





Article

Robust Negative Binomial Regression via the Kibria–Lukman Strategy: Methodology and Application

Adewale F. Lukman^{1,*}, Olayan Albalawi², Mohammad Arashi^{3,4}, Jeza Allohobi⁵, Abdulmajeed Atiah Alharbi⁵ and Rasha A. Farghali⁶

¹ Department of Mathematics and Statistics, University of North Dakota, Grand Forks, ND 58202, USA

² Department of Statistics, Faculty of Science, University of Tabuk, Tabuk 47512, Saudi Arabia; oalbalwi@ut.edu.sa

³ Department of Statistics, Faculty of Mathematical Sciences, Ferdowsi University of Mashhad, Mashhad 9177948974, Razavi Khorasan, Iran; arashi@um.ac.ir

⁴ Department of Statistics, Faculty of Natural and Agricultural Sciences, University of Pretoria, Pretoria 0002, South Africa

⁵ Department of Mathematics, Faculty of Science, Taibah University, Al-Madinah Al-Munawara 42353, Saudi Arabia; jlohobi@taibahu.edu.sa (J.A.); aahharbi@taibahu.edu.sa (A.A.A.)

⁶ Department of Mathematics, Insurance and Applied Statistics, Helwan University, Cairo 11795, Egypt

* Correspondence: adewale.lukman@und.edu

Abstract: Count regression models, particularly negative binomial regression (NBR), are widely used in various fields, including biometrics, ecology, and insurance. Over-dispersion is likely when dealing with count data, and NBR has gained attention as an effective tool to address this challenge. However, multicollinearity among covariates and the presence of outliers can lead to inflated confidence intervals and inaccurate predictions in the model. This study proposes a comprehensive approach integrating robust and regularization techniques to handle the simultaneous impact of multicollinearity and outliers in the negative binomial regression model (NBRM). We investigate the estimators' performance through extensive simulation studies and provide analytical comparisons. The simulation results and the theoretical comparisons demonstrate the superiority of the proposed robust hybrid KL estimator (M-NBKLE) with predictive accuracy and stability when multicollinearity and outliers exist. We illustrate the application of our methodology by analyzing a forestry dataset. Our findings complement and reinforce the simulation and theoretical results.

Keywords: negative binomial; multicollinearity; outliers; regularization; robust hybrid KL estimator; over-dispersion

MSC: 62J05; 62J07; 62J12



Citation: Lukman, A.F.; Albalawi, O.; Arashi, M.; Allohobi, J.; Alharbi, A.A.; Farghali, R.A. Robust Negative Binomial Regression via the Kibria–Lukman Strategy: Methodology and Application. *Mathematics* **2024**, *12*, 2929. <https://doi.org/10.3390/math12182929>

Academic Editor: Heng Lian

Received: 13 August 2024

Revised: 13 September 2024

Accepted: 15 September 2024

Published: 20 September 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Count regression models play an important role in modern applied statistics, with applications across fields such as biometrics, ecology, and insurance. These models are advantageous when the response variable, y , is a count data, as in studies that model the number of falls in medical research [1]. The relationship between the response variable y and a set of covariates $x = (1, x_1, \dots, x_p)^T$ is often expressed as $\mathbb{E}(y|x) = g^{-1}(x^T\beta)$, where $\beta = (\beta_0, \beta_1, \dots, \beta_p)^T$ are the regression coefficients and $g(\cdot)$ is the canonical link function within the framework of generalized linear model (GLM) [2]. This formulation provides a flexible approach to modeling count data while accounting for various covariates, making it widely applicable in many disciplines.

In recent years, the NBR has gained significant attention as an effective tool for modeling count data in the presence of overdispersion. The primary goal is to estimate β

accurately. However, multicollinearity among covariates (predictors) can lead to inflated confidence intervals and inaccurate predictions for the response variable [2,3]. Outlying observations can also significantly impact the estimates, making it essential to develop robust strategies to eliminate their misleading effects [3]. Unfortunately, in the GLM, multicollinearity and outliers are often considered separately, leading to incomplete analyses that fail to address the simultaneous impact of multicollinearity and outliers on the model's performance. Although there have been various studies addressing the issue of multicollinearity in the NBRM, such as the development of ridge and Liu estimators [4–6], and the Jackknifed Liu-type estimator [7], the robust inference in the presence of outliers is neglected.

Outliers are common occurrences in real-life applications and can often result in misleading results and conclusions in data science. Causes of outliers include human error, measurement error, and recording error, or can be simply inherent in the data. It is crucial to identify and remove such outliers using appropriate data-cleaning methods and employing methods that are not susceptible to outliers to ensure robustness in data science [8,9]. While considerable research has been conducted on robust regression modeling, relatively few studies have specifically addressed the issue of outliers in NBR [1,8–10]. For a comprehensive review of the robust inference for the GLMs, Medina and Ronchetti's [10] work is a valuable resource.

Lukman et al. [11,12] pointed out the simultaneous impact of multicollinearity and outliers on the Poisson and linear regression models. Roozbeh et al. [13,14] developed numerical algorithms to solve nonlinear equations resulting from the estimation of regression coefficients in the presence of multicollinearity and outliers. While both issues have received detailed attention in linear and Poisson regression models, more focus is needed in the NB framework. To the best of our knowledge, there is currently no available research handling multicollinearity and outliers in the context of NBRMs at the time of writing.

The primary objective of this article is to examine the difficulties associated with multicollinearity and outliers in NBR modeling and propose solutions to overcome these obstacles. To achieve this, we will integrate robust and regularization techniques for NBR modeling to effectively address the issues of multicollinearity and outliers. The main contribution of this work is the introduction of a new estimator for the NBRM that combines the shrinkage estimator with the robust regression to account for both multicollinearity and outliers in NBR.

The rest of the paper is as follows. Section 2 reviews some existing methods and develops new estimators. Additionally, this section includes analytical comparisons, where we theoretically compare the proposed methods with existing alternatives using performance measures. In Section 3, an extensive simulation study is considered, where we intensively study the effect of multicollinearity and outliers on the NBRM. The study is supported by a real-life application in Section 4, followed by the conclusion in Section 5. We have included the code used for the real-life application to facilitate replication of our results in Supplementary Material.

2. Methodology

2.1. Negative Binomial Regression Model

NBR is a generalization of Poisson regression designed to handle overdispersion, where the variance of the response variable exceeds the mean. NBR provides a more flexible alternative to Poisson regression for count data that exhibit greater dispersion by introducing an additional parameter to model this extra variability. Given a sample of n observations y_1, y_2, \dots, y_n , the probability mass function (p.m.f.) of the NB distribution is expressed as the following equation:

$$f(y_i|x_i) = \frac{\Gamma(\gamma^{-1} + y_i)}{\Gamma(\gamma^{-1}) \Gamma(1 + y_i)} \left(\frac{\gamma^{-1}}{\gamma^{-1} + \mu_i} \right)^{\gamma^{-1}} \left(\frac{\mu_i}{\gamma^{-1} + \mu_i} \right)^{y_i} \quad y_i = 0, 1, 2, \dots, \quad i = 1, 2, \dots, n. \quad (1)$$

The overdispersion parameter γ is defined as $(\gamma = \frac{1}{\tau})$. μ_i denotes the mean of the distribution for the i -th observation and $\Gamma(\cdot)$ is the gamma function. The conditional mean and variance are calculated as follows:

$$E(y_i|\mu_i, \tau_i) = \mu_i. \tag{2}$$

$$\text{Var}(y_i|\mu_i, \tau_i) = \mu_i(1 + \gamma\mu_i). \tag{3}$$

The maximum likelihood (ML) method is commonly used to estimate the parameter vector β by maximizing the log-likelihood function. The mean parameter μ_i is often modelled as a function of the explanatory variables x_i through a log-link: $\log\mu_i = x_i^T\beta$. Thus, the log-likelihood function is given as follows:

$$\ell(\beta, \gamma) = \sum_{j=1}^n \left[\log\Gamma(\gamma^{-1} + y_i) - \log\Gamma(y_i + 1) - \log\Gamma(\gamma^{-1}) + \gamma^{-1} \log\left(\frac{\gamma^{-1}}{\gamma^{-1} + \mu_i}\right) + y_i \log\left(\frac{\mu_i}{\gamma^{-1} + \mu_i}\right) \right]. \tag{4}$$

The score function, which results from taking the first derivative of the log-likelihood function to β is given as follows:

$$\frac{\partial \ell(\beta, \gamma)}{\partial \beta_j} = \sum_{i=1}^n \left(\frac{y_i - \mu_i}{1 + \gamma\mu_i} \right) x_{ij} = 0 \quad j = 1, 2, \dots, p^*. \tag{5}$$

This system of equations is nonlinear and, as mentioned, does not have a closed-form solution. We employ numerical methods to find the parameter estimates. Specifically, we use iterative reweighted least squares (IRLS) to solve the nonlinear system. At each iteration r , the estimate $\hat{\beta}^{r+1}$ is as follows:

$$\hat{\beta}^{r+1} = \hat{\beta}^r + (I(\hat{\beta}^r))^{-1} S(\hat{\beta}^r), \tag{6}$$

where $S(\hat{\beta}^r)$ is the score function evaluated at the r -th iteration. $I(\hat{\beta}^r) = E\left(\frac{\partial^2 \ell(\beta, \gamma)}{\partial \beta \partial \beta'}\right) = X'W^rX$ is a $p^* \times p^*$ Fisher information matrix and it is evaluated at $\hat{\beta}^r$ and W^r is a $n \times n$ diagonal matrix; the iterations will end when the difference $\|\hat{\beta}^r - \hat{\beta}^{r+1}\| < \epsilon$, ϵ is a certain small value that usually equals 10^{-6} . Thus, the maximum likelihood estimates (NBMLE) are obtained iteratively as follows:

$$\hat{\beta}^{r+1} = (X'\hat{W}^rX)^{-1} X'\hat{W}^r z^r, \tag{7}$$

where z^r is $n \times 1$ vector of adjusted response variable such that $z_i^r = \log(\hat{\mu}_i^r) + \frac{y_i - \hat{\mu}_i^r}{\hat{\mu}_i^r}$, and the weight $\hat{W}^r = \text{diag}\left(\frac{\hat{\mu}_i^r}{1 + \theta\hat{\mu}_i^r}\right)$. The final equation at convergence is

$$\hat{\beta}_{NBMLE} = (X'\hat{W}X)^{-1} X'\hat{W}z^*, \tag{8}$$

where z^* is the final adjusted response vector such that $z_i^r = \log(\hat{\mu}_i) + \frac{y_i - \hat{\mu}_i}{\hat{\mu}_i}$ and $\hat{W} = \text{diag}\left(\frac{\hat{\mu}_i}{1 + \theta\hat{\mu}_i}\right)$. The var-covariance matrix can be obtained as follows:

$$\text{Cov}(\hat{\beta}_{NBMLE}) = (X'\hat{W}X)^{-1}. \tag{9}$$

2.2. Shrinkage Estimators

Biased estimators, such as the ridge regression estimator developed by Hoerl and Kennard [15], the Liu and Liu-type estimators [16,17], and the Kibria–Lukman estimator [18], are among those introduced to address the multicollinearity problem in linear regression. These estimators are also used in generalized linear models to account for multicollinearity in binary logistic regression [19,20], zero-inflated Poisson regression [21–23], Poisson regression [24–27], NBR [4,5,8,28,29] and Conway–Maxwell–Poisson [30], among others.

Månsson [4] introduced the ridge regression estimator (NBRRE) for the NBR as follows:

$$\hat{\beta}_{NBRRE} = (X'WX + kI)^{-1} X'WX \hat{\beta}_{NBMLE}, k > 0. \tag{10}$$

where k is the ridge parameter, defined in this study as: $k = \frac{1}{\max(\hat{\beta}_{NBMLE}^2)}$.

The asymptotic covariance matrix is calculated as follows:

$$\text{Cov}(\hat{\beta}_{NBRRE}) = (X'WX + kI)^{-1} X'WX (X'WX + kI)^{-1}. \tag{11}$$

Following the works of [15,16], Kibria and Lukman [18] developed the Kibria–Lukman estimator (KLE), a single-parameter biased estimator designed to address multicollinearity in linear regression models. They demonstrated that this estimator outperforms both the Ridge and Liu estimators in terms of estimation accuracy and stability. Thus, the KLE for the NBRM is given by:

$$\hat{\beta}_{NBKLE} = (X'WX + kI)^{-1} (X'WX - kI) \hat{\beta}_{NBMLE}, k > 0. \tag{12}$$

The asymptotic covariance matrix is calculated as follows:

$$\text{Cov}(\hat{\beta}_{NBKLE}) = (X'WX + kI)^{-1} (X'WX - kI) (X'WX)^{-1} (X'WX - kI) (X'WX + kI)^{-1}. \tag{13}$$

Under certain conditions of their simulation study, they proved that the KL estimator outperforms both the ridge regression and Liu estimators.

2.3. Shrinkage-Robust Estimators

The occurrence of outliers is a prevalent issue in regression analysis. Outliers are observations that stand out from the rest of the data [31]. Outliers were divided into vertical outliers and good and bad leverage points [32,33]. Vertical outliers are those observations that are not outlying in the space of explanatory variables (the x-dimension) but have outlying values for the related error term (on the y-direction). Outlying observations in the explanatory variables' space located near the regression line are good leverage points. Bad leverage points are observations that are both off-center in the explanatory variable space and distant from the actual regression line. Outliers exist in both ways (x-direction of the explanatory variables and y-direction of the response variable). Outliers can produce biased regression estimates and misleading regression inferences as well as overestimate the standard errors. As a result, while doing power analyses, overestimate the required sample size [1,34].

The most popular robust regression technique is M-estimation [35], which is almost as effective as the least squares estimator. The M-estimator minimizes the function of the standardized residual to minimize the sum of squared residuals as the goal. This approach performs better than the conventional least squares estimator because it assigns fewer weights to atypical observations. The letter M denotes that the estimation is of the maximum likelihood kind. This method allows for removing certain data, which is

sometimes not the best course of action, especially if the material being removed is crucial. The M-estimator is obtained by minimizing the residual function as follows:

$$\hat{\beta}_M = \text{Min}_{\beta} \sum_{i=1}^n \rho \left(\frac{y_i - \beta_j x_{ij}}{s} \right), \tag{14}$$

where s is an estimate of scale from a linear combination of the residuals and the function ρ assigns the contribution of each residual to the objective function. A system of normal equations is obtained by taking the first partial derivatives with respect to $\beta_j, j = 1, 2, \dots, p^*$ and setting them equal to zero.

Aeberhard et al. [1] introduced robust estimators for the parameters of the NB distribution using the M-estimator. They achieved robustness in the y -direction by bounding the Pearson residuals $r_i = (y_i - \mu_i) \text{Var}^{1/2}(y_i | \mu_i, \tau_i)$ in the response that appears in the score function and on the design by introducing weights $w(x_i)$. Thus, the Mallows quasi-likelihood estimator is obtained as follows:

$$\sum_{i=1}^n S(\beta)_{r_i} \text{Var}^{1/2}(y_i | \mu_i, \tau_i) \frac{\partial \mu_i}{\partial \varphi} w(x_i) x_i - E_i(\beta) = 0, \tag{15}$$

where $S(\cdot)_{r_i}$ is a continuous and bounded function that depends on few tuning constants, $w(x_i)$ is a weight controlling the impact of possible leverage points, and $E_i(\beta) = E[S(\beta)_{r_i} \text{Var}^{1/2}(y_i | \mu_i, \tau_i) \frac{\partial \mu_i}{\partial \varphi} w(x_i) x_i]$ is an adjustment term guaranteeing Fisher consistency at the model. The choice of the $S(\cdot)_{r_i}$ function that bounds the Pearson residual is crucial because the resulting estimator has a bounded effect function thanks to the $S(\cdot)_{r_i}$ function's boundedness criterion. In this study, we consider Tukey's bi-weight function that has the advantage of being tuned by a single constant. The Tukey weight is a weight function that can be used in robust regression methods, including M-estimators based on Pearson residuals, to down-weight influential observations. The Tukey weight function is given as follows:

$$S_{\text{Tukey}}(r_i, c) = \begin{cases} \left(\left(\frac{r_i}{c} \right)^2 - 1 \right)^2 r_i & |r_i| \leq c \\ 0 & |r_i| > c \end{cases}, \quad i = 1, 2, \dots, n, \tag{16}$$

where r_i is the Pearson residual for the i^{th} observation, $i = 1, 2, \dots, n$, c is a tuning parameter that determines the cut-off point for the weight function, and $S_{\text{Tukey}}(r_i, c)$ is the weight assigned to i^{th} observation, $i = 1, 2, \dots, n$. The Huber weight function and the Tukey weight function are comparable; however, the Tukey weight function provides a more seamless change from high weights to low weights for observations with significant Pearson residuals. The Tukey weight function gives observations with Pearson residuals less than or equal to a c weight of 1, and it gradually reduces that weight as the Pearson residuals rise over c . A weight of 0 is applied to observations with Pearson residuals higher than c .

Researchers have proposed robust-biased estimators as viable alternatives when dealing with a regression model that encounters both multicollinearity and outliers in linear and Poisson regression models. These estimators combined biased estimators with robust estimators such as M-estimators. For more in-depth information, refer to the following works [11,12,36–41].

Silvapulle [36] proposed a robust ridge regression estimator for a linear regression model as follows:

$$\hat{\beta}_{M-RRE} = (X'X + kI)^{-1} X'X \hat{\beta}_{M-OLS}, \quad k > 0. \tag{17}$$

Majid et al. [41] proposed a robust Kibria–Lukman estimator for a linear regression model as follows:

$$\hat{\beta}_{M-KLE} = (X'X + kI)^{-1} (X'X - kI) \hat{\beta}_{M-OLS}, \quad k > 0. \tag{18}$$

Following the work of Silvapulle [36], Majid et al. [41], Lukman et al. [11], and Arum et al. [40], we extended the robust ridge regression estimator and the robust KL estimator to the NBRM; the new estimators were found by combining the ridge regression estimator and the KL estimator with the M-estimator instead of the maximum likelihood estimators as follows:

$$\hat{\beta}_{M-NBRRE} = (X' \hat{W} X + kI)^{-1} X' \hat{W} X \hat{\beta}_{M-MLE} \quad k > 0. \tag{19}$$

$$\hat{\beta}_{M-NBKLE} = (X' \hat{W} X + kI)^{-1} (X' \hat{W} X - kI) \hat{\beta}_{M-MLE} \quad k > 0. \tag{20}$$

The asymptotic covariance matrices are calculated as follows:

$$Cov(\hat{\beta}_{M-NBRRE}) = (X' \hat{W} X + kI)^{-1} (X' \hat{W} X) Var(\hat{\beta}_{M-MLE}) (X' \hat{W} X) (X' \hat{W} X + kI)^{-1}. \tag{21}$$

$$Cov(\hat{\beta}_{M-NBKLE}) = (X' \hat{W} X + kI)^{-1} (X' \hat{W} X - kI) Var(\hat{\beta}_{M-MLE}) (X' \hat{W} X - kI) (X' \hat{W} X + kI)^{-1} \tag{22}$$

2.4. Theoretical Comparisons between Estimators

For convenience, we use the spectral decomposition of the estimated weighted information matrix $(X' \hat{W} X)$; assume that there exists a matrix G such that $G(X' \hat{W} X)G' = \Lambda = diag\{\lambda_j\}, j = 1, 2, \dots, p^*$, where $\lambda_1 \geq \lambda_2 \geq \Delta \geq \lambda_{p^*}$ are the ordered eigenvalues of $X' \hat{W} X$ and G orthogonal Matrix of order $(p^* \times p^*)$, its columns are the corresponding eigenvectors for $\lambda_1, \lambda_2, \Delta, \lambda_{p^*}$. Assume that $H = XG, \therefore G'X' \hat{W} XG = H'H = \Lambda$, then $\hat{\alpha}_{NBMLE} = G' \hat{\beta}_{NBMLE}$. Thus, Equations (8), (10), (12), (19) and (20) can be written in canonical form as follows:

$$\hat{\alpha}_{NBMLE} = \Lambda^{-1} H \hat{W} z^*. \tag{23}$$

$$\hat{\alpha}_{NBRRE} = (\Lambda + kI)^{-1} \Lambda \hat{\alpha}_{NBMLE}. \tag{24}$$

$$\hat{\alpha}_{NBKLE} = (\Lambda + kI)^{-1} (\Lambda - kI) \hat{\alpha}_{NBMLE}. \tag{25}$$

$$\hat{\alpha}_{M-NBRRE} = (\Lambda + kI)^{-1} \Lambda \hat{\alpha}_{M-MLE}. \tag{26}$$

$$\hat{\alpha}_{M-NBKLE} = (\Lambda + kI)^{-1} (\Lambda - kI) \hat{\alpha}_{M-MLE}. \tag{27}$$

Specifically, we replace $\hat{\alpha}_{NBMLE}$, which is not robust, with the robust version proposed by Aeberhard et al. [1]. We have defined this robust version in our study as $\hat{\alpha}_{M-MLE}$.

The performance of the $\tilde{\xi}$ of the regression parameter ξ is evaluated using the scalar mean squared error (SMSE), which is defined as follows:

$$SMSE(\tilde{\xi}) = tr[Cov(\tilde{\xi})] + bias(\tilde{\xi})' bias(\tilde{\xi}). \tag{28}$$

where $Cov(\tilde{\xi})$ represents the covariance matrix of $\tilde{\xi}$ and $bias(\tilde{\xi}) = E(\tilde{\xi}) - \xi$.

The SMSE of the mentioned estimators above are obtained as follows:

$$SMSE(\hat{\alpha}_{NBMLE}) = \sum_{j=1}^{p^*} \frac{1}{\lambda_j} \tag{29}$$

$$SMSE(\hat{\alpha}_{NBRRE}) = \sum_{j=1}^{p^*} \frac{\lambda_j}{(\lambda_j + k)^2} + \sum_{j=1}^{p^*} \frac{\alpha_j^2 k^2}{(\lambda_j + k)^2}. \tag{30}$$

$$SMSE(\hat{\alpha}_{NBKLE}) = \sum_{j=1}^{p^*} \frac{(\lambda_j - k)^2}{(\lambda_j + k)^2 \lambda_j} + 4 \sum_{j=1}^{p^*} \frac{\alpha_j^2 k^2}{(\lambda_j + k)^2}. \tag{31}$$

$$SMSE(\hat{\alpha}_{M-NBRRE}) = \sum_{j=1}^{p^*} \frac{\lambda_j^2}{(\lambda_j + k)^2} \psi_{jj} + \sum_{j=1}^{p^*} \frac{\alpha_j^2 k^2}{(\lambda_j + k)^2}. \tag{32}$$

$$SMSE(\hat{\alpha}_{M-NBKLE}) = \sum_{j=1}^{p^*} \frac{(\lambda_j - k)^2}{(\lambda_j + k)^2} \psi_{jj} + 4 \sum_{j=1}^{p^*} \frac{\alpha_j^2 k^2}{(\lambda_j + k)^2}, \tag{33}$$

where $\hat{\alpha}_{M-MLE} = \Psi$, and $SMSE(\hat{\alpha}_{M-MLE}) = \sum_{j=1}^{p^*} \psi_{jj}$. We assume the following conditions are satisfied:

- (I) Ψ is finite.
- (II) ψ_{jj} is skew-symmetric and nondecreasing.
- (III) The errors have a zero mean and finite variance.

Theorem 1. *The estimator $\hat{\alpha}_{M-NBKLE}$ is superior to the estimator $\hat{\alpha}_{MLE}$ in the sense of SMSE criterion, i.e., $SMSE(\hat{\alpha}_{M-NBKLE}) - SMSE(\hat{\alpha}_{MLE}) < 0$ if $\psi_{jj} < (\lambda_j + k)^2 - 4\alpha_j^2 k^2 / (\lambda_j - k)^2$.*

Proof. The difference between $SMSE(\hat{\alpha}_{M-NBKLE})$ and $SMSE(\hat{\alpha}_{MLE})$ is as follows:

$$SMSE(\hat{\alpha}_{M-NBKLE}) - SMSE(\hat{\alpha}_{MLE}) = \sum_{j=1}^{p^*} \frac{(\lambda_j - k)^2}{(\lambda_j + k)^2} \psi_{jj} + 4 \sum_{j=1}^{p^*} \frac{\alpha_j^2 k^2}{(\lambda_j + k)^2} - \sum_{j=1}^{p^*} \frac{1}{\lambda_j} \\ \sum_{j=1}^{p^*} \frac{(\lambda_j - k)^2 \lambda_j \psi_{jj} + 4\alpha_j^2 k^2 \lambda_j - \lambda_j (\lambda_j + k)^2}{\lambda_j (\lambda_j + k)^2} \tag{34}$$

The difference between $SMSE(\hat{\alpha}_{M-NBKLE})$ and $SMSE(\hat{\alpha}_{MLE})$ in Equation (34) will be less than zero if $\psi_{jj} < (\lambda_j + k)^2 - 4\alpha_j^2 k^2 / (\lambda_j - k)^2$; thus, it means that $\hat{\alpha}_{M-NBKLE}$ is better than $\hat{\alpha}_{MLE}$ since it has smaller scalar mean squared errors. □

Theorem 2. *The estimator $\hat{\alpha}_{M-NBKLE}$ is superior to the estimator $\hat{\alpha}_{NBRRE}$ in the sense of SMSE criterion, i.e., $SMSE(\hat{\alpha}_{M-NBKLE}) - SMSE(\hat{\alpha}_{NBRRE}) < 0$ if $\psi_{jj} < \lambda_j - 3\alpha_j^2 k^2 / (\lambda_j - k)^2$.*

Proof. The difference between $SMSE(\hat{\alpha}_{M-NBKLE})$ and $SMSE(\hat{\alpha}_{NBRRE})$ is as follows:

$$SMSE(\hat{\alpha}_{M-NBKLE}) - SMSE(\hat{\alpha}_{NBRRE}) = \sum_{j=1}^{p^*} \frac{(\lambda_j - k)^2}{(\lambda_j + k)^2} \psi_{jj} + 4 \sum_{j=1}^{p^*} \frac{\alpha_j^2 k^2}{(\lambda_j + k)^2} - \sum_{j=1}^{p^*} \frac{\lambda_j}{(\lambda_j + k)^2} - \sum_{j=1}^{p^*} \frac{\alpha_j^2 k^2}{(\lambda_j + k)^2} \\ \sum_{j=1}^{p^*} \frac{(\lambda_j - k)^2 \psi_{jj} + 4\alpha_j^2 k^2 - \alpha_j^2 k^2 - \lambda_j}{(\lambda_j + k)^2} \tag{35}$$

The difference between $SMSE(\hat{\alpha}_{M-NBKLE})$ and $SMSE(\hat{\alpha}_{NBRRE})$ in Equation (35) will be less than zero if $\psi_{jj} < \lambda_j - 3\alpha_j^2 k^2 / (\lambda_j - k)^2$; thus, it means that $\hat{\alpha}_{M-NBKLE}$ is better than $\hat{\alpha}_{NBRRE}$ since it has smaller scalar mean squared errors. □

Theorem 3. *The estimator $\hat{\alpha}_{M-NBKLE}$ is superior to the estimator $\hat{\alpha}_{M-NBRRE}$ in the sense of SMSE criterion, i.e., $SMSE(\hat{\alpha}_{M-NBKLE}) - SMSE(\hat{\alpha}_{M-NBRRE}) < 0$ if $\psi_{jj} < 3\alpha_j^2 k^2 / 2\lambda_j k - k^2$.*

Proof. The difference between $SMSE(\hat{\alpha}_{M-NBKLE})$ and $SMSE(\hat{\alpha}_{M-NBRRE})$ is as follows:

$$SMSE(\hat{\alpha}_{M-NBKLE}) - SMSE(\hat{\alpha}_{M-NBRRE}) = \sum_{j=1}^{p^*} \frac{(\lambda_j - k)^2}{(\lambda_j + k)^2} \psi_{jj} + 4 \sum_{j=1}^{p^*} \frac{\alpha_j^2 k^2}{(\lambda_j + k)^2} - \sum_{j=1}^{p^*} \frac{\lambda_j^2}{(\lambda_j + k)^2} \psi_{jj} - \sum_{j=1}^{p^*} \frac{\alpha_j^2 k^2}{(\lambda_j + k)^2} \\ \sum_{j=1}^{p^*} \frac{(\lambda_j - k)^2 \psi_{jj} + 4\alpha_j^2 k^2 - \lambda_j^2 \psi_{jj} - \alpha_j^2 k^2}{(\lambda_j + k)^2} \tag{36}$$

The difference between $SMSE(\hat{\alpha}_{M-NBKLE})$ and $SMSE(\hat{\alpha}_{M-NBRRE})$ in Equation (36) will be less than zero if $\psi_{jj} < 3\alpha_j^2 k^2 / 2\lambda_j k - k^2$; thus, it means that $\hat{\alpha}_{M-NBKLE}$ is better than $\hat{\alpha}_{M-NBRRE}$ since it has smaller scalar mean squared errors. \square

Theorem 4. *The estimator $\hat{\alpha}_{M-NBKLE}$ is superior to the estimator $\hat{\alpha}_{NBKLE}$ in the sense of SMSE criterion, i.e., $SMSE(\hat{\alpha}_{M-NBKLE}) - SMSE(\hat{\alpha}_{NBKLE}) < 0$ if $\psi_{jj} < 1/\lambda_j$.*

Proof. The difference between $SMSE(\hat{\alpha}_{M-NBKLE})$ and $SMSE(\hat{\alpha}_{NBKLE})$ is as follows:

$$SMSE(\hat{\alpha}_{M-NBKLE}) - SMSE(\hat{\alpha}_{NBKLE}) = \sum_{j=1}^{p^*} \frac{(\lambda_j - k)^2}{(\lambda_j + k)^2} \psi_{jj} + 4 \sum_{j=1}^{p^*} \frac{\alpha_j^2 k^2}{(\lambda_j + k)^2} - \sum_{j=1}^{p^*} \frac{(\lambda_j - k)^2}{\lambda_j (\lambda_j + k)^2} - 4 \sum_{j=1}^{p^*} \frac{\alpha_j^2 k^2}{(\lambda_j + k)^2} \\ \sum_{j=1}^{p^*} \frac{(\lambda_j - k)^2 \psi_{jj} \lambda_j - (\lambda_j - k)^2}{\lambda_j (\lambda_j + k)^2} \tag{37}$$

The difference between $SMSE(\hat{\alpha}_{M-NBKLE})$ and $SMSE(\hat{\alpha}_{NBKLE})$ in Equation (37) will be less than zero if $\psi_{jj} < 1/\lambda_j$; thus, it means that $\hat{\alpha}_{M-NBKLE}$ is better than $\hat{\alpha}_{NBKLE}$ since it has smaller scalar mean squared errors. \square

3. Simulation Study

To ensure a robust evaluation, we meticulously designed the simulation study, carefully specifying the variables to influence the characteristics of the proposed estimator. We selected a suitable metric, drawing inspiration from well-established references to evaluate the outcomes [11,12,41–44]. To generate the regressors, we followed the equation:

$$x_{ij} = (1 - \rho^2)^{1/2} m_{ij} + \rho m_{i,p^*+1}, \quad i = 1, 2, \dots, n, \quad j = 1, 2, 3, \dots, p^* \tag{38}$$

Here, m_{ij} represents independent standard normal pseudo-random numbers. We considered various scenarios by varying the number of regressors ($p^* = 3$ and 5) and the level of multicollinearity ($\rho = 0.8, 0.9, 0.99, 0.999$). The response variable y was generated using the “rnbinom” function, which simulates NB random variables. The mean of y was computed based on the exponential of the linear combination of predictors ($\mu_i = \exp(x_i^T, \beta)$) where the over-dispersion parameter (σ) is set to 5. The length of y was varied among $n = 30, 50, 100,$ and 200 , ensuring that the regression parameters β were chosen such that $\beta' \beta = 1$ [42–44].

The line y [sample ($n, n \cdot 0.08$)] = 50 was used to introduce outliers in the response variable. The sample function randomly selects a subset of $n \cdot b$ indices, where $n \cdot b$ represents $b\%$ of the total number of observations in the response variable. The percentage of outliers (b) was chosen arbitrarily as 1 and 8, respectively. The selected indices corresponded to the positions of the outliers, and their values were then assigned to a predefined outlier value of 50. Consequently, 1% or 8% of the observations in the response variable were replaced with the outlier value of 50.

The model is estimated using the various methods discussed in this study. The “robNB” package is useful when dealing with count data with outliers, where the number of occurrences of an event is non-negative and discrete. It produced robust standard errors and other robust inferential procedures.

To ensure the reliability of our results, we conducted 2000 replications of the experiment. To evaluate the performance of the estimators, we calculated the estimated mean squared error (MSE) using the following formula:

$$MSE = \frac{1}{2000} \sum_{i=1}^{2000} \sum_{j=1}^{p^*} (\hat{\beta}_{ij} - \beta_j)^2 \tag{39}$$

Here, $\hat{\beta}_{ij}$ represents the estimated value of the j^{th} parameter in the i^{th} replication, and β_j denotes the corresponding true parameter value. The estimated MSE values for the proposed estimator, with other estimators, can be found in Tables 1 and 2.

Table 1. Estimated MSE values for $p^* = 3$ with 1% and 8% outliers.

n	1% Outliers				8%			
	30	50	100	200	30	50	100	200
	$\rho = 0.8$				$\rho = 0.8$			
$\hat{\alpha}_{MLE}$	0.2433	0.1759	0.1407	0.1371	1.7302	1.7117	1.3297	1.2726
$\hat{\alpha}_{NBRRE}$	0.2226	0.1489	0.1322	0.1312	1.5284	1.5106	1.2843	1.2561
$\hat{\alpha}_{M-NBRRE}$	0.1543	0.1480	0.1320	0.0274	0.4017	0.3218	0.1868	0.0971
$\hat{\alpha}_{NBKLE}$	0.2054	0.1338	0.1301	0.1299	1.3438	1.2856	1.2401	1.2397
$\hat{\alpha}_{M-NBKLE}$	0.1323	0.1258	0.1227	0.0255	0.3364	0.3178	0.1623	0.0933
	$\rho = 0.9$				$\rho = 0.9$			
$\hat{\alpha}_{MLE}$	2.2999	0.5132	0.3606	0.1671	4.3435	2.4649	1.4263	1.3409
$\hat{\alpha}_{NBRRE}$	2.1946	0.4154	0.2872	0.1494	3.6460	2.2617	1.3372	1.3218
$\hat{\alpha}_{M-NBRRE}$	1.3213	0.3346	0.2128	0.0439	2.0874	1.4598	0.4295	0.2087
$\hat{\alpha}_{NBKLE}$	2.1408	0.3481	0.2285	0.1341	3.0200	2.0920	1.2539	1.2630
$\hat{\alpha}_{M-NBKLE}$	1.2524	0.2508	0.1895	0.0404	1.5043	1.3688	0.3516	0.1169
	$\rho = 0.99$				$\rho = 0.99$			
$\hat{\alpha}_{MLE}$	4.2888	2.0952	2.0325	1.2622	13.6250	9.1668	2.8536	2.0822
$\hat{\alpha}_{NBRRE}$	3.4781	1.2561	1.3010	0.8275	9.6663	5.4477	1.9129	1.7719
$\hat{\alpha}_{M-NBRRE}$	2.7175	1.1227	0.5659	0.3005	5.9970	5.0101	1.4041	0.3160
$\hat{\alpha}_{NBKLE}$	3.4430	0.6316	0.5857	0.5342	6.9971	3.1266	1.6077	1.5263
$\hat{\alpha}_{M-NBKLE}$	2.3795	0.5089	0.4302	0.1149	3.7144	3.0711	0.8117	0.1274
	$\rho = 0.999$				$\rho = 0.999$			
$\hat{\alpha}_{MLE}$	13.6544	14.3600	7.7466	6.7521	149.2937	70.4487	33.4742	11.6880
$\hat{\alpha}_{NBRRE}$	5.6521	7.4897	2.4514	1.5667	104.4286	46.4188	23.5842	7.8009
$\hat{\alpha}_{M-NBRRE}$	3.8032	5.1630	1.7471	1.3126	67.5508	24.1480	2.1384	1.6055
$\hat{\alpha}_{NBKLE}$	5.0185	3.9674	1.2086	1.1580	71.7484	32.0596	17.3464	5.9552
$\hat{\alpha}_{M-NBKLE}$	3.1395	2.6155	1.1361	0.4985	32.8514	7.9232	1.4756	0.5311

Table 2. Estimated MSE values for $p^* = 5$ with 1% and 8% outliers.

n	1% Outliers				8%			
	30	50	100	200	30	50	100	200
	$\rho = 0.8$				$\rho = 0.8$			
$\hat{\alpha}_{MLE}$	0.3745	0.3413	0.3336	0.3259	2.3125	2.0120	1.6164	1.5833
$\hat{\alpha}_{NBRRE}$	0.2878	0.2794	0.2485	0.1849	2.2610	1.7329	1.3588	1.2773
$\hat{\alpha}_{M-NBRRE}$	0.2791	0.2633	0.2046	0.0458	1.0064	0.8254	0.3955	0.1100
$\hat{\alpha}_{NBKLE}$	0.2506	0.2460	0.2387	0.1468	2.2164	1.6370	1.2932	1.2714
$\hat{\alpha}_{M-NBKLE}$	0.2088	0.2001	0.1958	0.0400	1.0028	0.7087	0.3395	0.1082

Table 2. Cont.

n	1% Outliers				8%			
	30	50	100	200	30	50	100	200
	$\rho = 0.9$				$\rho = 0.9$			
$\hat{\alpha}_{MLE}$	3.7542	0.6788	0.6353	0.5962	5.5100	2.8125	1.6546	3.3937
$\hat{\alpha}_{NBRRE}$	3.5119	0.4856	0.4647	0.4254	4.3696	2.4532	1.5253	3.3180
$\hat{\alpha}_{M-NBRRE}$	2.5476	0.4323	0.4217	0.0565	3.0110	2.0634	0.5418	0.2955
$\hat{\alpha}_{NBKLE}$	3.3406	0.5366	0.5246	0.4618	4.2623	2.3850	1.4056	3.2442
$\hat{\alpha}_{M-NBKLE}$	2.3775	0.5032	0.2481	0.0462	3.0106	2.0042	0.4404	0.2774
	$\rho = 0.99$				$\rho = 0.99$			
$\hat{\alpha}_{MLE}$	6.8841	5.1915	4.6733	3.8550	14.3885	11.5578	7.9738	9.2243
$\hat{\alpha}_{NBRRE}$	4.1880	3.2344	3.1333	2.7374	12.8431	7.7036	6.2574	7.5123
$\hat{\alpha}_{M-NBRRE}$	3.8932	3.4318	1.6121	1.4313	6.0214	5.5507	2.1496	0.3618
$\hat{\alpha}_{NBKLE}$	3.5452	2.0011	1.8565	1.8070	10.6653	7.0699	4.8587	4.1347
$\hat{\alpha}_{M-NBKLE}$	3.1188	2.1364	1.4639	0.2391	5.0450	4.4454	2.0490	0.3477
	$\rho = 0.999$				$\rho = 0.999$			
$\hat{\alpha}_{MLE}$	71.1404	45.7179	38.5937	32.2868	245.5499	100.2213	87.3481	26.0989
$\hat{\alpha}_{NBRRE}$	43.6115	29.3162	12.7618	10.2636	128.8131	62.3445	50.5375	15.0965
$\hat{\alpha}_{M-NBRRE}$	39.3370	28.9090	2.8173	2.6831	71.0531	45.0481	4.4200	3.1227
$\hat{\alpha}_{NBKLE}$	26.5437	19.3226	10.4217	7.7887	86.0936	59.1129	47.7824	12.0998
$\hat{\alpha}_{M-NBKLE}$	23.4875	18.5239	2.3113	1.0858	70.7459	43.2366	3.6519	2.7898

Tables 1 and 2 present the estimated mean squared error (MSE) values for the proposed estimator and other estimators considered in this study. Based on the results obtained from the simulation, several important conclusions have been drawn:

i. Impact of Multicollinearity: To investigate the effect of multicollinearity on the estimates of regression parameters, we examined the correlation coefficients $\rho = 0.8, 0.9, 0.99,$ and 0.999 . It was observed that as the correlation between explanatory variables increases, the MSE of both classical estimators and robust hybrid estimators also increases (Figure 1). The MSE values for robust hybrid estimators, denoted with $\hat{\alpha}_{M-NBRRE}$ and $\hat{\alpha}_{M-NBKLE}$, show a reduction in magnitude compared to their non-hybrid counterparts.

ii. Influence of Sample Size: Comparing the performance of estimators for different sample sizes ($n = 30, 50, 100,$ and 200), it becomes evident that the MSE decreases as the sample size increases. Detailed graphical representations can be found in Figure 2. For instance, with $\rho = 0.8$, the MSE decreases from $n = 30$ to $n = 200$ for all estimators. Similarly, for other correlation coefficients, there is a consistent pattern of decreasing MSE values as the sample size grows.

iii. Impact of Explanatory Variables: The total number of explanatory variables has a considerable influence on the MSE values of estimators. The MSE values tend to be higher for all estimators when there are more explanatory variables.

iv. Impact of Outliers: The percentage of outliers in the model increased the estimated MSE values. In Tables 1 and 2 ($p^* = 3$ and $p^* = 5$), we observe that as the percentage of outliers increases from 1% to 8%, the MSE values for all estimators also increase significantly. The presence of outliers introduces more variability and influences the accuracy of the estimators, leading to higher MSE values. Overall, the MLE and NBRRE estimators typically exhibit higher MSE values compared to the robust hybrid estimators (M-NBRRE and M-NBKLE) across different sample sizes and correlation coefficients. Among the robust hybrid estimators, M-NBKLE generally outperforms M-NBRRE in terms of lower MSE values.

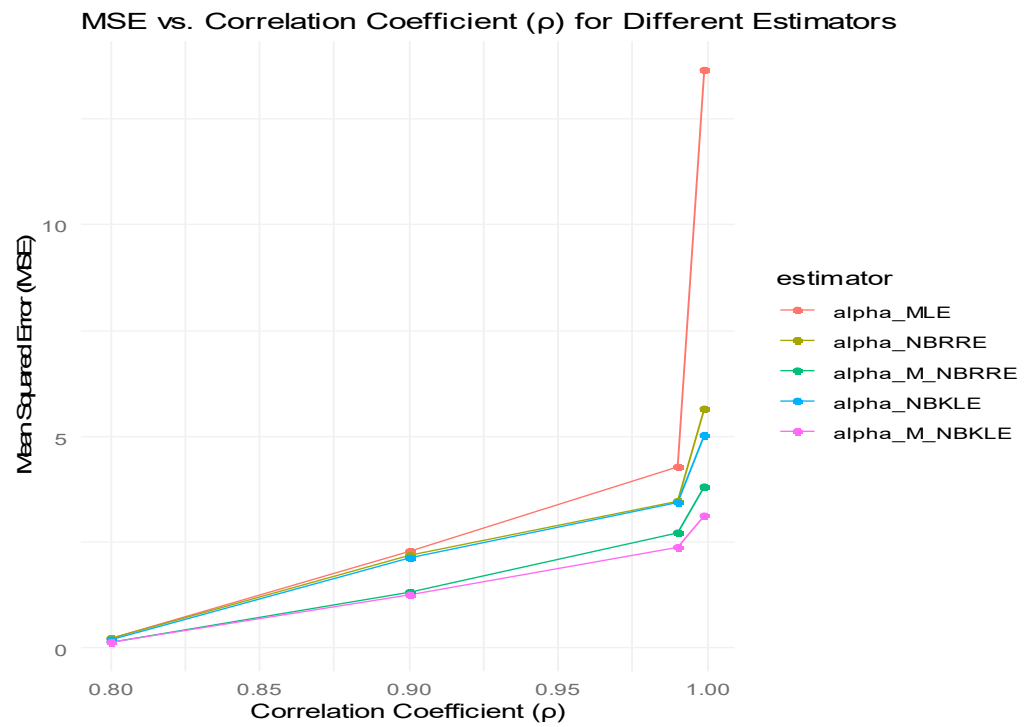


Figure 1. Graph of MSE against the level of correlation for each estimator when the outlier is 1%.

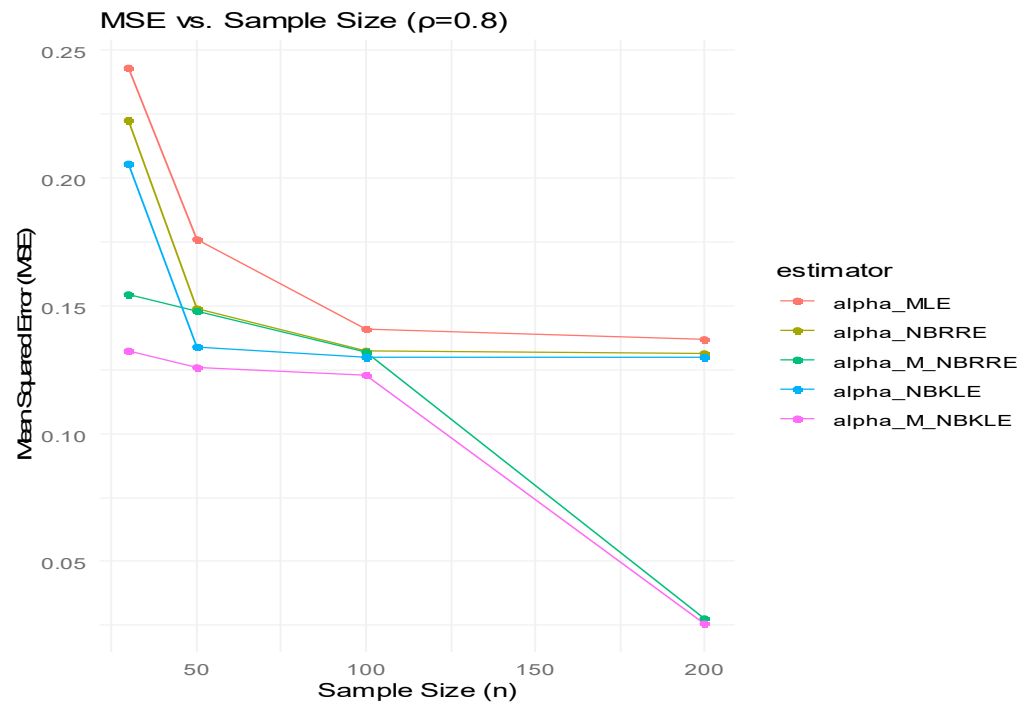


Figure 2. Graph of MSE against the sample size for the estimators corresponding to $\rho = 0.8$ when the outlier is 1%.

4. Application

In this application, we conducted a comprehensive analysis of the nuts dataset, as documented in Hilbe [45] and recently adopted by Algamal [46]. The dataset is available in the library COUNT. The dataset comprises fifty-two (52) observations and seven (7) explanatory variables, focusing on the behavior of squirrels and various forest characteristics across different plots within Scotland’s Abernathy Forest. Specifically, the response variable

represents the number of cones stripped by squirrels. The explanatory variables include the number of trees per plot (x_1), number of DBH (diameter at breast height) per plot (x_2), mean tree height per plot (x_3), canopy closure (as a percentage) (x_4), standardized number of trees per plot (x_5), standardized mean tree height per plot (x_6), and standardized canopy closure (as a percentage) (x_7).

Due to challenges encountered during the modeling process, we excluded variable x_7 from the analysis. We proceeded to fit Poisson and NBR models to identify the most appropriate model for the dataset. See Table 3 for the outcome. Notably, the Akaike Information Criterion (AIC) and residual variance from Table 3 unequivocally indicate that the NBRM provides an excellent fit to the data. Moreover, the dispersion test revealed the presence of overdispersion in the model, rendering the Poisson regression model unsuitable for this dataset.

Table 3. Model adequacy check.

Metrics	Poisson	Negative Binomial
Residual Variance	661.2	661.2
AIC	873.1	397.9

Additionally, the estimated variance inflation factors are 4.474, 16.332, 39.61, and 40.89, indicating a clear presence of correlated regressors in the model. Consequently, the model suffers from both multicollinearity and overdispersion. In our further examination, we conducted outlier diagnostics. The residual plots revealed the existence of outliers in the model. Specifically, the standardized Pearson residual plot against the leverage identified cases 2, 11, and 25 as outliers (See Figure 3). To tackle these challenges, we proceeded with model fitting using both non-robust estimators (MLE, NBRRE, and NBKLE) and robust estimators (M-NBRRE and M-NBKLE). See Table 4 for the results, which will aid in understanding the impact of these estimators on our model and its interpretations.

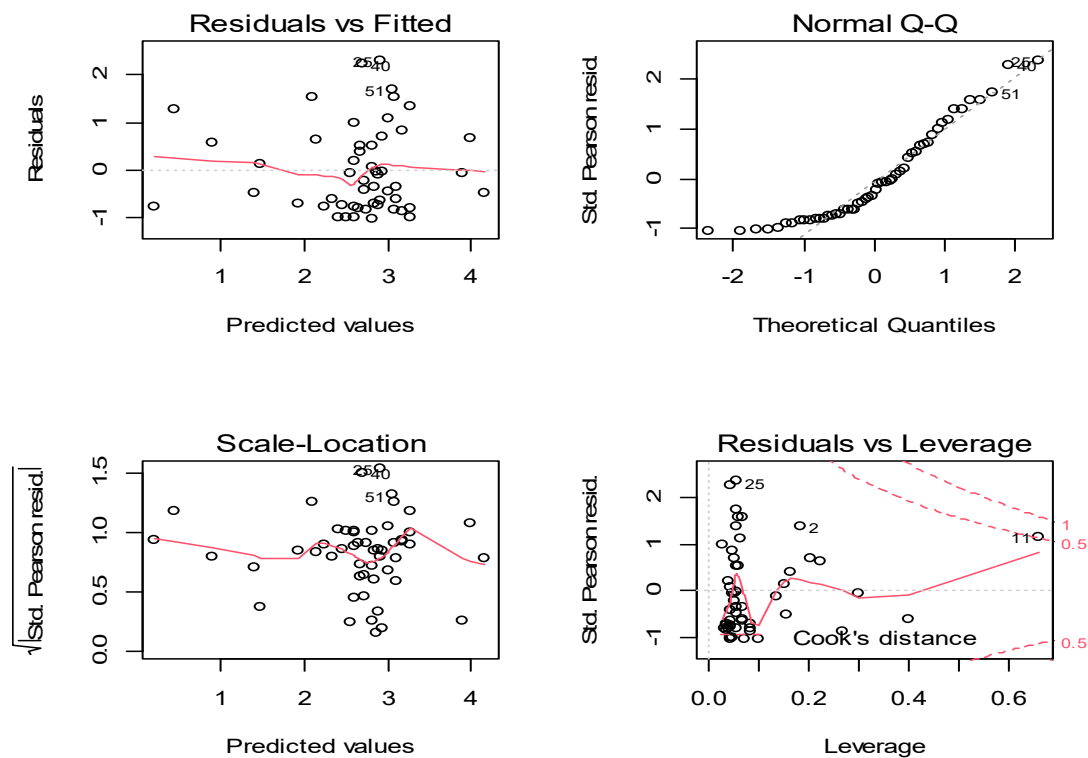


Figure 3. Diagnostic plot of squirrel dataset NB modelling.

Table 4. Estimated regression coefficients for the squirrel dataset.

Coef.	Estimators				
	$\hat{\alpha}_{MLE}$	$\hat{\alpha}_{NBRRE}$	$\hat{\alpha}_{M-NBRRE}$	$\hat{\alpha}_{NBKLE}$	$\hat{\alpha}_{M-NBKLE}$
x_1	0.0267 (0.0128)	0.0235 (0.0120)	−2.7247 (0.0122)	0.0203 (0.0115)	−2.7247 (0.0119)
x_2	1.9640 (1.8877)	1.0210 (1.3625)	0.0043 (1.5516)	0.0780 (0.9834)	0.0065 (1.2754)
x_3	0.0228 (0.0456)	0.0339 (0.0429)	0.0267 (0.0438)	0.0451 (0.0413)	0.0267 (0.0425)
x_4	0.0139 (0.0101)	0.0152 (0.0100)	−0.0350 (0.0100)	0.0165 (0.0100)	−0.0350 (0.0100)
x_5	0.4493 (0.1659)	0.4350 (0.1646)	0.0490 (0.1651)	0.4207 (0.1637)	0.0488 (0.1645)
SMSE	3.5934	1.8855	1.6560	3.5927	0.4692

Table 4 presents the estimated regression coefficients obtained from different estimators considered in this study. The estimators used comprise non-robust estimators (MLE, NBRRE, and NBKLE) and robust estimators (M-NBRRE and M-NBKLE). The values in parentheses next to each coefficient represent the standard errors of the estimates, providing insight into the precision of the coefficient estimates. Smaller standard errors indicate greater confidence in the estimate’s accuracy, while larger standard errors suggest more variability in the regression estimate.

Upon examination, we observe intriguing differences in the coefficient signs for x_1 and x_4 between the non-robust and robust estimators. This contrast highlights the sensitivity of non-robust estimators to outliers, potentially leading to biased coefficient estimates. In contrast, robust estimators, designed to withstand the impact of influential observations, demonstrate more consistent and reliable coefficient estimates.

To gauge the performance of these estimators, we rely on the SMSE. The results in Table 4 revealed that the SMSE values for the robust estimators, M-NBRRE and M-NBKLE, are substantially lower than those for the non-robust estimators (MLE, NBRRE, and NBKLE). This substantial difference suggests that the robust estimators exhibit superior predictive accuracy and resilience to the influence of outliers, making them more suitable for this dataset.

Particularly noteworthy is the robust hybrid KL estimator (M-NBKLE), which emerges as the most promising method among all the estimators. Not only does it provide accurate coefficient estimates, but it also achieves the lowest SMSE and has relatively small standard errors, indicating that its estimates are both reliable and precise. With its integration of robustness and superior predictive power, the M-NBKLE estimator dominates other approaches, providing researchers with a reliable tool to obtain accurate and stable coefficient estimates even in the presence of extreme observations.

5. Conclusions

NBR is widely used for modeling non-negative integer data, particularly when overdispersion is present. However, interpreting the model becomes more challenging when assumptions, such as the absence of multicollinearity among predictors, are violated. In practice, meeting these assumptions can be difficult, and multicollinearity can significantly reduce the efficiency of maximum likelihood estimators. Ridge and Kibria–Lukman estimators have been applied to address multicollinearity, but their performance often deteriorates in the presence of outliers. Robust estimation techniques have been proposed to handle the impact of outliers in NBR. Recent studies have shown that the combined effects of outliers and multicollinearity can significantly degrade model performance, a problem

well-documented in linear and Poisson regression models. This insight has motivated us to propose a new method to address these issues, specifically, in NBR.

This study addresses the challenges of multicollinearity and outliers in NBR by integrating Ridge and Kibria–Lukman estimators with robust estimators, resulting in new hybrid estimators termed M-NBRRE and M-NBKLE. We evaluated their performance theoretically and through simulations, with the results confirming the robustness and effectiveness of the hybrid estimators. Notably, M-NBKLE demonstrated superior performance, particularly by achieving a lower estimated mean squared error. A real-world application was also conducted, focusing on squirrel behavior and forest characteristics across plots in Scotland’s Abernathy Forest, further supporting the dominance of M-NBKLE.

Our findings contribute to the growing body of research on count regression models and offer valuable tools for researchers needing robust data analysis techniques across various fields. This work represents a significant step toward enhancing the accuracy and reliability of NBR, ensuring robust inferences, and enabling sound decision-making in practical data analysis. In future research, we plan to incorporate additional evaluation metrics, such as the Pearson chi-square statistic, which could provide further insights into model fit and performance, especially for count data. We also aim to refine these estimators by incorporating techniques like principal component regression or partial least squares to enhance estimator performance.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/math12182929/s1>. We have included the code used for the real-life application to facilitate replication of our results.

Author Contributions: Conceptualization, A.F.L., M.A. and R.A.F.; methodology, A.F.L. and R.A.F.; software, A.F.L.; formal analysis, A.F.L.; resources, All authors; writing—original draft preparation, All authors; writing—review and editing, All authors; visualization, R.A.F.; supervision, O.A.; funding acquisition, O.A., J.A. and A.A.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data will be made available upon request from the corresponding author.

Acknowledgments: We want to express our sincere gratitude to the Editorial team and reviewers for their meticulous review and invaluable feedback, which have greatly enhanced the quality of our manuscript. Mohammad Arashi’s work was supported in part by the Iran National Science Foundation (INSF) under grant No. 4015320.

Conflicts of Interest: There are no conflicts of interest to declare in this study.

References

1. Aeberhard, W.H.; Cantoni, E.; Heritier, S. Robust inference in the negative binomial regression model with an application to falls data. *Biometrics* **2014**, *70*, 920–931. [[CrossRef](#)]
2. McCullagh, P.; Nelder, J.A. *Generalized Linear Models*, 2nd ed.; Chapman & Hall: London, UK, 1989.
3. Algamil, Z.Y.; Abonazel, M.R.; Awwad, F.A.; Eldin, E.T. Modified jackknife ridge estimator for the Conway-Maxwell-Poisson model. *Sci. Afr.* **2023**, *19*, e01543. [[CrossRef](#)]
4. Månsson, K. On ridge estimators for the negative binomial regression model. *Econ. Model.* **2012**, *29*, 178–184. [[CrossRef](#)]
5. Mansson, K. Developing a Liu estimator for the negative binomial regression model: Method and application. *J. Stat. Comput. Simul.* **2013**, *83*, 1773–1780. [[CrossRef](#)]
6. Alobaidi, N.N.; Shamany, R.E.; Algamil, Z.Y. A new ridge estimator for the negative binomial regression model. *Thail. Stat.* **2021**, *19*, 116–125.
7. Jabur, D.; Rashad, N.; Algamil, Z. Jackknifed Liu-type estimator in the negative binomial regression model. *Int. J. Nonlinear Anal. Appl.* **2022**, *13*, 2675–2684.
8. Abonazel, M.R.; El-sayed, S.M.; Saber, O.M. Performance of robust count regression estimators in the case of overdispersion, zero inflated, and outliers: Simulation study and application to German health data. *Commun. Math. Biol. Neurosci.* **2021**, *2021*, 55.
9. Tüzen, F.; Erbaş, S.; Olmuş, H. A simulation study for count data models under varying degrees of outliers and zeros. *Commun. Stat.-Simul. Comput.* **2020**, *49*, 1078–1088. [[CrossRef](#)]

10. Medina, M.A.; Ronchetti, E. Robust statistics: A selective overview and new directions. *WIREs Comput. Stat.* **2015**, *7*, 372–393. [[CrossRef](#)]
11. Lukman, A.F.; Arashi, M.; Prokaj, V. Robust biased estimators for Poisson regression model: Simulation and applications. *Concurr. Comput. Pract. Exp.* **2023**, *35*, e7594. [[CrossRef](#)]
12. Lukman, A.F.; Farghali, R.A.; Kibria, B.M.G.; Oluyemi, O.A. Robust-stein estimator for overcoming outliers and multicollinearity. *Sci. Rep.* **2023**, *13*, 9066. [[CrossRef](#)] [[PubMed](#)]
13. Roozbeh, M.; Babaie-Kafaki, S.; Aminifard, Z. Two penalized mixed-integer nonlinear programming approaches to tackle multicollinearity and outliers effects in linear regression models. *J. Ind. Manag. Optim.* **2021**, *17*, 3475–3491. [[CrossRef](#)]
14. Roozbeh, M.; Babaie-Kafaki, S.; Maanavi, M. A heuristic algorithm to combat outliers and multicollinearity in regression model analysis. *Iran. J. Numer. Anal. Optim.* **2022**, *12*, 173–186.
15. Hoerl, A.E.; Kennard, R.W. Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics* **1970**, *12*, 55–67. [[CrossRef](#)]
16. Liu, K. A new class of biased estimate in linear regression. *Commun. Stat.* **1993**, *22*, 393–402.
17. Liu, K. Using Liu-type estimator to combat collinearity. *Commun. Stat.-Theory Methods* **2003**, *32*, 1009–2003. [[CrossRef](#)]
18. Kibria, B.M.G.; Lukman, A.F. A New Ridge-Type Estimator for the Linear Regression Model: Simulations and Applications. *Scientifica* **2020**, *2020*, 9758378. [[CrossRef](#)]
19. Asar, Y.; Genc, A. New shrinkage parameters for the liu-type logistic estimators. *Commun. Stat.-Simul. Comput.* **2016**, *45*, 1094–1103. [[CrossRef](#)]
20. Asar, Y.; Genç, A. Two-parameter ridge estimator in the binary logistic regression. *Commun. Stat.-Simul. Comput.* **2017**, *46*, 7088–7099. [[CrossRef](#)]
21. Kibria, B.G.; Månsson, K.; Shukur, G. Some ridge regression estimators for the zero-inflated Poisson model. *J. Appl. Stat.* **2013**, *40*, 721–735. [[CrossRef](#)]
22. Kibria, B.G.; Månsson, K.; Shukur, G. *A Ridge Regression Estimator for the Zero-Inflated Poisson Model*; Royal Institute of Technology, CESIS-Centre of Excellence for Science and Innovation Studies: Stockholm, Sweden, 2011.
23. Al-Taweel, Y.; Algamal, Z. Almost unbiased ridge estimator in the zero-inflated Poisson regression model. *TWMS J. Appl. Eng. Math.* **2022**, *12*, 235–246.
24. Månsson, K.; Shukur, G. A Poisson ridge regression estimator. *Econ. Model.* **2011**, *28*, 1475–1481. [[CrossRef](#)]
25. Asar, Y.; Genç, A. A new two-parameter estimator for the Poisson regression model. *Iran. J. Sci. Technol. Trans. Sci.* **2018**, *42*, 793–803. [[CrossRef](#)]
26. Lukman, A.F.; Aladeitan, B.; Ayinde, K.; Abonazel, M.R. Modified ridge-type for the Poisson Regression Model: Simulation and Application. *J. Appl. Stat.* **2021**, *49*, 2124–2136. [[CrossRef](#)]
27. Lukman, A.F.; Adewuyi, E.; Månsson, K.; Kibria, B.M.G. A new estimator for the multicollinear poisson regression model: Simulation and application. *Sci. Rep.* **2021**, *11*, 3732. [[CrossRef](#)]
28. Huang, J.; Yang, H. A two-parameter estimator in the negative binomial regression model. *J. Stat. Comput. Simul.* **2014**, *84*, 124–134. [[CrossRef](#)]
29. Çetinkaya, M.K.; Kaçiranlar, S. Improved two-parameter estimators for the negative binomial and Poisson regression models. *J. Stat. Comput. Simul.* **2019**, *89*, 2645–2660. [[CrossRef](#)]
30. Abonazel, M.R.; Saber, A.A.; Awwad, F.A. Kibria–Lukman estimator for the Conway–Maxwell Poisson regression model: Simulation and applications. *Sci. Afr.* **2023**, *19*, e01553. [[CrossRef](#)]
31. Barnett, V.; Lewis, T. *Outliers in Statistical Data*; Wiley: New York, NY, USA, 1994.
32. Chatterjee, S.; Hadi, A.S. Influential observations, high leverage points, and outliers in linear regression. *Stat. Sci.* **1986**, *1*, 379–416.
33. Rousseeuw, P.J.; Leroy, A.M. *Robust Regression and Outlier Detection*; Series in Applied Probability and Statistics; Wiley Interscience: New York, NY, USA, 1987; 329p. [[CrossRef](#)]
34. Wasim, D.; Suhail, M.; Albalawi, O.; Shabbir, M. Weighted penalized m-estimators in robust ridge regression: An application to gasoline consumption data. *J. Stat. Comput. Simul.* **2024**, *1*–30. [[CrossRef](#)]
35. Huber, P.J. Robust Regression: Asymptotics, Conjectures and Monte Carlo. *Ann. Stat.* **1973**, *1*, 799–821. [[CrossRef](#)]
36. Silvapulle, M. Robust ridge regression based on an M-estimator. *Aust. J. Stat.* **1991**, *33*, 319–333. [[CrossRef](#)]
37. Ertas, H. A modified ridge M-estimator for linear regression model with multicollinearity and outliers. *Commun. Stat.-Simul. Comput.* **2018**, *47*, 1240–1250. [[CrossRef](#)]
38. Dawoud, I.; Abonazel, M. Robust Dawoud–Kibria estimator for handling multicollinearity and outliers in the linear regression model. *J. Stat. Comput. Simul.* **2021**, *91*, 3678–3692. [[CrossRef](#)]
39. Abonazel, M.; Dawoud, I. Developing robust ridge estimators for Poisson regression model. *Concurr. Comput. Pract. Exp.* **2022**, *34*, e6979. [[CrossRef](#)]
40. Arum, K.; Ugwuowo, F.; Oranye, H.; Alakija, T.; Ugah, T.; Asogwa, O. Combating outliers and multicollinearity in linear regression model using robust Kibria-Lukman mixed with principal component estimator, simulation and computation. *Sci. Afr.* **2023**, *19*, e01566. [[CrossRef](#)]
41. Majid, A.; Ahmed, S.; Aslam, M.; Kashif, M. A robust Kibria–Lukman estimator for linear regression model to combat multicollinearity and outliers. *Concurr. Comput. Pract. Exp.* **2023**, *35*, e7533. [[CrossRef](#)]
42. Kibria, B.M.G. Performance of some new ridge regression estimators. *Commun. Stat.-Simul. Comput.* **2003**, *32*, 419–435. [[CrossRef](#)]

43. Dawoud, I.; Lukman, A.F.; Haadi, A. A new biased regression estimator: Theory, simulation and application. *Sci. Afr.* **2022**, *15*, e01100. [[CrossRef](#)]
44. Kibria, B.M.G. More than hundred (100) estimators for estimating the shrinkage parameter in a linear and generalized linear ridge regression models. *J. Econom. Stat.* **2022**, *2*, 233–252.
45. Hilbe, J.M. *Negative Binomial Regression*, 2nd ed.; Cambridge University Press: Cambridge, MA, USA, 2011.
46. Algarni, Z. Variable Selection in Count Data Regression Model based on Firefly Algorithm. *Stat. Optim. Inf. Comput.* **2019**, *7*, 520–529. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.