$$\hat{b}R = \hat{\rho}R\sqrt{\frac{SyR}{SxR}} \tag{10}$$

Where:

$$\hat{\rho}R = \frac{SxyR}{\sqrt{SxR \times SyR}}$$

And:

$$\hat{a}R = \bar{y}R - \hat{b}R \times \bar{x}\,R \tag{11}$$

- Finally, the proposed estimation for the shape and scale parameters of the Weibull distribution is:

$$\hat{\beta} = \hat{b}R \tag{12}$$

$$\hat{\eta} = e^{-\hat{a}R/\hat{b}R} \tag{13}$$

D. Mean Squares Error

The efficacy of the estimated parameters may be assessed using the mean squared error (MSE), which yields the minimal value for the estimated parameters $\hat{\theta}$ with optimal efficiency, calculable by the following formula [24]

$$MSE = \frac{\sum_{i=1}^{m}\left(\hat{\theta}_i - \theta\right)^2}{m} \tag{14}$$

where θ represents real parameters (proposed using simulation) and $\theta = [\beta\ \eta]$, and m is the number of samples.

E. Application Aspect

To prove the efficiency of the proposed method (RRRY) in estimating the shape and scale parameters of the Weibull distribution and to compare it with some classical methods (MLE and RRY), simulations were used in addition to real data, as follows:

F. Simulation Study

Data having a Weibull distribution with the shape and scale parameters (for several different values) and for several sample sizes (50, 100, 200, 300, 400, and 500) were generated using a program designed for this purpose in the MATLAB language, in addition, creating a program that calculates the cumulative function of the median ranks at any sample size based on the beta and *F* distribution approach (Appendix). Also, outliers have been added to the generated data.

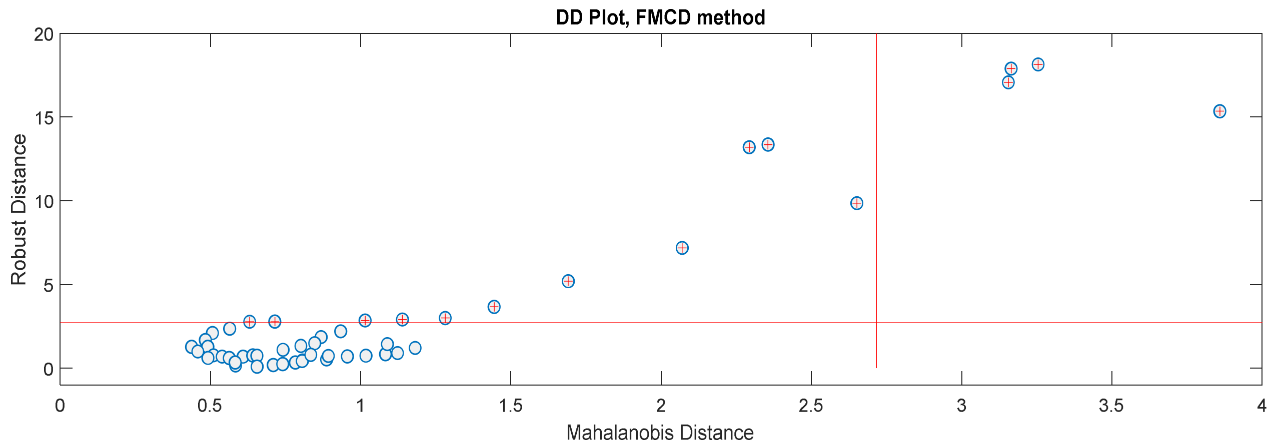


Fig. 1 Mahalanobis distance for the first simulation experiment

The initial simulation experiment utilizing Mahalanobis distance values is presented in Figure 1 for the FMCD method, revealing a total of 15 outliers among the 50 generated data points. To compare the proposed method (RRRY) and classical methods (MLE and RRY), the simulation experiments were repeated (1000) times, the parameters of the Weibull distribution were estimated, and the MSE average (for MSE(Beta) and MSE(Eta)) and Error Mean were calculated. The results were summarized in Tables 1-4:

TABLE I
MSE AVERAGE WHEN BETA = 2 & ETA = 20

Sample Size	Criterion	MLE		RRY		RRRY	
		MSE	Parameters	MSE	Parameters	MSE	Parameters
50	MSE(Beta)	1.1937	3.0926	0.6833	2.8266	**0.1428**	**2.3779**
	MSE(Eta)	138.6521	31.7751	123.9135	31.1316	**8.7389**	**22.9562**
	Error Mean	69.9229		62.2984		**4.4408**	
100	MSE(Beta)	0.9520	2.9757	0.3373	2.5808	**0.0726**	**2.2694**
	MSE(Eta)	35.9755	25.9980	28.5627	25.3444	**2.5572**	**21.5991**
	Error Mean	18.4638		14.4500		**1.3149**	
200	MSE(Beta)	0.7108	2.8431	0.1355	2.3681	**0.0310**	**2.1761**
	MSE(Eta)	10.1083	23.1794	7.2261	22.6881	**0.9234**	**20.9609**
	Error Mean	5.4096		3.6808		**0.4772**	
300	MSE(Beta)	0.5857	2.7653	0.0747	2.2733	**0.0217**	**2.1473**
	MSE(Eta)	4.7816	22.1867	3.2973	21.8158	**0.5456**	**20.7386**
	Error Mean	2.6837		1.6860		**0.2836**	
400	MSE(Beta)	0.5023	2.7087	0.0482	2.2195	**0.0148**	**2.1217**
	MSE(Eta)	2.7358	21.6540	1.8789	21.3707	**0.3984**	**20.6312**
	Error Mean	1.6191		0.9635		**0.2066**	
500	MSE(Beta)	0.4394	2.6629	0.0339	2.1841	**0.0127**	**2.1127**
	MSE(Eta)	1.7962	21.3402	1.2483	21.1173	**0.3029**	**20.5504**
	Error Mean	1.1178		0.6411		**0.1578**	

TABLE II
MSE AVERAGE WHEN BETA = 5 & ETA = 20

Sample Size	Criterion	MLE		RRY		RRRY	
		MSE	Parameters	MSE	Parameters	MSE	Parameters
50	MSE(Beta)	16.1022	9.0128	14.5530	8.8148	0.8855	**5.9410**
	MSE(Eta)	216.0218	34.6977	265.9051	36.3066	1.2746	**21.1290**
	Error Mean	116.0620		140.2290		1.0801	
100	MSE(Beta)	15.0043	8.8735	11.6083	8.4071	0.4561	**5.6754**
	MSE(Eta)	67.4139	28.2106	75.1310	28.6678	0.3925	**20.6265**
	Error Mean	41.2091		43.3696		0.4243	
200	MSE(Beta)	13.7319	8.7057	7.9472	7.8191	0.1941	**5.4406**
	MSE(Eta)	23.3910	24.8364	21.1394	24.5978	0.1442	**20.3797**
	Error Mean	18.5615		14.5433		0.1692	

Sample Size	Criterion	MLE		RRY		RRRY	
		MSE	Parameters	MSE	Parameters	MSE	Parameters
300	MSE(Beta)	12.9651	8.6007	5.8134	7.4111	0.1364	**5.3693**
	MSE(Eta)	12.8897	23.5902	9.9037	23.1470	0.0856	**20.2926**
	Error Mean	12.9274		7.8586		0.1110	
400	MSE(Beta)	12.3987	8.5212	4.4329	7.1054	0.0935	**5.3058**
	MSE(Eta)	8.3988	22.8981	5.6917	22.3857	0.0627	**20.2504**
	Error Mean	10.3987		5.0623		0.0781	
500	MSE(Beta)	11.9515	8.4571	3.4865	6.8672	0.0799	**5.2827**
	MSE(Eta)	6.0489	22.4595	3.7043	21.9247	0.0481	**20.2193**
	Error Mean	9.0002		3.5954		0.0640	

TABLE III

MSE AVERAGE WHEN BETA = 2 & ETA = 30

Sample Size	Criterion	MLE		RRY		RRRY	
		MSE	Parameters	MSE	Parameters	MSE	Parameters
50	MSE(Beta)	1.1937	3.0926	0.6833	2.8266	**0.1428**	**2.3779**
	MSE(Eta)	311.9672	47.6626	278.8053	36.6975	**9.6624**	**33.1084**
	Error Mean	156.5805		139.7443		**4.9026**	
100	MSE(Beta)	0.9520	2.9757	0.3373	2.5808	**0.0726**	**2.2694**
	MSE(Eta)	80.9450	38.9969	64.2662	38.0166	**5.7538**	**32.3987**
	Error Mean	40.9485		32.3017		**2.9132**	
200	MSE(Beta)	0.7108	2.8431	0.1355	2.3681	**0.0310**	**2.1761**
	MSE(Eta)	22.7436	34.7690	16.2588	34.0322	**2.0777**	**31.4414**
	Error Mean	11.7272		8.1972		**1.0544**	
300	MSE(Beta)	0.5857	2.7653	0.0747	2.2733	**0.0217**	**2.1473**
	MSE(Eta)	10.7586	33.2800	7.4188	32.7237	**1.2275**	**31.1079**
	Error Mean	5.6722		3.7468		**0.6246**	
400	MSE(Beta)	0.5023	2.7087	0.0482	2.2195	**0.0148**	**2.1217**
	MSE(Eta)	6.1555	32.4810	4.2275	32.0561	**0.8965**	**30.9468**
	Error Mean	3.3289		2.1378		**0.4556**	
500	MSE(Beta)	0.4394	2.6629	0.0339	2.1841	**0.0127**	**2.1127**
	MSE(Eta)	4.0414	32.0103	2.8086	31.6759	**0.6815**	**30.8255**
	Error Mean	2.2404		1.4213		**0.3471**	

TABLE IV

MSE AVERAGE WHEN BETA = 5 & ETA = 30

Sample Size	Criterion	MLE		RRY		RRRY	
		MSE	Parameters	MSE	Parameters	MSE	Parameters
50	MSE(Beta)	16.1022	9.0128	14.5530	8.8148	**0.8855**	**5.9410**
	MSE(Eta)	486.0491	52.0465	598.2864	54.4599	**2.8679**	**31.6935**
	Error Mean	251.0756		306.4197		**1.8767**	
100	MSE(Beta)	15.0043	8.8735	11.6083	8.4071	**0.4561**	**5.6754**
	MSE(Eta)	151.6813	42.3159	169.0447	43.0017	**0.8832**	**30.9398**
	Error Mean	83.3428		90.3265		**0.6696**	
200	MSE(Beta)	13.7319	8.7057	7.9472	7.8191	**0.1941**	**5.4406**
	MSE(Eta)	52.6298	37.2546	47.5637	36.8966	**0.3244**	**30.5696**
	Error Mean	33.1809		27.7555		**0.2593**	
300	MSE(Beta)	12.9651	8.6007	5.8134	7.4111	**0.1364**	**5.3693**
	MSE(Eta)	29.0017	35.3853	22.2834	34.7205	**0.1926**	**30.4389**
	Error Mean	20.9834		14.0484		**0.1645**	
400	MSE(Beta)	12.3987	8.5212	4.4329	7.1054	**0.0935**	**5.3058**
	MSE(Eta)	18.8972	34.3471	12.8062	33.5786	**0.1411**	**30.3756**
	Error Mean	15.6479		8.6196		**0.1173**	
500	MSE(Beta)	11.9515	8.4571	3.4865	6.8672	**0.0799**	**5.2827**
	MSE(Eta)	13.6100	33.6892	8.3346	32.8870	**0.1081**	**30.3288**
	Error Mean	12.7807		5.9105		**0.0940**	

III. RESULTS AND DISCUSSION

The results of Tables 1-4 show the following:

- The proposed (FMCD) method (RRRY) outperformed the classical methods (MLE and RRY), depending on MSE and Error Mean for all simulation cases.
- The accuracy of the shape parameter estimates was better than the scale parameter estimates for the three methods and all simulation cases.
- The accuracy of estimating Weibull distribution parameters increases when the sample size increases.
- The accuracy of estimating the Weibull distribution shape parameter decreases as its assumed value increases for all simulation cases.

- The accuracy of the scale parameter estimation of the Weibull distribution decreases as its assumed value increases for all simulation cases.
- The estimate of the shape parameter remains unaffected by an increase in the scale parameter, whereas the estimate of the scale parameter is influenced by an increase in the shape parameter.
- The Error Mean indicates the superiority of the proposed method (RRRY) over the classical methods (RRY and MLE), while the classical method (RRY) is superior to the classical method (MLE).

Figure 2 shows the error means from 100 repeated experiments for the three methods. The green line represents the proposed method (RRRY), and the red line represents the classical method (RRY). The blue line represents the classical method (MLE), which shows the efficiency of the proposed method compared to classical methods; the classical method (RRY) outperformed the classical method (MLE).

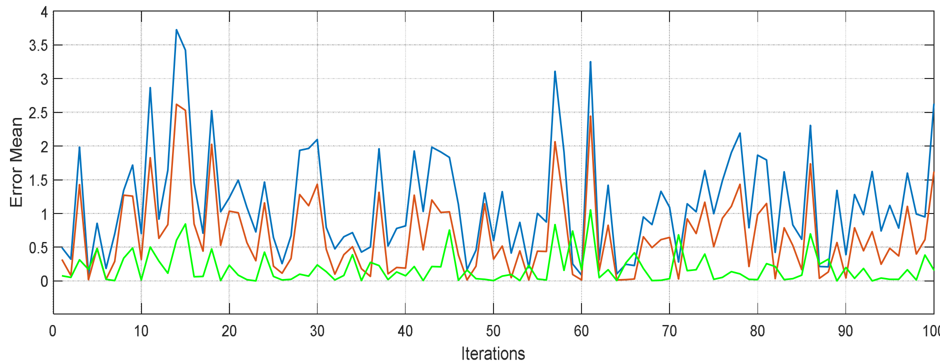


Fig. 2 Error Mean for the three methods

A. Real data

The real data set examined the survival times, measured in days, of guinea pigs infected with virulent tubercle bacilli, as summarized by Bjerkedal. The information can be found in the Appendix. The Mahalanobis distance values are presented in Figure 3 for the FMCD method, revealing that 30 outliers were identified from a total of 74 generated data points.

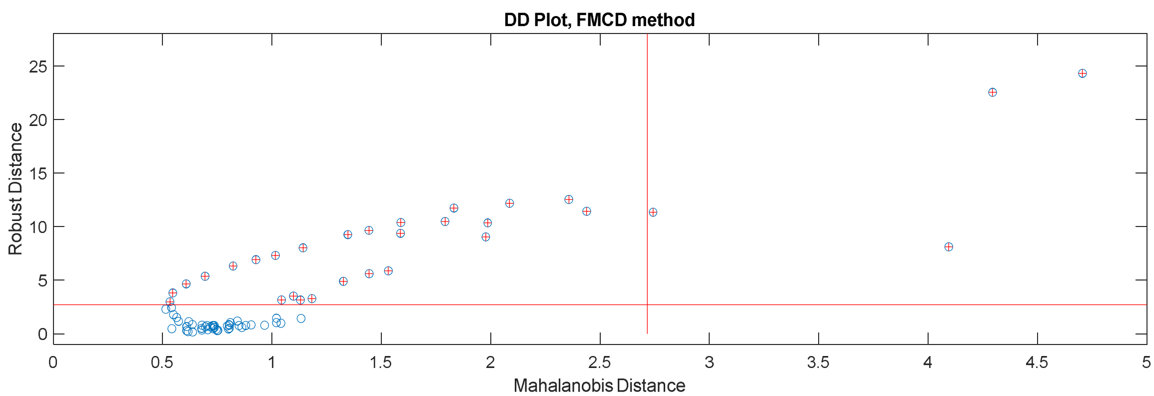


Fig. 3 Mahalanobis distance for the real data

Following the estimation of the shape and scale parameters of the Weibull distribution using the three methods, these parameters are applied to estimate survival times (expected values). Subsequently, a goodness-of-fit test, specifically the Chi-Square test, is employed to evaluate the efficiency of the estimated models. This evaluation measures the model's overall adequacy by contrasting the observed data against the predicted values generated by the model. The p-value derived from this test quantifies the statistical significance of the discrepancies between the observed and predicted values. By conducting statistical tests, we can thoroughly assess the validity of the proposed model. The model demonstrating the optimal fit is identified by the lowest Chi-Square values and non-significant p-values derived from the test statistic, thereby rendering it the most appropriate choice for the specified data set. Table 5 summarizes the estimation and testing results for the three methods, and shows that the proposed method (RRRY) was better than classical methods (RRY and MLE) because the value of the test statistic was equal to (10.627), which is less than the critical value (12.592) under significance level (0.05) and the degrees of freedom (6), (also, it less than test statistics 23.978 and 13.969 for classical models), and this is confirmed by the p-value (0.101), which was not significant, indicating the efficiency of the proposed model. While the classical models (RRY and MLE) were inefficient, because the value of the test statistic was equal to (23.978 and 13.969) respectively, which is more than the critical value (11.070) under significance level (0.05), and the degrees of freedom (5), and this is confirmed by the p-value (0.000 and 0.016) respectively, which was significant.

TABLE V

RESULTS OF ANALYSIS FOR THE REAL DATA

Method	Shape parameter (beta)	Scale parameter (eta)	Chi-Square Statistic	p-value	Critical Value	Degrees of freedom
MLE	1.3222	2.2383	23.978	0.000	11.070	5
RRY	1.6635	2.2141	13.969	0.016	11.070	5
RRRY	1.7123	1.9129	10.627	0.101	12.592	6

Figure 4 shows the probability density function of the Weibull distribution using the proposed method (RRRY).

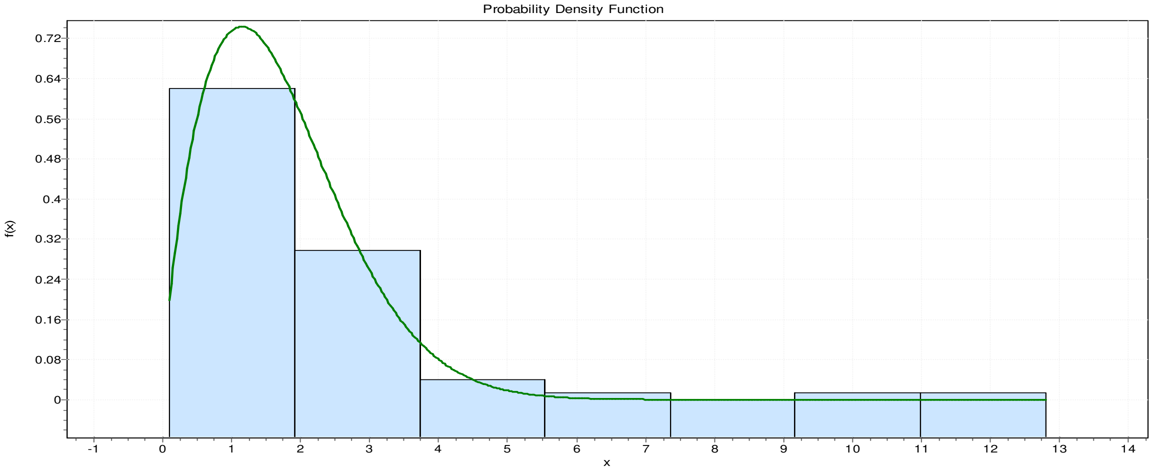


Fig. 4 The probability density function of the Weibull distribution

IV. CONCLUSION

The robust proposed (Fast Minimum Covariance Determinant) method (RRRY) outperformed the classical method (MLE and RRY), depending on MSE and Error Mean for all simulation cases and real data when there are outliers. The accuracy of the shape parameter estimates was better than that of the scale parameter estimates for the three methods and all simulation cases with outliers. The accuracy of estimating Weibull distribution parameters increases with sample size across all simulation cases. The accuracy of estimating the Weibull distribution shape parameter decreases as its assumed value increases, and the accuracy of estimating the Weibull distribution parameter decreases as its assumed value increases across all simulation cases. The estimate of the shape parameter remains accurate regardless of an increase in the scale parameter, whereas the estimate of the scale parameter is influenced by an increase in the shape parameter. The proposed method (RRRY) is superior to the classical methods (RRY and MLE), while the classical method (RRY) is superior to the classical method (MLE), depending on the Error Mean for all simulation cases when there are outliers.

Using the robust rank regression method to estimate a two-parameter Weibull distribution when there are outliers. Executing a prospective study employing the robust rank regression technique to estimate the three-parameter Weibull distribution. Conducting a prospective study using the robust rank regression method to estimate parameters' normal and exponential distribution.

REFERENCES

[1] T. H. Ali and J. R. Qadir, "Using wavelet shrinkage in the Cox proportional hazards regression model (simulation study)," *Iraqi J. Stat. Sci.*, vol. 19, no. 1, pp. 17–29, 2022, doi:10.33899/iqjoss.2022.0174328.

[2] M. S. R. M. S. Raza, T. H. A. T. H. Ali, and T. A. H. T. A. Hassan, "Using mixed distribution for gamma and exponential to estimate of survival function (brain stroke)," *Polytech. J.*, vol. 8, no. 1, 2018, doi:10.25156/ptj.2018.8.1.120.

[3] H. T. M. Ali and F. N. Abdulahad, "An application of Dirac's delta function in bivariate beta-gamma distribution," in *Proc. 8th Int. Conf. Contemp. Inf. Technol. Math. (ICCITM)*, Mosul, Iraq, 2022, pp. 478–481, doi: 10.1109/ICCITM56309.2022.10031844.

[4] K. Boudt, D. Caliskan, and C. Croux, "Robust explicit estimators of Weibull parameters," Metrika, vol. 73, no. 2, pp. 187–209, Mar. 2011, doi: 10.1007/s00184-009-0272-1.

[5] V. H. Coria, S. Maximov, F. Rivas-Dávalos, and C. L. Melchor-Hernández, "Perturbative method for maximum likelihood estimation of the Weibull distribution parameters," *SpringerPlus*, vol. 5, no. 1, p. 1802, Dec. 2016, doi: 10.1186/s40064-016-3500-y.

[6] S. K. Sinha, "Robust estimation in accelerated failure time models," *Lifetime Data Anal.*, vol. 25, no. 1, pp. 52–78, Jan. 2019, doi: 10.1007/s10985-018-9421-z.

[7] Z. Roohanizadeh, E. B. Jamkhaneh, and E. Deiri, "Parameters and reliability estimation for the Weibull distribution based on intuitionistic fuzzy lifetime data," *Complex Intell. Syst.*, vol. 8, no. 6, pp. 4881–4896, Dec. 2022, doi: 10.1007/s40747-022-00720-x.

[8] Y. M. Gómez *et al.*, "An in-depth review of the Weibull model with a focus on various parameterizations," *Mathematics*, vol. 12, no. 1, p. 56, Dec. 2023, doi: 10.3390/math12010056.

[9] M. Hubert, "An interview with Peter Rousseeuw," *Int. Stat. Rev.*, vol. 93, no. 1, pp. 1–17, 2025, doi: 10.1111/insr.12587.

[10] T. H. Ali, "Using proposed nonparametric regression models for clustered data (a simulation study)," *J. Humanit. Sci.*, vol. 29, no. 2, pp. 78–87, 2017, doi: 10.21271/ZJPAS.29.2.8.

[11] R. W. Butler, P. L. Davies, and M. Jhun, "Asymptotics for the minimum covariance determinant estimator," *Ann. Stat.*, vol. 21, no. 3, pp. 1385–1400, 1993, doi: 10.1214/aos/1176349264.

[12] T. H. Ali, B. S. Sedeeq, D. M. Saleh, and A. G. Rahim, "Robust multivariate quality control charts for enhanced variability monitoring," *Qual. Reliab. Eng. Int.*, vol. 40, no. 3, pp. 1369–1381, Apr. 2024, doi: 10.1002/qre.3472.

[13] C. Omar, T. H. Ali, and K. Hassn, "Using Bayes weights to remedy the heterogeneity problem of random error variance in linear models," *Iraqi J. Stat. Sci.*, vol. 17, no. 2, pp. 58–67, 2020, doi:10.33899/iqjoss.2020.167391.

[14] K. A. Murali, "Life data analysis reference," Reli. Corp. Worldw. Hqrs., Tucson, AZ, USA, 2016.
[15] M. M. Al-Kassab and S. Ibrahim, "Using ridge regression to estimate factors affecting the number of births. A comparative study," in *Advances in Artificial Intelligence, Software and Systems Engineering*, Cham, Switzerland: Springer, 2023, pp. 183–194, doi:10.1007/978-981-99-0447-1_15.
[16] D. J. Smith, *Reliability Engineering*. London, U.K.: Pitman, 1972.
[17] N. S. Kareem, T. H. Ali, and A. M. Shahab, "De-noise data by using multivariate wavelets in the path analysis with application," *Kirkuk Univ. J. Adm. Econ. Sci.*, vol. 10, no. 1, pp. 268–294, 2020, doi:10.32894/1913-010-001-012.
[18] T. H. Ali, H. A. A.-M. Hayawi, and D. S. I. Botani, "Estimation of the bandwidth parameter in Nadaraya-Watson kernel non-parametric regression based on universal threshold level," *Commun. Stat., Simul. Comput.*, vol. 52, no. 4, pp. 1476–1489, 2023, doi:10.1080/03610918.2021.1884719.
[19] B. S. Sedeeq, Z. A. Muhammad, I. M. Ali, and T. H. Ali, "Construction robust T²-chart and compare it with Hotelling's T²-chart," *Zanco J. Hum. Sci.*, vol. 28, pp. 140–157, 2024.
[20] T. H. Ali and D. M. Saleh, "Proposed hybrid method for wavelet shrinkage with robust multiple linear regression model: With simulation study," *Qalaai Zanist Sci. J.*, vol. 7, no. 1, pp. 920–937, 2022, doi: 10.25212/lfu.qzj.7.1.36.
[21] H. A. Taha, "Construction robust simple linear regression profile monitoring (A simulation study)," *J. Kirkuk Univ. Adm. Econ. Sci.*, vol. 9, no. 1, 2019.
[22] T. H. Ali, "Modification of the adaptive Nadaraya-Watson kernel method for nonparametric regression (simulation study)," *Commun. Stat., Simul. Comput.*, vol. 51, no. 2, pp. 391–403, 2022, doi:10.1080/03610918.2019.1652319.
[23] D. C. Montgomery, *Introduction to Statistical Quality Control*, 8th ed. Hoboken, NJ, USA: Wiley, 2020.
[24] T. H. Ali, N. H. S. Albarwari, and D. L. Ramadhan, "Using the hybrid proposed method for quantile regression and multivariate wavelet in estimating the linear model parameters," *Iraqi J. Stat. Sci.*, vol. 20, no. 1, pp. 9–24, 2023, doi: 10.33899/iqjoss.2023.0178679.

APPENDIX

0.10	0.74	1.00	1.08	1.16	1.30	1.53	1.71	1.97	2.23	2.54	3.47	10.53
0.33	0.77	1.02	1.08	1.20	1.34	1.59	1.72	2.02	2.31	2.54	3.61	12.80
0.44	0.92	1.05	1.09	1.21	1.36	1.60	1.76	2.13	2.40	2.78	4.02	
0.56	0.93	1.07	1.12	1.22	1.39	1.63	1.83	2.15	2.45	2.93	4.32	
0.59	0.96	1.07	1.13	1.22	1.44	1.63	1.95	2.16	2.51	3.27	4.58	
0.72	1.00	1.08	1.15	1.24	1.46	1.68	1.96	2.22	2.53	3.42	5.55	

THE PROGRAM

```
clc
clear all
% rng default
n=500;beta=5;eta=30;EX=[beta eta];
for j=1:1000
T = wblrnd(eta,beta,n,1); % Simulated strengths;
noise = randperm(n,5); T(noise) = T(noise)*10; T=sort(T); x=log(T);
% Compute F(t)
for i=1:n
   v1=2*(n-i+1); v2=2*i; f(i)=finv(.5,v1,v2); F(i)=1/(1+((n-i+1)/i)*f(i));
end
y=log(-log(1-F));y=y';h=[x y]; [C M]=robustcov(h,"Method","fmcd");
SYY=C(2,2); SXX=C(1,1); SXY=C(1,2); RhoR=SXY/sqrt(SXX*SYY);
b=sqrt(SYY/SXX)*RhoR; a=M(1,2)-M(1,1)*b; betarrry=b; etarrry=exp(-(a/b)); C =cov(h);
SYY=C(2,2); SXX=C(1,1); SXY=C(1,2); Rho=SXY/sqrt(SXX*SYY);
b=sqrt(SYY/SXX)*Rho; a=mean(y)-mean(x)*b; betarry=b; etarry=exp(-(a/b));
para= wblfit(T); etamle=para(1,1); betamle=para(1,2); parametermle=[betamle etamle];
EMLE(j,:)=abs(parametermle-EX); ERRY(j,:)=abs([betarry etarry]-EX);
ERRRY(j,:)=abs([betarrry etarrry]-EX); E1(j,:)=EMLE(j,:).^2; E2(j,:)=ERRY(j,:).^2;
E3(j,:)=ERRRY(j,:).^2;
end
MSEMLE=sum(E1)/1000, MSERRY=sum(E2)/1000, MSERRRY=sum(E3)/1000
E11= mean(E1'); E22= mean(E2'); E33= mean(E3'); MMLE=mean(MSEMLE)
MRRY=mean(MSERRY), MRRRY=mean(MSERRRY)
```