



OPEN

A bias-reduced estimator for generalized Poisson regression with application to carbon dioxide emission in Canada

Fatimah M. Alghamdi¹, Ahmed M. Gemeay², Gamal A. Abd-Elmougod³,
Ehab M. Almetwally⁴, M. A. El-Qurashi⁵, Getachew Tekle Mekiso⁶✉ & Ali T. Hammad²

The generalized Poisson regression model (GPRM) provides a flexible framework for modeling count data, especially those exhibiting over- or underdispersion. Although the generalized Poisson maximum likelihood estimator is considered the standard method for estimating the parameters of this model, its reliability and accuracy are severely affected by the presence of multicollinearity among explanatory variables. Multicollinearity inflates the variance of parameter estimates, undermining the validity of statistical inference and ultimately leading to unstable and unreliable estimators. To mitigate these problems, this study presents the ridge estimator as a robust alternative within the GPRM framework. Several new strategies are proposed for selecting the optimal value of the ridge parameter. The statistical properties of the proposed ridge estimator were theoretically studied. Theoretical comparisons and extensive Monte Carlo simulations demonstrated a clear and significant superiority of the ridge estimator under multicollinearity conditions, confirming its robustness and efficiency. To demonstrate the scientific and practical relevance of the proposed estimator, it was applied to a real-world case study modeling carbon dioxide emissions in Canada. The results of this experimental application conclusively confirmed the simulation and theoretical comparison results, with the ridge estimator providing more stable and interpretable results than the conventional method, making it a valuable tool for researchers and decision makers in analyzing multicollinear environmental and economic data.

Keywords Biased estimator, Generalized Poisson regression model, Multicollinearity, Ridge estimator, CO₂ emissions

The Poisson regression model (PRM) is a fundamental tool for analyzing count data, where the outcome variable consists of non-negative integers representing the frequency of events. A core assumption of the PRM is equidispersion—the equality of the mean and variance¹. However, empirical data frequently violate this assumption, displaying either overdispersion, where variance exceeds the mean, or the less common underdispersion, where variance is constrained below the mean. These patterns often emerge from unobserved heterogeneity, dependencies between events, or other latent factors, potentially leading to inefficient estimates and misleading inferences. To overcome these limitations, numerous alternative count models have been developed. These include the negative binomial regression model (NBRM), which incorporates a dispersion parameter to handle overdispersion^{2,3}; the geometric regression model (GERM) for underdispersed data⁴; and the Conway–Maxwell Poisson regression model (CMPRM), which offers flexibility for both types of dispersion⁵. Additional approaches include the double Poisson model (DPRM), which explicitly corrects for deviations from equidispersion⁶; the Bell regression model (BRM) for overdispersed counts; and the Poisson-inverse Gaussian model (PIGRM) for data with severe overdispersion and heavy tails⁷. Among these, the generalized Poisson

¹Department of Mathematical Sciences, College of Science, Princess Nourah bint Abdulrahman University, P.O. Box 84428, Riyadh 11671, Saudi Arabia. ²Department of Mathematics, Faculty of Science, Tanta University, Tanta 31527, Egypt. ³Department of Mathematics, Faculty of Science, Islamic University of Madinah, Madinah 42351, Saudi Arabia. ⁴Department of Mathematics and Statistics, College of Science, Imam Mohammad Ibn Saud Islamic University (IMSIU), Riyadh 11432, Saudi Arabia. ⁵Institute of Basic and Applied Science, College of Engineering and Technology, Arab Academy for Science, Technology and Maritime Transport, P.O. Box 1029, Abu Quir Campus, Egypt. ⁶Department of Statistics, Wachemo University, Hosaina, Ethiopia. ✉email: getachewtekle@wcu.edu.et

regression model (GPRM) stands out for its ability to directly adjust the mean-variance relationship, effectively addressing both over- and underdispersion, making it a versatile alternative when the standard Poisson assumptions are untenable⁸.

The GPRM effectively handles both over- or underdispersed count data, making it valuable in numerous fields⁹. For example, it has been used to study child stunting in Indonesia¹⁰, neonatal mortality in Ethiopia¹¹, and COVID-19 spread in Ghana¹². The GPRM has also inspired new methods, such as a stochastic process for signed integers¹³. These applications demonstrate the model's flexibility in addressing real-world data challenges.

The presence of multicollinearity, a high degree of intercorrelation among explanatory variables, poses a significant challenge to regression analysis. This condition detrimentally inflates the variance of parameter estimates derived from both ordinary least squares (OLS) and maximum likelihood estimation (MLE)¹⁴. The issue is particularly acute within the context of the GPRM, where MLE can produce excessively high variance in the presence of correlated predictors. To mitigate the adverse effects of multicollinearity, several techniques are commonly employed. These include variable selection methods like Lasso regression and the use of biased estimation techniques, such as Ridge regression¹⁵. Under conditions of multicollinearity, such biased estimators often outperform MLE by trading a small amount of bias for a substantial reduction in variance, thereby yielding a lower scalar mean squared error (MSE) and more reliable estimates. The ridge estimator, formulated as $\hat{\beta}_{RE} = (X'X + kI)^{-1}X'Y$ where X is the design matrix, Y the response vector, I the identity matrix, and $k > 0$ the ridge parameter, can be analyzed through its spectral properties to derive its MSE. The selection of an optimal k value is critical, and numerous methods have been developed for this purpose, building upon the foundational work of Hoerl and Kennard^{16,17} for nonorthogonal problems. Subsequent advancements include the alternative estimators proposed by Kibria¹⁸ and further methodological elaborations by Muniz and Kibria¹⁹.

Ridge regression has been widely used in generalized linear models (GLMs), with Segerstedt²⁰ being one of the first to apply it in GLMs. This technique has since been used in various models to deal with multicollinearity. For instance, Månsson and Shukur²¹, Schaefer et al.²², Rady et al.²³, Månsson²⁴, Tharshan et al.²⁵, Almulhim et al.²⁶, Abonazel et al.²⁷, Sami et al.²⁸, Algamal et al.²⁹, Dawoud³⁰, El-Alosey et al.³¹, Akram et al.³², Shahzad et al.³³, and Ashraf et al.³⁴.

A substantial body of literature has explored the development and application of ridge estimators and other biasing parameters across a range of statistical frameworks, from classical linear regression to various count and GLMs. Despite these extensive developments, the integration of ridge regression methodology within the GPRM remains notably understudied. This research aims to address this gap by systematically evaluating the efficacy of ridge regression in ameliorating the dual challenges of multicollinearity and overdispersion within the GPRM framework and derivation and evaluation of optimal estimators for the ridge parameter k . The performance of the resulting ridge estimators will be rigorously compared against the conventional generalized Poisson maximum likelihood estimates (GPMLE). Theoretical comparisons and Monte Carlo simulations will be used to study the performance of the proposed estimator, and these findings will be further validated through a real-world application.

This paper is organized in the following: Section "Generalized Poisson regression model" provides the statistical formulation of the Generalized Poisson (GP) distribution and its corresponding regression model. Section "Generalized Poisson ridge regression estimator" presents the methodological framework for addressing multicollinearity within the GPRM through the application of ridge regression. The performance criteria used to evaluate the proposed estimators are defined in Sect. "The superiority of the GPRRE over the GPMLE". Section "Selection of the biasing parameters" is devoted to a discussion of the methods for selecting the optimal ridge parameter, k . The efficacy of the proposed methodology is then rigorously assessed via an extensive Monte Carlo simulation study in Sect. "Monte Carlo simulation". To demonstrate its practical utility, the approach is applied to empirical datasets in Sect. "Application". Finally, the principal findings, along with concluding remarks and potential avenues for future research, are summarized in Sect. "Conclusion".

Generalized Poisson regression model

The GP distribution initially introduced by Consul and Jain³⁵, which has parameters θ and ν . This distribution's probability mass function (PMF) is defined as:

$$P(Y = y; \theta, \nu) = \frac{\theta}{y!} (\theta + \nu y)^{y-1} e^{-(\theta + \nu y)}, \quad (1)$$

where $y = 0, 1, 2, \dots$, $\theta > 0$, and $\max(-1, -\theta/4) \leq \nu \leq 1$. The GP distribution was initially introduced as an approximation to the generalized negative binomial distribution. Its properties were subsequently extensively developed by Consul³⁶, whose foundational work led to it being commonly referred to as Consul's GP distribution. The mean and variance of the distribution are given by:

$$E(Y) = \frac{\theta}{1 - \nu},$$

and

$$\text{Var}(Y) = \frac{\theta}{(1 - \nu)^3}.$$

Famoye³⁷ suggested a more appropriate parametrization of the GP distribution for regression models by reparameterizing the original formulation in Eq. (1). Specifically, let $\mu = \frac{\theta}{1 - \nu}$ and $\varphi = \frac{\nu}{\theta}$. This transformation

leads to the relationships $\theta = \frac{\mu}{1+\varphi\mu}$ and $\nu = \frac{\varphi\mu}{1+\varphi\mu}$. Under this reparameterization, the PMF for the GP distribution, denoted as $GP(\mu, \varphi)$, is given by:

$$P(Y = y; \mu, \varphi) = \frac{(1 + \varphi y)^{y-1}}{y!} \left(\frac{\mu}{1 + \varphi \mu} \right)^y \exp \left(-\frac{\mu(1 + \varphi y)}{1 + \varphi \mu} \right), \quad \mu > 0, y = 0, 1, 2, \dots \quad (2)$$

This parameterization is very good for regression since it is easy to understand and fits well with modeling frameworks. The GP distribution is great for analyzing count data because it can handle over- or underdispersion. The mean and variance of $GP(\mu, \varphi)$ are:

$$E(Y|x_i) = \mu_i,$$

and

$$\text{Var}(Y|x_i) = \mu_i(1 + \varphi\mu_i)^2.$$

The GPRM provides a flexible extension of the standard PRM, designed to handle count data characterized by either over- or underdispersion. The model is constructed within the GLM framework. In this specification, the mean of the response variable, denoted by μ_i , is connected to a linear combination of predictors through a logarithmic link function, expressed as:

$$\log \mu_i = \eta_i = x_i' \beta, \quad (3)$$

where μ_i denotes the expected count for the i -th observation. This mean is modeled as a function of a $p \times 1$ vector of explanatory variables, x_i , and a corresponding p -dimensional vector of unknown regression coefficients, β . The coefficients quantify the relationship between the explanatory variables and the expected value of the response variable. The MLE method is commonly used to estimate the parameters of the GPRM. In this approach, we maximize the likelihood function $L(\beta, \varphi)$, which expresses the probability of observing the data given the model parameters. The log-likelihood function $\ell(\beta, \varphi)$ is given by:

$$\ell(\beta, \varphi) = \sum_{i=1}^n \left[y_i [\log(\mu_i) - \log(1 + \varphi\mu_i)] + (y_i - 1) \log(1 + \varphi y_i) - \log(y_i!) - \frac{\mu_i(1 + \varphi y_i)}{1 + \varphi\mu_i} \right]. \quad (4)$$

Eq.(4) represent the GPMLE for the parameters φ (dispersion parameter) and β_r (regression coefficients) in a GPRM:

$$\frac{\partial \ell}{\partial \varphi} = \sum_{i=1}^n \left(\frac{-y_i \mu_i}{1 + \varphi \mu_i} + \frac{y_i(y_i - 1)}{1 + \varphi y_i} - \frac{\mu_i(y_i - \mu_i)}{(1 + \varphi \mu_i)^2} \right) = 0, \quad (5)$$

$$\frac{\partial \ell}{\partial \beta_r} = \sum_{i=1}^n \frac{y_i - \mu_i}{\mu_i(1 + \varphi \mu_i)^2} \frac{\partial \mu_i}{\partial \beta_r} = 0. \quad (6)$$

Since the likelihood equation for the regression coefficients, β , is nonlinear, the Iterative Weighted Least Squares (IWLS) algorithm (also known as the Fisher Scoring method) proposed by Dutang³⁸ is employed to derive the MLE. Let $\beta^{(s-1)}$ denote the vector of regression coefficients estimated at the $(s-1)$ -th iteration. The coefficient vector is subsequently updated at the s -th iteration according to the following rule:

$$\beta^{(s)} = \beta^{(s-1)} + I^{-1}(\beta^{(s-1)}) S(\beta^{(s-1)}), \quad (7)$$

where $S(\beta^{(s-1)})$ is the score function evaluated at $\beta^{(s-1)}$ and $I(\beta^{(s-1)})$ is the Fisher information matrix. In the final step of the algorithm, the GPMLE for the regression coefficients, $\hat{\beta}_{\text{GPMLE}}$, is given by:

$$\hat{\beta}_{\text{GPMLE}} = A^{-1} X' \hat{W} \hat{z}, \quad (8)$$

where $A = X' \hat{W} X$, $\hat{z}$ is the adjusted response vector, and $\hat{W}$ is a diagonal weight matrix with diagonal elements $\hat{w}_i$. The diagonal elements of $\hat{W}$ are $\hat{w}_i = \frac{\hat{\mu}_i}{(1 + \varphi \hat{\mu}_i)^2}$, and the elements of the adjusted response vector $\hat{z}$ are $\hat{z}_i = \log(\hat{\mu}_i) + \frac{y_i - \hat{\mu}_i}{\hat{\mu}_i}$.

The asymptotic covariance matrix, matrix mean squared error (MMSE), and the MSE for GPMLE are given by:

$$\text{Cov}(\hat{\beta}_{\text{GPMLE}}) = \left[-E \left(\frac{\partial^2 \ell}{\partial \beta \partial \beta'} \right) \right] = \hat{\varphi} Q \Lambda^{-1} Q', \quad (9)$$

$$\begin{aligned} \text{MMSE}(\hat{\beta}_{\text{GPMLE}}) &= E[(\hat{\beta}_{\text{GPMLE}} - \beta)(\hat{\beta}_{\text{GPMLE}} - \beta)'] \\ &= \text{Cov}(\hat{\beta}_{\text{GPMLE}}) + (E(\hat{\beta}_{\text{GPMLE}}) - \beta)(E(\hat{\beta}_{\text{GPMLE}}) - \beta)' = \hat{\varphi} Q \Lambda^{-1} Q', \end{aligned} \quad (10)$$

$$\text{MSE}(\hat{\beta}_{\text{GPMLE}}) = \text{Tr}(\text{MMSE}(\hat{\beta}_{\text{GPMLE}})) = \hat{\varphi} \sum_{j=1}^p \frac{1}{\lambda_j}. \quad (11)$$

where Tr is the trace of the matrix, $\hat{\varphi}$ is the estimated dispersion parameter, the matrix A is expressed as $A = Q \Lambda Q'$, $\Lambda = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_p) = Q A Q'$ with Q is an orthogonal matrix whose columns represent the eigenvectors of $Q A Q'$, and λ_j is the j th eigenvalue of the A matrix. When the explanatory variables in the GPRM are highly correlated, the weighted cross-product matrix A becomes unstable. This leads to inefficient estimates from the GPMLE, with large variances. As a result, the estimated coefficients are often too large, making them difficult to interpret.

Generalized Poisson ridge regression estimator

Segerstedt²⁰ introduced the ridge estimator for GLMs as a solution to multicollinearity, building on the foundational work of Hoerl and Kennard^{16,17}. When the explanatory variables in the GPRM are highly correlated, the GPMLE produces inefficient estimates characterized by a large MSE. Following the contributions of Månsson and Shukur²¹, Sami et al.²⁸, Shahzad et al.³³, and Ashraf et al.³⁴, this paper introduces a ridge estimator extended to the GPRM, referred to as the Generalized Poisson Ridge Regression Estimator (GPRRE). Its formulation is expressed as:

$$\hat{\beta}_{k^*} = (A + k^* I_p)^{-1} S \hat{\beta}_{\text{GPMLE}}, \quad (12)$$

where $k^* (k^* > 0)$ is the ridge parameter, I_p is the identity matrix, and if $k^* = 0$ then the GPRRE is reduced to GPMLE.

The bias vector and variance-covariance matrix of the GPRRE are given by:

$$\text{Bias}(\hat{\beta}_{k^*}) = E(\hat{\beta}_{k^*}) - \beta = -k^* Q \Lambda_{k^*}^{-1} \alpha, \quad (13)$$

$$\text{Cov}(\hat{\beta}_{k^*}) = E[(\hat{\beta}_{k^*} - E(\hat{\beta}_{k^*}))(\hat{\beta}_{k^*} - E(\hat{\beta}_{k^*}))'] = \hat{\varphi} Q \Lambda_{k^*}^{-1} \Lambda \Lambda_{k^*}^{-1} Q'. \quad (14)$$

The MMSE and MSE for the GPRRE can be computed using Eqs. (13) and (14) as follows:

$$\begin{aligned} \text{MMSE}(\hat{\beta}_{k^*}) &= E[(\hat{\beta}_{k^*} - \beta)(\hat{\beta}_{k^*} - \beta)'] = \text{Cov}(\hat{\beta}_{k^*}) + \text{Bias}(\hat{\beta}_{k^*}) \text{Bias}'(\hat{\beta}_{k^*}) \\ &= \hat{\varphi} Q \Lambda_{k^*}^{-1} \Lambda \Lambda_{k^*}^{-1} Q' + k^{*2} Q \Lambda_{k^*}^{-1} \alpha \alpha' \Lambda_{k^*}^{-1} Q', \end{aligned} \quad (15)$$

$$\text{MSE}(\hat{\beta}_{k^*}) = \text{Tr}[\text{MMSE}(\hat{\beta}_{k^*})] = \hat{\varphi} \sum_{j=1}^p \frac{\lambda_j}{(\lambda_j + k^*)^2} + k^{*2} \sum_{j=1}^p \frac{\alpha_j^2}{(\lambda_j + k^*)^2}, \quad (16)$$

where the vector $\alpha = Q^T \beta$, and $\Lambda_{k^*} = \text{diag}(\lambda_1 + k^*, \lambda_2 + k^*, \dots, \lambda_p + k^*) = Q(A + kI)Q'$.

The superiority of the GPRRE over the GPMLE

To assess the superiority of the GPRRE compared to GPMLE, Hoerl and Kennard¹⁶ proposed theoretical results regarding the properties of the MSE for ridge regression estimators in the linear regression model. In this study, we demonstrate that these results are also applicable to the GPRM. Based on these theorems, we will investigate the superiority of the GPRRE over the GPMLE.

Theorem 1 The variance $D_1(k^*)$ and the squared bias $D_2(k^*)$ are continuous functions of k^* , where $D_1(k^*)$ is monotonically decreasing and $D_2(k^*)$ is monotonically increasing, provided that $k^* > 0$ and $\lambda_j > 0$.

Proof Using Eq. 16, we are given the following expressions for the variance and squared bias:

$$D_1(k^*) = \hat{\varphi} \sum_{j=1}^p \frac{\lambda_j}{(\lambda_j + k^*)^2}, \quad D_2(k^*) = k^{*2} \sum_{j=1}^p \frac{\alpha_j^2}{(\lambda_j + k^*)^2}.$$

1. Monotonicity of $D_1(k^*)$: The derivative of $D_1(k^*)$ with respect to k^* is:

$$\frac{dD_1(k^*)}{dk^*} = -\hat{\varphi} \sum_{j=1}^p \frac{2\lambda_j}{(\lambda_j + k^*)^3},$$

since $\lambda_j > 0$ and $k^* > 0$, we conclude that $\frac{dD_1(k^*)}{dk^*} < 0$, implying $D_1(k^*)$ is monotonically decreasing.

2. Monotonicity of $D_2(k^*)$: The derivative of $D_2(k^*)$ with respect to k^* is:

$$\frac{dD_2(k^*)}{dk^*} = 2k^* \sum_{j=1}^p \frac{\alpha_j^2}{(\lambda_j + k^*)^3} \lambda_j,$$

since $\lambda_j > 0$ and $k^* > 0$, we have $\frac{dD_2(k^*)}{dk^*} > 0$, implying $D_2(k^*)$ is monotonically increasing.

Thus, $D_1(k^*)$ is monotonically decreasing and $D_2(k^*)$ is monotonically increasing for $k^* > 0$. $\square$

Theorem 2 For the GPRM, the GPRRE is more efficient than the GPMLE if

$$\text{MSE}(\hat{\beta}_{\text{GPMLE}}) - \text{MSE}(\hat{\beta}_{k^*}) > 0 \quad \text{for all } k^* > 0 \quad \text{and } \lambda_j > 0.$$

Proof For $D_1(k^*)$ when $k^* = 0$, we have:

$$D_1(k^*) = \hat{\varphi} \sum_{j=1}^p \frac{1}{\lambda_j},$$

which equals $\text{MSE}(\hat{\beta}_{k^*})$. The difference between $\text{MSE}(\hat{\beta}_{k^*})$ and $\text{MSE}(\hat{\beta}_{\text{GPMLE}})$ is:

$$\Delta = \hat{\varphi} \sum_{j=1}^p \frac{1}{\lambda_j} - \left(\hat{\varphi} \sum_{j=1}^p \frac{\lambda_j}{(\lambda_j + k^*)^2} + \sum_{j=1}^p k^{*2} \frac{\alpha_j^2}{(\lambda_j + k^*)^2} \right) = \sum_{j=1}^p \frac{k^* (\hat{\varphi} k^* + 2\hat{\varphi} \lambda - k^* \lambda \alpha_j^2)}{\lambda_j (\lambda_j + k^*)^2},$$

for any $k^* > 0$, then $\Delta > 0$ if and only if $\hat{\varphi} k^* + 2\hat{\varphi} \lambda - k^* \lambda \alpha_j^2 > 0$. Consequently, $\text{MSE}(\hat{\beta}_{\text{GPMLE}}) - \text{MSE}(\hat{\beta}_{k^*}) > 0$ holds under the same condition, i.e., $\hat{\varphi} k^* + 2\hat{\varphi} \lambda - k^* \lambda \alpha_j^2 > 0$. $\square$

Selection of the biasing parameters

The ridge parameter (k) is a critical component of the ridge regression estimator, as its value directly governs the degree of shrinkage and bias introduced to stabilize the coefficient estimates. Consequently, the selection of an optimal value for this shrinkage parameter has become a central challenge in the application of the ridge methodology. This is particularly vital in the presence of multicollinearity, where high correlations among explanatory variables can severely degrade the performance of standard estimators. In response, a significant body of research has been dedicated to developing methods for estimating the optimal k across diverse regression frameworks, such as those of Månsson and Shukur²¹, Schaefer et al.²², Rady et al.²³, Tharshan and Wijekoon²⁵, Algarnal et al.²⁹, Akram et al.³².

The foundational work on this technique was established by^{16,17}, who first proposed ridge estimation to mitigate multicollinearity in linear regression models. Their approach has since been successfully generalized to a wider array of models, including the gamma regression model³² and the zero-inflated CMPRM³⁴. Within the context of the GPRM, the resulting estimator is designated as the GPRRE. For our analysis, following the works of^{16,17} and²¹, we adopt the following values for k^* :

$$\hat{k}_1^* = \frac{\hat{\varphi}}{\sum_{j=1}^p \hat{\alpha}_j^2}, \quad (17)$$

$$\hat{k}_2^* = \frac{\hat{\varphi}}{\max(\hat{\alpha}_j^2)}, \quad (18)$$

$$\hat{k}_3^* = \min \left(\frac{\hat{\varphi}}{\hat{\alpha}_j^2} \right). \quad (19)$$

Following Ashraf et al.³⁴, we use the following values for k^* :

$$q_j = \frac{\hat{\varphi}}{2\lambda_j \hat{\alpha}_j^2 + \hat{\varphi}}, \quad (20)$$

$$\hat{k}_4^* = \text{median}(q_j), \quad (21)$$

$$\hat{k}_5^* = \text{median} \left(\frac{\lambda_j}{1 + 2\lambda_j \hat{\alpha}_j^2} \right). \quad (22)$$

Following Shahzad et al.³³, we use the following values for k^* :

$$\hat{k}_6^* = \min \left(\frac{\lambda_j \hat{\varphi}}{\hat{\varphi} + \lambda_j \hat{\alpha}_j^2} \right), \quad (23)$$

$$\hat{k}_7^* = \frac{1}{\max(\hat{\alpha}_j^2)}. \quad (24)$$

Following Tharshan et al.²⁵, we use the following values for k^* :

$$\hat{k}_8^* = \text{median} \left(\sqrt{\frac{1}{\hat{\alpha}_j^2}} \right), \quad (25)$$

$$\hat{k}_9^* = \max \left(\frac{1}{\sqrt{\frac{1}{\hat{\alpha}_j^2}}} \right). \quad (26)$$

Following Sami et al.²⁸, we use the following value for k^* :

ρ	n	GPMLE	GPRRE														
		–	$\hat{k}_1^*$	$\hat{k}_2^*$	$\hat{k}_3^*$	$\hat{k}_4^*$	$\hat{k}_5^*$	$\hat{k}_6^*$	$\hat{k}_7^*$	$\hat{k}_8^*$	$\hat{k}_9^*$	$\hat{k}_{10}^*$	$\hat{k}_{11}^*$	$\hat{k}_{12}^*$	$\hat{k}_{13}^*$	$\hat{k}_{14}^*$	$\hat{k}_{15}^*$
0.80	50	0.2898	0.2751	0.2707	0.2707	0.2827	0.1991	0.2708	0.2671	0.2048	0.2684	0.2679	0.2213	0.1436	0.1718	0.1367	0.1551
	100	0.1324	0.1289	0.1283	0.1283	0.1309	0.0965	0.1283	0.1277	0.1051	0.1284	0.1275	0.1145	0.0728	0.0959	0.0830	0.0965
	150	0.0945	0.0912	0.0905	0.0905	0.0936	0.0745	0.0905	0.0919	0.0800	0.0924	0.0922	0.0849	0.0564	0.0728	0.0629	0.0711
	200	0.0617	0.0603	0.0601	0.0601	0.0613	0.0506	0.0600	0.0606	0.0543	0.0609	0.0607	0.0574	0.0440	0.0516	0.0463	0.0518
	300	0.0386	0.0379	0.0378	0.0378	0.0385	0.0326	0.0378	0.0382	0.0349	0.0383	0.0382	0.0369	0.0332	0.0343	0.0318	0.0350
	400	0.0342	0.0333	0.0332	0.0332	0.0341	0.0290	0.0332	0.0339	0.0311	0.0340	0.0339	0.0328	0.0294	0.0304	0.0283	0.0310
0.85	50	0.2582	0.2452	0.2419	0.2419	0.2521	0.1772	0.2421	0.2408	0.1848	0.2421	0.2412	0.2020	0.1251	0.1521	0.1247	0.1448
	100	0.1897	0.1778	0.1759	0.1759	0.1859	0.1327	0.1752	0.1793	0.1416	0.1803	0.1784	0.1535	0.0935	0.1199	0.0962	0.1111
	150	0.1105	0.1067	0.1058	0.1058	0.1093	0.0858	0.1058	0.1068	0.0915	0.1074	0.1070	0.0968	0.0575	0.0810	0.0674	0.0767
	200	0.0828	0.0789	0.0783	0.0783	0.0820	0.0648	0.0784	0.0808	0.0703	0.0811	0.0808	0.0748	0.0464	0.0640	0.0547	0.0618
	300	0.0529	0.0518	0.0516	0.0516	0.0526	0.0422	0.0516	0.0521	0.0465	0.0522	0.0521	0.0495	0.0334	0.0453	0.0413	0.0460
	400	0.0422	0.0412	0.0410	0.0410	0.0420	0.0338	0.0410	0.0417	0.0375	0.0418	0.0417	0.0400	0.0275	0.0364	0.0332	0.0372
0.90	50	0.5974	0.5289	0.5035	0.5035	0.5654	0.3804	0.5099	0.4957	0.3481	0.4899	0.4994	0.3668	0.2488	0.2360	0.1436	0.1604
	100	0.2662	0.2446	0.2383	0.2383	0.2582	0.1802	0.2387	0.2439	0.1848	0.2450	0.2465	0.1980	0.1037	0.1344	0.0884	0.1071
	150	0.1761	0.1651	0.1634	0.1634	0.1730	0.1237	0.1634	0.1665	0.1326	0.1675	0.1654	0.1430	0.0782	0.1093	0.0836	0.0984
	200	0.1346	0.1236	0.1216	0.1216	0.1321	0.1008	0.1218	0.1287	0.1072	0.1295	0.1292	0.1137	0.0555	0.0863	0.0648	0.0760
	300	0.0818	0.0778	0.0772	0.0772	0.0810	0.0631	0.0773	0.0796	0.0689	0.0800	0.0799	0.0735	0.0386	0.0619	0.0513	0.0596
	400	0.0562	0.0539	0.0536	0.0536	0.0557	0.0430	0.0536	0.0552	0.0483	0.0554	0.0552	0.0520	0.0276	0.0450	0.0387	0.0452
0.95	50	1.5628	1.3330	1.1923	1.1923	1.3788	0.8473	1.2329	1.0755	0.6101	0.8662	0.9604	0.7141	0.4779	0.3230	0.1252	0.1377
	100	0.5249	0.4493	0.4227	0.4227	0.4934	0.3271	0.4282	0.4398	0.3032	0.4357	0.4428	0.3208	0.1785	0.1624	0.0857	0.1013
	150	0.3151	0.2744	0.2631	0.2631	0.3017	0.2107	0.2649	0.2820	0.2095	0.2826	0.2848	0.2201	0.1094	0.1287	0.0739	0.0887
	200	0.2129	0.1914	0.1872	0.1872	0.2065	0.1414	0.1879	0.1977	0.1497	0.1985	0.1979	0.1623	0.0777	0.1079	0.0685	0.0868
	300	0.1507	0.1386	0.1363	0.1363	0.1477	0.1100	0.1365	0.1435	0.1171	0.1442	0.1435	0.1247	0.0570	0.0910	0.0633	0.0774
	400	0.1203	0.1129	0.1115	0.1115	0.1183	0.0891	0.1116	0.1153	0.0957	0.1160	0.1155	0.1021	0.0447	0.0775	0.0560	0.0684
0.99	50	5.1331	3.6900	3.1279	3.1279	3.2843	2.2495	3.5360	2.9300	0.8933	0.7646	1.4590	1.6130	0.6255	0.2144	0.0904	0.0887
	100	2.7809	2.1335	1.7621	1.7621	2.1783	1.3616	1.9195	1.6389	0.6881	0.8592	1.1961	0.9519	0.5569	0.1459	0.0399	0.0436
	150	1.6278	1.2167	1.0004	1.0004	1.3530	0.8740	1.0816	1.0614	0.5628	0.8179	0.9808	0.6623	0.3522	0.1422	0.0357	0.0415
	200	1.3280	0.9515	0.7765	0.7765	1.0993	0.7049	0.8530	0.8800	0.4879	0.7348	0.8498	0.5532	0.3204	0.1201	0.0341	0.0386
	300	0.8750	0.6573	0.5644	0.5644	0.7635	0.4781	0.5920	0.6484	0.3825	0.6009	0.6419	0.4224	0.2215	0.1168	0.0321	0.0422
	400	0.5920	0.4872	0.4472	0.4472	0.5456	0.3519	0.4551	0.4792	0.3147	0.4665	0.4805	0.3326	0.1908	0.1351	0.0445	0.0578

Table 1. Estimated MSE values for different estimators at $p = 4$ and $\varphi = 0.1$. Bold values indicate the minimum MSE.

ρ	n	GPMLE	GPRRE														
		–	$\hat{k}_1^*$	$\hat{k}_2^*$	$\hat{k}_3^*$	$\hat{k}_4^*$	$\hat{k}_5^*$	$\hat{k}_6^*$	$\hat{k}_7^*$	$\hat{k}_8^*$	$\hat{k}_9^*$	$\hat{k}_{10}^*$	$\hat{k}_{11}^*$	$\hat{k}_{12}^*$	$\hat{k}_{13}^*$	$\hat{k}_{14}^*$	$\hat{k}_{15}^*$
0.80	50	0.8220	0.6225	0.5491	0.5491	0.6965	0.4756	0.5687	0.6162	0.4041	0.5756	0.5123	0.4362	0.2693	0.2601	0.3483	0.3246
	100	0.3983	0.3277	0.2989	0.2989	0.3729	0.2712	0.3011	0.3409	0.2592	0.3459	0.3110	0.2682	0.1690	0.1639	0.1814	0.1659
	150	0.2482	0.2157	0.2067	0.2067	0.2375	0.1909	0.1923	0.2218	0.1877	0.2282	0.2124	0.1875	0.1295	0.1300	0.1527	0.1330
	200	0.1677	0.1443	0.1398	0.1398	0.1622	0.1332	0.1333	0.1559	0.1350	0.1591	0.1512	0.1373	0.0991	0.0997	0.1131	0.0982
	300	0.1015	0.0912	0.0912	0.0912	0.0992	0.0789	0.0870	0.0980	0.0837	0.0987	0.0951	0.0890	0.0845	0.0683	0.0718	0.0668
	400	0.0893	0.0833	0.0819	0.0819	0.0881	0.0741	0.0806	0.0859	0.0770	0.0870	0.0843	0.0798	0.0642	0.0669	0.0603	0.0604
0.85	50	0.7522	0.5727	0.5013	0.5013	0.6468	0.4471	0.5216	0.5748	0.3918	0.5483	0.4836	0.4156	0.2700	0.2654	0.3620	0.3329
	100	0.5080	0.4406	0.4153	0.4153	0.4740	0.3322	0.4014	0.4186	0.3062	0.4210	0.3671	0.3161	0.2097	0.1855	0.1924	0.1818
	150	0.3109	0.2527	0.2347	0.2347	0.2915	0.2280	0.2233	0.2683	0.2184	0.2769	0.2527	0.2173	0.1308	0.1310	0.1552	0.1398
	200	0.2312	0.2034	0.1965	0.1965	0.2225	0.1655	0.1922	0.2104	0.1673	0.2135	0.1976	0.1748	0.1064	0.1147	0.1014	0.1006
	300	0.1510	0.1435	0.1406	0.1406	0.1484	0.1167	0.1386	0.1418	0.1202	0.1440	0.1367	0.1245	0.0759	0.0954	0.0724	0.0799
	400	0.1084	0.0992	0.0985	0.0985	0.1061	0.0867	0.0918	0.1031	0.0903	0.1048	0.1011	0.0941	0.0610	0.0718	0.0654	0.0640
0.90	50	1.9967	1.4032	1.1688	1.1688	1.4414	0.9476	1.3027	1.2470	0.6316	0.8582	0.7814	0.7623	0.4169	0.2837	0.3792	0.3590
	100	0.8377	0.5733	0.4856	0.4856	0.6904	0.4598	0.5255	0.6172	0.3709	0.5764	0.4973	0.4136	0.1839	0.1449	0.1861	0.1694
	150	0.4564	0.3838	0.3504	0.3504	0.4249	0.3077	0.3477	0.3725	0.2819	0.3811	0.3312	0.2825	0.1764	0.1520	0.1361	0.1330
	200	0.3680	0.2873	0.2644	0.2644	0.3421	0.2633	0.2627	0.3107	0.2462	0.3193	0.2871	0.2443	0.1286	0.1239	0.1204	0.1130
	300	0.2240	0.1854	0.1740	0.1740	0.2137	0.1693	0.1683	0.2015	0.1672	0.2069	0.1938	0.1691	0.0810	0.0986	0.0819	0.0792
	400	0.1560	0.1385	0.1318	0.1318	0.1520	0.1196	0.1316	0.1455	0.1235	0.1486	0.1406	0.1282	0.0665	0.0904	0.0681	0.0730
0.95	50	3.8194	2.3711	1.9112	1.9112	2.2518	1.5623	2.3487	2.1922	0.7851	0.8564	0.8715	1.2078	0.5174	0.2696	0.4781	0.4510
	100	1.5045	1.0422	0.8315	0.8315	1.1737	0.8046	0.9295	0.9611	0.5416	0.7591	0.6290	0.6064	0.3356	0.1654	0.1980	0.1901
	150	1.0376	0.7071	0.5711	0.5711	0.8676	0.6168	0.6274	0.7187	0.4774	0.6637	0.5668	0.4887	0.2558	0.1391	0.1530	0.1493
	200	0.5757	0.4185	0.3763	0.3763	0.5057	0.3272	0.3916	0.4630	0.2894	0.4551	0.3885	0.3214	0.1439	0.1129	0.1000	0.1037
	300	0.3579	0.2688	0.2417	0.2417	0.3292	0.2534	0.2434	0.3023	0.2365	0.3110	0.2776	0.2339	0.1136	0.1033	0.0816	0.0804
	400	0.3077	0.2214	0.1952	0.1952	0.2841	0.2230	0.1979	0.2667	0.2128	0.2742	0.2484	0.2124	0.0919	0.0952	0.0676	0.0660
0.99	50	16.7179	9.1220	7.6334	7.6334	3.3114	4.5110	9.9153	8.9474	0.7863	0.3171	0.5814	4.0202	0.3304	0.1879	0.6243	0.5402
	100	8.6364	3.9314	3.0897	3.0897	2.7907	2.7088	4.6309	4.7197	0.7608	0.4671	0.6669	2.3045	0.2862	0.0996	0.3660	0.3227
	150	4.7889	2.2297	1.6376	1.6376	2.2234	1.9964	2.4581	2.6431	0.7892	0.6876	0.7425	1.4200	0.2910	0.0792	0.2268	0.2109
	200	4.2318	1.9942	1.4255	1.4255	2.1155	1.7478	2.1377	2.3026	0.7569	0.7006	0.7325	1.2572	0.4574	0.0667	0.1819	0.1743
	300	2.3594	1.1437	0.7755	0.7755	1.4394	1.1820	1.1636	1.3182	0.6514	0.7770	0.6705	0.7877	0.3348	0.0626	0.0917	0.0910
	400	1.6393	0.9453	0.6527	0.6527	1.2012	0.9106	0.8500	0.9920	0.5957	0.7701	0.6449	0.6344	0.3390	0.0794	0.0489	0.0491

Table 2. Estimated MSE values for different estimators at $p = 4$ and $\varphi = 0.5$. Bold values indicate the minimum MSE.

$$\hat{k}_{10}^* = \min \left(\frac{1}{\sqrt{\frac{\max(\lambda_j \hat{\varphi})}{(n-p-1)\hat{\varphi} + \max(\lambda_j)\hat{\alpha}_j^2}}} \right).$$

(27)

Following Amin et al.³⁹, we use the following values for k^* :

$$\hat{k}_j = \frac{1}{\hat{\alpha}_j^2},$$

(28)

$$\hat{k}_{11}^* = \frac{p}{\sum_{j=1}^p \frac{1}{k_j}}.$$

(29)

Building upon the previous works, we propose the following values for k^* :

$$\hat{k}_{12}^* = \frac{\hat{\varphi}}{\sum_{j=1}^p \sqrt{\frac{1}{k_j}}},$$

(30)

$$\hat{k}_{13}^* = \sum_{j=1}^p (pq_j),$$

(31)

ρ	n	GPMLE	GPRRE														
		–	$\hat{k}_1^*$	$\hat{k}_2^*$	$\hat{k}_3^*$	$\hat{k}_4^*$	$\hat{k}_5^*$	$\hat{k}_6^*$	$\hat{k}_7^*$	$\hat{k}_8^*$	$\hat{k}_9^*$	$\hat{k}_{10}^*$	$\hat{k}_{11}^*$	$\hat{k}_{12}^*$	$\hat{k}_{13}^*$	$\hat{k}_{14}^*$	$\hat{k}_{15}^*$
0.80	50	3.6489	1.7840	1.5776	1.5776	1.4195	0.9476	2.0133	1.8727	0.6849	0.8922	0.7117	0.9642	0.5797	0.6368	0.9644	0.8964
	100	0.9571	0.6492	0.5691	0.5691	0.7714	0.5420	0.5889	0.6802	0.4448	0.6126	0.4506	0.4743	0.3659	0.3461	0.6071	0.5475
	150	0.5211	0.4269	0.4043	0.4043	0.4650	0.3448	0.3491	0.4144	0.3162	0.4171	0.3340	0.3172	0.3521	0.2613	0.4741	0.4124
	200	0.3610	0.3069	0.3052	0.3052	0.3308	0.2511	0.2481	0.3084	0.2418	0.3135	0.2635	0.2457	0.2823	0.1921	0.3709	0.3081
	300	0.2400	0.2221	0.2288	0.2288	0.2276	0.1769	0.1823	0.2165	0.1763	0.2208	0.1910	0.1810	0.2164	0.1308	0.2199	0.1776
	400	0.1715	0.1610	0.1617	0.1617	0.1666	0.1337	0.1469	0.1600	0.1355	0.1619	0.1470	0.1386	0.1408	0.1117	0.1413	0.1176
0.85	50	4.0078	1.5357	1.3720	1.3720	1.2153	0.8870	1.6035	1.6477	0.6760	0.8763	0.6878	0.8689	0.6007	0.6739	1.0125	0.9435
	100	1.1519	0.8641	0.7409	0.7409	0.9168	0.6134	0.7679	0.7883	0.4786	0.6855	0.4900	0.5219	0.5047	0.3725	0.5372	0.4799
	150	0.7483	0.5325	0.4773	0.4773	0.6375	0.4525	0.4510	0.5614	0.3912	0.5462	0.4183	0.4047	0.3171	0.2654	0.5123	0.4593
	200	0.5378	0.4299	0.3874	0.3874	0.4894	0.3439	0.3799	0.4228	0.3091	0.4245	0.3257	0.3116	0.2277	0.1866	0.2567	0.2326
	300	0.3416	0.2922	0.2702	0.2702	0.3250	0.2404	0.2701	0.2934	0.2308	0.3000	0.2503	0.2323	0.1575	0.1513	0.1454	0.1342
	400	0.2135	0.1843	0.1826	0.1826	0.2040	0.1601	0.1651	0.1956	0.1615	0.1984	0.1770	0.1642	0.1348	0.1156	0.1636	0.1359
0.90	50	8.5537	4.2505	3.7717	3.7717	2.1770	1.5578	4.8827	4.3108	0.8267	0.9296	0.7978	1.8232	0.6627	0.6689	1.0485	0.9703
	100	2.4749	1.4253	1.1802	1.1802	1.4381	0.9700	1.4313	1.4723	0.5925	0.7810	0.5282	0.8176	0.3843	0.3213	0.6477	0.6005
	150	1.0471	0.7850	0.6814	0.6814	0.8836	0.5627	0.7050	0.7439	0.4383	0.6591	0.4707	0.4947	0.3500	0.2567	0.3700	0.3248
	200	0.7964	0.5680	0.4881	0.4881	0.6836	0.4783	0.4960	0.5802	0.3985	0.5649	0.4164	0.4059	0.2599	0.1994	0.3289	0.2982
	300	0.4590	0.3283	0.2951	0.2951	0.4101	0.3073	0.2846	0.3714	0.2820	0.3807	0.3066	0.2783	0.1728	0.1421	0.2338	0.2061
	400	0.3178	0.2638	0.2490	0.2490	0.3004	0.2110	0.2468	0.2792	0.2096	0.2829	0.2376	0.2189	0.1496	0.1292	0.1337	0.1236
0.95	50	12.6088	6.2636	5.5080	5.5080	2.7403	2.2383	7.2900	6.4603	0.8256	0.7356	0.7104	2.6804	0.6402	0.6548	1.1497	1.0466
	100	3.7182	2.1997	1.7574	1.7574	2.0520	1.4411	2.1849	2.1261	0.7059	0.8197	0.5638	1.1552	0.5030	0.3381	0.6570	0.6037
	150	2.2661	1.2740	0.9779	0.9779	1.4555	1.0220	1.2351	1.3215	0.5961	0.8031	0.5244	0.7618	0.3775	0.2287	0.4988	0.4650
	200	1.4026	0.8972	0.7106	0.7106	1.0706	0.7598	0.8126	0.9046	0.5243	0.7493	0.4729	0.5770	0.3159	0.1718	0.3062	0.2874
	300	0.8113	0.5183	0.4321	0.4321	0.6760	0.4787	0.4579	0.5973	0.3942	0.5804	0.4209	0.4096	0.2264	0.1426	0.2343	0.2138
	400	0.6150	0.3725	0.3283	0.3283	0.5135	0.3682	0.3364	0.4776	0.3181	0.4711	0.3507	0.3344	0.1830	0.1164	0.2174	0.1886
0.99	50	73.1927	33.8598	30.9481	30.9481	2.3796	7.8464	41.4792	35.5031	0.7683	0.3623	0.4318	12.3884	0.4705	0.6210	1.3286	1.1616
	100	22.4155	9.5273	8.0536	8.0536	2.8515	5.2309	12.1525	12.4546	0.7808	0.2865	0.3609	5.6033	0.3500	0.3376	1.0869	0.9433
	150	10.7908	4.3459	3.4415	3.4415	2.6203	3.1024	5.4814	5.8097	0.7263	0.3877	0.3648	2.7016	0.3189	0.1933	0.7474	0.6645
	200	9.2668	4.1048	3.2091	3.2091	2.9107	3.0169	4.8104	5.0379	0.7521	0.4261	0.3924	2.4307	0.3128	0.1537	0.6041	0.5397
	300	4.9456	2.0607	1.4914	1.4914	2.0414	1.9294	2.4068	2.6616	0.7074	0.6138	0.4203	1.3836	0.2586	0.0926	0.3879	0.3532
	400	3.4664	1.5695	1.1006	1.1006	1.8078	1.5584	1.6910	1.9099	0.6938	0.7355	0.4782	1.0584	0.2535	0.0779	0.2325	0.2178

Table 3. Estimated MSE values for different estimators at $p = 4$ and $\varphi = 1$. Bold values indicate the minimum MSE.

$$\hat{k}_{14}^* = \text{median} \left(\frac{\hat{\varphi} p}{\sqrt{\hat{\alpha}_j^2}} \right), \quad (32)$$

$$\hat{k}_{15}^* = \frac{1}{p} \text{mean} \left(\frac{\lambda_j}{\hat{\varphi} + 2\lambda_j \hat{\alpha}_j^2} \right). \quad (33)$$

Monte Carlo simulation

This section presents Monte Carlo simulations to evaluate the performance of the proposed estimator, including the simulation design, results, and a comparison of relative efficiency.

Simulation design

This section describes the Monte Carlo simulation study conducted to evaluate the performance of different estimators in the GPRM under multicollinearity. The response variable (y) was generated from a GP distribution^{40,41}, with the mean ($\mu_i = \exp(x_i \beta)$) for $i = 1, \dots, n$, β representing the vector of coefficients, and x_i being the i th row of the design matrix X contains the explanatory variables. The explanatory variables were simulated using the formula^{42,34}:

$$x_{ij} = \sqrt{1 - \rho^2} e_{ij} + \rho e_{i,p+1}, \quad i = 1, \dots, n; \quad j = 2, \dots, p, \quad (34)$$

where ρ determines the correlation between explanatory variables and e_{ij} is drawn from a standard normal distribution. Multicollinearity was analyzed for ρ values of 0.80, 0.85, 0.90, 0.95, and 0.99. Models were tested with 4, 7, and 10 explanatory variables. The intercept (β_0) was set to 1, and the dispersion parameter φ was varied

ρ	n	GPMLE	GPRRE														
		–	$\hat{k}_1^*$	$\hat{k}_2^*$	$\hat{k}_3^*$	$\hat{k}_4^*$	$\hat{k}_5^*$	$\hat{k}_6^*$	$\hat{k}_7^*$	$\hat{k}_8^*$	$\hat{k}_9^*$	$\hat{k}_{10}^*$	$\hat{k}_{11}^*$	$\hat{k}_{12}^*$	$\hat{k}_{13}^*$	$\hat{k}_{14}^*$	$\hat{k}_{15}^*$
0.80	50	0.7017	0.6444	0.6154	0.6154	0.6604	0.3936	0.6197	0.6171	0.4078	0.6343	0.6358	0.4367	0.2191	0.2180	0.2076	0.2228
	100	0.3298	0.3015	0.2923	0.2923	0.3177	0.2013	0.2930	0.3095	0.2254	0.3169	0.3136	0.2425	0.1176	0.1225	0.1247	0.1353
	150	0.1965	0.1848	0.1824	0.1824	0.1924	0.1234	0.1825	0.1901	0.1471	0.1927	0.1912	0.1627	0.0810	0.0926	0.0962	0.1152
	200	0.1341	0.1276	0.1263	0.1263	0.1319	0.0862	0.1263	0.1310	0.1049	0.1323	0.1315	0.1167	0.0661	0.0733	0.0750	0.0907
	300	0.0932	0.0900	0.0892	0.0892	0.0922	0.0626	0.0892	0.0917	0.0764	0.0924	0.0920	0.0842	0.0543	0.0576	0.0579	0.0702
	400	0.0729	0.0720	0.0718	0.0718	0.0724	0.0497	0.0718	0.0720	0.0609	0.0724	0.0721	0.0674	0.0473	0.0506	0.0518	0.0623
0.85	50	1.2653	1.0494	0.9347	0.9347	1.0894	0.6325	0.9771	0.9987	0.5796	1.0120	1.0642	0.6262	0.2558	0.2067	0.2081	0.2047
	100	0.3712	0.3442	0.3359	0.3359	0.3573	0.2148	0.3366	0.3455	0.2405	0.3538	0.3490	0.2640	0.1145	0.1207	0.1236	0.1459
	150	0.2772	0.2553	0.2495	0.2495	0.2677	0.1606	0.2503	0.2620	0.1860	0.2674	0.2645	0.2066	0.0851	0.0928	0.0931	0.1160
	200	0.2035	0.1883	0.1842	0.1842	0.1982	0.1232	0.1844	0.1960	0.1483	0.1991	0.1972	0.1657	0.0717	0.0822	0.0831	0.1047
	300	0.1128	0.1110	0.1106	0.1106	0.1116	0.0706	0.1106	0.1105	0.0885	0.1116	0.1108	0.0994	0.0504	0.0640	0.0671	0.0852
	400	0.0950	0.0937	0.0934	0.0934	0.0942	0.0601	0.0934	0.0934	0.0765	0.0942	0.0936	0.0855	0.0416	0.0578	0.0605	0.0759
0.90	50	1.5898	1.2730	1.0845	1.0845	1.3115	0.7799	1.1584	1.1887	0.6522	1.1785	1.2469	0.7106	0.2788	0.1995	0.2179	0.2051
	100	0.7877	0.6630	0.5990	0.5990	0.7172	0.4503	0.6133	0.6738	0.4392	0.7014	0.6974	0.4528	0.1824	0.1297	0.1320	0.1321
	150	0.4141	0.3818	0.3690	0.3690	0.3954	0.2415	0.3696	0.3819	0.2635	0.3920	0.3864	0.2824	0.1157	0.1012	0.0918	0.1144
	200	0.3465	0.3258	0.3166	0.3166	0.3335	0.2043	0.3166	0.3237	0.2287	0.3317	0.3289	0.2492	0.0988	0.0920	0.0842	0.1133
	300	0.2028	0.1963	0.1941	0.1941	0.1984	0.1200	0.1941	0.1948	0.1451	0.1980	0.1963	0.1623	0.0617	0.0759	0.0747	0.1051
	400	0.1471	0.1390	0.1368	0.1368	0.1443	0.0904	0.1368	0.1433	0.1131	0.1449	0.1439	0.1262	0.0485	0.0645	0.0620	0.0853
0.95	50	3.0292	2.5059	2.1398	2.1398	2.3637	1.4310	2.2886	2.0990	1.0119	1.7262	1.9217	1.1955	0.5672	0.2407	0.2173	0.2051
	100	1.5290	1.2809	1.1169	1.1169	1.3177	0.8152	1.1629	1.1641	0.6822	1.1740	1.1731	0.6972	0.3574	0.1220	0.1229	0.1210
	150	0.9041	0.8188	0.7549	0.7549	0.8253	0.4985	0.7570	0.7512	0.4610	0.7785	0.7770	0.4748	0.2069	0.1030	0.0804	0.0948
	200	0.5995	0.5424	0.5127	0.5127	0.5593	0.3406	0.5132	0.5326	0.3461	0.5487	0.5404	0.3610	0.1527	0.0807	0.0665	0.0860
	300	0.3722	0.3314	0.3166	0.3166	0.3540	0.2198	0.3168	0.3469	0.2432	0.3548	0.3510	0.2625	0.0939	0.0732	0.0631	0.0874
	400	0.2792	0.2584	0.2516	0.2516	0.2697	0.1657	0.2516	0.2648	0.1919	0.2700	0.2671	0.2106	0.0757	0.0723	0.0662	0.0965
0.99	50	20.0532	13.7990	11.1530	11.1530	5.4951	6.0691	13.4736	12.3782	1.6942	1.0565	3.3032	5.4084	0.9910	0.1263	0.3738	0.3176
	100	6.0251	4.6956	3.7970	3.7970	3.7630	2.7080	4.2466	3.7683	1.2478	1.7902	2.6554	1.8935	0.8647	0.0770	0.1005	0.0921
	150	3.9783	2.9370	2.2153	2.2153	2.6677	1.8769	2.5665	2.5536	1.0776	1.7923	2.3125	1.3552	0.6347	0.0512	0.0705	0.0647
	200	2.9540	2.0199	1.4693	1.4693	2.0084	1.4063	1.7843	1.9342	0.9099	1.6049	1.8928	1.0373	0.4822	0.0425	0.0595	0.0548
	300	1.7375	1.2677	0.9978	0.9978	1.3714	0.8765	1.1011	1.2850	0.6920	1.2546	1.3292	0.7225	0.3339	0.0400	0.0332	0.0339
	400	1.2668	1.0563	0.8980	0.8980	1.0803	0.6704	0.9089	0.9955	0.5735	1.0112	1.0309	0.5779	0.3073	0.0434	0.0272	0.0329

Table 4. Estimated MSE values for different estimators at $p = 7$ and $\varphi = 0.1$. Bold values indicate the minimum MSE.

at 0.01, 0.5, and 1⁴¹. The slope coefficients were set such that $\sum_{j=2}^p \beta_j^2 = 1$, with equal values for $\beta_1, \dots, \beta_{p-1}$. Simulations were conducted for sample sizes of 50, 100, 150, 200, 300, and 400.

The simulations were implemented in the R software (R version 4.4.1). For each iteration, the estimated MSE of the estimators was calculated as follows^{43,44}:

$$\text{MSE}(\beta^*) = \frac{\sum_{l=1}^{1000} (\hat{\beta}_l - \beta)^T (\hat{\beta}_l - \beta)}{1000}, \quad (35)$$

where β_l denotes the vector of estimated coefficients from the l -th simulation run for a specific estimator (such as the GPMLE or a GPRRE employing a particular ridge parameter). The estimator associated with the smallest MSE was deemed optimal for alleviating the effects of multicollinearity within the GPRM framework.

Simulation results

Simulation Tables 1, 2, 3, 4, 5, 6, 7, 8 and 9 provide a detailed comparison of the MSE for the GPMLE and different versions of the GPRRE under various experimental conditions. These conditions include different levels of multicollinearity (ρ), sample sizes (n), dimensions (p), and shrinkage parameters (φ). The tables highlight the best-performing estimator in each scenario by marking the lowest MSE values in bold. Main factors affecting simulation:

1. Effect of multicollinearity:

- The degree of multicollinearity (ρ) emerged as the most critical factor influencing estimator performance. As expected, the MSE of the GPMLE becomes increasingly severe with higher ρ values.
- Under severe multicollinearity, the GPMLE's MSE becomes prohibitively large, often by an order of magnitude or more compared to the best-performing GPRRE. The ridge estimators, particularly $\hat{k}_{13}^*$, $\hat{k}_{14}^*$, and

ρ	n	GPMLE	GPRRE														
		–	$\hat{k}_1^*$	$\hat{k}_2^*$	$\hat{k}_3^*$	$\hat{k}_4^*$	$\hat{k}_5^*$	$\hat{k}_6^*$	$\hat{k}_7^*$	$\hat{k}_8^*$	$\hat{k}_9^*$	$\hat{k}_{10}^*$	$\hat{k}_{11}^*$	$\hat{k}_{12}^*$	$\hat{k}_{13}^*$	$\hat{k}_{14}^*$	$\hat{k}_{15}^*$
0.80	50	3.2325	2.1434	1.8052	1.8052	1.6820	0.9828	2.1016	2.0620	0.7769	1.3945	1.2490	0.9881	0.4157	0.5200	0.6765	0.6007
	100	0.9011	0.7451	0.6440	0.6440	0.7907	0.4867	0.6595	0.7081	0.4550	0.7571	0.6465	0.4606	0.2551	0.3169	0.3945	0.3410
	150	0.5856	0.4978	0.4583	0.4583	0.5396	0.3332	0.4570	0.5061	0.3416	0.5332	0.4781	0.3537	0.1767	0.1896	0.2426	0.2058
	200	0.3863	0.3329	0.3127	0.3127	0.3664	0.2349	0.3125	0.3519	0.2555	0.3661	0.3394	0.2672	0.1393	0.1488	0.1836	0.1585
	300	0.2693	0.2401	0.2253	0.2253	0.2592	0.1693	0.2258	0.2511	0.1905	0.2597	0.2449	0.2027	0.1075	0.1111	0.1232	0.1136
	400	0.1980	0.1895	0.1827	0.1827	0.1938	0.1277	0.1828	0.1873	0.1479	0.1932	0.1849	0.1595	0.0948	0.0959	0.0898	0.1024
0.85	50	7.8202	5.3076	4.7770	4.7770	2.4711	1.6676	4.7630	5.1508	1.0122	1.6948	1.6988	1.7863	0.5035	0.5652	0.8339	0.7490
	100	1.1477	0.9346	0.7998	0.7998	0.9725	0.5824	0.8213	0.8691	0.5162	0.9124	0.7630	0.5363	0.2580	0.2931	0.3774	0.3305
	150	0.8676	0.6950	0.5965	0.5965	0.7608	0.4585	0.6205	0.6954	0.4312	0.7366	0.6335	0.4446	0.1868	0.1879	0.2488	0.2139
	200	0.5302	0.4221	0.3749	0.3749	0.4855	0.3101	0.3790	0.4563	0.3177	0.4872	0.4375	0.3270	0.1554	0.1703	0.2464	0.1960
	300	0.3154	0.2963	0.2826	0.2826	0.3043	0.1914	0.2828	0.2906	0.2137	0.3020	0.2823	0.2279	0.1094	0.1104	0.1030	0.1137
	400	0.2779	0.2654	0.2563	0.2563	0.2702	0.1700	0.2564	0.2595	0.1945	0.2682	0.2525	0.2090	0.0925	0.0970	0.0869	0.1085
0.90	50	5.9029	3.7296	2.9772	2.9772	2.4956	1.7694	3.6752	3.5626	1.0339	1.6834	1.5897	1.5697	0.5516	0.5983	0.8830	0.8126
	100	2.2916	1.6298	1.2516	1.2516	1.6366	1.0745	1.4578	1.5253	0.7825	1.3880	1.1410	0.8446	0.3784	0.3644	0.5540	0.5051
	150	1.3322	1.0759	0.8690	0.8690	1.1143	0.6926	0.9237	0.9682	0.5796	1.0251	0.8445	0.5832	0.2622	0.1983	0.2576	0.2309
	200	1.0541	0.8927	0.7354	0.7354	0.9238	0.5883	0.7592	0.8054	0.5188	0.8695	0.7508	0.5136	0.2350	0.1540	0.1785	0.1607
	300	0.6192	0.5526	0.4958	0.4958	0.5725	0.3532	0.4983	0.5264	0.3511	0.5590	0.4978	0.3606	0.1481	0.1062	0.0994	0.1001
	400	0.4097	0.3451	0.3045	0.3045	0.3844	0.2443	0.3086	0.3667	0.2628	0.3862	0.3557	0.2770	0.1083	0.0917	0.0990	0.0972
0.95	50	10.5843	7.0806	5.9125	5.9125	3.8011	2.8498	6.9864	6.4531	1.3367	1.7615	1.8616	2.6896	0.8338	0.5889	0.9031	0.8234
	100	3.9217	2.7952	2.1136	2.1136	2.4471	1.7547	2.5055	2.4785	1.0480	1.7062	1.3859	1.2694	0.6122	0.3425	0.5792	0.5387
	150	2.6906	2.1520	1.6390	1.6390	1.9883	1.3176	1.8054	1.7437	0.8875	1.5214	1.2302	0.9509	0.4792	0.1986	0.2902	0.2718
	200	1.7517	1.4372	1.1145	1.1145	1.4214	0.9307	1.1920	1.2053	0.7097	1.2315	0.9964	0.7079	0.3597	0.1378	0.1806	0.1665
	300	1.0475	0.8044	0.6329	0.6329	0.8934	0.5876	0.6839	0.8076	0.5138	0.8666	0.7370	0.5092	0.2225	0.0966	0.1284	0.1143
	400	0.8092	0.6313	0.5098	0.5098	0.7156	0.4700	0.5394	0.6577	0.4393	0.7082	0.6198	0.4408	0.1808	0.0805	0.0957	0.0885
0.99	50	63.1640	39.9406	33.0709	33.0709	4.2545	12.5910	40.6951	37.4582	1.5280	0.4250	1.1941	13.5719	0.7797	0.5791	1.2742	1.1157
	100	19.6837	13.3269	10.4364	10.4364	4.9210	6.0585	12.7453	11.8404	1.5031	0.9308	1.4385	5.0314	0.9795	0.2323	0.6602	0.5736
	150	12.7634	8.3207	6.2424	6.2424	3.9310	4.4185	7.9100	7.6234	1.3548	1.1547	1.4043	3.2993	0.8687	0.1587	0.4670	0.4159
	200	8.7063	5.0538	3.5677	3.5677	3.1355	3.3136	4.9629	5.1701	1.2420	1.4011	1.3747	2.3261	0.7758	0.1285	0.4192	0.3815
	300	5.3728	3.3092	2.2601	2.2601	2.7869	2.4599	3.0419	3.2663	1.2150	1.7226	1.4843	1.6332	0.7372	0.0702	0.2124	0.2002
	400	4.0594	2.8981	1.9895	1.9895	2.5001	1.9594	2.4366	2.4998	1.1080	1.6942	1.3881	1.3062	0.7798	0.0526	0.1050	0.1003

Table 5. Estimated MSE values for different estimators at $p = 7$ and $\varphi = 0.5$. Bold values indicate the minimum MSE.

$\hat{k}_{15}^*$, demonstrate remarkable robustness, maintaining stable and low MSE by effectively shrinking the coefficients and controlling variance, even when the correlation between predictors approaches 0.99.

- While the performance gap narrows under moderate multicollinearity, the GPRRE variants still consistently achieve a lower MSE than the GPMLE. This advantage is most evident for smaller sample sizes ($n = 50, 100$), where the data provides less information to stabilize the MLE.

2. Effect of sample size:

- The benefits of the ridge approach are most acute in “small n ” situations, which are common in modern statistical applications.
- For small $n = 50, 100$, the GPMLE is highly unstable. The GPRRE provides dramatic improvements in these settings, often reducing the MSE by half or more. This confirms that ridge regression is an essential tool for preventing overfitting when data is scarce.
- As the sample size increases ($n = 300, 400$), the performance of all estimators improves, and the relative advantage of the GPRRE diminishes. This is consistent with theoretical expectations, as the GPMLE is asymptotically unbiased. However, even with $n = 400$, the GPRRE often retains a slight edge, especially under high multicollinearity.

3. Effect of number of explanatory variables:

- The benefits of the ridge approach are most acute in “large p ” situations, which are common in modern statistical applications.
- The challenge of estimation increases with the number of explanatory variables. The GPRRE shows a clear and growing advantage over the GPMLE as p increases from 4 to 10, effectively managing the added complexity and severe multicollinearity.

ρ	n	GPMLE	GPRRE														
		–	$\hat{k}_1^*$	$\hat{k}_2^*$	$\hat{k}_3^*$	$\hat{k}_4^*$	$\hat{k}_5^*$	$\hat{k}_6^*$	$\hat{k}_7^*$	$\hat{k}_8^*$	$\hat{k}_9^*$	$\hat{k}_{10}^*$	$\hat{k}_{11}^*$	$\hat{k}_{12}^*$	$\hat{k}_{13}^*$	$\hat{k}_{14}^*$	$\hat{k}_{15}^*$
0.80	50	16.1732	8.0728	7.2716	7.2716	2.7418	1.8034	9.4390	8.3400	1.2007	1.8469	1.5786	2.6585	0.8037	1.0956	1.3595	1.2575
	100	2.4058	1.7802	1.4759	1.4759	1.5857	0.9128	1.6242	1.6113	0.7244	1.3013	0.8854	0.8341	0.5529	0.7998	0.9848	0.9005
	150	1.5092	1.1371	0.9427	0.9427	1.1621	0.6966	1.0098	1.0701	0.5971	1.0394	0.7412	0.6214	0.3683	0.5256	0.7076	0.6377
	200	0.8858	0.7169	0.6257	0.6257	0.7779	0.4843	0.6261	0.6950	0.4576	0.7438	0.5679	0.4548	0.2949	0.3949	0.5446	0.4707
	300	0.5869	0.4888	0.4318	0.4318	0.5395	0.3496	0.4355	0.4928	0.3513	0.5295	0.4362	0.3482	0.2164	0.2833	0.3714	0.3060
	400	0.4116	0.3742	0.3509	0.3509	0.3920	0.2557	0.3510	0.3677	0.2694	0.3855	0.3367	0.2731	0.1835	0.2126	0.2331	0.1925
0.85	50	44.5253	19.2054	17.7375	17.7375	3.2841	4.8726	87.8744	20.5468	1.3894	1.7603	1.8163	5.6880	0.9071	1.1960	1.4980	1.3740
	150	3.2801	2.2984	1.8786	1.8786	1.9107	1.1435	2.1526	2.1340	0.8296	1.4693	1.0039	1.0353	0.5275	0.7466	0.9901	0.9066
	150	2.0796	1.4872	1.1718	1.1718	1.4862	0.9178	1.3304	1.3993	0.6979	1.2701	0.8470	0.7608	0.3865	0.5428	0.7493	0.6796
	200	1.1694	0.8629	0.7306	0.7306	0.9703	0.6121	0.7587	0.8777	0.5458	0.9263	0.6919	0.5439	0.3351	0.4779	0.6799	0.6020
	300	0.6897	0.6049	0.5356	0.5356	0.6328	0.3946	0.5422	0.5693	0.3864	0.6110	0.4928	0.3847	0.2163	0.2621	0.3011	0.2546
	400	0.5908	0.5395	0.4927	0.4927	0.5541	0.3477	0.4946	0.5049	0.3484	0.5356	0.4408	0.3511	0.1758	0.1916	0.1957	0.1690
0.90	50	26.6443	12.8660	11.3970	11.3970	3.1424	2.7380	15.2469	13.2616	1.2917	1.6274	1.5598	4.0495	0.8801	1.1796	1.4996	1.3943
	150	6.3980	4.2367	3.4625	3.4625	2.7218	1.7854	4.1330	3.9512	1.0399	1.6638	1.2103	1.6834	0.6607	0.8795	1.1795	1.0958
	150	3.1826	2.3064	1.7719	1.7719	1.9998	1.2619	2.0536	2.0021	0.8280	1.4781	0.9561	0.9837	0.4810	0.5551	0.7791	0.7209
	200	2.3122	1.7464	1.3433	1.3433	1.6688	1.0504	1.5099	1.5241	0.7569	1.3561	0.9151	0.8133	0.4105	0.4528	0.6259	0.5738
	300	1.4348	1.1587	0.9133	0.9133	1.1909	0.7569	0.9774	1.0312	0.6226	1.0807	0.7633	0.6187	0.2995	0.2789	0.3712	0.3313
	400	0.8575	0.6629	0.5523	0.5523	0.7507	0.4830	0.5780	0.6862	0.4488	0.7352	0.5730	0.4392	0.2056	0.2186	0.3167	0.2687
0.95	50	38.7926	20.4570	17.9864	17.9864	3.9621	4.2892	23.1129	20.2791	1.4289	1.4173	1.5132	5.9982	0.9520	1.1485	1.5305	1.4264
	150	8.8604	5.7931	4.5739	4.5739	3.2445	2.6658	5.6083	5.3304	1.1368	1.5757	1.1548	2.2290	0.6718	0.8339	1.2101	1.1164
	150	6.3144	4.4583	3.4314	3.4314	3.0500	2.1912	4.0834	3.8861	1.0754	1.6363	1.1061	1.7558	0.6657	0.5983	0.9183	0.8451
	200	3.9766	3.0092	2.2939	2.2939	2.5367	1.6674	2.6174	2.4998	0.9872	1.6719	1.0705	1.2511	0.6195	0.4113	0.6419	0.5853
	300	2.2919	1.6210	1.1845	1.1845	1.6415	1.1040	1.4012	1.5027	0.7759	1.3902	0.9011	0.8148	0.3986	0.2815	0.4622	0.4198
	400	1.6893	1.1671	0.8611	0.8611	1.2913	0.8693	1.0021	1.1899	0.6771	1.1990	0.8267	0.6797	0.3163	0.2070	0.3557	0.3181
0.99	50	293.3510	144.5770	130.8878	130.8878	3.1598	20.2619	169.9600	148.5945	1.6232	0.6247	0.9731	38.5042	0.7641	1.1852	1.7468	1.6240
	150	56.0313	35.2408	29.2588	29.2588	3.9917	11.0812	36.0232	33.4683	1.3765	0.4765	0.6987	11.9138	0.6885	0.6918	1.3683	1.2176
	150	35.4648	22.4000	17.9599	17.9599	4.1469	8.7160	22.5259	21.1925	1.3732	0.5730	0.6925	8.2405	0.6718	0.5411	1.1911	1.0531
	200	20.6019	11.5381	8.6184	8.6184	3.4313	6.1718	11.9510	12.2228	1.3063	0.7933	0.7121	4.8790	0.5420	0.4630	1.0999	0.9923
	300	12.3012	7.5286	5.4996	5.4996	3.7441	4.4425	7.2676	7.4877	1.3711	1.2109	0.8663	3.3271	0.8401	0.2577	0.7083	0.6403
	400	9.2597	6.1726	4.4768	4.4768	3.6194	3.4974	5.6342	5.5786	1.2850	1.4149	0.9554	2.5114	0.9182	0.1848	0.4890	0.4398

Table 6. Estimated MSE values for different estimators at $p = 7$ and $\varphi = 1$. Bold values indicate the minimum MSE.

4. Effect of dispersion parameter:

- The value of the dispersion parameter φ influences the scale of the MSE but does not alter the fundamental ranking of the estimators. The relative performance of the different GPRRE variants remains consistent across values of φ . Among the fifteen evaluated ridge estimators, $\hat{k}_{13}^*$, $\hat{k}_{14}^*$, and $\hat{k}_{15}^*$ consistently emerge as top performers. Their success is attributed to a more effective calibration of the shrinkage intensity, optimally balancing the introduced bias against the reduction in variance to minimize the total MSE.

The results consistently demonstrate that the proposed GPRRE outperforms the conventional GPMLE across virtually all simulated scenarios. The reduction in MSE is particularly pronounced, underscoring the efficacy of introducing a bias-variance trade-off to manage the adverse effects of multicollinearity. The GPMLE, which relies on asymptotic properties that are violated in the presence of high correlation among predictors and finite samples, exhibits significantly inflated variance. In contrast, the GPRRE successfully stabilizes the coefficient estimates, leading to a substantial decrease in MSE.

In summary, the simulation study provides robust empirical evidence that the GPRRE is a superior alternative to the traditional maximum likelihood estimator in the presence of multicollinearity. Its performance is particularly strong in finite samples, with high-dimensional data, and under severe correlation among regressors. The proposed estimators $\hat{k}_{13}^*$ and $\hat{k}_{14}^*$ are recommended as reliable choices for practitioners, as they consistently provide the most accurate and stable estimates across a wide range of challenging data conditions. This demonstrates that the GPRRE is not merely a theoretical exercise but a practical and necessary enhancement to the regression toolkit for overdispersed and multicollinear count data.

ρ	n	GPMLE	GPRRE														
		–	$\hat{k}_1^*$	$\hat{k}_2^*$	$\hat{k}_3^*$	$\hat{k}_4^*$	$\hat{k}_5^*$	$\hat{k}_6^*$	$\hat{k}_7^*$	$\hat{k}_8^*$	$\hat{k}_9^*$	$\hat{k}_{10}^*$	$\hat{k}_{11}^*$	$\hat{k}_{12}^*$	$\hat{k}_{13}^*$	$\hat{k}_{14}^*$	$\hat{k}_{15}^*$
0.80	50	1.4062	1.2383	1.1235	1.1235	1.2388	0.6644	1.1528	1.1623	0.6739	1.1984	1.2199	0.7064	0.2753	0.2754	0.2726	0.2673
	150	0.5650	0.5371	0.5205	0.5205	0.5418	0.3041	0.5206	0.5219	0.3507	0.5406	0.5335	0.3652	0.1569	0.1562	0.1454	0.1653
	150	0.3287	0.3063	0.2990	0.2990	0.3185	0.1768	0.2994	0.3146	0.2216	0.3217	0.3174	0.2441	0.1061	0.1123	0.1166	0.1388
	200	0.2295	0.2246	0.2230	0.2230	0.2256	0.1322	0.2230	0.2232	0.1693	0.2266	0.2245	0.1858	0.0873	0.0910	0.0973	0.1299
	300	0.1451	0.1428	0.1420	0.1420	0.1433	0.0835	0.1421	0.1424	0.1114	0.1440	0.1428	0.1248	0.0654	0.0666	0.0753	0.1014
	400	0.1218	0.1201	0.1196	0.1196	0.1206	0.0741	0.1196	0.1200	0.0975	0.1210	0.1203	0.1078	0.0620	0.0622	0.0696	0.0909
0.85	50	2.0705	1.6912	1.4560	1.4560	1.6372	0.8495	1.5733	1.5867	0.7884	1.5602	1.6396	0.8685	0.2874	0.2766	0.2978	0.2729
	150	0.8047	0.6826	0.6201	0.6201	0.7378	0.4146	0.6358	0.7126	0.4474	0.7488	0.7420	0.4556	0.1640	0.1708	0.1797	0.1669
	150	0.4192	0.4033	0.3963	0.3963	0.4050	0.2236	0.3963	0.3973	0.2718	0.4076	0.4005	0.2938	0.1117	0.1067	0.1063	0.1416
	200	0.3553	0.3400	0.3333	0.3333	0.3445	0.1947	0.3334	0.3385	0.2395	0.3467	0.3419	0.2571	0.0981	0.0961	0.0959	0.1264
	300	0.2062	0.2022	0.2009	0.2009	0.2027	0.1146	0.2009	0.2007	0.1510	0.2037	0.2018	0.1678	0.0683	0.0719	0.0836	0.1189
	400	0.1581	0.1556	0.1548	0.1548	0.1561	0.0904	0.1548	0.1550	0.1215	0.1567	0.1554	0.1350	0.0574	0.0639	0.0757	0.1054
0.90	50	3.7407	2.9990	2.4765	2.4765	2.4930	1.4104	2.6973	2.5932	1.1048	2.1467	2.4423	1.6389	0.3976	0.2697	0.2970	0.2646
	150	1.4160	1.1803	0.9990	0.9990	1.1983	0.6501	1.0485	1.1210	0.6055	1.1888	1.1956	0.6220	0.1938	0.1718	0.1781	0.1597
	150	0.6022	0.5618	0.5397	0.5397	0.5715	0.3259	0.5399	0.5572	0.3695	0.5756	0.5686	0.3820	0.1397	0.1029	0.0986	0.1183
	200	0.5225	0.4938	0.4772	0.4772	0.4986	0.2806	0.4773	0.4855	0.3241	0.5019	0.4946	0.3384	0.1232	0.0912	0.0887	0.1160
	300	0.3405	0.3295	0.3245	0.3245	0.3311	0.1852	0.3245	0.3250	0.2304	0.3327	0.3282	0.2479	0.0881	0.0733	0.0812	0.1206
	400	0.2261	0.2213	0.2196	0.2196	0.2221	0.1239	0.2196	0.2194	0.1637	0.2229	0.2207	0.1804	0.0650	0.0648	0.0784	0.1172
0.95	50	10.2149	8.5428	7.5293	7.5293	4.5904	6.7328	8.0403	7.5675	1.8907	2.8028	4.4658	5.4586	0.7120	0.2463	0.3533	0.3033
	150	2.5728	2.1707	1.7531	1.7531	2.0278	1.1813	1.8526	1.8837	0.9390	1.8733	1.9465	0.9753	0.3553	0.1398	0.1595	0.1400
	150	1.5011	1.3211	1.1547	1.1547	1.3100	0.7579	1.1703	1.2217	0.6938	1.2833	1.2854	0.6816	0.2641	0.0995	0.0990	0.0944
	200	1.0722	0.9911	0.9175	0.9175	0.9797	0.5535	0.9183	0.9184	0.5464	0.9661	0.9382	0.5399	0.2004	0.0744	0.0682	0.0808
	300	0.6670	0.6326	0.6110	0.6110	0.6320	0.3568	0.6110	0.6107	0.3948	0.6326	0.6199	0.4028	0.1509	0.0624	0.0633	0.0926
	400	0.4621	0.4422	0.4314	0.4314	0.4440	0.2481	0.4314	0.4331	0.2927	0.4463	0.4391	0.3079	0.1086	0.0585	0.0653	0.1031
0.99	50	32.4039	24.4047	19.7646	19.7646	8.7644	10.0134	22.9465	21.6112	2.9838	1.8059	6.4482	9.1490	1.6363	0.1968	0.5256	0.4411
	150	12.8321	9.5508	7.0341	7.0341	5.1040	5.1113	8.4529	8.1270	1.9224	2.3267	4.4613	3.4463	1.1347	0.1257	0.2843	0.2505
	150	7.7078	6.9505	5.5602	5.5602	4.3766	3.2532	9.8280	5.5190	1.6118	2.7408	3.7907	2.2650	1.1277	0.0763	0.1257	0.1069
	200	5.2463	4.4771	3.4217	3.4217	3.5625	2.3493	3.6828	3.4454	1.3930	2.6048	3.1553	1.6674	0.8815	0.0548	0.0761	0.0667
	300	2.6472	2.3487	1.9468	1.9468	2.1480	1.2806	1.9737	1.9370	0.9957	1.9349	1.9984	0.9788	0.5707	0.0373	0.0359	0.0352
	400	2.2951	2.0695	1.7679	1.7679	1.9347	1.1649	1.7822	1.7425	0.9505	1.7800	1.8098	0.9134	0.5115	0.0327	0.0287	0.0310

Table 7. Estimated MSE values for different estimators at $p = 10$ and $\varphi = 0.1$. Bold values indicate the minimum MSE.

Relative efficiency

Relative Efficiency (RE) is used to compare the performance of statistical estimators by measuring their precision and reliability. This comparison relies on the MSE, which combines bias and variance, with a lower MSE indicating better performance. The formula for RE is:

$$RE(\beta_{k^*_i}) = \frac{MSE(\beta_{GPMLE})}{MSE(\beta_{k^*_i})}, \tag{36}$$

where $\beta_{k^*_i}$ represents the MSE of GPRRE with each parameter. The reference estimator, $MSE(\beta_{GPMLE})$, is often used as a benchmark due to its strong asymptotic properties.

Figures 1, 2, 3 and 4 present a comprehensive evaluation of RE was conducted to rigorously assess the performance of the GPRRE under different shrinkage parameters, with RE plotted as a function of key statistical parameters: sample size (n), population correlation (ρ), the number of predictor variables (p), and a measure of dispersion (φ). The results demonstrate that the proposed GPRRE estimator consistently achieved the highest relative efficiency across the vast majority of the investigated scenarios. This superior performance manifests as a high RR, indicating that the GPRRE provides estimates with greater precision and stability that is, a smaller variance and reduced susceptibility to bias compared to its competitors. The empirical evidence thus robustly confirms that the GPRRE is the most efficient estimator within the defined class of models under study. This dominance was particularly pronounced when compared to the estimator denoted as $\hat{k}_{13}^*$, $\hat{k}_{14}^*$, and $\hat{k}_{15}^*$, which was consistently outperformed, often by a significant margin.

Application

This study investigates CO₂ emissions from plug-in hybrid electric vehicles (PHEVs) sold in Canada between 2020 and 2025. The primary dataset focuses on vehicles from this five-year period, though additional data for

ρ	n	GPMLE	GPRRE														
		–	$\hat{k}_1^*$	$\hat{k}_2^*$	$\hat{k}_3^*$	$\hat{k}_4^*$	$\hat{k}_5^*$	$\hat{k}_6^*$	$\hat{k}_7^*$	$\hat{k}_8^*$	$\hat{k}_9^*$	$\hat{k}_{10}^*$	$\hat{k}_{11}^*$	$\hat{k}_{12}^*$	$\hat{k}_{13}^*$	$\hat{k}_{14}^*$	$\hat{k}_{15}^*$
0.80	50	10.0413	6.5984	5.6933	5.6933	3.3771	1.9264	7.5501	6.4572	1.3362	2.5071	3.7315	8.6925	0.6424	0.7781	0.9773	0.8511
	150	1.7898	1.5391	1.2726	1.2726	1.4750	0.7910	1.3252	1.3288	0.7133	1.4068	1.1779	0.7170	0.3250	0.4996	0.5227	0.4449
	150	1.0541	0.8711	0.7401	0.7401	0.9291	0.5143	0.7716	0.8520	0.5187	0.9363	0.8098	0.5084	0.2273	0.3764	0.4125	0.3334
	200	0.6941	0.6362	0.5880	0.5880	0.6478	0.3576	0.5902	0.6117	0.3941	0.6507	0.5853	0.3949	0.1795	0.2471	0.2206	0.1835
	300	0.4091	0.3878	0.3694	0.3694	0.3944	0.2267	0.3698	0.3782	0.2687	0.3959	0.3691	0.2794	0.1394	0.1691	0.1459	0.1401
	400	0.3359	0.3194	0.3040	0.3040	0.3250	0.1912	0.3042	0.3134	0.2296	0.3273	0.3096	0.2408	0.1225	0.1448	0.1254	0.1254
0.85	50	11.2624	6.9740	5.9049	5.9049	3.3051	2.0109	7.2177	6.8487	1.3386	2.4649	2.4821	2.2982	0.5955	0.8259	1.0479	0.9365
	150	3.0234	2.0703	1.5728	1.5728	2.0213	1.2002	1.9317	2.0745	0.9207	1.9218	1.6007	1.0067	0.3553	0.5881	0.7833	0.6891
	150	1.2829	1.1406	0.9953	0.9953	1.1259	0.6032	1.0124	1.0332	0.5844	1.1077	0.9241	0.5861	0.2482	0.3447	0.3312	0.2698
	200	1.1445	1.0240	0.8864	0.8864	1.0280	0.5791	0.8971	0.9324	0.5718	1.0175	0.8745	0.5568	0.2259	0.2840	0.2793	0.2294
	300	0.6448	0.6033	0.5651	0.5651	0.6072	0.3361	0.5656	0.5751	0.3725	0.6087	0.5514	0.3796	0.1492	0.1621	0.1385	0.1317
	400	0.4413	0.4207	0.4007	0.4007	0.4251	0.2426	0.4008	0.4060	0.2871	0.4263	0.3963	0.2971	0.1249	0.1355	0.1170	0.1248
0.90	50	21.2886	14.8412	13.7362	13.7362	4.8245	3.2693	13.9964	14.1815	1.7533	2.6035	3.1982	4.5273	0.8007	0.8366	1.1202	0.9984
	150	5.2755	3.7706	2.8606	2.8606	2.8511	1.8235	3.4791	3.4300	1.1671	2.3215	1.9960	1.4843	0.4349	0.6061	0.8008	0.7206
	150	2.2256	1.8223	1.4441	1.4441	1.7593	0.9923	1.5524	1.6225	0.8258	1.6793	1.3836	0.8397	0.3248	0.3312	0.3886	0.3278
	200	1.6677	1.4419	1.1726	1.1726	1.4070	0.8078	1.2136	1.2558	0.7098	1.3659	1.1189	0.6885	0.2836	0.2831	0.3001	0.2494
	300	0.9812	0.8908	0.7858	0.7858	0.8934	0.5094	0.7892	0.8168	0.5111	0.8894	0.7711	0.4933	0.1997	0.1767	0.1638	0.1402
	400	0.6979	0.6552	0.6080	0.6080	0.6576	0.3736	0.6085	0.6134	0.4056	0.6560	0.5917	0.4035	0.1561	0.1244	0.1083	0.1108
0.95	50	46.2000	29.1804	25.2659	25.2659	5.4837	5.8644	30.2338	28.4016	2.0015	1.9057	3.2731	21.5943	1.0508	0.8383	1.2761	1.1310
	150	7.7996	5.9213	4.5311	4.5311	3.7790	2.7474	5.2810	5.0166	1.4398	2.5307	2.1785	2.0708	0.7427	0.5808	0.8271	0.7487
	150	4.9747	3.9420	2.9481	2.9481	3.1353	2.0870	3.3786	3.2766	1.3085	2.4907	2.0358	1.5181	0.6213	0.3795	0.5363	0.4824
	200	2.7902	2.3894	1.8350	1.8350	2.1707	1.3487	1.9560	1.9215	0.9977	1.9515	1.4803	0.9913	0.4586	0.2909	0.3492	0.3010
	300	1.9752	1.7565	1.4396	1.4396	1.6631	0.9894	1.4677	1.4550	0.8212	1.5677	1.2487	0.7804	0.3680	0.1622	0.1709	0.1474
	400	1.3726	1.2224	1.0014	1.0014	1.2074	0.7226	1.0184	1.0661	0.6598	1.1884	1.0039	0.6259	0.2731	0.1312	0.1324	0.1143
0.99	50	224.6266	133.1416	117.7330	117.7330	6.4006	23.3219	145.6016	134.0307	2.7253	0.5511	2.4014	38.1833	1.0479	0.8056	1.5100	1.3295
	150	45.6679	32.1013	25.0650	25.0650	5.4541	12.3144	30.2648	29.2505	2.0631	0.9091	1.8923	10.6451	1.2642	0.5692	1.1834	1.0402
	150	26.6899	21.1364	16.5910	16.5910	6.3364	8.4746	18.6381	17.0394	2.0629	1.4275	2.0042	6.5258	1.7247	0.3591	0.7869	0.6792
	200	16.5518	12.9521	9.6456	9.6456	5.4292	5.9350	11.1317	10.3195	1.9041	1.9134	2.0424	4.1712	1.2661	0.2514	0.5495	0.4835
	300	8.5828	7.1642	5.3534	5.3534	4.6558	3.6995	5.9379	5.4118	1.7028	2.5801	2.1512	2.4129	1.3575	0.1367	0.2615	0.2337
	400	6.7322	5.6428	4.1643	4.1643	3.9672	2.9926	4.6288	4.2021	1.5484	2.5519	2.0452	1.9070	1.2166	0.1068	0.1880	0.1731

Table 8. Estimated MSE values for different estimators at $p = 10$ and $\varphi = 0.5$. Bold values indicate the minimum MSE.

other years is available through [Open Canada](#). The dataset includes 245 vehicles, with CO₂ emissions in grams per kilometer (y) as the response variable and six explanatory variables: motor power in kilowatts (x_1), engine size in liters (x_2), number of cylinders (x_3), city fuel consumption in liters per 100 kilometers (x_4), highway fuel consumption in liters per 100 kilometers (x_5), and combined fuel consumption in liters per 100 kilometers (x_6). Fuel consumption metrics are provided separately for city and highway driving, along with a combined rating (55% city and 45% highway) expressed in both liters per 100 kilometers and miles per gallon. CO₂ emissions are derived from combined city and highway driving data and are reported in grams per kilometer.

Initially, the fit of the data to the chosen model is carefully evaluated using commonly used criteria such as log-likelihood (LL), Akaike Information Criterion (AIC), and Bayesian Information Criterion (BIC). The best model is selected based on having the highest LL value and the lowest values for both AIC and BIC. Based on the results presented in Table 10, the GPRM model performs best in modeling this data, achieving the highest LL value and the lowest values for both AIC and BIC.

The correlation plot (Fig. 5) indicated strong correlation among the variables, suggesting the presence of multicollinearity. To assess this formally, variance inflation factors (VIFs) and the condition number (CN) were calculated. The CN, defined as the ratio of the largest to the smallest eigenvalue, was 4557.855. The computed VIF values were 1.28, 8.79, 10.38, 2873.80, 804.81, and 6496.13, which clearly demonstrate severe multicollinearity. These results confirm that the variables are highly correlated, raising concerns about the stability and reliability of subsequent regression analyses.

Table 11 compares the GPMLE and the GPRRE for a dataset with seven regression coefficients (β_0 to β_6). The coefficients are estimated using Eqs. (8) and (12), while the MSE is computed using Eqs. (11) and (16), respectively. The GPMLE yields an MSE of 29.2930, indicating poor performance likely due to multicollinearity. In contrast, the GPRRE, evaluated using 15 ridge parameters (k_1^* to k_{15}^*), demonstrates significantly lower MSE values, with the best performance achieved by k_{14}^* (MSE = 1.3728). This substantial improvement highlights the effectiveness of the ridge estimator in mitigating multicollinearity by stabilizing coefficient estimates and

ρ	n	GPMLE	GPRRE														
		–	$\hat{k}_1^*$	$\hat{k}_2^*$	$\hat{k}_3^*$	$\hat{k}_4^*$	$\hat{k}_5^*$	$\hat{k}_6^*$	$\hat{k}_7^*$	$\hat{k}_8^*$	$\hat{k}_9^*$	$\hat{k}_{10}^*$	$\hat{k}_{11}^*$	$\hat{k}_{12}^*$	$\hat{k}_{13}^*$	$\hat{k}_{14}^*$	$\hat{k}_{15}^*$
0.80	50	43.9419	24.0414	22.8368	22.8368	4.4846	3.5615	31.7292	25.2534	1.8394	2.7311	2.8121	14.8916	1.0697	1.3603	1.5806	1.4420
	150	6.4898	4.4968	3.7773	3.7773	2.6913	1.4813	4.3489	4.0301	1.0811	2.1588	1.4694	1.4894	0.6453	1.0880	1.2134	1.1180
	150	2.9094	2.2065	1.7685	1.7685	1.9340	1.0646	1.9890	1.9819	0.8720	1.7604	1.1886	0.9321	0.4685	0.8918	1.0132	0.9200
	200	1.8857	1.5842	1.2773	1.2773	1.5143	0.8172	1.3579	1.3641	0.7241	1.4320	0.9814	0.7210	0.3704	0.6912	0.7456	0.6464
	300	0.8863	0.7920	0.6941	0.6941	0.8091	0.4554	0.7041	0.7332	0.4707	0.8046	0.6434	0.4555	0.2771	0.4992	0.4913	0.3961
	400	0.7087	0.6395	0.5684	0.5684	0.6584	0.3868	0.5734	0.6065	0.4111	0.6590	0.5551	0.3968	0.2388	0.4170	0.4034	0.3224
0.85	50	51.2366	26.2680	24.0040	24.0040	4.1655	507.8358	29.1396	27.7856	1.7760	2.5333	2.6531	8.7187	1.0204	1.3795	1.6148	1.4918
	150	12.1771	7.7588	6.5989	6.5989	3.4575	2.2944	7.9762	7.5613	1.3622	2.4683	1.8194	2.5646	0.7606	1.1845	1.4128	1.3114
	150	3.6171	2.8546	2.2705	2.2705	2.2510	1.2269	2.5159	2.4169	0.9150	1.9513	1.1892	1.0698	0.4788	0.8718	0.9816	0.8710
	200	2.6577	2.1716	1.6957	1.6957	1.9945	1.1324	1.8540	1.8306	0.8936	1.8000	1.1825	0.9035	0.4247	0.7397	0.8374	0.7483
	300	1.5448	1.3575	1.1086	1.1086	1.3302	0.7505	1.1457	1.1546	0.6767	1.2751	0.9190	0.6463	0.2990	0.4809	0.5020	0.4232
	400	0.9851	0.8918	0.7834	0.7834	0.8962	0.5042	0.7919	0.8124	0.5098	0.8872	0.7003	0.4933	0.2457	0.4019	0.3841	0.3105
0.90	50	85.4574	45.4213	42.0775	42.0775	5.2160	5.6672	52.0209	47.3936	2.0616	2.6698	3.0893	13.0073	1.2579	1.3704	1.6454	1.5057
	150	17.4779	11.2047	9.4451	9.4451	3.9037	2.8437	11.4164	10.6959	1.4504	2.3276	1.8647	3.4581	0.7999	1.1961	1.4560	1.3494
	150	7.8280	5.7391	4.6553	4.6553	3.4554	2.1673	5.3750	5.1165	1.3159	2.4686	1.5645	1.9888	0.6329	0.8680	1.0941	1.0019
	200	4.1110	3.2349	2.4796	2.4796	2.5920	1.5697	2.8052	2.7231	1.0702	2.1852	1.3371	1.2183	0.5024	0.7535	0.9032	0.8181
	300	2.3133	1.9535	1.5203	1.5203	1.8224	1.0549	1.6262	1.6314	0.8440	1.6991	1.1110	0.8301	0.3724	0.5407	0.6110	0.5298
	400	1.4976	1.3322	1.1042	1.1042	1.2996	0.7510	1.1277	1.1343	0.6735	1.2566	0.9227	0.6314	0.2849	0.3776	0.3831	0.3212
0.95	50	202.3284	108.4413	114.3752	114.3752	5.3887	10.8222	118.8485	114.3699	2.4270	1.5071	3.0951	136.3603	1.1029	1.3907	1.7543	1.6204
	150	24.0414	16.0191	13.2113	13.2113	5.0109	5.0292	15.8748	14.6336	1.6450	2.0407	1.6839	4.7445	0.9850	1.1917	1.5089	1.4045
	150	13.2234	9.6023	7.6636	7.6636	4.3308	3.4118	8.9754	8.4137	1.5173	2.3241	1.6254	3.0560	0.9286	0.9334	1.2421	1.1462
	200	6.2824	5.0350	3.8973	3.8973	3.4608	2.2541	4.3478	4.0954	1.2274	2.4656	1.3977	1.7045	0.6826	0.7811	0.9907	0.8818
	300	4.6646	3.8960	2.9911	2.9911	3.0593	1.9392	3.2716	3.0550	1.2173	2.3697	1.3611	1.3896	0.6735	0.5267	0.6691	0.6067
	400	2.8593	2.3807	1.7980	1.7980	2.1443	1.3153	1.9456	1.9745	0.9732	1.9688	1.2805	0.9731	0.4700	0.4314	0.5197	0.4561
0.99	50	1016.4337	521.0672	479.1481	479.1481	4.5520	54.7318	597.0681	545.1915	3.1899	0.7824	1.6398	138.3801	0.8867	1.3948	1.8611	1.7616
	150	148.3388	94.5453	79.9782	79.9782	3.7671	21.8796	96.8118	89.2753	2.0481	0.5560	0.9642	25.0264	0.6954	1.1844	1.7162	1.5920
	150	70.4286	51.4349	41.6901	41.6901	5.5590	16.0473	48.1100	43.1926	2.0389	0.6651	0.9603	14.1791	1.0333	0.9111	1.4824	1.3405
	200	46.2591	34.5278	27.4815	27.4815	5.4063	11.9639	31.7283	29.2801	1.9306	0.9042	0.9515	10.2964	1.2540	0.7319	1.3018	1.1534
	300	21.6349	17.2964	13.5154	13.5154	5.9929	6.9744	15.1198	13.7596	1.8851	1.6330	1.1961	5.3072	1.8774	0.4880	0.9212	0.8144
	400	14.3064	11.4588	8.5342	8.5342	5.1806	5.2026	9.6996	8.7160	1.7196	1.9678	1.3351	3.4466	1.7150	0.3963	0.7279	0.6543

Table 9. Estimated MSE values for different estimators at $p = 10$ and $\varphi = 1$. Bold values indicate the minimum MSE.

reducing overfitting. The consistent performance of GPRRE across various ridge parameters underscores its robustness, making it a superior choice for datasets affected by multicollinearity.

Figure (6) compares the performance of the GPMLE and GPRRE estimators based on MSE values across a range of ridge parameters (k^*). The graph illustrates that GPRRE significantly outperforms GPMLE in reducing MSE, particularly as k^* increases. The MSE associated with the GPRRE demonstrates a substantial reduction, achieving its minimum at higher values of the ridge parameter k^* . In contrast, the MSE for the Generalized Poisson Maximum Likelihood Estimator (GPMLE) remains consistently larger across the entire range of k^* values. This performance improvement is attributed to the ridge estimator's mechanism for mitigating multicollinearity, which stabilizes coefficient estimates and reduces their variance through the introduction of a penalty term contingent upon k^* . Furthermore, the accompanying figure delineates the bias-squared and variance components of the GPRRE, illustrating how the estimator successfully negotiates a trade-off between these two elements. Although the squared bias experiences a marginal increase, it is offset by a considerable reduction in variance, culminating in a net decrease in the overall MSE. These findings underscore the efficacy of the GPRRE in enhancing predictive performance, particularly in contexts where the data exhibit pronounced multicollinearity.

Table 12 presents the results of the verification process for the necessary condition associated with Theorem 2 under the GPRRE. The analysis evaluates this condition across a range of proposed ridge parameters (k_i^*) and for each of the seven coefficients (λ_1 to λ_7). The computed values of the condition are consistently positive for all combinations of k^* and λ_j , thereby confirming that the requisite condition is satisfied. This uniform positivity demonstrates the efficacy of the GPRRE in stabilizing the model and reducing estimation variance, even in the presence of significant multicollinearity. Consequently, these results underscore the capability of the GPRRE to manage complex data structures while simultaneously enhancing model performance and reliability.

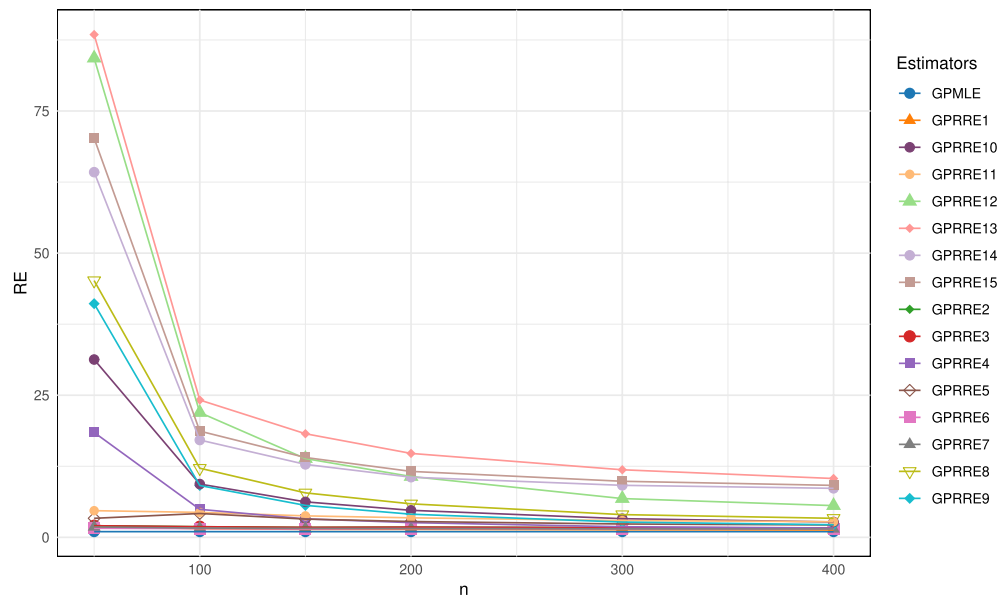


Fig. 1. Estimated MSE of the GPRRE under different parameter settings and sample sizes.

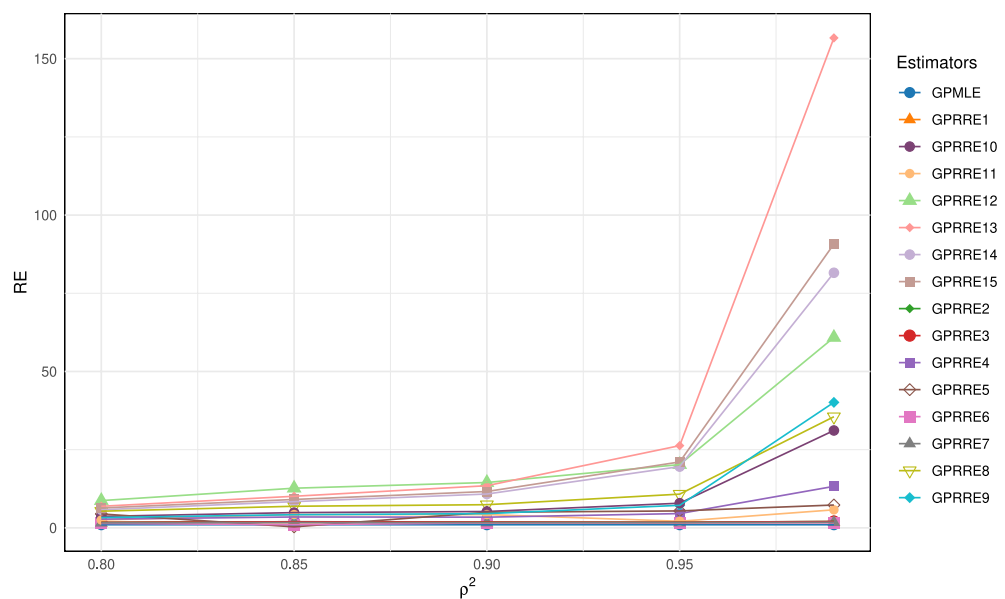


Fig. 2. Estimated MSE of the GPRRE under different parameter settings and multicollinearity level.

Conclusion

The GPRM serves as a robust framework for analyzing count data, particularly in cases of overdispersion or underdispersion. While the GPMLE is widely employed, its efficacy is compromised in the presence of multicollinearity among explanatory variables. To mitigate this issue, this study introduces ridge estimators within the generalized GPRRE framework and explores methodologies for optimal ridge parameter selection. The theoretical properties of the ridge estimator are rigorously derived, and its performance is assessed using both MMSE and MSE criteria. A detailed Monte Carlo simulation study is conducted, examining various parametric configurations, including sample sizes, numbers of predictor variables, dispersion levels, and degrees of multicollinearity. Furthermore, the proposed estimators are applied to a real-world dataset concerning carbon dioxide emissions. The results from both the simulation study and the empirical application consistently demonstrate that the GPRRE, particularly when paired with an optimally selected ridge parameter, significantly outperforms the GPMLE in reducing MSE under conditions of severe multicollinearity, especially for parameter values k_{13}^* and k_{14}^* . These findings underscore the GPRRE as a superior estimation technique for addressing multicollinearity in overdispersed count data. Despite the positive results of this study, it is not without limitations. The performance of the GPRRE method depends primarily on selecting the optimal value

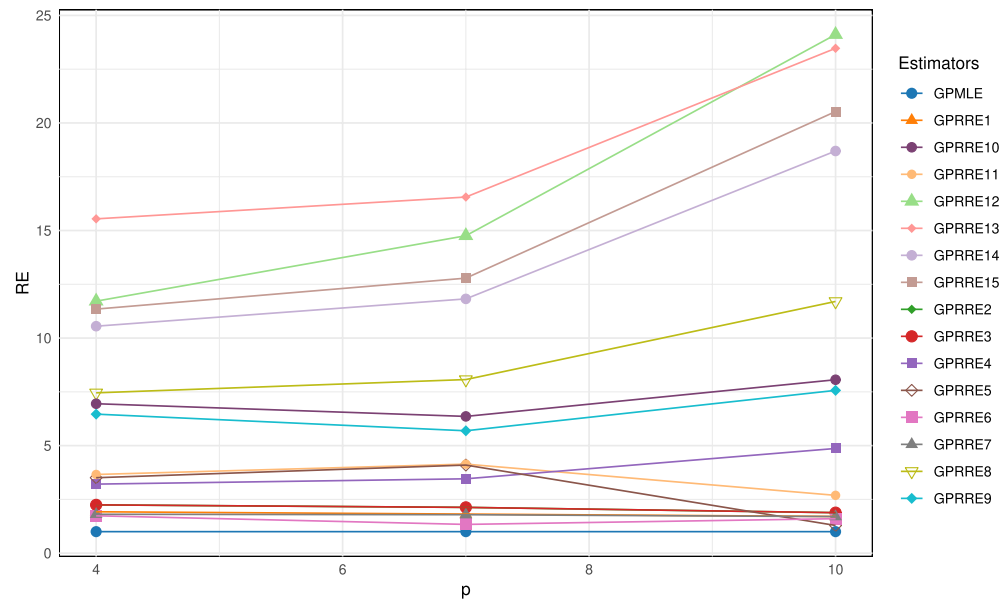


Fig. 3. Estimated MSE of the GPRRE under different parameter settings and number of explanatory variables.

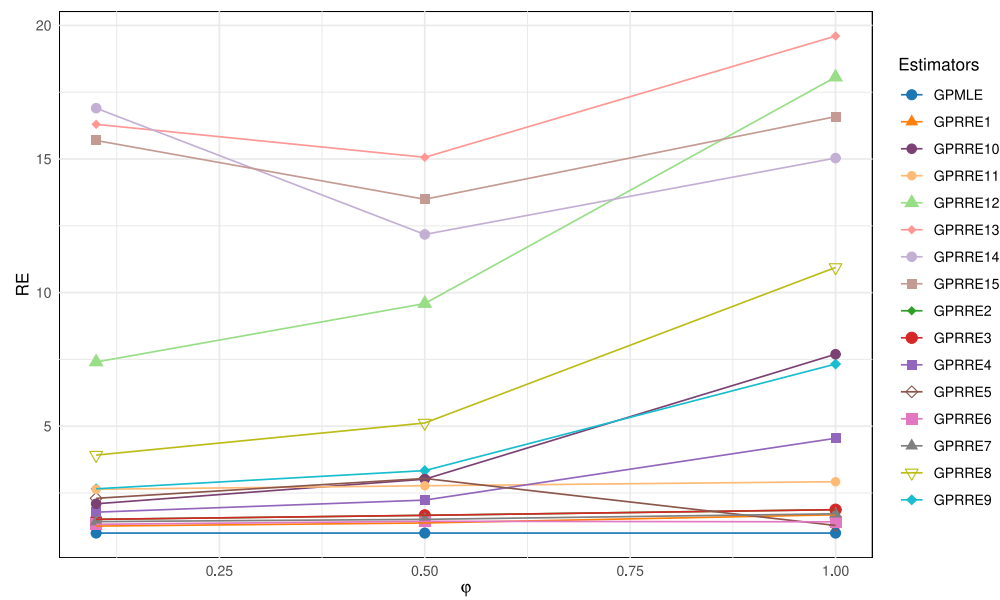


Fig. 4. Estimated MSE of the GPRRE under different parameter settings and dispersion parameter.

Criteria	GPRM	PRM	NBRM	PIGRM	DPORM	GERM	CMPRM
AIC	2278.56	3533.44	2288.82	2301.36	2328.97	2730.21	5531.51
BIC	2310.07	3561.45	2320.33	2332.87	2360.48	2758.22	5563.02
LL	−1130.28	−1758.72	−1135.41	−1141.68	−1155.49	−1357.10	−2756.75

Table 10. Comparison of model performance for the dataset.

for the shrinkage parameter (k^*), which is highly sensitive and balances bias and variance. Furthermore, the high dimensionality of the data increases the complexity of the calculations and poses significant challenges that require further research and improvement. Additionally, the current application of this method is limited to addressing multicollinearity in the Generalized Poisson Regression Model (GPRM), which opens promising

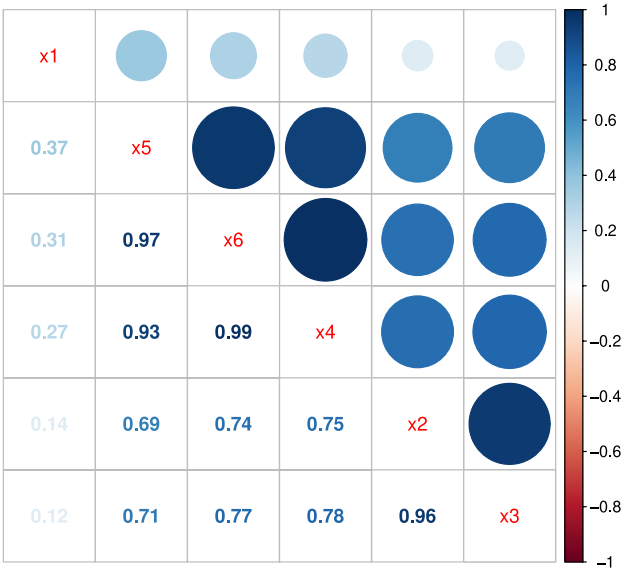


Fig. 5. Correlation matrix illustrating the relationships between explanatory variables in the dataset.

Estimator		β_0	β_1	β_2	β_3	β_4	β_5	β_6	MSE
GPMLE	–	3.1606	– 0.0034	– 0.0161	– 0.0147	0.0864	0.1021	0.0128	29.2930
	$\hat{k}_1^*$	3.0143	– 0.0034	– 0.0273	0.0007	0.0528	0.1127	0.0487	2.7943
GPRRE	$\hat{k}_2^*$	2.9869	– 0.0034	– 0.0292	0.0035	0.0491	0.1167	0.0508	2.3287
	$\hat{k}_3^*$	2.9869	– 0.0034	– 0.0292	0.0035	0.0491	0.1167	0.0508	2.3287
	$\hat{k}_4^*$	3.1597	– 0.0034	– 0.0162	– 0.0146	0.0859	0.1019	0.0136	28.4057
	$\hat{k}_5^*$	3.1080	– 0.0034	– 0.0202	– 0.0092	0.0687	0.1014	0.0359	8.0755
	$\hat{k}_6^*$	3.1073	– 0.0034	– 0.0203	– 0.0091	0.0686	0.1014	0.0360	7.9869
	$\hat{k}_7^*$	3.1576	– 0.0034	– 0.0163	– 0.0144	0.0848	0.1015	0.0152	26.3976
	$\hat{k}_8^*$	3.0658	– 0.0034	– 0.0235	– 0.0047	0.0606	0.1057	0.0433	4.4871
	$\hat{k}_9^*$	3.0900	– 0.0034	– 0.0216	– 0.0073	0.0650	0.1030	0.0396	6.0952
	$\hat{k}_{10}^*$	3.1605	– 0.0034	– 0.0161	– 0.0147	0.0863	0.1020	0.0129	29.1616
	$\hat{k}_{11}^*$	3.1432	– 0.0034	– 0.0175	– 0.0129	0.0786	0.1003	0.0240	17.1243
	$\hat{k}_{12}^*$	2.8888	– 0.0034	– 0.0360	0.0136	0.0370	0.1319	0.0563	1.5296
	$\hat{k}_{13}^*$	2.9429	– 0.0034	– 0.0323	0.0081	0.0435	0.1234	0.0535	1.8580
	$\hat{k}_{14}^*$	2.8458	– 0.0034	– 0.0387	0.0179	0.0320	0.1388	0.0581	1.3728
	$\hat{k}_{15}^*$	2.9580	– 0.0034	– 0.0313	0.0065	0.0454	0.1211	0.0527	1.9912

Table 11. Estimated MSE and coefficient of GPMLE and GPRRE for the dataset.

avenues for future development to address other issues, such as outliers, building upon previous research, such as the study by Dawoud et al.⁴⁵, Abonazel and Dawoud⁴⁶, Mohammad et al.⁴⁷, and Alghamdi et al.⁴⁸.

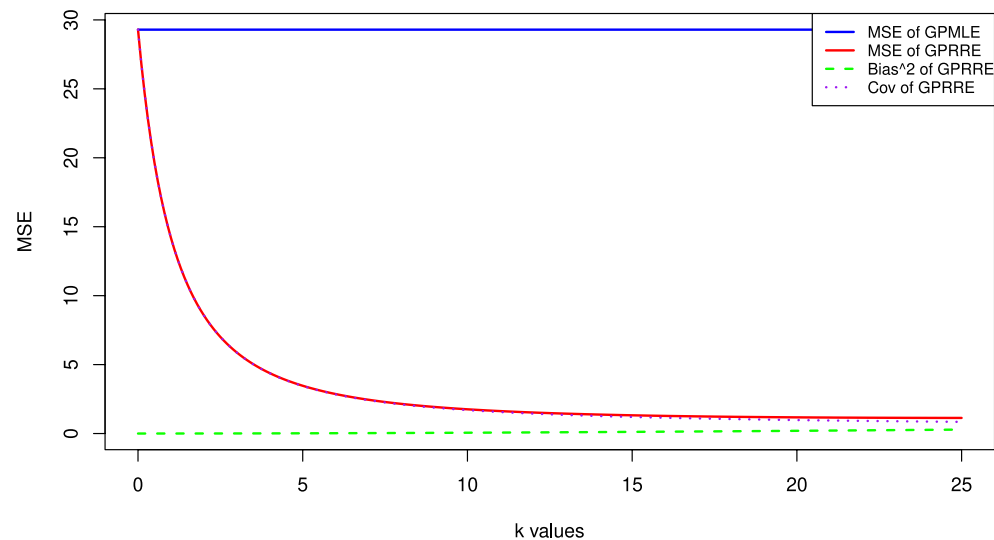


Fig. 6. Estimated MSE with different values of ridge parameter.

Estimator	Condition	λ_1	λ_2	λ_3	λ_4	λ_5	λ_6	λ_7
GPRRE	$\hat{k}_1^*$ $\hat{\varphi}k_1^* + 2\hat{\varphi}\lambda - k_1^*\lambda\alpha_j^2$	5.53E+09	2.30E+07	5.26E+05	2.46E+05	2.17E+04	8.61E+03	6.44E+02
	$\hat{k}_2^*$ $\hat{\varphi}k_2^* + 2\hat{\varphi}\lambda - k_2^*\lambda\alpha_j^2$	5.53E+09	2.30E+07	5.26E+05	2.45E+05	2.15E+04	7.52E+03	7.19E+02
	$\hat{k}_3^*$ $\hat{\varphi}k_3^* + 2\hat{\varphi}\lambda - k_3^*\lambda\alpha_j^2$	5.53E+09	2.30E+07	5.26E+05	2.45E+05	2.15E+04	7.52E+03	7.19E+02
	$\hat{k}_4^*$ $\hat{\varphi}k_4^* + 2\hat{\varphi}\lambda - k_4^*\lambda\alpha_j^2$	5.54E+09	2.31E+07	5.27E+05	2.51E+05	2.25E+04	1.41E+04	2.69E+02
	$\hat{k}_5^*$ $\hat{\varphi}k_5^* + 2\hat{\varphi}\lambda - k_5^*\lambda\alpha_j^2$	5.54E+09	2.31E+07	5.27E+05	2.49E+05	2.22E+04	1.22E+04	3.98E+02
	$\hat{k}_6^*$ $\hat{\varphi}k_6^* + 2\hat{\varphi}\lambda - k_6^*\lambda\alpha_j^2$	5.54E+09	2.31E+07	5.27E+05	2.49E+05	2.22E+04	1.22E+04	4.00E+02
	$\hat{k}_7^*$ $\hat{\varphi}k_7^* + 2\hat{\varphi}\lambda - k_7^*\lambda\alpha_j^2$	5.54E+09	2.31E+07	5.27E+05	2.50E+05	2.25E+04	1.40E+04	2.74E+02
	$\hat{k}_8^*$ $\hat{\varphi}k_8^* + 2\hat{\varphi}\lambda - k_8^*\lambda\alpha_j^2$	5.53E+09	2.30E+07	5.27E+05	2.47E+05	2.20E+04	1.06E+04	5.07E+02
	$\hat{k}_9^*$ $\hat{\varphi}k_9^* + 2\hat{\varphi}\lambda - k_9^*\lambda\alpha_j^2$	5.53E+09	2.30E+07	5.27E+05	2.48E+05	2.21E+04	1.15E+04	4.44E+02
	$\hat{k}_{10}^*$ $\hat{\varphi}k_{10}^* + 2\hat{\varphi}\lambda - k_{10}^*\lambda\alpha_j^2$	5.54E+09	2.31E+07	5.27E+05	2.51E+05	2.25E+04	1.41E+04	2.67E+02
	$\hat{k}_{11}^*$ $\hat{\varphi}k_{11}^* + 2\hat{\varphi}\lambda - k_{11}^*\lambda\alpha_j^2$	5.54E+09	2.31E+07	5.27E+05	2.50E+05	2.24E+04	1.35E+04	3.09E+02
	$\hat{k}_{12}^*$ $\hat{\varphi}k_{12}^* + 2\hat{\varphi}\lambda - k_{12}^*\lambda\alpha_j^2$	5.53E+09	2.29E+07	5.26E+05	2.41E+05	2.08E+04	3.41E+03	1.00E+03
	$\hat{k}_{13}^*$ $\hat{\varphi}k_{13}^* + 2\hat{\varphi}\lambda - k_{13}^*\lambda\alpha_j^2$	5.53E+09	2.30E+07	5.26E+05	2.43E+05	2.12E+04	5.71E+03	8.43E+02
	$\hat{k}_{14}^*$ $\hat{\varphi}k_{14}^* + 2\hat{\varphi}\lambda - k_{14}^*\lambda\alpha_j^2$	5.53E+09	2.29E+07	5.25E+05	2.39E+05	2.05E+04	1.51E+03	1.13E+03
	$\hat{k}_{15}^*$ $\hat{\varphi}k_{15}^* + 2\hat{\varphi}\lambda - k_{15}^*\lambda\alpha_j^2$	5.53E+09	2.30E+07	5.26E+05	2.44E+05	2.13E+04	6.33E+03	8.00E+02

Table 12. Results confirming the necessary condition for Theorem 2.

Data availability

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

Received: 23 June 2025; Accepted: 10 October 2025
Published online: 10 November 2025

References

1. Meraou, M. A., Al-Kandari, N. M. & Mohammad, R. Z. Fundamental properties of the characteristic function using the compound Poisson distribution as the sum of the gamma model. *Modern J. Stat.* **1**(1), 49–57 (2025).
2. Gemeay, A. M., Moakofi, T., Balogun, O. S., Ozkan, E. & Hossain, M. M. Analyzing real data by a new heavy-tailed statistical model. *Modern J. Stat.* **1**(1), 1–24 (2025).
3. Hilbe, J. M. *Negative Binomial Regression* (Cambridge University Press, 2011).
4. Bring, J. A geometric approach to compare variables in a regression model. *Am. Stat.* **50**(1), 57–62 (1996).

5. Sellers, K. F. & Premeaux, B. Conway–Maxwell–Poisson regression models for dispersed count data. *Wiley Interdiscip. Rev. Comput. Stat.* **13**(6), e1533 (2021).
6. Efron, B. Double exponential families and their use in generalized linear regression. *J. Am. Stat. Assoc.* **81**(395), 709–721 (1986).
7. Putri, G. N., Nurrohman, S. & Fithriani, I. Comparing Poisson-inverse Gaussian model and negative binomial model on case study: Horseshoe crabs data. *J. Phys. Conf. Ser.* **1442**(1), 012028 (2020).
8. Consul, P. & Famoye, F. Generalized Poisson regression model. *Commun. Stat. Theory Methods* **21**(1), 89–109 (1992).
9. Yadav, B. et al. Can generalized Poisson model replace any other count data models? An evaluation. *Clin. Epidemiol. Global Health* **11**, 100774 (2021).
10. Laís, M. F., Atti, A., Pangaribuan, R. M. & Guntur, R. D. Model generalized Poisson regression (gpr) pada kasus stunting di provinsi Nusa Tenggara timur. *J. Difer.* **5**(2), 68–75 (2023).
11. Getaneh, F. B. et al. A generalized Poisson regression analysis of determinants of early neonatal mortality in Ethiopia using 2019 Ethiopian mini demographic health survey. *Sci. Rep.* **14**(1), 2784 (2024).
12. Odoi, B., Ofori, R. A. & William, K. A generalised Poisson regression analysis of covid-19 cases in Ghana. *Int. J. Stat. Appl. Math.* **9**(2), 131–136 (2024).
13. Carallo, G., Casarin, R. & Robert, C. P. Generalized Poisson difference autoregressive processes. *Int. J. Forecast.* **40**(4), 1359–1390 (2024).
14. Dawoud, I. & Eledum, H. Detection of influential observations for the regression model in the presence of multicollinearity: Theory and methods. *Communications in Statistics-Theory and Methods* 1–26 (Taylor & Francis, 2025).
15. Algamal, Z. Y. & Lee, M. H. Adjusted adaptive lasso in high-dimensional Poisson regression model. *Mod. Appl. Sci.* **9**(4), 170 (2015).
16. Hoerl, A. E. & Kennard, R. W. Ridge regression: Applications to nonorthogonal problems. *Technometrics* **12**(1), 69–82 (1970).
17. Hoerl, A. E. & Kennard, R. W. Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics* **12**(1), 55–67 (1970).
18. Kibria, B. G. Performance of some new ridge regression estimators. *Commun. Stat. Simul. Comput.* **32**(2), 419–435 (2003).
19. Muniz, G. & Kibria, B. G. On some ridge regression estimators: An empirical comparisons. *Commun. Stat. Simul. Comput.* **38**(3), 621–630 (2009).
20. Segerstedt, B. On ordinary ridge regression in generalized linear models. *Commun. Stat. Theory Methods* **21**(8), 2227–2246 (1992).
21. Månsson, K. & Shukur, G. A Poisson ridge regression estimator. *Econ. Model.* **28**(4), 1475–1481 (2011).
22. Schaefer, R. L., Roi, L. D. & Wolfe, R. A. A ridge logistic estimator. *Commun. Stat. Theory Methods* **13**(1), 99–113 (1984).
23. Rady, E. A., Abonazel, M. R., & Taha, I. M. Ridge estimators for the negative binomial regression model with application. in *The 53rd Annual Conference on Statistics, Computer Science, and Operation Research* 3–5 (2018).
24. Månsson, K. On ridge estimators for the negative binomial regression model. *Econ. Model.* **29**(2), 178–184 (2012).
25. Tharshan, R. & Wijekoon, P. Ridge estimator in a mixed Poisson regression model. *Commun. Stat. Simul. Comput.* **53**(7), 3253–3270 (2024).
26. Almulhim, F. A. et al. Development of the generalized ridge estimator for the Poisson-inverse Gaussian regression model with multicollinearity. *Sci. Rep.* **15**(1), 31162 (2025).
27. Abonazel, M. R. et al. Developing ridge estimators for the extended Poisson-tweedie regression model: Method, simulation, and application. *Sci. Afr.* **23**, e02006 (2024).
28. Sami, F., Amin, M. & Butt, M. M. On the ridge estimation of the Conway–Maxwell Poisson regression model with multicollinearity: Methods and applications. *Concurr. Comput. Pract. Exp.* **34**(1), e6477 (2022).
29. Algamal, Z. Y., Lukman, A. F., Abonazel, M. R. & Awwad, F. A. Performance of the ridge and Liu estimators in the zero-inflated bell regression model. *J. Math.* **2022**(1), 9503460 (2022).
30. Dawoud, I. New biased estimators for the Conway–Maxwell–Poisson model. *J. Stat. Comput. Simul.* **95**(1), 117–136 (2025).
31. El-Alosey, A. R., Hammad, A. T., & Gemeay, A. M. A novel zero-inflated regression model for overdispersed count data with enhancing its estimation for multicollinearity in medical data. *Statistics* 1–32 (Taylor & Francis, 2025).
32. Akram, M. N., Kibria, B. G., Abonazel, M. R. & Afzal, N. On the performance of some biased estimators in the gamma regression model: Simulation and applications. *J. Stat. Comput. Simul.* **92**(12), 2425–2447 (2022).
33. Shahzad, A., Amin, M., Emam, W. & Faisal, M. New ridge parameter estimators for the quasi-Poisson ridge regression model. *Sci. Rep.* **14**(1), 8489 (2024).
34. Ashraf, B., Amin, M. & Akram, M. N. New ridge parameter estimators for the zero-inflated Conway Maxwell Poisson ridge regression model. *J. Stat. Comput. Simul.* **94**(8), 1814–1840 (2024).
35. Consul, P. C. & Jain, G. C. A generalization of the Poisson distribution. *Technometrics* **15**(4), 791–799 (1973).
36. Consul, P. C. *Generalized Poisson distributions: Properties and applications* (Marcel Dekker, New York, 1989).
37. Famoye, F. Restricted generalized Poisson regression model. *Commun. Stat. Theory Methods* **22**(5), 1335–1354 (1993).
38. Dutang, C. Some explanations about the IWLS algorithm to fit generalized linear models. *HAL Open Science* (2017).
39. Amin, M., Akram, M. N. & Majid, A. On the estimation of bell regression model using ridge estimator. *Commun. Stat. Simul. Comput.* **52**(3), 854–867 (2023).
40. Orji, G. O. et al. A new odd reparameterized exponential transformed-x family of distributions with applications to public health data. *Innov. Stat. Prob.* **1**(1), 88–118 (2025).
41. Rigby, R. A., Stasinopoulos, M. D., Heller, G. Z. & De Bastiani, F. *Distributions for Modeling Location, Scale, and Shape: Using GAMLSS in R* (Chapman and Hall/CRC, 2019).
42. Abonazel, M. R. A new biased estimation class to combat the multicollinearity in regression models: Modified two-parameter Liu estimator. *Comput. J. Math. Stat. Sci.* **4**(1), 316–347 (2025).
43. Dawoud, I. & Abonazel, M. R. Robust Dawoud–Kibria estimator for handling multicollinearity and outliers in the linear regression model. *J. Stat. Comput. Simul.* **91**(17), 3678–3692 (2021).
44. Hammad, A. T. et al. New modified Liu estimators to handle the multicollinearity in the beta regression model: Simulation and applications. *Modern J. Stat.* **1**(1), 58–79 (2025).
45. Dawoud, I., Awwad, F. A., Tag Eldin, E. & Abonazel, M. R. New robust estimators for handling multicollinearity and outliers in the Poisson model: Methods, simulation and applications. *Axioms* **11**(11), 612 (2022).
46. Abonazel, M. R. & Dawoud, I. Developing robust ridge estimators for Poisson regression model. *Concurr. Comput. Pract. Exp.* **34**(15), e6979 (2022).
47. Mohammad, H. H. et al. New robust two-parameter estimator for overcoming outliers and multicollinearity in Poisson regression model. *Sci. Rep.* **15**(1), 27445 (2025).
48. Alghamdi, F. M. et al. On robust and non-robust modified Liu estimation in Poisson regression model with multicollinearity and outliers. *Internat. J. Uncertain. Fuzziness Knowl.-Based Syst.* **33**(06), 787–823 (2025).

Acknowledgements

Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2025R735), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

Author contributions

Fatimah M. Alghamdi: Conceptualization, Validation, Methodology, Formal analysis, Data curation, Software, Writing -original draft, Writing - review & editing. Ahmed M. Gemeay: Conceptualization, Validation, Methodology, Formal analysis, Data curation, Software, Writing -original draft, Writing - review & editing. Gamal A. Abd-Elmougod: Conceptualization, Validation, Methodology, Formal analysis, Data curation, Software, Writing -original draft, Writing - review & editing. Ehab M. Almetwally: Conceptualization, Validation, Methodology, Formal analysis, Data curation, Software, Writing -original draft, Writing - review & editing. M. A. El-Qurashi: Conceptualization, Validation, Methodology, Formal analysis, Data curation, Software, Writing -original draft, Writing - review & editing. Getachew Tekle Mekiso: Conceptualization, Validation, Methodology, Formal analysis, Data curation, Software, Writing -original draft, Writing - review & editing. Ali T. Hammad: Conceptualization, Validation, Methodology, Formal analysis, Data curation, Software, Writing -original draft, Writing - review & editing.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to G.T.M.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025