Electronic Journal of Applied Statistical Analysis
EJASA, Electron. J. App. Stat. Anal.
http://siba-ese.unisalento.it/index.php/ejasa/index
e-ISSN: 2070-5948
DOI: 10.1285/i20705948v17n2p436

Kibria-Lukman Hybrid Estimator for the Conway–Maxwell–Poisson Regression Model

By Alrweili

15 October 2024

Electronic Journal of Applied Statistical Analysis
Vol. 17, Issue 02, October 2024, 436-471
DOI: 10.1285/i20705948v17n2p436

Kibria-Lukman Hybrid Estimator for the Conway–Maxwell–Poisson Regression Model

Hleil Alrweili *[a]

[a]*Department of Mathematics, College of Science, Northern Border University, Arar, Saudi Arabia; Hleil.Alrweili@nbu.edu.sa*

15 October 2024

The Conway-Maxwell-Poisson regression (CMPR) model provides a flexible framework for analyzing count data in cases of over- and under-dispersion. Estimating the parameter in CMPR typically relies on the maximum likelihood estimator (MLE), which can be challenging, mainly when multicollinearity exists. In such cases, many estimators offer alternatives to MLE, but often with a more considerable bias. This paper introduces a new hybrid estimator, combining the modified ridge-type estimator's robustness with the Kibria-Lukman estimator's efficiency, named the Kibria-Lukman hybrid estimator (KLHE). We propose that KLHE address multicollinearity in CMPR, demonstrating its performance through Monte Carlo simulations. The effectiveness of KLHE is highlighted by its ability to handle multicollinearity, resulting in improved estimation accuracy compared to other estimators. We illustrate the practical application of KLHE using a real dataset, demonstrating its potential to enhance parameter estimation in CMPR models, particularly in settings with prevalent multicollinearity. KLHE is a valuable addition to the statistical toolkit, providing researchers with a robust and efficient means to address multicollinearity in CMPR modeling.

keywords: Conway-Maxwell-Poisson regression; multicollinearity; ridge estimator; Liu estimator; Kibria-Lukman estimator; Kibria-Lukman hybrid estimator; modified ridge-type estimator; CMPMLE; MSE

*Corresponding author: Hleil.Alrweili@nbu.edu.sa

ISSN: 2070-5948
http://siba-ese.unisalento.it/index.php/ejasa/index

1 Introduction

Count data regression models are essential in various research fields, including the Poisson, negative binomial (NB), and Conway-Maxwell-Poisson (CMPR) models. The Poisson model is widely used but assumes equal mean and variance, limiting its applicability to over- and under-dispersed data. The NB regression model is more flexible than the Poisson model for over-dispersed data. However, the CMPR model offers even greater flexibility, accommodating both over- and under-dispersion Sellers and Shmueli (2010); Shmueli et al. (2005).

The Poisson regression model's single parameter restricts its ability to address dispersion issues, leading to the development of alternative models. For example, the NB and bell regression models are adequate for handling over-dispersed data but inadequate for under-dispersion. The Conway-Maxwell-Poisson distribution, introduced by Conway and Maxwell (1962), is suitable for real count data demonstrating both over- and under-dispersion. As a result, the CMPR model, proposed by Sellers and Shmueli (2010), offers a comprehensive solution for count data modeling by addressing both types of dispersion.

Multicollinearity, which occurs when the explanatory variables in a regression model are highly correlated, poses a challenge for accurate parameter estimation. In the context of the CMPR model, multicollinearity can lead to inefficient estimates by the maximum likelihood estimator (MLE). A high correlation among explanatory variables in the CMPR model indicates the presence of multicollinearity. Consequently, in such cases, the estimates' standard error (SE) increases significantly, making the MLE less effective in providing accurate estimates. To address this issue, researchers have proposed various biased estimation methods. To address multicollinearity, the ridge regression estimation method was first introduced by Hoerl and Kennard (1970), which offers more accurate estimates for model parameters. Additionally, other notable biased estimators have been developed, including the Liu estimator (Kejian, 1993), the Liu-type (Liu, 2003), the two-parameter estimator (Asar and Genç, 2018), the modified ridge-type estimator (Lukman et al., 2019), the Kibria-Lukman estimator (Kibria, B. M. G. and Lukman et al., 2020), and the Kibria-Lukman hybrid estimator (Shewa and Ugwuowo, 2023). Numerous researchers have proposed various biased estimation approaches for the count and other response models to manage multicollinearity. Månsson and Shukur (2011) introduced the ridge estimator for the Poisson regression model. Månsson (2012) developed the ridge estimator for the NB regression model. Amin et al. (2023) defined the ridge estimator for the Bell regression model. Sami et al. (2022b) described the ridge estimator for the CMPR model. Månsson et al. (2012) extended the Liu estimator for the Poisson regression model. Akram et al. (2022) explained the ridge estimator for the CMPR model. Şiray et al. (2015) defined the Liu estimator for the logistic regression model. Tanış and Asar (2024) introduced a Liu-type estimator for the CMPR model. Akay and Ertan (2022) improved a Liu-type estimator for Poisson regression models. Zandi et al. (2022) extended Liu-type shrinkage strategies in zero-inflated NB models. Lukman et al. (2022) expressed modified ridge-type for the Poisson regression model. Akram et al. (2023b) developed a modified ridge-type estimator for the zero-inflated NB regression model. Algamal et al. (2023) defined a modified jackknife ridge estimator for the CMPR model. Akram et al. (2023a) improved the Kibria-Lukman estimator for the zero-inflated NB regression model. Lukman et al. (2021) extended the Kibria-Lukman estimator for the Poisson regression model. Sami et al. (2022a) developed a modified one-parameter Liu estimator for the CMPR model. Abonazel et al. (2023b) extended the Kibria–Lukman estimator for the CMPR model. Sami et al. (2023b) defined Two parameter estimators for the CMPR model. Sami et al. (2023a) introduced an almost unbiased ridge estimator for the CMPR model.

Our review of existing literature reveals no Kibria-Lukman hybrid estimator (KLHE) in the context of the Conway-Maxwell-Poisson Regression (CMPR) model for addressing multicollinearity. So, we introduce a novel estimator, the Conway-Maxwell-Poisson Kibria-Lukman hybrid estimator (CMPKLHE), which is based on two parameters. This novel estimator is designed to reduce the impact of multicollinearity among the explanatory variables in the CMPR model. To evaluate the effectiveness of the proposed estimator, we employ several methods. Firstly, we conduct a theoretical comparison to assess its performance relative to existing estimators, then carry out a Monte Carlo simulation study under various scenarios to analyze its behavior in controlled environments and apply the estimator to real-world data to evaluate its practical performance.

The article organizes its content into sections. Section 2 delves into the methodology, reviewing the CMPR model, explaining biased estimators, and conducting a theoretical comparison. Section 3 outlines the Monte Carlo simulation design and presents its results. Section 4 explores the practical application of the estimator. Finally, Section 5 concludes the article with insightful remarks.

2 Methodology

This section delves into the background of the CMPR model, providing essential context for understanding the proposed estimator. We then present an estimation procedure for the CMPKLHE, followed by a detailed derivation of its bias and mean square error (MSE). Finally, we outline the procedure for comparing the proposed estimator theoretically with existing methods, providing a framework for evaluating its relative performance.

2.1 The Conway–Maxwell–Poisson Distribution

The Conway–Maxwell–Poisson (CMP) distribution offers remarkable flexibility in handling the dispersion often observed in count data. This flexibility comes from its additional parameter, φ, which allows for both overdispersion ($\varphi < 1$) and underdispersion ($\varphi > 1$). The CMP distribution is characterized by two parameters, δ and φ.

Interestingly, the CMP distribution encompasses several well-known discrete distributions as special cases:

- When $\varphi = 1$, the CMP distribution converges to the Poisson distribution, a widely used model for count data.

- When $\varphi = 0$ and $\delta < 1$, the CMP distribution transforms into the geometric distribution.
- As φ approaches infinity, the CMP distribution with probability $\delta/(1+\delta)$ converges to the Bernoulli distribution.

Formally, if y is a random variable following a CMP(δ, φ) distribution, its probability mass function is given by:

$$P(Y = y; \delta, \varphi) = \frac{1}{Z(\delta, \varphi)} \frac{\delta^y}{(y!)^\varphi}, \qquad y = 0, 1, 2, ..., \infty, \tag{1}$$

where φ ($\varphi \geq 0$) is the dispersion parameter Sellers and Shmueli (2010), δ ($\delta > 0$) is the mean parameter, and the CMP distribution's normalizing constant, Z(δ, φ), is defined as an infinite series:

$$Z(\delta, \varphi) = \sum_{n=0}^{\infty} \frac{\delta^n}{(n!)^\varphi}. \tag{2}$$

This infinite series lacks a closed-form solution, preventing the expression of the mean and its derivatives in a simple, closed-form manner. Generalized linear models (GLMs) are not directly applicable to the CMP distribution due to this limitation. Consequently, parameter estimation and inference for the CMP distribution typically rely on either numerical gradient-based methods (NGBM) or Markov chain Monte Carlo (MCMC) methods (Chatla and Shmueli, 2018).

The mean and variance of the CMP distribution can be approximated using an asymptotic expression for Z in Eq.(2). These approximations are as follows:

$$E(Y) \approx \delta^{\frac{1}{\varphi}} + \frac{1}{2\varphi} - \frac{1}{2}, \tag{3}$$

and

$$V(Y) \approx \frac{1}{\varphi}\delta^{\frac{1}{\varphi}}. \tag{4}$$

To obtain the reparametrized CMP function, we set $\mu = \delta^{\frac{1}{\varphi}}$ in Eq.(1) (Guikema and Goffelt, 2008). This leads to the following new formulation:

$$P(Y = y; \mu, \varphi) = \frac{1}{S(\mu, \varphi)} \left(\frac{\mu^y}{y!}\right)^\varphi, \qquad y = 0, 1, 2, ..., \infty, \tag{5}$$

where $S(\mu, \varphi) = \sum_{n=0}^{\infty} \left(\frac{\mu^n}{n!}\right)^\varphi$.

The mean and variance of Eq.(5) are derived as $E(Y) \approx \mu + \frac{1}{2\varphi} - \frac{1}{2}$ and $V(Y) = \frac{\mu}{\varphi}$ (Shmueli et al., 2005).

2.2 The Conway–Maxwell–Poisson Regression Model

The new formulation enables the establishment of a GLM framework. This framework facilitates the interpretation of the results through the use of link functions (Francis et al., 2012). The log-likelihood of Eq.(5) is given by:

$$l(y_i; \mu_i, \varphi) = \varphi y_i \log(\mu_i) - \varphi \log(y!) - \log(S(\mu_i, \varphi)). \tag{6}$$

Considering the linear predictor under log link, $\eta_i = \log(\mu_i) = x_i^t \beta$, the log-likelihood function of Eq.(6) is given by:

$$l(y_i; \beta, \varphi) = \varphi \sum_{i=1}^{n} y_i \eta_i - \varphi \sum_{i=1}^{n} \log(y_i!) - \sum_{i=1}^{n} \log(S(\eta_i, \varphi)). \tag{7}$$

To estimate the unknown parameters using the MLE method, we take the derivative of Eq.(7) with respect to β and φ.

$$\frac{\partial l(y_i; \beta, \varphi)}{\partial \beta_j} = \sum_{i=1}^{n} \left(y_i \varphi - \frac{\partial}{\partial \eta_i} \log S(\eta_i, \varphi) \right) x_{ij}, \tag{8}$$

$$\frac{\partial l(y_i; \beta, \varphi)}{\partial \varphi} = \sum_{i=1}^{n} \left(- \log(y_i!) - \frac{\partial}{\partial \varphi} \log S(\eta_i, \varphi) \right). \tag{9}$$

To estimate β, φ needs to be fixed. For a detailed explanation of the information matrix, consult Shmueli et al. (2005). In the CMPR model framework, the mean and variance are modeled using different sets of covariates, defined respectively as follows (Francis et al., 2012):

$$\log(\mu_i) = \beta_0 + \sum_{j=1}^{p} \beta_j x_{ij} = x_i^t \beta. \tag{10}$$

$$\log(\varphi_i) = \nu_0 + \sum_{k=1}^{q} \nu_k z_{ik} = z_i^t \nu. \tag{11}$$

In these equations, x_i and z_i denote the covariates for the mean and variance link functions, with p and q terms respectively (Francis et al., 2012). We can use the iterative reweighted least squares (IRLS) estimation method to solve Eqs. (8) and (9). This allows us to obtain the Conway-Maxwell-Poisson maximum likelihood estimator (CMPMLE) for the β vector as follows:

$$\hat{\beta}_{\text{CMPMLE}} = (X^t \hat{U} X)^{-1} X^T \hat{U} \hat{c}, \tag{12}$$

where $c = \log(\hat{\mu}) + \frac{y-\hat{\mu}}{\hat{\mu}^2}$ is a vector of the adjusted response variable, and $\hat{U}$ is a matrix of weights, i.e., $\hat{U} = \text{diag}(\hat{u}_i)$, where

$$\hat{u}_i = \frac{\hat{\tau}_i}{\hat{\varphi}} + \frac{\hat{\varphi}^2 - 1}{24\hat{\varphi}^3}\hat{\tau}_i^{-1} + \frac{\hat{\varphi}^2 - 1}{12\hat{\varphi}^4}\hat{\tau}_i^{-2} + \frac{\hat{\varphi}^2 - 1}{6\hat{\varphi}^5}\hat{\tau}_i^{-3} \tag{13}$$

with $\hat{\tau}_i = \frac{\hat{\mu}_i}{\hat{\varphi}}$. Both $\hat{U}$ and $\hat{c}$ are evaluated using the Fisher scoring procedure.

The MSE of $\hat{\beta}_{\text{CMPMLE}}$ is given as follows:

$$\text{MSE}(\hat{\beta}_{\text{CMPMLE}}) = E(\hat{\beta}_{\text{CMPMLE}} - \beta)^t(\hat{\beta}_{\text{CMPMLE}} - \beta) = \hat{\varphi}\text{tr}(QS^{-1}Q^t) = \hat{\varphi}\sum_{j=1}^{p}\frac{1}{s_j}, \tag{14}$$

where $\text{tr}(\cdot)$ denoted the trace of the matrix, $S = \text{diag}(s_1, s_2, \ldots, s_p)$, s_j is the j-th eigenvalue of $Q(X^t\hat{U}X)^{-1}Q^t$, Q is the orthogonal matrix whose columns are the eigenvectors of $(X^t\hat{U}X)^{-1}$, and $\hat{\varphi}$ is the maximum likelihood estimate of φ.

2.3 Conway-Maxwell-Poisson Ridge Estimator

To address the issue of multicollinearity, Sami et al. (2022b) proposed the Conway-Maxwell-Poisson ridge regression estimator (CMPRRE) as an alternative to the CMPMLE. The CMPRRE is formulated as follows:

$$\hat{\beta}_{\text{CMPRRE}} = (X^t\hat{U}X + k_R I_p)^{-1}X^t\hat{U}X\hat{\beta}_{\text{CMPMLE}}, \quad k_R > 0, \tag{15}$$

where k_R is the ridge parameter, and I_p is the identity matrix of order $p \times p$. The bias vector, covariance matrix, and matrix mean squared error (MMSE) of the $\hat{\beta}_{\text{CMPRRE}}$ are given as follows:

$$\text{Bias}(\hat{\beta}_{\text{CMPRRE}}) = -k_R Q S_k^{-1}\alpha, \tag{16}$$

$$\text{Cov}(\hat{\beta}_{\text{CMPRRE}}) = \hat{\varphi}\ QS_k^{-1}SS_k^{-1}Q^t, \tag{17}$$

$$\text{MMSE}(\hat{\beta}_{\text{CMPRRE}}) = \hat{\varphi}\ QS_k^{-1}SS_k^{-1}Q^t + k_R^2 QS_k^{-1}\alpha\alpha^t S_k^{-1}Q^t. \tag{18}$$

Finally, the MSE of the CMPRRE is derived by taking the trace of Eq.(18), expressed as follows:

$$\text{MSE}(\hat{\beta}_{\text{CMPRRE}}) = \hat{\varphi}\sum_{j=1}^{p}\frac{s_j}{(s_j + k_R)^2} + k_R^2\sum_{j=1}^{p}\frac{\alpha_j^2}{(s_j + k_R)^2}, \tag{19}$$

where α_j is the jth element of $Q^t\beta_{\text{CMPMLE}}$, and the matrix $S_k = \text{diag}(s_1 + k, s_2 + k, \ldots, s_p + k)$.

2.4 Conway-Maxwell-Poisson Liu Estimator

The Liu estimator is proposed to mitigate multicollinearity, showing better performance than the MLE method (Kejian, 1993). Akram et al. (2022) introduced a Liu estimator for the CMPR model, known as the Conway-Maxwell-Poisson Liu estimator (CMPLE), defined as:

$$\hat{\beta}_{\text{CMPLE}} = \left(X^t\hat{U}X + I_p\right)^{-1}\left(X^t\hat{U}X + d_L I_p\right)\hat{\beta}_{\text{CMPMLE}}, \qquad 0 < d_L < 1, \tag{20}$$

where d_L Liu shrinkage parameter. The following expressions represent the bias vector, covariance matrix, and MMSE for $\hat{\beta}_{\text{CMPLE}}$:

$$\text{Bias}(\hat{\beta}_{\text{CMPLE}}) = Q\ (S + I_p)^{-1}\alpha(d_L - 1), \tag{21}$$

$$\text{Cov}(\hat{\beta}_{\text{CMPLE}}) = \hat{\varphi}\ Q\ (S + I_p)^{-1}(S + d_L I_p)S^{-1}(S + d_L I_p)(S + I_p)^{-1}Q^t, \tag{22}$$

$$\begin{aligned}\text{MMSE}(\hat{\beta}_{\text{CMPLE}}) &= \hat{\varphi}\ Q\ (S + I_p)^{-1}(S + d_L I_p)S^{-1}(S + d_L I_p)(S + I_p)^{-1}Q^t \\ &\quad + (d_L - 1)^2 Q(S + I_p)^{-1}\alpha\alpha^t(S + I_p)^{-1}Q^t.\end{aligned} \tag{23}$$

Ultimately, the MSE of the CMPLE is determined by taking the trace of Eq. (23), which is expressed as follows:

$$\text{MSE}(\hat{\beta}_{\text{CMPLE}}) = \hat{\varphi}\sum_{j=1}^{p}\frac{(s_j + d_L)^2}{s_j(s_j + 1)^2} + (d_L - 1)^2\sum_{j=1}^{p}\frac{\alpha_j^2}{(s_j + 1)^2}. \tag{24}$$

2.5 Conway-Maxwell-Poisson Liu-Type Estimator

Building upon the seminal work of Liu (2003), Tanış and Asar (2024) proposed a novel Liu-type estimator specifically designed to address multicollinearity issues within the CMPR model. This innovative estimator extends the research conducted by Akay and Ertan (2022), Zandi et al. (2022), stat Alrwali (2024), and Ertan and Akay (2023). The formula for the proposed Conway-Maxwell-Poisson Liu-type estimator (CMPLTE) is presented below:

$$\hat{\beta}_{\text{CMPLTE}} = \left(X^t\hat{U}X + k_{LT}I_p\right)^{-1}\left(X^t\hat{U}X - d_{LT}I_p\right)\hat{\beta}_{\text{CMPMLE}}, \tag{25}$$

where $k_{LT}(k_{LT} > 0)$, and$d_{LT}(-\infty < d_{LT} < +\infty)$ are the Liu-type parameters. The bias vector, covariance matrix, and MMSE for $\hat{\beta}_{\text{CMPLTE}}$ are as follows:

$$\text{Bias}(\hat{\beta}_{\text{CMPLTE}}) = -(k_{LT} + d_{LT})QS_k^{-1}\alpha, \tag{26}$$

$$\text{Cov}(\hat{\beta}_{\text{CMPLTE}}) = \hat{\varphi}\ Q\ S_k^{-1}S_d S^{-1}S_k^{-1}S_d Q^t, \tag{27}$$

$$\text{MMSE}(\hat{\beta}_{\text{CMPLTE}}) = \hat{\varphi}\ Q\ S_k^{-1}S_d S^{-1}S_k^{-1}S_d Q^t + (k_{LT} + d_{LT})^2 QS_k^{-1}\alpha\alpha^t S_k^{-1}Q^t. \tag{28}$$

where the matrix $S_d = \text{diag}(s_1 + d, s_2 + d, \ldots, s_p + d)$. The MSE of the CMPLTE is determined by taking the trace of Eq. (28), which is given by:

$$\text{MSE}(\hat{\beta}_{\text{CMPLE}}) = \hat{\varphi} \sum_{j=1}^{p} \frac{(s_j + d_L)^2}{s_j (s_j + k_{LT})^2} + (d_{LT} + k_{LT})^2 \sum_{j=1}^{p} \frac{\alpha_j^2}{(s_j + k_{LT})^2}. \tag{29}$$

2.6 Conway-Maxwell-Poisson Kibria–Lukman Estimator

Abonazel et al. (2023b) introduced the Kibria-Lukman estimator for the CMPR model (CMPRKLE). This innovative estimator provides a powerful alternative to the traditional Liu and Ridge estimators in tackling the challenges of multicollinearity. The formula for the CMPRKLE is as follows:

$$\hat{\beta}_{\text{CMPKLE}} = (X^t \hat{U} X + k_{KL} I_p)^{-1} (X^t \hat{U} X - k_{KL} I_p) \hat{\beta}_{\text{CMPMLE}}, \quad k_{KL} > 0 \tag{30}$$

where k_{KL} is the Kibria-Lukman parameter. The bias vector, covariance matrix, and MMSE of the $\hat{\beta}_{\text{CMPKLE}}$ are given as follows:

$$\text{Bias}(\hat{\beta}_{\text{CMPKLE}}) = 2k_{KL} Q S_k^{-1} \alpha, \tag{31}$$

$$\text{Cov}(\hat{\beta}_{\text{CMPKLE}}) = \hat{\varphi}\, Q S_{-k} S_k^{-1} S^{-1} S_k^{-1} S_{-k} Q^t, \tag{32}$$

$$\text{MMSE}(\hat{\beta}_{\text{CMPKLE}}) = \hat{\varphi}\, Q S_{-k} S_k^{-1} S^{-1} S_k^{-1} S_{-k} Q^t + (2k_{KL})^2 Q S_k^{-1} \alpha \alpha^t S_k^{-1} Q^t. \tag{33}$$

where the matrix $S_{-k} = \text{diag}(s_1 - k, s_2 - k, \ldots, s_p - k)$. The MSE of the CMPKLE is derived by taking the trace of Eq.(33), expressed as follows:

$$\text{MSE}(\hat{\beta}_{\text{CMPKLE}}) = \hat{\varphi} \sum_{j=1}^{p} \frac{(s_j - k_{KL})^2}{s_j (s_j + k_{KL})^2} + (2k_{KL})^2 \sum_{j=1}^{p} \frac{\alpha_j^2}{(s_j + k_{KL})^2}, \tag{34}$$

2.7 Conway-Maxwell-Poisson Modified Ridge-type Estimator

Following Lukman et al. (2019) and extending the research of Lukman et al. (2020, 2022) and Akram et al. (2023b), we introduce a modified ridge-type estimator for the CMPR model (CMPMRTE). The formula for the CMPMRTE is as follows:

$$\hat{\beta}_{\text{CMPMRTE}} = \left(X^t \hat{U} X + k_M (1 + d_M) \right)^{-1} X^t \hat{U} X \hat{\beta}_{\text{CMPMLE}}, \tag{35}$$

where $k_M (k_M > 0)$, and$d_m (0 < d_M < 1)$ are the modified ridge-type parameters. The bias vector, covariance matrix, and MMSE for $\hat{\beta}_{\text{CMPMRTE}}$ are as follows:

$$\text{Bias}(\hat{\beta}_{\text{CMPMRTE}}) = -k_M (1 + d_M) Q S_k^{-1} \alpha, \tag{36}$$

$$\text{Cov}(\hat{\beta}_{\text{CMPMRTE}}) = \hat{\varphi}\, Q S_M^{-1} S S_M^{-1} Q^t, \tag{37}$$

$$\text{MMSE}(\hat{\beta}_{\text{CMPMRTE}}) = \hat{\varphi}\, Q S_M^{-1} S S_M^{-1} Q^t + [k_M (1 + d_M)]^2 Q S_M^{-1} \alpha \alpha^t S_M^{-1} Q^t. \tag{38}$$

where matrix $S_M = \text{diag}(s_1 + k(1+d), s_2 + k(1+d), \ldots, s_p + k(1+d))$. The MSE of the CMPLTE is determined by taking the trace of Eq. (38), which is given by:

$$\text{MSE}(\hat{\beta}_{\text{CMPMRTE}}) = \hat{\varphi} \sum_{j=1}^{p} \frac{s_j}{(s_j + k_M(1+d_M))^2} + [k_M(1+d_M)]^2 \sum_{j=1}^{p} \frac{\alpha_j^2}{(s_j + k_M(1+d_M))^2}. \quad (39)$$

2.8 Proposed Estimator

Following Shewa and Ugwuowo (2023), Alrweili (2024) presented a novel hybrid estimator for Poisson regression models. This innovative approach ingeniously combines the strengths of the Kibria-Lukman estimator and a modified ridge-type estimator (MRTE). The key idea is to replace the maximum likelihood estimator ($\hat{\beta}_{\text{MLE}}$) within the Kibria-Lukman estimator with the MRTE ($\hat{\beta}_{\text{MRTE}}$). This paper proposes a Kibria-Lukman hybrid estimator for the CMPR model. The formula for the Conway-Maxwell-Poisson Kibria-Lukman hybrid estimator (CMPKLHE) is given by:

$$\begin{aligned} \hat{\beta}_{\text{CMPKLHE}} &= (X^t\hat{U}X + k_* I_p)^{-1}(X^t\hat{U}X - k_* I_p)^{-1}\hat{\beta}_{\text{CMPMRTE}}, \\ &= (X^t\hat{U}X + k_* I_p)^{-1}(X^t\hat{U}X - k_* I_p)^{-1}(X^t\hat{U}X + k_*(1+d_*)I_p)^{-1}X^t\hat{U}X\hat{\beta}_{\text{CMPMLE}}, \end{aligned} \quad (40)$$

where $k_*(k_* > 0)$ and $d_*(0 < d_* < 1)$ are the Kibria Lukman hybrid parameters, and $\hat{\beta}_{\text{CMPKLHE}}$ returns to some estimator under the following conditions:

1. If $k_* = 0$, then $\hat{\beta}_{\text{CMPKLHE}} = \hat{\beta}_{\text{CMPMLE}}$.
2. If $d_* = 0$, then $\hat{\beta}_{\text{CMPKLHE}} = \hat{\beta}_{\text{CMPMKLE}}$.

The CMPKLHE estimator offers a new approach to address multicollinearity in CMPR models, potentially improving estimation accuracy in scenarios with correlated predictor variables. The bias vector and the covariance matrix of the CMPKLHE are expressed as

$$\begin{aligned} \text{Bias}(\hat{\beta}_{\text{CMPKLHE}}) &= E(\hat{\beta}_{\text{CMPKLHE}}) - \beta \\ &= -\left((3k_* + k_* d_*)S + k_*^2(1+d_*)\right) Q S_k^{-1} S_M^{-1} \alpha. \end{aligned} \quad (41)$$

$$\begin{aligned} \text{Cov}(\hat{\beta}_{\text{CMPKLHE}}) &= E\left(\left[\hat{\beta}_{\text{CMPKLHE}} - E(\hat{\beta}_{\text{CMPKLHE}})\right]\left[\hat{\beta}_{\text{CMPKLHE}} - E(\hat{\beta}_{\text{CMPKLHE}})\right]^t\right) \\ &= \hat{\varphi}\ Q S_k^{-1} S_{-k} S_M^{-1} S_M^{-1} S_{-k} S_k^{-1} Q^t. \end{aligned} \quad (42)$$

Thus, MMSE and MSE of the CMPKLHE are obtained using the covariance matrix and bias vector of CMPKLHE as follows:

$$\begin{aligned}\text{MMSE}(\hat{\beta}_{\text{CMPKLHE}}) =& \text{Cov}(\hat{\beta}_{\text{CMPKLHE}}) + \text{Bias}(\hat{\beta}_{\text{CMPKLHE}})\text{Bias}(\hat{\beta}_{\text{CMPKLHE}})^t \\ =& \hat{\varphi}\, QS_k^{-1}S_{-k}S_M^{-1}S_M^{-1}S_{-k}S_k^{-1}Q^t \\ &+ \left((3k_* + k_*d_*)S + k_*^2(1+d_*)\right)^2 QS_k^{-1}S_M^{-1}\alpha \\ &\times \alpha^t S_k^{-1}S_M^{-1}Q^t.\end{aligned} \tag{43}$$

$$\begin{aligned}\text{MSE}(\hat{\beta}_{\text{CMPKLHE}}) &= \text{tr}\left(\text{MMSE}(\hat{\beta}_{\text{CMPKLHE}})\right) \\ &= \hat{\varphi}\sum_{j=1}^{p}\frac{(s_j - k_*)^2}{(s_j+k_*)^2(s_j+k_*(1+d_*))^2} \\ &+ \sum_{j=1}^{p}\frac{k_*^2\left((3+d_*)s_j + k_*(1+d_*)\right)^2\alpha_j^2}{(s_j+k_*)^2(s_j+k_*(1+d_*))^2}.\end{aligned} \tag{44}$$

2.9 Estimating the Biasing Parameters

Drawing from the methodologies proposed by Sami et al. (2022b) and Abonazel et al. (2023b,a), Akram et al. (2022), Tanış and Asar (2024), Abonazel et al. (2023b), Lukman et al. (2019), Lukman et al. (2020, 2022), and Akram et al. (2023b), and Shewa and Ugwuowo (2023), we can determine the biasing parameters for various estimators in the Conway-Maxwell-Poisson regression model.

For the CMPRRE estimator, we use the following formula:

$$\hat{k}_R = \min\left(\frac{\hat{\varphi}}{\hat{\alpha}_j^2}\right). \tag{45}$$

For CMPLE, the estimator is given by:

$$\hat{d}_L = \min\left(\frac{\hat{\alpha}_j^2 - \hat{\varphi}}{\frac{\hat{\varphi}}{s_j} + \hat{\alpha}_j^2}\right). \tag{46}$$

The parameters for CMPLTE are defined as:

$$\hat{k}_{LT} = \hat{k}_R. \tag{47}$$

$$\hat{d}_{LT} = \frac{\sum_{j=1}^{p}\frac{\hat{\varphi}^{-1}-\hat{\alpha}_j^2}{(s_j+1)^2}}{\sum_{j=1}^{p}\frac{\frac{s_j}{\hat{\varphi}}+\hat{\alpha}_j^2}{(s_j+1)^2}}. \tag{48}$$

The CMPKLE parameter is obtained from CMPRRE $\hat{k}_{KL} = \hat{k}_R$.

For CMPMRTE, the estimators are given by:

$$\hat{k}_M = \min\left(\frac{\hat{\varphi}}{(1+\hat{d})\hat{\alpha}_j^2}\right). \tag{49}$$

$$\hat{d}_M = \min\left(\frac{\hat{\alpha}_j^2}{\hat{\alpha}_j^2 + \frac{\hat{\varphi}}{s_j}}\right). \tag{50}$$

Finally, for the proposed estimator CMPKLHE, we first use $\hat{d}_M$ as the value of $\hat{d}_*$. Then, we calculate $\hat{k}_*$ as follows:

$$\hat{k}_*^1 = \frac{1}{2}\min\left(\frac{d+1+\sqrt{2d^2+6d+4}s_j^2}{1+d}\right)^{\frac{1}{p}}. \tag{51}$$

$$\hat{k}_*^2 = \sqrt{\frac{p\hat{\varphi}}{\sum_{j=1}^{p}\hat{\alpha}_j^2}}. \tag{52}$$

$$\hat{k}_*^3 = \prod_{j=1}^{p}\left(\frac{d+1+\sqrt{2d^2+6d+4}\hat{\alpha}_j^2}{1+d}\right)^{\frac{1}{p}}. \tag{53}$$

$$\hat{k}_*^4 = \frac{1}{p}\max\left(\frac{d+1+\sqrt{2d^2+6d+4}\hat{\alpha}_j^2}{1+d}\right). \tag{54}$$

2.10 The superiority of the proposed estimator

To demonstrate the performance of the proposed estimator, we compare the MMSE and MSE of CMPKLHE with those of CMPMLE, CMPRRE, CMPLE, CMPLTE, CMPKLE, and CMPMRTE for the CMPR model, based on theoretical findings.

Lemma 1. *Given that D is a positive definite (p.d.) matrix, m is a positive constant, and δ is a vector of nonzero constants, the inequality $mD - \delta\delta^t > 0$ is satisfied iff $\delta D\delta^t < m$ (Farebrother, 1976).*

Theorem 1. *The estimator $\hat{\beta}_{CMPKLHE}$ outperforms $\hat{\beta}_{CMPMLE}$ iff MMSE$(\hat{\beta}_{CMPMLE})$ − MMSE$(\hat{\beta}_{CMPKLHE}) > 0$, where $k_* > 0$, $0 < d_* < 1$.*

Proof. The difference between MMSE$(\hat{\beta}_{\text{CMPMLE}})$ and MMSE$(\hat{\beta}_{\text{CMPKLHE}})$ is given by

$$\begin{aligned}\text{MMSE}(\hat{\beta}_{\text{CMPMLE}}) - \text{MMSE}(\hat{\beta}_{\text{CMPKLHE}}) =& \hat{\varphi}Q(S^{-1} - S_k^{-1}S_{-k}S_M^{-1}S_M^{-1}S_{-k}S_k^{-1})Q^t \\ &- \text{Bias}(\hat{\beta}_{\text{CMPKLHE}})\text{Bias}(\hat{\beta}_{\text{CMPKLHE}})^t.\end{aligned} \tag{55}$$

Eq. (55) can be expressed using the MSE as:

$$\begin{aligned}\mathrm{MSE}(\hat{\beta}_{\mathrm{MLE}}) - \mathrm{MSE}(\hat{\beta}_{\mathrm{CMPKLHE}}) =& \hat{\varphi} Q \,\mathrm{diag}\left(\frac{1}{s_j} - \frac{(s_j - k_*)^2}{(s_j + k_*)^2 (s_j + k_*(1 + d_*))^2}\right) Q^t \\ & - \mathrm{Bias}(\hat{\beta}_{\mathrm{CMPKLHE}})\,\mathrm{Bias}(\hat{\beta}_{\mathrm{CMPKLHE}})^t\end{aligned} \tag{56}$$

The matrix $(S^{-1} - S_k^{-1} S_{-k} S_M^{-1} S_M^{-1} S_{-k} S_k^{-1})$ is pd if $(s_j + k_*)^2 (s_j + k_*(1 + d_*))^2 - s_j (s_j - k_*)^2 > 0$, which is equivalent to $(s_j + k_*)^2 (s_j + k_*(1 + d_*))^2 > s_j (s_j - k_*)^2$ being non-negative. Therefore, if $k_* > 0$ and $0 < d_* < 1$, the proof is completed by Lemma 1. □

Theorem 2. *The estimator $\hat{\beta}_{CMPKLHE}$ outperforms $\hat{\beta}_{CMPRRE}$ iff MMSE($\hat{\beta}_{CMPRRE}$) − MMSE($\hat{\beta}_{CMPKLHE}$) > 0, where $k_*,\ k_R > 0$, and $0 < d_* < 1$.*

Proof. The difference between MMSE($\hat{\beta}_{\mathrm{CMPRRE}}$) and MMSE($\hat{\beta}_{\mathrm{CMPKLHE}}$) is given by

$$\begin{aligned}\mathrm{MMSE}(\hat{\beta}_{\mathrm{CMPRRE}}) - \mathrm{MMSE}(\hat{\beta}_{\mathrm{CMPKLHE}}) =& \hat{\varphi} Q (S_k^{-1} S S_k^{-1} - S_k^{-1} S_{-k} S_M^{-1} S_M^{-1} S_{-k} S_k^{-1}) Q^t \\ & + \mathrm{Bias}(\hat{\beta}_{\mathrm{CMPRRE}}) \times \mathrm{Bias}(\hat{\beta}_{\mathrm{CMPRRE}})^t \\ & - \mathrm{Bias}(\hat{\beta}_{\mathrm{CMPKLHE}})\mathrm{Bias}(\hat{\beta}_{\mathrm{CMPKLHE}})^t.\end{aligned} \tag{57}$$

Eq. (57) can be expressed using the MSE as:

$$\begin{aligned}\mathrm{MSE}(\hat{\beta}_{\mathrm{CMPRRE}}) - \mathrm{MSE}(\hat{\beta}_{\mathrm{CMPKLHE}}) =& \\ & \hat{\varphi} Q \mathrm{diag}\left(\frac{s_j}{(s_j + k_R)^2} - \frac{(s_j - k_*)^2}{(s_j + k_*)^2 (s_j + k_*(1 + d_*))^2}\right) Q^t \\ & + \mathrm{Bias}(\hat{\beta}_{\mathrm{CMPRRE}})\mathrm{Bias}(\hat{\beta}_{\mathrm{CMPRRE}})^t \\ & - \mathrm{Bias}(\hat{\beta}_{\mathrm{CMPKLHE}})\mathrm{Bias}(\hat{\beta}_{\mathrm{CMPKLHE}})^t.\end{aligned} \tag{58}$$

The matrix $(S_k^{-1} S S_k^{-1} - S_k^{-1} S_{-k} S_M^{-1} S_M^{-1} S_{-k} S_k^{-1})$ is pd if $s_j (s_j + k_*)^2 (s_j + k_*(1 + d_*))^2 - (s_j + k_R)^2 (s_j - k_*)^2 > 0$, which is equivalent to $s_j (s_j + k_*)^2 (s_j + k_*(1 + d_*))^2 > (s_j + k_R)^2 (s_j - k_*)^2$ being non-negative. Therefore, if $k_*,\ k_R > 0$, and $0 < d_* < 1$, the proof is completed by Lemma 1. □

Theorem 3. *The estimator $\hat{\beta}_{CMPKLHE}$ outperforms $\hat{\beta}_{CMPLE}$ iff MMSE($\hat{\beta}_{CMPLE}$) − MMSE($\hat{\beta}_{CMPKLHE}$) > 0, where $k_* > 0$, $0 < d_L < 1$, and $0 < d_* < 1$.*

Proof. The difference between MMSE($\hat{\beta}_{\text{CMPLE}}$) and MMSE($\hat{\beta}_{\text{CMPKLHE}}$) is given by

$$\begin{aligned}\text{MMSE}(\hat{\beta}_{\text{CMPLE}}) - \text{MMSE}(\hat{\beta}_{\text{CMPKLHE}}) \\ &= \hat{\varphi}Q((S+I_p)^{-1}(S+d_L I_p)S^{-1}(S+d_L I_p)(S+I_p)^{-1} \\ &- S_k^{-1}S_{-k}S_M^{-1}S_M^{-1}S_{-k}S_k^{-1})Q^t \\ &+ \text{Bias}(\hat{\beta}_{\text{CMPLE}})\text{Bias}(\hat{\beta}_{\text{CMPLE}})^t - \text{Bias}(\hat{\beta}_{\text{CMPKLHE}})\text{Bias}(\hat{\beta}_{\text{CMPKLHE}})^t.\end{aligned} \tag{59}$$

Eq. (59) can be expressed using the MSE as

$$\begin{aligned}\text{MSE}(\hat{\beta}_{\text{CMPLE}}) - \text{MSE}(\hat{\beta}_{\text{CMPKLHE}}) = \\ &\hat{\varphi}Q\text{diag}\left(\frac{(s_j+d_L)^2}{s_j(s_j+1)^2} - \frac{(s_j-k_*)^2}{(s_j+k_*)^2(s_j+k_*(1+d_*))^2}\right)Q^t \\ &+ \text{Bias}(\hat{\beta}_{\text{CMPLE}})\text{Bias}(\hat{\beta}_{\text{CMPLE}})^t \\ &- \text{Bias}(\hat{\beta}_{\text{CMPKLHE}})\text{Bias}(\hat{\beta}_{\text{CMPKLHE}})^t.\end{aligned} \tag{60}$$

The matrix $(S+I_p)^{-1}(S+d_L I_p)S^{-1}(S+d_L I_p)(S+I_p)^{-1} - S_k^{-1}S_{-k}S_M^{-1}S_M^{-1}S_{-k}S_k^{-1})$ is pd if $(s_j+d_L)^2(s_j+k_*)^2(s_j+k_*(1+d_*))^2 - s_j(s_j+1)^2(s_j-k_*)^2 > 0$, which is equivalent to $(s_j+d_L)^2(s_j+k_*)^2(s_j+k_*(1+d_*))^2 > s_j(s_j+1)^2(s_j-k_*)^2$ being non-negative. Therefore, if $k_* > 0$, $0 < d_* < 1$ and $0 > d_L > 1$, the proof is completed by Lemma 1. □

Theorem 4. *The estimator $\hat{\beta}_{CMPKLHE}$ outperforms $\hat{\beta}_{CMPLTE}$ iff MMSE($\hat{\beta}_{CMPLTE}$) − MMSE($\hat{\beta}_{CMPKLHE}$) > 0, where $k_*, k_{LT} > 0$, $0 < d_* < 1$, and $-\infty < d_{LT} < \infty$.*

Proof. The difference between MMSE($\hat{\beta}_{\text{CMPLTE}}$) and MMSE($\hat{\beta}_{\text{CMPKLHE}}$) is given by

$$\begin{aligned}\text{MMSE}(\hat{\beta}_{\text{CMPLTE}}) - \text{MMSE}(\hat{\beta}_{\text{CMPKLHE}}) = \\ &\hat{\varphi}Q(S_k^{-1}S_dS^{-1}S_dS_k^{-1} - S_k^{-1}S_{-k}S_M^{-1}S_M^{-1}S_{-k}S_k^{-1})Q^t \\ &+ \text{Bias}(\hat{\beta}_{\text{CMPLTE}}) \times \text{Bias}(\hat{\beta}_{\text{CMPLTE}})^t \\ &- \text{Bias}(\hat{\beta}_{\text{CMPKLHE}})\text{Bias}(\hat{\beta}_{\text{CMPKLHE}})^t.\end{aligned} \tag{61}$$

Eq. (61) can be expressed using the MSE as:

$$\begin{aligned}\text{MSE}(\hat{\beta}_{\text{CMPLTE}}) - \text{MSE}(\hat{\beta}_{\text{CMPKLHE}}) = \\ &\hat{\varphi}Q\text{diag}\left(\frac{(s_j-d_{LT})^2}{s_j(s_j+k_{LT})^2} - \frac{(s_j-k_*)^2}{(s_j+k_*)^2(s_j+k_*(1+d_*))^2}\right)Q^t \\ &+ \text{Bias}(\hat{\beta}_{\text{CMPLTE}})\text{Bias}(\hat{\beta}_{\text{CMPLTE}})^t \\ &- \text{Bias}(\hat{\beta}_{\text{CMPKLHE}})\text{Bias}(\hat{\beta}_{\text{CMPKLHE}})^t.\end{aligned} \tag{62}$$

The matrix $(S_k^{-1}S_dS^{-1}S_dS_k^{-1} - S_k^{-1}S_{-k}S_M^{-1}S_M^{-1}S_{-k}S_k^{-1})$ is pd if $(s_j - d_{LT})^2(s_j + k_*)^2(s_j + k_*(1+d_*))^2 - s_j(s_j + k_{LT})^2(s_j - k_*)^2 > 0$, which is equivalent to $(s_j - d_{LT})^2(s_j + k_*)^2(s_j + k_*(1 + d_*))^2 > s_j(s_j + k_{LT})^2(s_j - k_*)^2$ being non-negative. Therefore, if $k_*, k_{LT} > 0$, $0 < d_* < 1$, and $-\infty > d_{LT} > \infty$, the proof is completed by Lemma 1. □

Theorem 5. *The estimator $\hat{\beta}_{CMPKLHE}$ outperforms $\hat{\beta}_{CMPKLE}$ iff MMSE$(\hat{\beta}_{CMPKLE})$ − MMSE$(\hat{\beta}_{CMPKLHE}) > 0$, where $k_*, k_{KL} > 0$ and $0 < d_* < 1$.*

Proof. The difference between MMSE$(\hat{\beta}_{\text{CMPKLE}})$ and MMSE$(\hat{\beta}_{\text{CMPKLHE}})$ is given by

$$\begin{aligned}\text{MMSE}(\hat{\beta}_{\text{CMPKLE}}) - \text{MMSE}(\hat{\beta}_{\text{CMPKLHE}}) = &\\ &\hat{\varphi}Q(S_k^{-1}S_{-k}S^{-1}S_{-k}S_k^{-1} - S_k^{-1}S_{-k}S_M^{-1}S_M^{-1}S_{-k}S_k^{-1})Q^t\\ &+ \text{Bias}(\hat{\beta}_{\text{CMPKLE}}) \times \text{Bias}(\hat{\beta}_{\text{CMPKLE}})^t\\ &- \text{Bias}(\hat{\beta}_{\text{CMPKLHE}})\text{Bias}(\hat{\beta}_{\text{CMPKLHE}})^t.\end{aligned} \tag{63}$$

Eq. (63) can be expressed using the MSE as:

$$\begin{aligned}\text{MSE}(\hat{\beta}_{\text{CMPKLE}}) - \text{MSE}(\hat{\beta}_{\text{CMPKLHE}}) = &\\ &\hat{\varphi}Q\text{diag}\left(\frac{(s_j - k_{KL})^2}{s_j(s_j + k_{KL})^2} - \frac{(s_j - k_*)^2}{(s_j + k_*)^2(s_j + k_*(1 + d_*))^2}\right)Q^t\\ &+ \text{Bias}(\hat{\beta}_{\text{CMPKLE}})\text{Bias}(\hat{\beta}_{\text{CMPKLE}})^t\\ &- \text{Bias}(\hat{\beta}_{\text{CMPKLHE}})\text{Bias}(\hat{\beta}_{\text{CMPKLHE}})^t.\end{aligned} \tag{64}$$

The matrix $(S_k^{-1}S_{-k}S^{-1}S_{-k}S_k^{-1} - S_k^{-1}S_{-k}S_M^{-1}S_M^{-1}S_{-k}S_k^{-1})$ is pd if $(s_j - k_{KL})^2(s_j + k_*)^2(s_j + k_*(1 + d_*))^2 - s_j(s_j + k_{KL})^2(s_j - k_*)^2 > 0$, which is equivalent to $(s_j - k_{KL})^2(s_j + k_*)^2(s_j + k_*(1 + d_*))^2 > s_j(s_j + k_{KL})^2(s_j - k_*)^2$ being non-negative. Therefore, if $k_* > 0$, $0 < d_* < 1$, and $k_{KL} > 0$ the proof is completed by Lemma 1. □

Theorem 6. *The estimator $\hat{\beta}_{CMPKLHE}$ outperforms $\hat{\beta}_{CMPMRTE}$ iff MMSE$(\hat{\beta}_{CMPMRTE})$− MMSE$(\hat{\beta}_{CMPKLHE}) > 0$, where $k_*, k_M > 0$, $0 < d_M < 1$, and $0 < d_* < 1$.*

Proof. The difference between MMSE$(\hat{\beta}_{\text{CMPMRTE}})$ and MMSE$(\hat{\beta}_{\text{CMPKLHE}})$ is given by

$$\begin{aligned}\text{MMSE}(\hat{\beta}_{\text{CMPMRTE}}) - \text{MMSE}(\hat{\beta}_{\text{CMPKLHE}}) =& \hat{\varphi}Q(S_M^{-1}SS_M^{-1} - S_k^{-1}S_{-k}S_M^{-1}S_M^{-1}S_{-k}S_k^{-1})Q^t\\ &+ \text{Bias}(\hat{\beta}_{\text{CMPMRTE}}) \times \text{Bias}(\hat{\beta}_{\text{CMPMRTE}})^t\\ &- \text{Bias}(\hat{\beta}_{\text{CMPKLHE}})\text{Bias}(\hat{\beta}_{\text{CMPKLHE}})^t.\end{aligned} \tag{65}$$

Eq. (65) can be expressed using the MSE as:

$$\begin{aligned}\text{MSE}(\hat{\beta}_{\text{CMPMRTE}}) - \text{MSE}(\hat{\beta}_{\text{CMPKLHE}}) = \\ &\hat{\varphi} Q \text{diag}\left(\frac{s_j}{(s_j + k_{MTR}(d_M + 1))^2} - \frac{(s_j - k_*)^2}{(s_j + k_*)^2 (s_j + k_*(1 + d_*))^2}\right) Q^t \\ &+ \text{Bias}(\hat{\beta}_{\text{CMPMRTE}}) \text{Bias}(\hat{\beta}_{\text{CMPMRTE}})^t \\ &- \text{Bias}(\hat{\beta}_{\text{CMPKLHE}}) \text{Bias}(\hat{\beta}_{\text{CMPKLHE}})^t.\end{aligned} \tag{66}$$

The matrix $(S_M^{-1} S S_M^{-1} - S_k^{-1} S_{-k} S_M^{-1} S_M^{-1} S_{-k} S_k^{-1})$ is pd if $s_j (s_j + k_*)^2 (s_j + k_*(1 + d_*))^2 - (s_j + k_M(d_M + 1))^2 (s_j - k_*)^2 > 0$, which is equivalent to $s_j (s_j + k_*)^2 (s_j + k_*(1 + d_*))^2 > (s_j + k_M(d_M + 1))^2 (s_j - k_*)^2$ being non-negative. Therefore, if $k_*, k_M > 0$, $0 < d_* < 1$, and $0 > d_M > 1$, the proof is completed by Lemma 1. □

3 Monte Carlo Simulation Study

This section presents an in-depth numerical evaluation of the proposed estimator. We compare its performance with the existing estimators through a Monte Carlo simulation study and validate its effectiveness using two real-world applications.

3.1 Simulation Design

This section outlines the simulation study designed to evaluate the performance of the proposed estimator under varying conditions.

The study systematically varied four key factors: sample size (n), dispersion (φ), multicollinearity (ρ), number of explanatory variables (p), as described in Table 1. The response variable followed a CMP distribution with a mean (μ_i) and dispersion (φ). This can be expressed mathematically as:

$$Y_i \sim \text{CMP}(\mu_i, \varphi),$$

where Y_i is the response variable for the i-th observation and μ_i is the mean of the i-th observation, defined as:

$$\mu_i = \exp(\beta_0 + \beta_1 x_{i1} + \ldots + \beta_p x_{ip}),$$

where $\beta_0, \beta_1, ..., \beta_p$ are the regression coefficients that were chosen to satisfy the constraint
$\sum_{j=1}^{p} \beta_j^2 = 1$, $x_{i1}, x_{i2}, ..., x_{ip}$ are the explanatory variables for the i-th observation, and φ is the dispersion parameter.

Correlated explanatory variables were generated using the following formula:

$$x_{ij} = \sqrt{1 - \rho^2}\, z_{ij} + \rho\, z_{i(j+1)}, i = 1, ..., n, j = 1, ..., p,$$

where x_{ij} is the j-th explanatory variable for the i-th observation, z_{ij} are independent standard normal pseudo-random numbers, and ρ is the correlation parameter.

The estimator's performance was evaluated using the MSE, which was calculated as the average squared difference between the true parameter values and the estimated values across all replications. This can be expressed mathematically as:

$$MSE(\hat{\beta}) = \frac{1}{R}\sum_{i=1}^{R}(\hat{\beta}_i - \beta)^t(\hat{\beta}_i - \beta),$$

,

where $(\hat{\beta}_i - \beta)$ represents the difference between the true parameter and the estimated vectors of the proposed and other estimators in the i-th replication, and R=1000 is the number of replications. The R software with the `COMPoissonReg` package was used for all computations.

Table 1: Different factors in the simulation study.

Factor	**Notation**	**Values**
Sample size	n	20, 75, 150, 200, 300, 400
Dispersion parameter	φ	0.85, 1, 1.25
Degree of correlation	ρ	0.85, 0.90, 0.95, 0.99
Number of explanatory variables	p	4, 7, 10

Table 2: Average MSE values when p=4 and φ = 0.85.

ρ	n	CMPMLE	CMPRRE	CMPLE	CMPLTE	CMPMRTE	CMPKLE	CMPKLHE			
		-	$\hat{k}_R$	$\hat{d}_L$	$\hat{d}_{LT}$	$\hat{k}_M, \hat{d}_M$	$\hat{k}_{KL}$	$\hat{k}_*^1$	$\hat{k}_*^2$	$\hat{k}_*^3$	$\hat{k}_*^4$
0.85	20	2.8700	2.6014	2.1214	2.5965	2.6014	2.3825	1.5865	1.6045	0.9719	0.8784
	75	0.2388	0.2334	0.2265	0.2333	0.2334	0.2285	0.1893	0.2028	0.1822	0.1940
	150	0.0937	0.0926	0.0914	0.0926	0.0926	0.0917	0.0825	0.0862	0.0822	0.0856
	200	0.0682	0.0676	0.0669	0.0676	0.0676	0.0670	0.0614	0.0637	0.0617	0.0639
	300	0.0407	0.0405	0.0403	0.0405	0.0405	0.0403	0.0391	0.0392	0.0385	0.0392
	500	0.0238	0.0237	0.0236	0.0237	0.0237	0.0236	0.0234	0.0232	0.0229	0.0232
0.90	20	3.3503	2.8592	2.0888	2.8506	2.8592	2.4821	1.5587	1.5529	1.0460	0.9502
	75	0.2649	0.2537	0.2427	0.2535	0.2537	0.2441	0.1978	0.2091	0.1844	0.1997
	150	0.1143	0.1121	0.1102	0.1121	0.1121	0.1103	0.1012	0.1030	0.0970	0.1022
	200	0.0816	0.0804	0.0794	0.0804	0.0804	0.0794	0.0751	0.0755	0.0723	0.0754
	300	0.0517	0.0512	0.0507	0.0512	0.0512	0.0507	0.0495	0.0489	0.0474	0.0489
	500	0.0312	0.0311	0.0309	0.0311	0.0311	0.0309	0.0315	0.0303	0.0298	0.0303
0.95	20	4.6987	3.5911	2.1674	3.5756	3.5911	2.8255	1.6291	1.6488	1.2717	1.1564
	75	0.4053	0.3716	0.3458	0.3707	0.3716	0.3454	0.2830	0.2962	0.2826	0.2725
	150	0.1737	0.1672	0.1626	0.1670	0.1672	0.1621	0.1531	0.1530	0.1430	0.1470
	200	0.1213	0.1173	0.1149	0.1172	0.1173	0.1141	0.1104	0.1082	0.1015	0.1061
	300	0.0775	0.0756	0.0746	0.0756	0.0756	0.0741	0.0738	0.0708	0.0674	0.0705
	500	0.0474	0.0468	0.0464	0.0468	0.0468	0.0463	0.0484	0.0452	0.0437	0.0450
0.99	20	18.1521	11.0196	2.3399	11.0011	11.0196	7.0784	2.2442	2.2901	1.3652	1.4166
	75	1.5019	1.1301	0.8040	1.1125	1.1301	0.9049	0.6179	0.7184	0.9886	0.7169
	150	0.6641	0.5635	0.4842	0.5592	0.5635	0.4993	0.4036	0.4497	0.5322	0.3776
	200	0.4849	0.4250	0.3877	0.4224	0.4250	0.3861	0.3403	0.3678	0.3989	0.3132
	300	0.3039	0.2749	0.2583	0.2740	0.2749	0.2557	0.2385	0.2492	0.2506	0.2162
	500	0.1768	0.1660	0.1609	0.1658	0.1660	0.1589	0.1623	0.1595	0.1534	0.1458

Table 3: Average MSE values when p=4 and φ = 1.

ρ	n	CMPMLE	CMPRRE	CMPLE	CMPLTE	CMPMRTE	CMPKLE	CMPKLHE			
		-	$\hat{k}_R$	$\hat{d}_L$	$\hat{d}_{LT}$	$\hat{k}_M, \hat{d}_M$	$\hat{k}_{KL}$	$\hat{k}_*^1$	$\hat{k}_*^2$	$\hat{k}_*^3$	$\hat{k}_*^4$
0.85	20	3.3128	2.8306	2.1410	2.8249	2.8306	2.4663	1.5458	1.5109	0.9554	0.8578
	75	0.2883	0.2775	0.2677	0.2773	0.2775	0.2681	0.2193	0.2319	0.2053	0.2198
	150	0.1118	0.1098	0.1081	0.1097	0.1098	0.1080	0.0977	0.1006	0.0949	0.0997
	200	0.0828	0.0816	0.0806	0.0815	0.0816	0.0805	0.0741	0.0759	0.0728	0.0760
	300	0.0499	0.0495	0.0491	0.0495	0.0495	0.0491	0.0481	0.0476	0.0464	0.0475
	500	0.0294	0.0292	0.0291	0.0292	0.0292	0.0290	0.0289	0.0284	0.0280	0.0285
0.90	20	3.7636	2.9799	2.0913	2.9710	2.9799	2.4413	1.5486	1.5026	1.0732	0.9750
	75	0.3248	0.3031	0.2884	0.3027	0.3031	0.2856	0.2325	0.2433	0.2169	0.2284
	150	0.1433	0.1389	0.1363	0.1389	0.1389	0.1354	0.1248	0.1265	0.1177	0.1241
	200	0.1018	0.0995	0.0982	0.0995	0.0995	0.0976	0.0930	0.0928	0.0879	0.0920
	300	0.0643	0.0632	0.0627	0.0632	0.0632	0.0624	0.0615	0.0601	0.0578	0.0600
	500	0.0392	0.0389	0.0387	0.0389	0.0389	0.0387	0.0399	0.0380	0.0371	0.0379
0.95	20	5.9163	4.1006	2.2229	4.0874	4.1006	2.9847	1.6970	1.7275	1.2823	1.1922
	75	0.5324	0.4670	0.4311	0.4660	0.4670	0.4213	0.3464	0.3691	0.3714	0.3301
	150	0.2264	0.2127	0.2064	0.2125	0.2127	0.2029	0.1896	0.1934	0.1804	0.1812
	200	0.1566	0.1487	0.1457	0.1487	0.1487	0.1431	0.1396	0.1388	0.1296	0.1326
	300	0.1005	0.0968	0.0955	0.0967	0.0968	0.0940	0.0949	0.0914	0.0862	0.0893
	500	0.0612	0.0598	0.0594	0.0598	0.0598	0.0588	0.0622	0.0580	0.0557	0.0573
0.99	20	23.6670	13.1621	2.2663	13.1589	13.1621	7.6420	2.3283	2.3673	1.3211	1.3952
	75	2.0129	1.3756	0.9160	1.3587	1.3756	1.0476	0.7102	0.8596	1.1513	0.8950
	150	0.9102	0.7196	0.6011	0.7143	0.7196	0.6131	0.4849	0.5675	0.7052	0.4867
	200	0.6735	0.5561	0.4984	0.5531	0.5561	0.4894	0.4181	0.4755	0.5401	0.3925
	300	0.4151	0.3594	0.3339	0.3583	0.3594	0.3277	0.2980	0.3310	0.3441	0.2682
	500	0.2432	0.2214	0.2139	0.2211	0.2214	0.2089	0.2078	0.2166	0.2079	0.1871

Table 4: Average MSE values when p=4 and φ = 1.25.

ρ	n	CMPMLE	CMPRRE	CMPLE	CMPLTE	CMPMRTE	CMPKLE	CMPKLHE			
		-	$\hat{k}_R$	$\hat{d}_L$	$\hat{d}_{LT}$	$\hat{k}_M, \hat{d}_M$	$\hat{k}_{KL}$	$\hat{k}_*^1$	$\hat{k}_*^2$	$\hat{k}_*^3$	$\hat{k}_*^4$
0.85	20	3.8877	2.9836	2.1702	2.9783	2.9836	2.4036	1.5407	1.4350	1.0122	0.9155
	75	0.3734	0.3473	0.3342	0.3471	0.3473	0.3272	0.2684	0.2805	0.2464	0.2600
	150	0.1421	0.1369	0.1349	0.1369	0.1369	0.1330	0.1215	0.1240	0.1150	0.1217
	200	0.1044	0.1013	0.1003	0.1013	0.1013	0.0989	0.0924	0.0933	0.0882	0.0930
	300	0.0640	0.0629	0.0625	0.0629	0.0629	0.0621	0.0617	0.0603	0.0582	0.0599
	500	0.0377	0.0373	0.0372	0.0373	0.0373	0.0370	0.0376	0.0362	0.0354	0.0362
0.90	20	4.9317	3.4343	2.1852	3.4270	3.4343	2.5652	1.5907	1.5213	1.1621	1.0633
	75	0.4428	0.3921	0.3727	0.3918	0.3921	0.3573	0.2949	0.3133	0.2870	0.2802
	150	0.1895	0.1787	0.1760	0.1787	0.1787	0.1713	0.1597	0.1636	0.1496	0.1560
	200	0.1343	0.1287	0.1275	0.1287	0.1287	0.1249	0.1213	0.1216	0.1131	0.1177
	300	0.0867	0.0841	0.0836	0.0841	0.0841	0.0823	0.0823	0.0804	0.0761	0.0790
	500	0.0525	0.0516	0.0514	0.0516	0.0516	0.0510	0.0533	0.0505	0.0488	0.0499
0.95	20	7.9839	4.7678	2.3013	4.7615	4.7678	3.1436	1.8098	1.8416	1.3394	1.2711
	75	0.7438	0.5996	0.5494	0.5989	0.5996	0.5186	0.4298	0.4850	0.5195	0.4094
	150	0.3118	0.2803	0.2727	0.2802	0.2803	0.2629	0.2463	0.2653	0.2475	0.2315
	200	0.2180	0.1998	0.1969	0.1998	0.1998	0.1897	0.1871	0.1958	0.1802	0.1747
	300	0.1404	0.1313	0.1304	0.1313	0.1313	0.1259	0.1280	0.1281	0.1182	0.1197
	500	0.0863	0.0830	0.0828	0.0830	0.0830	0.0809	0.0860	0.0821	0.0772	0.0790
0.99	20	34.8159	17.1520	12.3461	17.1467	17.1520	9.2486	2.5159	2.5715	1.3793	1.5016
	75	3.0522	1.8226	1.6058	1.8166	1.8226	1.2955	0.8273	1.0373	1.2866	1.1254
	150	1.3669	0.9576	0.7671	0.9533	0.9576	0.7764	0.5982	0.7471	0.9455	0.6805
	200	0.9886	0.7350	0.6427	0.7328	0.7350	0.6231	0.5185	0.6351	0.7540	0.5215
	300	0.6147	0.4886	0.4497	0.4884	0.4886	0.4351	0.3840	0.4685	0.5106	0.3527
	500	0.3634	0.3110	0.3004	0.3110	0.3110	0.2887	0.2762	0.3162	0.3059	0.2492

Table 5: Average MSE values when p=7 and φ = 0.85.

ρ	n	CMPMLE	CMPRRE	CMPLE	CMPLTE	CMPMRTE	CMPKLE	CMPKLHE			
		-	$\hat{k}_R$	$\hat{d}_L$	$\hat{d}_{LT}$	$\hat{k}_M, \hat{d}_M$	$\hat{k}_{KL}$	$\hat{k}_*^1$	$\hat{k}_*^2$	$\hat{k}_*^3$	$\hat{k}_*^4$
0.85	20	3.7047	3.5835	3.2886	3.5364	3.5835	3.4755	1.5192	0.9708	1.3634	1.3898
	75	0.4278	0.4161	0.4024	0.4156	0.4161	0.4050	0.3504	0.3369	0.3220	0.3642
	150	0.1476	0.1456	0.1437	0.1456	0.1456	0.1438	0.1320	0.1315	0.1310	0.1396
	200	0.1015	0.1005	0.0996	0.1005	0.1005	0.0996	0.0931	0.0933	0.0933	0.0977
	300	0.0615	0.0611	0.0607	0.0611	0.0611	0.0607	0.0573	0.0577	0.0579	0.0600
	500	0.0346	0.0345	0.0344	0.0345	0.0345	0.0344	0.0331	0.0334	0.0335	0.0341
0.90	20	6.4915	4.9909	4.2362	4.9909	4.9909	3.7585	1.8899	1.1409	1.5113	1.7495
	75	0.5097	0.4870	0.4655	0.4859	0.4870	0.4663	0.3931	0.3713	0.3456	0.4124
	150	0.1782	0.1744	0.1716	0.1743	0.1744	0.1709	0.1549	0.1530	0.1514	0.1659
	200	0.1352	0.1328	0.1312	0.1328	0.1328	0.1306	0.1193	0.1187	0.1182	0.1280
	300	0.0775	0.0766	0.0761	0.0766	0.0766	0.0759	0.0713	0.0716	0.0715	0.0751
	500	0.0451	0.0448	0.0446	0.0448	0.0448	0.0445	0.0423	0.0426	0.0427	0.0442
0.95	20	10.4043	5.9996	4.3579	5.9996	5.9996	4.5360	2.6095	2.0121	1.6500	1.8105
	75	0.8229	0.7410	0.6833	0.7362	0.7410	0.6710	0.5277	0.4742	0.4286	0.5662
	150	0.3026	0.2875	0.2799	0.2869	0.2875	0.2742	0.2393	0.2308	0.2227	0.2662
	200	0.2094	0.2015	0.1979	0.2013	0.2015	0.1945	0.1738	0.1707	0.1669	0.1914
	300	0.1268	0.1238	0.1225	0.1237	0.1238	0.1211	0.1122	0.1121	0.1108	0.1205
	500	0.0753	0.0742	0.0737	0.0742	0.0742	0.0731	0.0690	0.0693	0.0690	0.0730
0.99	20	27.9655	17.5544	17.0542	7.5544	17.5544	4.3713	3.1908	3.3070	3.5710	3.9805
	75	3.2797	2.3791	1.5946	2.2896	2.3791	1.7335	0.9474	0.8448	1.0070	0.9380
	150	1.2885	1.0436	0.9256	1.0163	1.0436	0.8572	0.6552	0.5870	0.5757	0.7453
	200	0.9294	0.7877	0.7282	0.7748	0.7877	0.6765	0.5411	0.4952	0.4661	0.6335
	300	0.5709	0.5046	0.4832	0.5001	0.5046	0.4511	0.3788	0.3583	0.3363	0.4494
	500	0.3234	0.2989	0.2926	0.2979	0.2989	0.2786	0.2447	0.2398	0.2290	0.2830

Table 6: Average MSE values when p=7 and φ = 1.

ρ	n	CMPMLE	CMPRRE	CMPLE	CMPLTE	CMPMRTE	CMPKLE	CMPKLHE			
		-	$\hat{k}_R$	$\hat{d}_L$	$\hat{d}_{LT}$	$\hat{k}_M, \hat{d}_M$	$\hat{k}_{KL}$	$\hat{k}_*^1$	$\hat{k}_*^2$	$\hat{k}_*^3$	$\hat{k}_*^4$
0.85	20	4.9210	3.7133	3.6573	3.6552	3.7133	3.5346	1.9072	1.3096	1.6775	1.6097
	75	0.5241	0.4997	0.4807	0.4988	0.4997	0.4774	0.4064	0.3754	0.3585	0.4228
	150	0.1828	0.1788	0.1762	0.1787	0.1788	0.1750	0.1592	0.1559	0.1558	0.1698
	200	0.1255	0.1234	0.1222	0.1234	0.1234	0.1215	0.1126	0.1116	0.1119	0.1192
	300	0.0780	0.0771	0.0766	0.0771	0.0771	0.0762	0.0715	0.0714	0.0720	0.0755
	500	0.0442	0.0439	0.0438	0.0439	0.0439	0.0437	0.0420	0.0422	0.0423	0.0434
0.90	20	7.2548	3.9785	3.8265	3.9565	4.0215	3.7655	2.0215	1.7236	1.9566	1.8659
	75	0.6464	0.5993	0.5696	0.5973	0.5993	0.5584	0.4657	0.4227	0.3925	0.4880
	150	0.2315	0.2234	0.2198	0.2232	0.2234	0.2162	0.1952	0.1898	0.1876	0.2109
	200	0.1718	0.1668	0.1648	0.1667	0.1668	0.1623	0.1476	0.1447	0.1443	0.1599
	300	0.1020	0.1001	0.0994	0.1001	0.1001	0.0984	0.0920	0.0916	0.0916	0.0978
	500	0.0595	0.0587	0.0585	0.0587	0.0587	0.0580	0.0549	0.0550	0.0553	0.0579
0.95	20	12.1489	6.1205	5.9878	6.1879	6.1205	5.2456	2.5462	2.0155	2.0015	2.3555
	75	1.0832	0.9209	0.8434	0.9131	0.9209	0.7913	0.6276	0.5456	0.5043	0.6640
	150	0.4118	0.3797	0.3706	0.3785	0.3797	0.3527	0.3087	0.2909	0.2791	0.3471
	200	0.2830	0.2660	0.2620	0.2655	0.2660	0.2516	0.2250	0.2173	0.2115	0.2513
	300	0.1759	0.1691	0.1677	0.1689	0.1691	0.1632	0.1506	0.1491	0.1469	0.1642
	500	0.1021	0.0996	0.0991	0.0996	0.0996	0.0974	0.0915	0.0913	0.0910	0.0979
0.99	20	42.2549	20.2565	19.1545	19.2156	20.2549	18.5145	4.2565	4.0256	4.6542	4.3216
	75	4.5666	3.0281	1.8078	2.9412	3.0281	2.0213	1.0076	0.9519	1.1883	1.0261
	150	1.8181	1.3494	1.1591	1.3073	1.3494	1.0241	0.7680	0.6843	0.7120	0.8348
	200	1.3165	1.0345	0.9426	1.0120	1.0345	0.8310	0.6608	0.5933	0.5738	0.7591
	300	0.8043	0.6683	0.6400	0.6593	0.6683	0.5670	0.4783	0.4442	0.4187	0.5720
	500	0.6606	0.5747	0.5843	0.5711	0.5747	0.5056	0.4692	0.4117	0.4047	0.5212

Table 7: Average MSE values when p=7 and φ = 1.25.

ρ	n	CMPMLE	CMPRRE	CMPLE	CMPLTE	CMPMRTE	CMPKLE	CMPKLHE			
		-	$\hat{k}_R$	$\hat{d}_L$	$\hat{d}_{LT}$	$\hat{k}_M, \hat{d}_M$	$\hat{k}_{KL}$	$\hat{k}^1_*$	$\hat{k}^2_*$	$\hat{k}^3_*$	$\hat{k}^4_*$
0.85	20	8.1257	4.5587	4.1314	4.5645	4.5681	4.3557	2.5141	1.6542	0.9994	1.3081
	75	0.6674	0.6097	0.5875	0.6081	0.6097	0.5610	0.4789	0.4200	0.4012	0.4981
	150	0.2347	0.2250	0.2227	0.2248	0.2250	0.2166	0.1976	0.1896	0.1898	0.2130
	200	0.1775	0.1715	0.1701	0.1714	0.1715	0.1662	0.1523	0.1475	0.1487	0.1643
	300	0.1101	0.1078	0.1073	0.1077	0.1078	0.1057	0.0992	0.0980	0.0988	0.1051
	500	0.0591	0.0583	0.0582	0.0583	0.0583	0.0576	0.0550	0.0549	0.0554	0.0576
0.90	20	11.3754	8.5003	5.5856	8.5003	8.5003	4.6943	3.1507	2.2450	1.1024	1.4396
	75	0.8560	0.7330	0.7025	0.7296	0.7330	0.6359	0.5393	0.4601	0.4321	0.5783
	150	0.3307	0.3080	0.3049	0.3075	0.3080	0.2894	0.2613	0.2472	0.2427	0.2872
	200	0.2457	0.2314	0.2300	0.2312	0.2314	0.2192	0.1977	0.1878	0.1870	0.2199
	300	0.1506	0.1453	0.1449	0.1452	0.1453	0.1407	0.1314	0.1288	0.1291	0.1416
	500	0.0837	0.0819	0.0818	0.0819	0.0819	0.0804	0.0766	0.0763	0.0766	0.0809
0.95	20	19.7331	9.1622	6.3861	9.1622	9.1622	5.1005	3.6495	2.2862	1.1851	1.8357
	75	1.5586	1.1866	1.0828	1.1758	1.1866	0.9290	0.7597	0.6580	0.6375	0.7799
	150	0.5646	0.4882	0.4841	0.4856	0.4882	0.4316	0.3888	0.3621	0.3440	0.4441
	200	0.4161	0.3743	0.3736	0.3732	0.3743	0.3429	0.3145	0.3021	0.2901	0.3542
	300	0.2734	0.2532	0.2536	0.2529	0.2532	0.2373	0.2191	0.2125	0.2077	0.2450
	500	0.1490	0.1423	0.1429	0.1423	0.1423	0.1371	0.1303	0.1297	0.1283	0.1409
0.99	20	64.4664	25.2547	8.1996	25.2547	25.2547	20.8927	5.5373	4.4490	1.8550	2.5855
	75	6.5722	3.7224	1.9132	3.6762	3.7224	2.1640	1.0263	1.1228	1.3369	1.1671
	150	2.8211	1.8493	1.4859	1.8060	1.8493	1.2616	0.8983	0.8250	0.9291	0.9084
	200	2.0537	1.4154	1.2574	1.3850	1.4154	1.0145	0.8005	0.7118	0.7402	0.8721
	300	1.2515	0.9354	0.8952	0.9199	0.9354	0.7320	0.6234	0.5761	0.5450	0.7274
	500	0.7278	0.5935	0.5940	0.5883	0.5935	0.5009	0.4530	0.4279	0.3980	0.5433

Table 8: Average MSE values when p=10 and φ = 0.85.

ρ	n	CMPMLE	CMPRRE	CMPLE	CMPLTE	CMPMRTE	CMPKLE	CMPKLHE			
		-	$\hat{k}_R$	$\hat{d}_L$	$\hat{d}_{LT}$	$\hat{k}_M, \hat{d}_M$	$\hat{k}_{KL}$	$\hat{k}_*^1$	$\hat{k}_*^2$	$\hat{k}_*^3$	$\hat{k}_*^4$
0.85	20	6.6448	5.8746	4.5067	5.8551	5.8746	5.2290	3.6184	2.9785	2.0732	2.3349
	75	0.6651	0.6462	0.6235	0.6452	0.6462	0.6281	0.5626	0.5004	0.5041	0.5838
	150	0.2079	0.2051	0.2025	0.2051	0.2051	0.2025	0.1909	0.1818	0.1863	0.1995
	200	0.1315	0.1302	0.1290	0.1301	0.1302	0.1289	0.1227	0.1184	0.1214	0.1280
	300	0.0808	0.0802	0.0797	0.0802	0.0802	0.0796	0.0767	0.0751	0.0766	0.0794
	500	0.0426	0.0425	0.0423	0.0425	0.0425	0.0423	0.0413	0.0409	0.0414	0.0422
0.90	20	7.7033	6.5352	4.6256	6.5057	6.5352	5.5845	3.5048	2.7963	1.9848	2.3300
	75	0.7847	0.7449	0.7088	0.7425	0.7449	0.7080	0.6152	0.5179	0.5178	0.6555
	150	0.2648	0.2588	0.2545	0.2586	0.2588	0.2531	0.2350	0.2180	0.2247	0.2507
	200	0.1749	0.1720	0.1701	0.1720	0.1720	0.1692	0.1598	0.1517	0.1561	0.1688
	300	0.1007	0.0996	0.0989	0.0996	0.0996	0.0985	0.0943	0.0914	0.0935	0.0985
	500	0.0562	0.0558	0.0556	0.0558	0.0558	0.0555	0.0538	0.0530	0.0538	0.0555
0.95	20	13.1495	9.8964	5.3407	9.8453	9.8964	7.5164	3.7080	3.0618	2.0257	2.4181
	75	1.2686	1.1377	1.0428	1.1276	1.1377	1.0228	0.8404	0.6476	0.6138	0.9181
	150	0.4258	0.4055	0.3946	0.4045	0.4055	0.3868	0.3501	0.3080	0.3161	0.3868
	200	0.3061	0.2948	0.2894	0.2944	0.2948	0.2843	0.2614	0.2373	0.2446	0.2866
	300	0.1701	0.1659	0.1642	0.1658	0.1659	0.1619	0.1524	0.1436	0.1480	0.1637
	500	0.0945	0.0931	0.0926	0.0930	0.0931	0.0917	0.0880	0.0852	0.0872	0.0925
0.99	20	59.1709	38.2687	26.2499	38.2451	38.2365	24.7601	5.2548	4.9528	5.2155	5.1255
	75	5.3046	3.8161	3.4613	3.6755	3.8161	2.7161	1.4328	1.0239	1.0958	1.4278
	150	1.7124	1.3729	1.2274	1.3316	1.3729	1.1026	0.8920	0.6446	0.6160	1.1197
	200	1.2523	1.0578	0.9859	1.0376	1.0578	0.8965	0.7556	0.5629	0.5465	0.9485
	300	0.7526	0.6678	0.6466	0.6620	0.6678	0.5957	0.5328	0.4338	0.4391	0.6458
	500	0.4311	0.4013	0.3955	0.4000	0.4013	0.3752	0.3472	0.3066	0.3165	0.3981

Table 9: Average MSE values when p=10 and φ = 1.

ρ	n	CMPMLE	CMPRRE	CMPLE	CMPLTE	CMPMRTE	CMPKLE	CMPKLHE			
		-	$\hat{k}_R$	$\hat{d}_L$	$\hat{d}_{LT}$	$\hat{k}_M, \hat{d}_M$	$\hat{k}_{KL}$	$\hat{k}_*^1$	$\hat{k}_*^2$	$\hat{k}_*^3$	$\hat{k}_*^4$
0.85	20	7.7064	4.1180	3.3088	4.1011	4.1180	3.6260	2.5815	1.9555	1.4824	1.8330
	75	0.7608	0.7225	0.6935	0.7205	0.7225	0.6872	0.6099	0.5077	0.5203	0.6394
	150	0.2599	0.2532	0.2495	0.2531	0.2532	0.2469	0.2299	0.2097	0.2204	0.2454
	200	0.1664	0.1635	0.1619	0.1635	0.1635	0.1607	0.1524	0.1438	0.1492	0.1603
	300	0.1028	0.1017	0.1011	0.1017	0.1017	0.1006	0.0969	0.0937	0.0962	0.1007
	500	0.0555	0.0551	0.0549	0.0551	0.0551	0.0547	0.0532	0.0522	0.0532	0.0548
0.90	20	8.4698	4.4239	3.3204	4.3937	4.4239	3.6152	2.4514	1.8123	1.4843	1.8938
	75	1.0155	0.9323	0.8805	0.9278	0.9323	0.8581	0.7394	0.5779	0.5806	0.7921
	150	0.3555	0.3418	0.3359	0.3414	0.3418	0.3292	0.3041	0.2705	0.2828	0.3293
	200	0.2326	0.2259	0.2235	0.2258	0.2259	0.2197	0.2066	0.1900	0.1986	0.2214
	300	0.1376	0.1349	0.1340	0.1348	0.1349	0.1323	0.1262	0.1195	0.1240	0.1334
	500	0.0790	0.0782	0.0779	0.0782	0.0782	0.0774	0.0752	0.0736	0.0750	0.0778
0.95	20	15.3639	7.3490	4.3174	7.3114	7.3490	5.2457	2.8273	2.2108	1.6945	2.0337
	75	1.6537	1.3934	1.2667	1.3767	1.3934	1.1787	0.9728	0.7011	0.6676	1.0736
	150	0.5903	0.5434	0.5306	0.5412	0.5434	0.5023	0.4575	0.3795	0.3939	0.5177
	200	0.3880	0.3650	0.3599	0.3642	0.3650	0.3445	0.3200	0.2799	0.2924	0.3564
	300	0.2487	0.2388	0.2370	0.2386	0.2388	0.2299	0.2170	0.1988	0.2070	0.2362
	500	0.1379	0.1347	0.1343	0.1347	0.1347	0.1319	0.1271	0.1217	0.1251	0.1344
0.99	20	60.7415	24.8832	18.4094	24.8497	24.8832	16.0194	6.7262	5.6664	6.6028	6.7921
	75	7.0346	4.5711	3.6307	4.4286	4.5711	2.9205	1.3941	1.0820	1.2404	1.3668
	150	2.5468	1.8607	1.6108	1.7881	1.8607	1.3574	1.0818	0.7430	0.7373	1.3416
	200	1.8155	1.4105	1.2931	1.3703	1.4105	1.0992	0.9210	0.6379	0.6141	1.1769
	300	1.0876	0.9042	0.8756	0.8901	0.9042	0.7577	0.6837	0.5199	0.5182	0.8593
	500	0.6259	0.5559	0.5524	0.5525	0.5559	0.4976	0.4661	0.3878	0.4023	0.5563

Table 10: Average MSE values when p=10 and φ = 1.25.

ρ	n	CMPMLE	CMPRRE	CMPLE	CMPLTE	CMPMRTE	CMPKLE	CMPKLHE			
		-	$\hat{k}_R$	$\hat{d}_L$	$\hat{d}_{LT}$	$\hat{k}_M, \hat{d}_M$	$\hat{k}_{KL}$	$\hat{k}_*^1$	$\hat{k}_*^2$	$\hat{k}_*^3$	$\hat{k}_*^4$
0.85	20	10.1643	7.5814	5.1314	7.5625	7.5814	5.7755	3.6414	2.6042	1.6383	2.0208
	75	1.1073	0.9974	0.9515	0.9932	0.9974	0.9017	0.7930	0.5891	0.6084	0.8305
	150	0.3446	0.3289	0.3254	0.3285	0.3289	0.3148	0.2957	0.2587	0.2760	0.3180
	200	0.2519	0.2433	0.2416	0.2432	0.2433	0.2355	0.2232	0.2015	0.2139	0.2383
	300	0.1524	0.1490	0.1484	0.1489	0.1490	0.1458	0.1399	0.1312	0.1374	0.1474
	500	0.0857	0.0846	0.0844	0.0846	0.0846	0.0835	0.0812	0.0785	0.0808	0.0842
0.90	20	14.0353	9.7170	5.5106	9.6923	9.7170	6.8674	3.7328	2.7805	1.6999	2.0843
	75	1.4286	1.2203	1.1553	1.2118	1.2203	1.0487	0.9236	0.6548	0.6637	0.9952
	150	0.5216	0.4816	0.4774	0.4804	0.4816	0.4465	0.4176	0.3428	0.3695	0.4656
	200	0.3385	0.3202	0.3190	0.3198	0.3202	0.3041	0.2891	0.2545	0.2704	0.3151
	300	0.2041	0.1965	0.1962	0.1964	0.1965	0.1897	0.1817	0.1665	0.1756	0.1953
	500	0.1175	0.1148	0.1148	0.1148	0.1148	0.1124	0.1088	0.1035	0.1074	0.1146
0.95	20	24.4707	14.4941	5.9701	14.4723	14.4941	8.8713	3.8095	3.1791	1.7065	2.0262
	75	2.6657	1.9877	1.7725	1.9621	1.9877	1.4889	1.2557	0.8327	0.8043	1.3524
	150	0.9129	0.7834	0.7740	0.7781	0.7834	0.6786	0.6347	0.4834	0.5032	0.7425
	200	0.6221	0.5510	0.5523	0.5485	0.5510	0.4921	0.4693	0.3755	0.4010	0.5455
	300	0.3741	0.3471	0.3486	0.3465	0.3471	0.3244	0.3124	0.2752	0.2901	0.3478
	500	0.2042	0.1955	0.1963	0.1953	0.1955	0.1880	0.1827	0.1710	0.1772	0.1966
0.99	20	108.1876	58.9710	50.0928	58.9653	58.8617	34.4399	7.9421	7.4041	7.6145	7.5124
	75	11.6028	6.8704	4.9880	6.8016	6.8704	3.9910	1.3565	1.2399	1.4212	1.3422
	150	4.2550	2.7359	2.1487	2.6573	2.7359	1.7398	1.2445	0.8584	0.9141	1.3821
	200	2.9788	2.0283	1.7915	1.9654	2.0283	1.3807	1.1386	0.7508	0.7457	1.4067
	300	1.8794	1.3723	1.3391	1.3404	1.3723	1.0111	0.9488	0.6470	0.6303	1.2444
	500	0.9933	0.8027	0.8188	0.7932	0.8027	0.6600	0.6507	0.5011	0.5151	0.8203

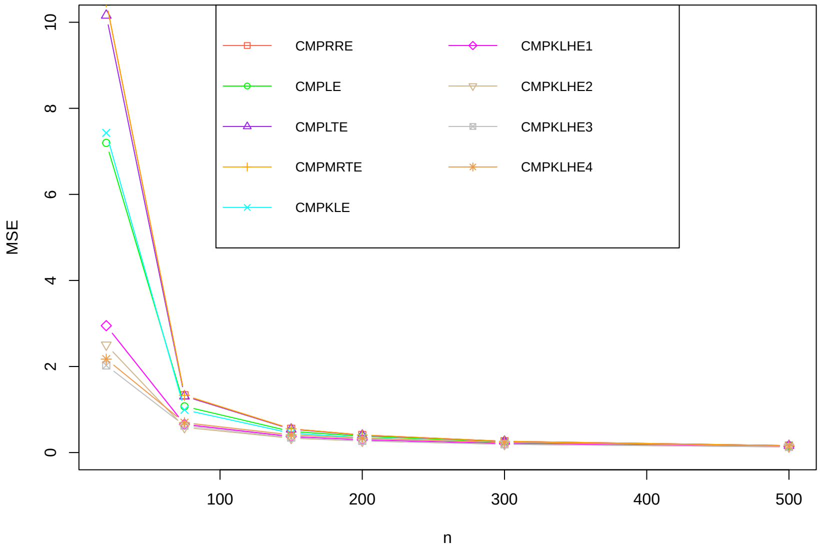

Figure 1: MSE values for different estimators by sample size.

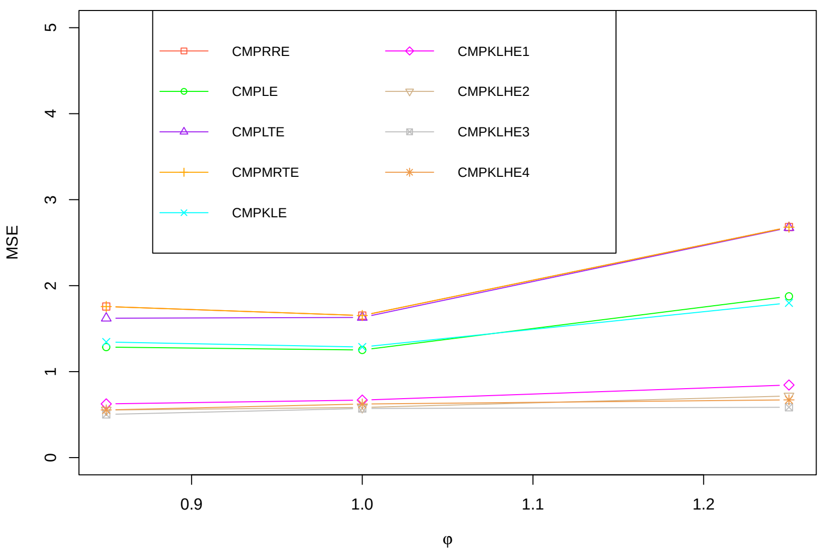

Figure 2: MSE values for different estimators by dispersion parameter.

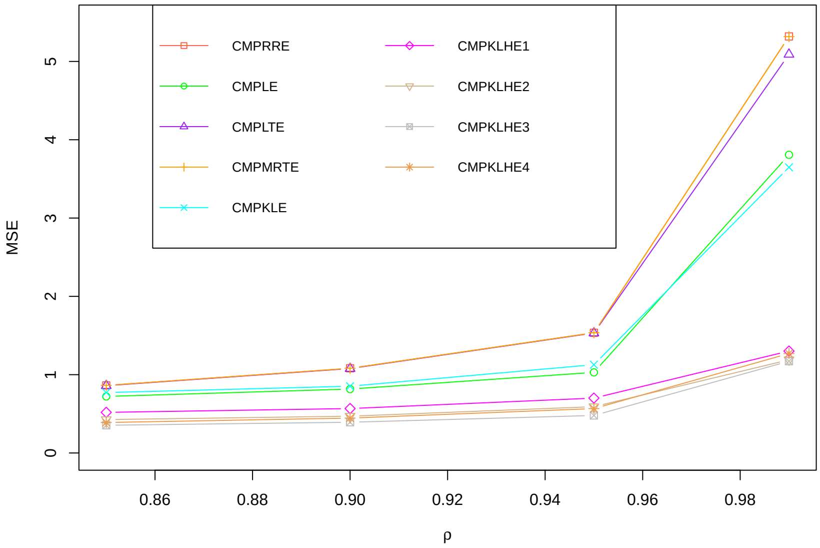

Figure 3: MSE values for different estimators by levels of correlation.

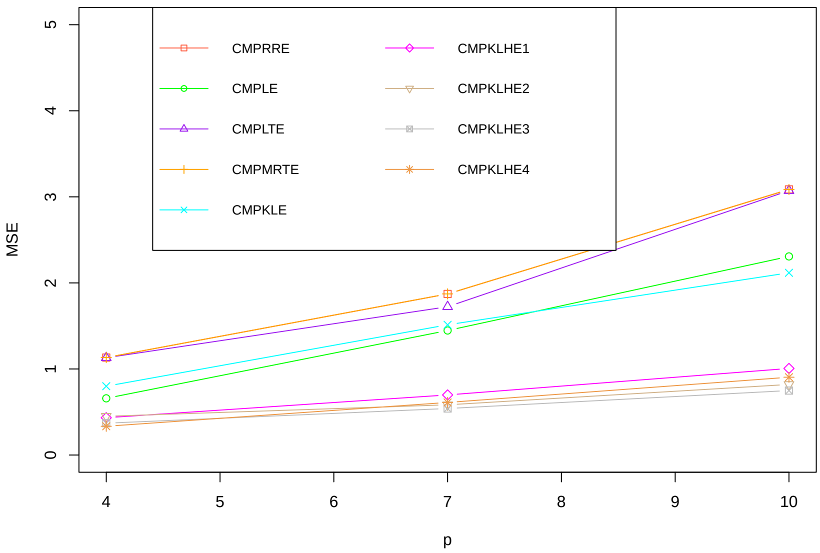

Figure 4: MSE values for different estimators by number of explanatory variables.

3.2 Simulation Result

Tables 2-10 demonstrate that the CMPKLHE estimator with four parameters consistently achieves the lowest MSE compared to other estimators in each case, making it the most robust and efficient choice, particularly when dealing with multicollinearity and higher values of ρ (correlation between explanatory variables). It consistently outperforms other estimators in all evaluated scenarios, with the CMPKLE estimator ranking second and the CMPMLE performing the weakest due to its sensitivity to multicollinearity. Increasing the sample size generally reduces MSE for all estimators, while a higher number of explanatory variables, larger dispersion parameters, and a high degree of multicollinearity typically result in higher MSEs. Among the CMPKLHE parameters, $\hat{k}_*^3$ consistently yields the lowest MSE, a trend observed across different values of p (number of explanatory variables), n (sample size), and φ (dispersion parameter). Choosing $\hat{k}_*^3$ as the CMPKLHE parameter is recommended for optimal performance. However, caution should be exercised when interpreting results with many explanatory variables or high dispersion, as these factors can inflate MSE.

Figures 1–4 present the MSE results for the estimators CMPRRE, CMPLE, CMPLTE, CMPMRTE, CMPKLE, CMPKLHE1, CMPKLHE2, CMPKLHE3, and CMPKLHE4 across different combinations of n, φ, ρ, and p. These figures demonstrate that the proposed estimator (CMPKLHE), with its four parameter values, consistently achieves the lowest MSE among all the estimators evaluated.

4 Applications

4.1 Sweden Football Data

In this analysis, we investigate the performance of various estimators using data from the 2023 season of the Allsvenskan league, Sweden's premier professional football division. The dataset comprises detailed match statistics, offering a comprehensive context for our study. The data set can be found in `https://www.football-data.co.uk/` and for more details about it, see Qasim et al. (2020a,b). Our main goal is to develop a model that predicts the number of goals scored by the home team (HG, y). The model considers a range of predictors related to match outcomes, including Pinnacle home win odds (PH, x_1), Pinnacle draw odds (PD, x_2), Pinnacle away win odds (PA, x_3), maximum home win odds (MaxH, x_4), market maximum draw odds (MaxD, x_5), maximum away win odds (MaxA, x_6), market average home win odds (AvgH, x_7), market average draw odds (AvgD, x_8), and market average away win odds (AvgA, x_9). The dataset contains 242 observations, each representing a match from the Allsvenskan league.

We initially fit several regression models for count data, including the Poisson, Negative Binomial (NB), and CMP distributions. We evaluated the models based on their log-likelihood (LL) and the Akaike Information Criteria (AIC) to identify the most suitable model. The AIC (LL) values for the Poisson, NB, and CMP models were 720.11 (-350.058), 722.12 (-350.059), and 718.42 (-348.21), respectively. The CMP model exhibited the lowest AIC value, indicating its superior fit to our dataset. Furthermore, the estimated dispersion value was 1.319, indicating underdispersion in the data. We examined the correlation coefficients between the explanatory variables and the condition number (CN) to assess multicollinearity. Table 11 shows high correlation coefficients among the explanatory variables, and the CN value is 443.91. These findings collectively indicate the presence of multicollinearity.

Table 11: Correlation matrix for explanatory variables in the Sweden football data.

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9
x_1	1.0000								
x_2	-0.1947	1.0000							
x_3	-0.6237	0.8384	1.0000						
x_4	0.9967	-0.1833	-0.6159	1.0000					
x_5	-0.1940	0.9958	0.8429	-0.1827	1.0000				
x_6	-0.5783	0.8573	0.9924	-0.5714	0.8639	1.0000			
x_7	0.9982	-0.1979	-0.6279	0.9978	-0.1972	-0.5829	1.0000		
x_8	-0.1905	0.9967	0.8347	-0.1787	0.9962	0.8506	-0.1937	1.0000	
x_9	-0.6286	0.8400	0.9978	-0.6214	0.8456	0.9913	-0.6334	0.8385	1.0000

Table 12: Estimated coefficients and MSEs of the different estimators for the Sweden football data.

Estimator	Parameter	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$	$\hat{\beta}_6$	$\hat{\beta}_7$	$\hat{\beta}_8$	$\hat{\beta}_9$	MSE
CMPMLE	-	-0.2980	-0.9868	1.2751	-0.4289	0.5334	-1.7917	0.0639	0.0398	1.1985	0.2670	8.3377
CMPRRE	$\hat{k}_R$	-0.2120	-0.7338	1.1477	-0.3925	0.4721	-1.4419	0.0237	-0.1641	0.9438	0.2709	4.4133
CMPLE	$\hat{d}_L$	-0.1297	-0.4783	0.8772	-0.3258	0.3002	-0.8834	-0.0347	-0.2470	0.6142	0.2615	3.2634
CMPLTE	$\hat{k}_{LT}/\hat{d}_{LT}$	-0.2120	-0.7338	1.1477	-0.3925	0.4721	-1.4419	0.0237	-0.1641	0.9438	0.2709	4.4133
CMPMRTE	$\hat{k}_M/\hat{d}_M$	-0.2120	-0.7338	1.1477	-0.3925	0.4721	-1.4419	0.0237	-0.1641	0.9438	0.2709	4.4133
CMPKLE	$\hat{k}_{KL}$	-0.1259	-0.4807	1.0204	-0.3561	0.4108	-1.0922	-0.0166	-0.3680	0.6890	0.2747	3.0657
CMPKLHE	$\hat{k}_*^1$	-0.1018	-0.4126	0.9143	-0.3323	0.3320	-0.8767	-0.0393	-0.3529	0.5636	0.2738	2.8899
	$\hat{k}_*^2$	-0.1119	-0.4429	0.9481	-0.3403	0.3546	-0.9478	-0.0317	-0.3455	0.6063	0.2743	2.8718
	$\hat{k}_*^3$	-0.0979	-0.4008	0.9003	-0.3290	0.3226	-0.8479	-0.0423	-0.3551	0.5466	0.2734	2.9178
	$\hat{k}_*^4$	-0.1048	-0.4215	0.9245	-0.3347	0.3389	-0.8980	-0.0370	-0.3510	0.5762	0.2740	2.8767

Table 12 presents the estimated coefficients and MSEs for different estimators applied to the Sweden football data. The CMPMLE estimator shows the highest MSE (8.3377), indicating relatively poor performance. The CMPRRE ($\hat{k}_R$), CMPLTE ($\hat{k}_{LT}/\hat{d}_{LT}$), and CMPMRTE ($\hat{k}_M/\hat{d}_M$) estimators have significantly lower MSEs (4.4133), reflecting better performance. The CMPLE ($\hat{d}_L$) and CMPKLE ($\hat{k}_{KL}$) estimators show further improvements with MSEs of 3.2634 and 3.0657, respectively. Among the CMPKLHE estimators, $\hat{k}_*^2$ achieves the lowest MSE (2.8718), indicating the best overall performance, closely followed by $\hat{k}_*^4$ (2.8767). These results suggest that the CMPKLHE estimator, particularly $\hat{k}_*^2$, provides the most accurate parameter estimates and effectively handles multicollinearity in the dataset.

4.2 Aircraft Damage Data

In this study, we evaluate the performance of the proposed estimator using real-world aircraft damage data. This dataset, originally analyzed by Myers et al. (2012) and later utilized by Abonazel et al. (2023b) and Tanış and Asar (2024), provides insights into damage incidents involving the Grumman A-6 Intruder and McDonnell Douglas A-4 Skyhawk aircraft. With a sample size of $n = 30$ observations, the data set includes a response variable, representing the number of damaged locations (y) and three explanatory variables: aircraft type (x_1), bomb load in tons (x_2) and aircrew experience in months (x_3). This dataset offers a comprehensive context for our analysis.

Abonazel et al. (2023b) and Tanış and Asar (2024) employed a COMP regression model with an intercept using the `glm.cmp` function from the `COMPoissonReg` package in R. The dispersion parameter ν was estimated to be 1.4578, indicating under-dispersion in the data. The condition number, calculated as the square root of the ratio of the maximum eigenvalue to the minimum eigenvalue of the matrix $X^t\hat{U}X$, was found to be 285.5439, highlighting a serious multicollinearity problem in the data set. The eigenvalues of $X^t\hat{U}X$ were 77,920.2111, 127.4488, 3.5176, and 0.9557.

Table 13: Estimated coefficients and MSEs of the different estimators for the aircraft damage data.

Estimator	Parameter	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	MSE
CMPMLE	-	-0.2110	0.6302	0.2215	-0.0180	1.8765
CMPRRE	$\hat{k}_R$	-0.0798	0.3351	0.2388	-0.0190	0.2819
CMPLE	$\hat{d}_L$	-0.1407	0.5033	0.2286	-0.0185	0.6351
CMPLTE	$\hat{k}_{LT}/\hat{d}_{LT}$	-0.0576	0.2852	0.2417	-0.0191	0.2292
CMPMRTE	$\hat{k}_M/\hat{d}_M$	-0.0798	0.3351	0.2388	-0.0190	0.2819
CMPKLE	$\hat{k}_{KL}$	0.0515	0.0400	0.2560	-0.0200	0.8825
CMPKLHE	$\hat{k}_*^1$	-0.0463	0.2771	0.2422	-0.0193	0.1717
	$\hat{k}_*^2$	-0.1019	0.4337	0.2325	-0.0188	0.1703
	$\hat{k}_*^3$	-0.0772	0.3700	0.2365	-0.0190	0.1338
	$\hat{k}_*^4$	-0.1038	0.4384	0.2323	-0.0188	0.1763

Table 13 presents the estimated coefficients and MSEs for different estimators applied to the aircraft damage data, evaluated for their performance in handling multicollinearity and providing accurate parameter estimates. The Maximum Likelihood Estimator (MLE) exhibits the highest MSE (1.87648), indicating poor performance in multicollinearity. The CMPRRE ($\hat{k}_R$) shows significant improvement with a much lower MSE (0.28189), indicating better handling of multicollinearity. The CMPLE ($\hat{d}_L$) performs better than the MLE with an MSE of 0.63514, although it is less effective than the

CMPRRE. The CMPLTE ($\hat{k}_{LT}/\hat{d}_{LT}$) demonstrates superior performance with the lowest MSE (0.22924) among MLE, CMPRRE, and CMPLE, suggesting high effectiveness in mitigating multicollinearity. The CMPMRTE ($\hat{k}_M/\hat{d}_M$) shows similar performance to the CMPRRE with an MSE of 0.28189, indicating good handling of multicollinearity. The CMPKLE ($\hat{k}_{KL}$) has a moderate MSE (0.88250), showing some effectiveness in addressing multicollinearity. Among all estimators, the CMPKLHE $\hat{k}_*^3$ achieves the lowest MSE (0.1338), indicating the most accurate parameter estimates and excellent handling of multicollinearity. Other parameters for CMPKLHE ($\hat{k}_*^1$, $\hat{k}_*^2$, and $\hat{k}_*^4$) show lower MSEs than other estimators. Finally, the CMPKLHE $\hat{k}_*^3$ is the best-performing estimator for the aircraft damage data, providing the most accurate parameter estimates and effectively addressing multicollinearity. At the same time, $\hat{k}_*^2$ is also a strong alternative among the CMPKLHE parameters.

5 Conclusion

This study addresses two critical challenges in count data modeling: dispersion and multicollinearity. The Conway-Maxwell-Poisson Regression (CMPR) model offers a concurrent solution to both issues. Specifically, the novel Kibria-Lukman Hybrid Estimator (KLHE) for the CMPR model demonstrates superior performance, yielding the lowest Mean Squared Error (MSE) compared to alternative estimators. Our simulation studies reveal that the level of multicollinearity, sample size, and the number of explanatory variables significantly influence the efficacy of the CMPKLHE and other estimators. Furthermore, analyses of real-world datasets consistently support the CMPKLHE's effectiveness. In conclusion, both simulation and empirical data underscore the CMPKLHE's superiority over conventional methods such as CMPMLE, CMPRRE, CMPLE, CMPLTE, CMPMRTE, and CMPKLE. Therefore, practitioners confronted with multicollinearity in CMPR modeling are advised to adopt the CMPKLHE for improved estimation accuracy and model performance. Future research could explore several promising avenues. One key area is the development of advanced methods for selecting shrinkage parameters within the CMPKLHE estimator, leveraging insights from Uslu et al. (2014) and Inan et al. (2017). Additionally, there is potential for applying these methods to other regression models, such as the negative binomial, zero-inflated Poisson, and Bell regression models. Another critical direction is to address the current limitations of the estimator in managing multicollinearity. Future studies could focus on incorporating robust biased estimators capable of simultaneously handling both multicollinearity and outliers in the CMPR model. This would build upon the methodologies proposed by Awwad et al. (2022) and Dawoud and Abonazel (2021).

Data Availability Statement

The data that supports the findings of this study are available within the article.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgements

The authors thank the peer reviewers for their support.

6 Appendix

Table 14: Validation of the prerequisite conditions in theorems for aircraft damage data.

Condition	s_1	s_2	s_3	s_4	s_5
Theorem 2.1	0.0000046	0.0015912	0.1230154	0.2748498	1.0514686
Theorem 2.2	0.0000046	0.0015771	0.0621030	0.0720548	0.0670022
Theorem 2.3	0.0000046	0.0015861	0.0944035	0.1583330	0.2492476
Theorem 2.4	0.0000046	0.0015745	0.0531481	0.0480335	0.0143464
Theorem 2.5	0.0000046	0.0015771	0.0621030	0.0720548	0.0670022
Theorem 2.6	0.0000046	0.0015630	0.0200745	0.0133013	0.2551654

The results in Table 14 present the validation of the prerequisite conditions as outlined in theorems for the aircraft damage data. The successful validation of these conditions is crucial for confirming the efficiency of the proposed estimator.

- **Theorem 2.1:** The necessary condition is $(s_j + k_*)^2(s_j + k_*(1 + d_*))^2 - s_j(s_j - k_*)^2 > 0$.
- **Theorem 2.2:** The necessary condition is $s_j(s_j + k_*)^2(s_j + k_*(1 + d_*))^2 - (s_j + k_R)^2(s_j - k_*)^2 > 0$.
- **Theorem 2.3:** The necessary condition is $(s_j + d_L)^2(s_j + k_*)^2(s_j + k_*(1 + d_*))^2 - s_j(s_j + 1)^2(s_j - k_*)^2 > 0$.
- **Theorem 2.4:** The necessary condition is $(s_j - d_{LT})^2(s_j + k_*)^2(s_j + k_*(1 + d_*))^2 - s_j(s_j + k_{LT})^2(s_j - k_*)^2 > 0$.
- **Theorem 2.5:** The necessary condition is $(s_j - k_{KL})^2(s_j + k_*)^2(s_j + k_*(1 + d_*))^2 - s_j(s_j + k_{KL})^2(s_j - k_*)^2 > 0$.
- **Theorem 2.6:** The necessary condition is $s_j(s_j + k_*)^2(s_j + k_*(1 + d_*))^2 - (s_j + k_M(d_M + 1))^2(s_j - k_*)^2 > 0$.

The results displayed in the table indicate that each theorem's condition has been met for all variables (s_1 to s_5), thereby validating the applicability of the proposed estimator under the specified conditions. This validation is an essential step in ensuring the robustness and reliability of the model's performance when applied to the aircraft damage data.

Table 15: Validation of the prerequisite conditions in theorems for Sweden football data.

Condition	s_1	s_2	s_3	s_4	s_5	s_6	s_7	s_8	s_9	s_10
Theorem 2.1	0.0000162	0.0002138	0.0052192	0.0242674	0.1013757	0.1192597	0.5194582	0.5363856	1.1786050	3.1249065
Theorem 2.2	0.0000162	0.0002137	0.0052059	0.0239704	0.0955319	0.1109606	0.3323631	0.3382872	0.5288923	0.9548507
Theorem 2.3	0.0000162	0.0002137	0.0051646	0.0230757	0.0799439	0.0894579	0.0641151	0.0596901	0.0209043	0.1223680
Theorem 2.4	0.0000162	0.0002137	0.0052059	0.0239704	0.0955319	0.1109606	0.3323631	0.3382872	0.5288923	0.9548507
Theorem 2.5	0.0000162	0.0002137	0.0052059	0.0239704	0.0955319	0.1109606	0.3323631	0.3382872	0.5288923	0.9548507
Theorem 2.6	0.0000162	0.0002137	0.0051926	0.0236752	0.0898447	0.1029268	0.1744286	0.1719876	0.0724026	0.0066890

Table 15 presents the results of validating the prerequisite conditions for various theorems using the Sweden football data. Each condition (`Theorem 2.1` through `Theorem 2.6`) corresponds to a different theorem, and the values in the table represent the computed statistics for these conditions across 10 different variables of s_j.
These values must satisfy specific inequalities to ensure the validity of the theorems being applied. The successful validation of these conditions indicates that the assumptions underlying the theorems hold for the Sweden football data, thus affirming the This structure provides a clear and professional presentation of your results and their interpretation.

References

Abonazel, M. R., Awwad, F. A., Tag Eldin, E., Kibria, B. M. G., and Khattab, I. G. (2023a). Developing a two-parameter liu estimator for the com–poisson regression model: application and simulation. *Frontiers in Applied Mathematics and Statistics*, 9:956963.

Abonazel, M. R., Saber, A. A., and Awwad, F. A. (2023b). Kibria–lukman estimator for the conway–maxwell poisson regression model: Simulation and applications. *Scientific African*, 19:e01553.

Akay, K. U. and Ertan, E. (2022). A new improved liu-type estimator for poisson regression models. *Hacettepe Journal of Mathematics and Statistics*, pages 1–20.

Akram, M. N., Amin, M., Afzal, N., and Kibria, B. M. G. (2023a). Kibria–lukman estimator for the zero inflated negative binomial regression model: theory, simulation and applications. *Communications in Statistics-Simulation and Computation*, pages 1–17.

Akram, M. N., Amin, M., Sami, F., Mastor, A. B., Egeh, O. M., Muse, A. H., et al. (2022). A new conway maxwell–poisson liu regression estimator—method and application. *Journal of Mathematics*, 2022.

Akram, M. N., Amin, M., Ullah, M. A., and Afzal, S. (2023b). Modified ridge-type estimator for the inverse gaussian regression model. *Communications in Statistics-Theory and Methods*, 52(10):3314–3332.

Algamal, Z. Y., Abonazel, M. R., Awwad, F. A., and Eldin, E. T. (2023). Modified jackknife ridge estimator for the conway-maxwell-poisson model. *Scientific African*, 19:e01543.

Alrweili, H. (2024). Kibria–lukman hybrid estimator for handling multicollinearity in poisson regression model: Method and application. *International Journal of Mathematics and Mathematical Sciences*, 2024(1):1053397.

Amin, M., Akram, M. N., and Majid, A. (2023). On the estimation of bell regression model using ridge estimator. *Communications in Statistics-Simulation and Computation*, 52(3):854–867.

Asar, Y. and Genç, A. (2018). A new two-parameter estimator for the poisson regression model. *Iranian Journal of Science and Technology, Transactions A: Science*, 42:793–803.

Awwad, F. A., Dawoud, I., and Abonazel, M. R. (2022). Development of robust özkale–kaçiranlar and yang–chang estimators for regression models in the presence of multicollinearity and outliers. *Concurrency and Computation: Practice and Experience*, 34(6):e6779.

Chatla, S. B. and Shmueli, G. (2018). Efficient estimation of com–poisson regression and a generalized additive model. *Computational Statistics & Data Analysis*, 121:71–88.

Conway, R. and Maxwell, W. (1962). Wl. a queuing model with state dependent service rates. *J Ind Eng*, 12:132–136.

Dawoud, I. and Abonazel, M. R. (2021). Robust dawoud–kibria estimator for handling multicollinearity and outliers in the linear regression model. *Journal of Statistical Computation and Simulation*, 91(17):3678–3692.

Ertan, E. and Akay, K. U. (2023). A new class of poisson ridge-type estimator. *Scientific Reports*, 13(1):4968.

Farebrother, R. W. (1976). Further results on the mean square error of ridge regression. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 248–250.

Francis, R. A., Geedipally, S. R., Guikema, S. D., Dhavala, S. S., Lord, D., and LaRocca, S. (2012). Characterizing the performance of the conway-maxwell poisson generalized linear model. *Risk Analysis: An International Journal*, 32(1):167–183.

Guikema, S. D. and Goffelt, J. P. (2008). A flexible count data regression model for risk analysis. *Risk Analysis: An International Journal*, 28(1):213–223.

Hoerl, A. E. and Kennard, R. W. (1970). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, 12(1):55–67.

Inan, D., Egrioglu, E., Sarica, B., Askin, O. E., and Tez, M. (2017). Particle swarm optimization based liu-type estimator. *Communications in Statistics-Theory and Methods*, 46(22):11358–11369.

Kejian, L. (1993). A new class of biased estimate in linear regression. *Communications* *in Statistics-Theory and Methods*, 22(2):393–402.

Kibria, B. M. G. and Lukman, A. F. et al. (2020). A new ridge-type estimator for the linear regression model: Simulations and applications. *Scientifica*, 2020.

Liu, K. (2003). Using liu-type estimator to combat collinearity. *Communications in Statistics-Theory and Methods*, 32(5):1009–1020.

Lukman, A. F., Kibria, B. M. G., Ayinde, K., and Jegede, S. L. (2020). Modified one-parameter liu estimator for the linear regression model. *Modelling and Simulation in Engineering*, 2020:1–17.

Lukman, A. F., Adewuyi, E., Månsson, K., and Kibria, B. M. G. (2021). A new estimator for the multicollinear poisson regression model: simulation and application. *Scientific Reports*, 11(1):3732.

Lukman, A. F., Aladeitan, B., Ayinde, K., and Abonazel, M. R. (2022). Modified ridge-type for the poisson regression model: simulation and application. *Journal of Applied Statistics*, 49(8):2124–2136.

Lukman, A. F., Ayinde, K., Binuomote, S., and Clement, O. A. (2019). Modified ridge-type estimator to combat multicollinearity: Application to chemical data. *Journal of Chemometrics*, 33(5):e3125.

Månsson, K. (2012). On ridge estimators for the negative binomial regression model. *Economic Modelling*, 29(2):178–184.

Månsson, K., Kibria, B. M. G., Sjölander, P., and Shukur, G. (2012). Improved liu estimators for the poisson regression model. *International Journal of Statistics and Probability*, 1(1):2.

Månsson, K. and Shukur, G. (2011). A poisson ridge regression estimator. *Economic Modelling*, 28(4):1475–1481.

Myers, R. H., Montgomery, D. C., Vining, G. G., and Robinson, T. J. (2012). *Generalized linear models: with applications in engineering and the sciences*. John Wiley & Sons.

Qasim, M., Kibria, B., Månsson, K., and Sjölander, P. (2020a). A new poisson liu regression estimator: method and application. *Journal of Applied Statistics*, 47(12):2258–2271.

Qasim, M., Månsson, K., Amin, M., Golam Kibria, B. M., and Sjölander, P. (2020b). Biased adjusted poisson ridge estimators-method and application. *Iranian Journal of Science and Technology, Transactions A: Science*, 44:1775–1789.

Sami, F., Amin, M., Akram, M. N., Butt, M. M., and Ashraf, B. (2022a). A modified one parameter liu estimator for conway-maxwell poisson response model. *Journal of Statistical Computation and Simulation*, 92(12):2448–2466.

Sami, F., Amin, M., and Butt, M. M. (2022b). On the ridge estimation of the conway-maxwell poisson regression model with multicollinearity: Methods and applications. *Concurrency and Computation: Practice and Experience*, 34(1):e6477.

Sami, F., Amin, M., Butt, M. M., and Yasin, S. (2023a). An almost unbiased ridge estimator for the conway–maxwell–poisson regression model. *Iranian Journal of Science*, 47(4):1209–1219.

Sami, F., Butt, M. M., and Amin, M. (2023b). Two parameter estimators for the conway–maxwell–poisson regression model. *Journal of Statistical Computation and Simulation*, 93(13):2137–2157.

Sellers, K. F. and Shmueli, G. (2010). A flexible regression model for count data. *The Annals of Applied Statistics*, pages 943–961.

Shewa, G. A. and Ugwuowo, F. I. (2023). A new hybrid estimator for linear regression model analysis: computations and simulations. *Scientific African*, 19:e01441.

Shmueli, G., Minka, T. P., Kadane, J. B., Borle, S., and Boatwright, P. (2005). A useful distribution for fitting discrete data: revival of the conway–maxwell–poisson distribution. *Journal of the Royal Statistical Society Series C: Applied Statistics*, 54(1):127–142.

Şiray, G. Ü., Toker, S., and Kaçiranlar, S. (2015). On the restricted liu estimator in the logistic regression model. *Communications in Statistics-Simulation and Computation*, 44(1):217–232.

stat Alrwali, H. (2024). Liu-type estimator for the poisson-inverse gaussian regression model: Simulation and practical applications. *Statistics, Optimization & Information Computing*.

Tanış, C. and Asar, Y. (2024). Liu-type estimator in conway–maxwell–poisson regression model: theory, simulation and application. *Statistics*, pages 1–22.

Uslu, V. R., Egrioglu, E., and Bas, E. (2014). Finding optimal value for the shrinkage parameter in ridge regression via particle swarm optimization. *American Journal of Intelligent Systems*, 4(4):142–147.

Zandi, Z., Arabi B., R., and Bevrani, H. (2022). Liu-type shrinkage strategies in zero-inflated negative binomial models with application to expenditure and default data. *Communications in Statistics-Simulation and Computation*, pages 1–31.