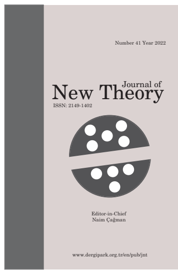
Journal of New Theory
ISSN: 2149-1402

41 (2022) 1-17
Journal of New Theory
https://dergipark.org.tr/en/pub/jnt
Open Access



Kibria-Lukman Estimator for General Linear Regression Model with AR(2) Errors: A Comparative Study with Monte Carlo Simulation

Tuğba Söküt Açar[1]

Article Info
Received: 04 Jul 2022
Accepted: 11 Nov 2022
Published: 31 Dec 2022
doi:10.53570/jnt.1139885
Research Article

Abstract — The sensitivity of the least-squares estimation in a regression model is impacted by multicollinearity and autocorrelation problems. To deal with the multicollinearity, Ridge, Liu, and Ridge-type biased estimators have been presented in the statistical literature. The recently proposed Kibria-Lukman estimator is one of the Ridge-type estimators. The literature has compared the Kibria-Lukman estimator with the others using the mean square error criterion for the linear regression model. It was achieved in a study conducted on the Kibria-Lukman estimator's performance under the first-order autoregressive erroneous autocorrelation. When there is an autocorrelation problem with the second-order, evaluating the performance of the Kibria-Lukman estimator according to the mean square error criterion makes this paper original. The scalar mean square error of the Kibria-Lukman estimator under the second-order autoregressive error structure was evaluated using a Monte Carlo simulation and two real examples, and compared with the Generalized Least-squares, Ridge, and Liu estimators. The findings revealed that when the variance of the model was small, the mean square error of the Kibria-Lukman estimator gave very close values with the popular biased estimators. As the model variance grew, Kibria-Lukman did not give fairly similar values with popular biased estimators as in the model with small variance. However, according to the mean square error criterion the Kibria-Lukman estimator outperformed the Generalized Least-Squares estimator in all possible cases.

Keywords — *Autocorrelation, multicollinearity, second-order autoregressive errors, Kibria-Lukman estimator*

Mathematics Subject Classification (2020) — 62J07, 62M10

1. Introduction

Regression analysis is widely used to create a functional model based on the relationship between an observed dependent variable (response) and one or more observed independent variables (regressors). A linear regression model is one in which the variables are included in the model as a first-order polynomial. The model is called a simple linear regression model if there is only one independent variable, and a multiple linear regression model if there is more than one independent variable. The following is a formula for a multiple linear regression model with p independent variables:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_p x_p + \varepsilon \tag{1}$$

It takes on the matrix form as

$$y = X\beta + \varepsilon$$

[1]t.sokut@comu.edu.tr (Corresponding Author)
[1]Department of Statistics, Faculty of Science, Çanakkale Onsekiz Mart University, Çanakkale, Türkiye

where y is an $n \times 1$ vector of observed response variable, $X = [1, x]$, $x = (x_1, x_2, .., x_p)$ with $x_j = (x_{1j}, x_{2j}, ..., x_{nj})'$ for $j = 1, ..., p$, is an $n \times (p+1)$ vector of known regressor matrix whose first column equals to one, β is an $(p+1) \times 1$ vector of unknown regression parameters and ε is an $n \times 1$ vector of errors with properties $E(\varepsilon) = 0$ and $E(\varepsilon\varepsilon') = \sigma^2 I_n$. The unknown regression parameters must be calculated to determine the functional relationship between the dependent and independent variables. If the following assumptions are met, the Ordinary Least-Squares (OLS) can be used to estimate unknown parameters in a regression model: -Regressor matrix is a non-stochastic matrix - Regressor matrix is a full column rank - Response is a linear function of regressors - The error term is normally distributed with zero mean and constant variance. OLS estimation procedure can be applied which is based on the minimization of the sum of squares error

$$\min_{\beta} \left\{ (Y - X\beta)' (Y - X\beta) \right\}$$

as

$$\hat{\beta}_{OLS} = \left(X'X\right)^{-1} X'Y$$

In these circumstances, OLS is the best linear unbiased estimator (BLUE) with $E(\hat{\beta}_{OLS}) = \beta$ and $cov(\hat{\beta}_{OLS}) = \sigma^2 \left(X'X\right)^{-1}$.

Multicollinearity is a term used in data analytics to indicate the occurrence of two regressors that are shown to be associated in a linear regression model. If the matrix $X'X$ is not linearly independent, it will not be full column rank. In this case, the matrix $X'X$ becomes ill-conditioned. The condition number, correlation coefficient, and variance inflation factor are utilized to determine the multicollinearity in a dataset. The condition number, κ, is a value calculated from the eigenvalues of the $X'X$ matrix's characteristic roots or eigenvalues. κ, including $\lambda_1, \lambda_2, ..., \lambda_n$ are the eigenvalues of $X'X$, is determined by $\kappa = \sqrt{\frac{\lambda_{\max}}{\lambda_{\min}}}$. According to Belsley et al. [1], there is no substantial problem with multicollinearity if the κ value is less than 10, moderate to strong collinearity if the κ value is between 30 and 100, and severe multicollinearity if the κ value is greater than 100. It is also a sign that the two variables generate multicollinearity when the correlation coefficient between any two independent variables is close to 1 in absolute value or statistically significant. Multicollinearity can be regarded as a result of these variables. In this scenario, some biased estimators defined in Equation 1 to deal with multicollinearity are presented below. Hoerl et al. [2] suggested a ridge estimator under Equation 1 based on the solution of

$$\min_{\beta} \left\{ (Y - X\beta)' (Y - X\beta) + k \left(\beta'\beta - c\right) \right\}$$

where c is a constant and k is a lagrangian multiplier called the biasing parameter as

$$\hat{\beta}_{ridge} = \left(X'X + kI_p\right)^{-1} X'Y, k > 0$$

It is obvious that the ridge regression's effectiveness will change depending on the k-biasing parameter. As a result, the statistical literature includes the proposed biasing parameters according to various criteria (see, [2–4]). Condition number of $X'X + kI_p$ is a decreasing function of k. As a result, as k increases, the condition number decreases dramatically [5]. In practice, however, k is small, and this may not be enough to solve the ill-conditioned.

Liu [6] proposed a liu estimator which is an alternative to ridge based on the solution of

$$\min_{\beta} \left\{ (Y - X\beta)' (Y - X\beta) + (\beta - d\hat{\beta})(\beta - d\hat{\beta})' \right\}$$

as

$$\hat{\beta}_{liu} = \left(X'X + I_p\right)^{-1} \left(X'X + dI_p\right) \hat{\beta}, 0 < d < 1$$

Özkale and Kaçıranlar [7] stated that the liu estimator is more advantageous than the ridge because it is a linear function of d. Numerous studies on the choice of the d parameter in the liu estimator can be found in the literature (see, [6–8]). Kibria and Lukman [9] proposed a new biased estimator called as Kibria-Lukman (KL) estimator to cope with the multicollinearity which is based on the solving of

$$\min_{\beta}\left\{(Y-X\beta)'(Y-X\beta)+k\left[(\beta+\hat{\beta})'(\beta+\hat{\beta})-c\right]\right\}$$

as

$$\hat{\beta}_{kl}=\left(X'X+kI_p\right)^{-1}\left(X'X-kI_p\right)\hat{\beta}, k>0$$

The user of biased estimators must select a biasing parameter (k or d) in order to see improvements in the estimates [10]. There have been numerous studies on the biasing parameter selection processes (see, [3, 7, 9]).

Many KL estimators have been described, each based on a different distribution (inverse Gaussian regression model, Gamma regression model, Poisson regression model, distributed lag model). (see, [11–15]). The goal of the paper is to apply the mean square error (MSE) criterion to extend the KL estimator's performance from non-autoregressive or first-order autoregressive processes which are in the statistical literature to second-order autoregressive process.

The article is structured as follows: The general linear regression model, error structures, and estimators are provided in Section 2. The method for calculating the MSE, which is used to assess model performance for any estimator, is provided in the next section. Section 4 discusses the Monte Carlo simulation's layout and findings. In Section 5, the performance of the KL estimator in the second-order autoregressive model is examined over two real datasets. The paper's findings and recommendations are presented in the final section.

2. General Linear Regression Model, Error Structure and Estimators

When the variance-covariance matrix of the errors is not diagonal form that is $E(\varepsilon\varepsilon')=\sigma_{\varepsilon}^2 V$, $V\neq I_n$ it is called as general linear regression (GLR) model. There is a violation of the assumption "The error term is normally distributed with zero mean and constant variance", and in this case, the autocorrelation problem arises. Therefore, the errors are correlated. Since V $n\times n$ matrix assumed that known is symmetric and positive definite, then there exits a non-singular $n\times n$ matrix P such that $V^{-1}=P'P$. Premultiplying both sides of Equation 1 by P gives the transformed model as

$$Py=PX\beta+P\varepsilon \tag{2}$$

In the transformed model new error terms has covariance matrix as $E\left[P\varepsilon\left(P\varepsilon\right)'\right]=\sigma^2 I_n$. Therefore, by applying the least-squares estimation procedure which based on $\min_{\beta}\left\{(PY-PX\beta)'(PY-X\beta)\right\}$ and by solving the normal equations, generalized least-squares estimator (GLS) obtained as

$$\hat{\beta}_{GLS}=(X'V^{-1}X)^{-1}X'V^{-1}y$$

GLS is a BLUE estimator with $E[\hat{\beta}_{GLS}]=\beta$ and $cov(\hat{\beta}_{GLS})=\sigma^2(X'V^{-1}X)^{-1}$ under Equation 2. In models with an autocorrelation problem, multicollinearity can also occur. Many biased estimators proposed under the linear regression model has also been extended to GLR models. Ridge Regression (RR) estimator for Equation 2 given by Trenkler [16] as

$$\hat{\beta}_{RR}=(X'V^{-1}X+kI_p)^{-1}X'V^{-1}y, k>0$$

RR is a biased estimator with

$$bias(\hat{\beta}_{RR},\beta)=-k\left(X'V^{-1}X+kI_n\right)^{-1}\beta$$

and variance-covariance matrix is

$$cov(\hat{\beta}_{RR},\beta)=\sigma^2\left(X'V^{-1}X+kI_p\right)^{-1}X'V^{-1}X\left(X'V^{-1}X+kI_p\right)^{-1}$$

Noted that, because β and σ^2 population parameters are unknown in real-world applications, GLS estimations which are the best estimators of these parameters under Equation 2, are utilized. Liu estimator for Equation 2 given by Kaçiranlar [17] as

$$\hat{\beta}_{Liu}=(X'V^{-1}X+I_p)^{-1}(X'V^{-1}y+d\hat{\beta}_{GLS}), 0<d<1$$

This is a biased estimator also, and the expected value and the variance-covariance matrix as follows:

$$bias(\hat{\beta}_{Liu}, \beta) = (d-1)(X'V^{-1}X + I_p)^{-1}\beta$$

and

$$cov(\hat{\beta}_{Liu}, \beta) = \sigma^2 \left(X'V^{-1}X + I_p\right)^{-1}(X'V^{-1}X + dI_p)(X'V^{-1}X)^{-1}(X'V^{-1}X + dI_p)\left(X'V^{-1}X + I_p\right)^{-1}$$

In this study, two different structures of autocorrelation, which is another problem apart from multicollinearity, were examined. The first-order autoregressive (AR(1)) model is one of them, while the second-order autoregressive model (AR(2)) is the other. The error terms for the AR(1) process are satisfied

$$\varepsilon_i = \rho\varepsilon_{i-1} + u_i \tag{3}$$

where $E(u_i) = 0$, $E(u_i^2) = \sigma_u^2$ and $E(u_i u_j) = 0$, $i \neq j$. This process will be stationary if $|\rho| < 1$. The matrix P for the AR(1) process is (see [18–20])

$$P = \begin{bmatrix} \sqrt{1-\rho^2} & 0 & 0 & \dots & 0 & 0 \\ -\rho & 1 & 0 & \dots & 0 & 0 \\ 0 & -\rho & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & 0 & \dots & -\rho & 1 \end{bmatrix} \tag{4}$$

In the AR(2) process the errors generated by

$$\varepsilon_i = \phi_1\varepsilon_{i-1} + \phi_2\varepsilon_{i-2} + u_i \tag{5}$$

where $E(u_i) = 0$, $E(u_i^2) = \sigma_u^2$, and $E(u_i u_j) = 0, i \neq j$. For the AR(2) process to be stationary, the parameters ϕ_1 and ϕ_2 must take values such that $\phi_1 + \phi_2 < 1$, $\phi_2 - \phi_1 < 1$ and $-1 < \phi_2 < 1$. The P matrix under the AR(2) structure is given as Judge et al. [21]

$$P = \begin{bmatrix} q_{11} & 0 & 0 & 0 & \dots & 0 & 0 \\ -\rho_1\sqrt{1-\phi_2^2} & \sqrt{1-\phi_2^2} & 0 & 0 & \dots & 0 & 0 \\ -\phi_2 & -\phi_1 & 1 & 0 & \dots & 0 & 0 \\ 0 & -\phi_2 & -\phi_1 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & 0 & 0 & \dots & -\phi_1 & 1 \end{bmatrix} \tag{6}$$

where $q_{11} = \left\{\frac{(1+\phi_2)[(1-\phi_2)^2-\phi_1^2]}{1-\phi_2}\right\}^{1/2}$ and $\rho_1 = \frac{\phi_1}{1-\phi_2}$.

KL estimator for the GLR model with AR(1) structure is given by Zubari and Adenomon [22] as

$$\hat{\beta}_{KL_{AR1}} = (X'V^{-1}X + kI_p)^{-1}(X'V^{-1}X - kI_p)X'V^{-1}y, k > 0$$

This is a biased estimator also, and the expected value and the variance-covariance matrix as follows:

$$bias(\hat{\beta}_{KL_{AR1}}, \beta) = -2k(X'V^{-1}X + kI_p)^{-1}\beta$$

and

$$cov(\hat{\beta}_{KL_{AR1}}, \beta) = \sigma^2 \left(X'V^{-1}X + kI_p\right)^{-1}\left(X'V^{-1}X - kI_p\right)\left(X'V^{-1}X\right)^{-1}\left(X'V^{-1}X - kI_p\right)\left(X'V^{-1}X + kI_p\right)^{-1}$$

In the paper, KL is presented as an estimator for the second-order autoregressive model as

$$\hat{\beta}_{KL_{AR2}} = (X'P'PX + kI_p)^{-1}(X'P'PX - kI_p)X'P'Py, \quad k > 0$$

Unlike $\hat{\beta}_{KL_{AR1}}$ based on the P matrix given in Equaiton 4, $\hat{\beta}_{KL_{AR2}}$ based on the P matrix is in the form of the matrix given Equation 6. It is to be noted that, the estimator denoted by KL in applications is $\hat{\beta}_{KL_{AR2}}$.

In the paper, the performance of the KL estimator for the P matrix given by Equation 6 is compared with the performance of the alternative estimators. MSE criterion was used to compare the performance. As a result, the following section explains how to calculate the MSE matrix of any estimators.

3. MSE Criterion to Determine the Best Model

MSE of an estimator measures the average of the squares of the errors that is, the average squared difference between the estimated and actual values. MSE is a risk function that represents expected value of squared error loss. MSE is defined for $\tilde{\beta}$ being any estimator as

$$MSE(\tilde{\beta}, \beta) = cov(\tilde{\beta}) + bias(\tilde{\beta})bias(\tilde{\beta})'$$

It should be noted that if $\tilde{\beta}$ is unbiased, the MSE will be equal to the variance-covariance matrix of $\tilde{\beta}$. The scalar mean square error (sMSE) value is equal to the sum of the diagonal elements of the MSE matrix, namely its trace. Let the two estimator be $\tilde{\beta}_1$ and $\tilde{\beta}_2$. For $\tilde{\beta}_2$ to be superior than $\tilde{\beta}_1$ according to the sMSE criterion, the necessary and sufficient condition is

$$\triangle\left(\tilde{\beta}_1, \tilde{\beta}_2\right) = trace(MSE(\tilde{\beta}_1, \beta)) - trace(MSE(\tilde{\beta}_2, \beta)) > 0$$

In other words, it defines that the estimator with the smaller sMSE value performs better according to this criterion. The statistical literature has extensively researched the comparison of biased estimators to unbiased estimators using the sMSE criterion for both linear regression and GLR models with AR(1) errors (see, [16,22,23]). In the following section, the biasing parameters (k/d) of the biased estimators were coded to minimize the sMSE in each cycle while evaluating the estimators' performance using the sMSE criterion.

4. Monte Carlo Simulation Study

The Monte Carlo simulation is created based on a number of criteria, which are listed in Table 1.

Table 1. Assumed values of various components for Monte Carlo simulation

Factor	**Notation**	**Values**
Sample Size	n	30, 100
Number of the independent variable	p	3
Degree of multicollinearity	γ^2	0.70, 0.80, 0.90
Dispersion parameters	σ	0.1, 0.5, 1
Number of replicates	MCN	1000

Settings: The goal of Monte Carlo simulation research evaluated by Matlab is to examine the GLR model with AR(2) errors fitted by the GLS, RR, Liu, and KL estimators. The sMSE criterion was used to assess the estimators' performance. The correlated regressor variables were generated from McDonald and Galarneau [24] as

$$x_{ij} = \left(1 - \gamma^2\right)^{1/2} z_{ij} + \gamma z_{ip}, \quad j = 1, 2, ..., p, \ i = 1, 2, ..., n \tag{7}$$

where z_{ij} are independent standard normal pseudo-random numbers. γ is specified so that the correlation between any two explanatory variables is given by γ^2 [25]. The regressor matrix are centralized and standardized after x_{ij} was produced, so that the $X'X$ becomes the correlation form. The final regressor matrix is written as $Z = (ones(n, 1) \ \ X)$ to correspond to the constant parameter. β, $p + 1$

vector was written with eigenvectors corresponding to the largest value of $(X'X)$ except the constant parameter which is taken as 0.5.

The simulation loop started with the derivation of the u_i error terms from standard normal distribution. Then, the error terms, ε_i, are generated from Equation 5. Here, to satisfy the stationarity condition different ϕ_1 and ϕ_2 values are as in Table 2.

Table 2. Researcher's parameters for the AR(2) model

	ϕ_1			
$\phi_2 = -0.9$	-1.5	-0.5	0.5	1.5
$\phi_2 = -0.7$	-1.5	-0.5	0.5	1.5
$\phi_2 = 0.7$	-0.2	-0.1	0.1	0.2
$\phi_2 = 0.9$	-0.05	-0.025	0.025	0.05

Finally, observation of the response variable is generated from

$$y_i = \beta_0 + \sum_{j=1}^{p} \beta_j x_{ij} + \varepsilon_i$$

P matrix in AR(2) structure (Equation 6) was created for different ϕ_1 and ϕ_2 values, and transformed response vector and regressor matrix were obtained. While obtaining the biased estimators, k/d values that minimize the related sMSE were obtained with a Matlab code and assigned as optimum k/d. The experiments were replicated 1000 times and the sMSE of the estimators was calculated for each replicate using the following Equation:

$$sMSE(\tilde{\beta}, \beta) = \frac{1}{MCN} \sum_{r=1}^{MCN} \left(\tilde{\beta}_r - \beta\right)' \left(\tilde{\beta}_r - \beta\right)$$

where $\tilde{\beta}$ is any of the earlier estimators. The estimators with the lowest sMSE has been regarded the best. The simulation results are presented in Tables 3–18.

Table 3. $\phi_1 = -1.5$ and $\phi_2 = -0.9$

γ^2	σ_u	n=30				n=100			
		GLS	RR	Liu	KL	GLS	RR	Liu	KL
0.7	0.1	0.0140	0.0126	0.0131	0.0126	0.0120	0.0109	0.0112	0.0109
	0.5	0.3491	0.2159	0.2109	0.2397	0.2996	0.1876	0.1868	0.2046
	1	1.3964	0.7919	0.7759	0.9395	1.1984	0.7004	0.6987	0.7984
0.8	0.1	0.0188	0.0151	0.0162	0.0149	0.0161	0.0130	0.0137	0.0130
	0.5	0.4702	0.2633	0.2524	0.3107	0.4020	0.2273	0.2261	0.2651
	1	1,8809	0.9991	0.9592	1.2607	1.6082	0.8724	0.8687	1.0785
0.9	0.1	0.0340	0.0219	0.0233	0.0228	0.0292	0.0187	0.0207	0.0195
	0.5	0.8502	0.4557	0.4236	0.5603	0.7312	0.3918	0.3892	0.4895
	1	3,4009	1.7908	1.6631	2.2566	2.9249	1.5447	1.5326	1.9698

Table 4. $\phi_1 = -0.5$ and $\phi_2 = -0.9$

γ^2	σ_u	n=30				n=100			
		GLS	RR	Liu	KL	GLS	RR	Liu	KL
0.7	0.1	0.0259	0.0214	0.0233	0.0210	0.0220	0.0193	0.0203	0.0191
	0.5	0.6471	0.3911	0.3662	0.4566	0.5498	0.3357	0.3355	0.3739
	1	2.5884	1.4329	1.3368	1.8662	2.1992	1.2093	1.2053	1.4852
0.8	0.1	0.0353	0.0252	0.0287	0.0250	0.0294	0.0225	0.0248	0.0222
	0.5	0.8820	0.5178	0.4709	0.6310	0.7360	0.4239	0.4227	0.4927
	1	3.5280	1.9790	1.7800	2.5801	2.9439	1.5975	1.5916	1.9872
0.9	0.1	0.0647	0.0393	0.0426	0.0439	0.0535	0.0329	0.0380	0.0348
	0.5	1.6176	0.9747	0.8604	1.1614	1.3385	0.7777	0.7792	0.8979
	1	6.4703	3.8719	3.3753	4.6652	5.3542	3.0742	3.0668	3.5956

Table 5. $\phi_1 = 0.5$ and $\phi_2 = -0.9$

γ^2	σ_u	n=30				n=100			
		GLS	RR	Liu	KL	GLS	RR	Liu	KL
0.7	0.1	0.0195	0.0171	0.0181	0.0170	0.0198	0.0184	0.0189	0.0184
	0.5	0.4880	0.2988	0.2892	0.3301	0.4945	0.3236	0.3250	0.3497
	1	1.9520	1.0851	1.0491	1.3234	1.9780	1.1201	1.1184	1.3304
0.8	0.1	0.0264	0.0199	0.0222	0.0195	0.0260	0.0216	0.0233	0.0213
	0.5	0.6588	0.3777	0.3572	0.4426	0.6502	0.3845	0.3853	0.4273
	1	2.6354	1.4311	1.3508	1.7991	2.6010	1.4201	1.4155	1.7331
0.9	0.1	0.0481	0.0296	0.0326	0.0313	0.0467	0.0302	0.0349	0.0303
	0.5	1.2024	0.6965	0.6383	0.8264	1.1685	0.6752	0.6772	0.7756
	1	4.8098	2.7518	2.5044	3.3293	4.6739	2.6510	2.6356	3.1279

Table 6. $\phi_1 = 1.5$ and $\phi_2 = -0.9$

γ^2	σ_u	n=30				n=100			
		GLS	RR	Liu	KL	GLS	RR	Liu	KL
0.7	0.1	0.0115	0.0111	0.0112	0.0111	0.0106	0.0102	0.0103	0.0102
	0.5	0.2871	0.2029	0.2038	0.2051	0.2643	0.1925	0.1923	0.1991
	1	1.1483	0.7347	0.7359	0.7460	1.0573	0.6928	0.6925	0.7308
0.8	0.1	0.0146	0.0133	0.0137	0.0133	0.0137	0.0123	0.0127	0.0122
	0.5	0.3653	0.2302	0.2278	0.2420	0.3416	0.2163	0.2161	0.2304
	1	1.4613	0.8565	0.8486	0.9618	1.3664	0.8095	0.8090	0.9091
0.9	0.1	0.0247	0.0184	0.0201	0.0180	0.0240	0.0171	0.0186	0.0168
	0.5	0.6182	0.3449	0.3277	0.4127	0.6000	0.3297	0.3294	0.3957
	1	2.4729	1.3291	1.2673	1.6831	2.3998	1.2815	1.2787	1.6109

Table 7. $\phi_1 = -1.5$ and $\phi_2 = -0.7$

γ^2	σ_u	n=30				n=100			
		GLS	RR	Liu	KL	GLS	RR	Liu	KL
0.7	0.1	0.0152	0.0138	0.0142	0.0137	0.0130	0.0117	0.0120	0.0117
	0.5	0.3801	0.2320	0.2267	0.2559	0.3246	0.2007	0.2000	0.2206
	1	1.5202	0.8480	0.8284	0.9912	1.2983	0.7471	0.7461	0.8672
0.8	0.1	0.0205	0.0163	0.0175	0.0160	0.0174	0.0140	0.0148	0.0139
	0.5	0.5116	0.2828	0.2707	0.3291	0.4359	0.2443	0.2438	0.286
	1	2.0462	1.0732	1.0248	1.3316	1.7434	0.9361	0.9336	1.1744
0.9	0.1	0.0370	0.0233	0.0251	0.0241	0.0317	0.0201	0.0222	0.0210
	0.5	0.9246	0.4940	0.4535	0.5954	0.7933	0.4263	0.4231	0.5261
	1	3.6986	1.9345	1.7780	2.4064	3.1731	1.6759	1.6630	2.1160

Table 8. $\phi_1 = -0.5$ and $\phi_2 = -0.7$

γ^2	σ_u	n=30				n=100			
		GLS	RR	Liu	KL	GLS	RR	Liu	KL
0.7	0.1	0.0309	0.0251	0.0276	0.0247	0.0260	0.0224	0.0238	0.0222
	0.5	0.7713	0.4623	0.4308	0.5341	0.6510	0.3927	0.3940	0.4406
	1	3.0850	1.6880	1.5651	2.1917	2.6041	1.4154	1.4111	1.7402
0.8	0.1	0.0420	0.0295	0.0339	0.0294	0.0349	0.0260	0.0291	0.0257
	0.5	1.0503	0.6114	0.5601	0.7357	0.8723	0.5002	0.5044	0.5738
	1	4.2012	2.3400	2.1313	2.9897	3.4892	1.9013	1.9038	2.3344
0.9	0.1	0.0770	0.0462	0.0507	0.0516	0.0635	0.0383	0.0449	0.0406
	0.5	1.9244	1.1569	1.0462	1.3541	1.5878	0.9298	0.9390	1.0531
	1	7.6976	4.5945	4.1141	5.4314	6.3511	3.6694	3.6843	4.2093

Table 9. $\phi_1 = 0.5$ and $\phi_2 = -0.7$

γ^2	σ_u	n=30				n=100			
		GLS	RR	Liu	KL	GLS	RR	Liu	KL
0.7	0.1	0.0231	0.0212	0.0217	0.0197	0.0233	0.0215	0.0221	0.0214
	0.5	0.5769	0.3482	0.3353	0.3909	0.5835	0.3778	0.3796	0.4082
	1	2.3077	1.2645	1.2091	1.5874	2.3339	1.3017	1.2964	1.5537
0.8	0.1	0.0311	0.0231	0.0261	0.0227	0.0307	0.0250	0.0273	0.0247
	0.5	0.7779	0.4447	0.4179	0.5260	0.7670	0.4513	0.4546	0.4988
	1	3.1117	1.6864	1.5731	2.1493	3.0681	1.6690	1.6616	2.0148
0.9	0.1	0.0567	0.0344	0.0385	0.0368	0.0551	0.0349	0.0412	0.0352
	0.5	1.4180	0.8198	0.7542	0.9683	1.3780	0.7984	0.8004	0.8964
	1	5.6719	3.2416	2.9440	3.9004	5.5122	3.1417	3.1128	3.6184

Table 10. $\phi_1 = 1.5$ and $\phi_2 = -0.7$

γ^2	σ_u	n=30				n=100			
		GLS	RR	Liu	KL	GLS	RR	Liu	KL
0.7	0.1	0.0182	0.0177	0.0177	0.0177	0.0133	0.0129	0.0130	0.0129
	0.5	0.4551	0.2951	0.2959	0.2992	0.3318	0.2428	0.2424	0.2465
	1	1.8204	0.9413	0.9421	0.8856	1.3272	0.8491	0.8488	0.8637
0.8	0.1	0.0217	0.0207	0.0208	0.0207	0.0167	0.0154	0.0158	0.0154
	0.5	0.5415	0.3250	0.3230	0.3350	0.4163	0.2688	0.2683	0.2794
	1	2.1658	1.0766	1.0703	1.0996	1.6654	0.9768	0.9760	1.0813
0.9	0.1	0.0328	0.0287	0.0297	0.0286	0.0279	0.0213	0.0233	0.0208
	0.5	0.8205	0.4528	0.4356	0.4988	0.6985	0.3945	0.3937	0.4582
	1	3.2821	1.6075	1.5529	1.9361	2.7938	1.5073	1.4994	1.8897

Table 11. $\phi_1 = -0.2$ and $\phi_2 = 0.7$

γ^2	σ_u	n=30				n=100			
		GLS	RR	Liu	KL	GLS	RR	Liu	KL
0.7	0.1	0.0310	0.0267	0.0286	0.0264	0.0271	0.0230	0.0247	0.0228
	0.5	0.7756	0.4712	0.4560	0.5107	0.6767	0.3977	0.3980	0.4365
	1	3.1024	1.6768	1.6115	2.0099	2.7069	1.4271	1.4161	1.7668
0.8	0.1	0.0411	0.0304	0.0351	0.0297	0.0363	0.0266	0.0303	0.0259
	0.5	1.0268	0.5891	0.5642	0.6590	0.9085	0.5041	0.5026	0.5785
	1	4.1071	2.2158	2.0976	2.6590	3.6341	1.8971	1.8795	2.3674
0.9	0.1	0.0732	0.0438	0.0510	0.0457	0.0662	0.0388	0.0462	0.0404
	0.5	1.8299	1.0488	0.9871	1.1816	1.6554	0.9320	0.9295	1.0625
	1	7.3198	4.1652	3.8566	4.7536	6.6215	3.6887	3.6340	4.2780

Table 12. $\phi_1 = -0.1$ and $\phi_2 = 0.7$

γ^2	σ_u	n=30				n=100			
		GLS	RR	Liu	KL	GLS	RR	Liu	KL
0.7	0.1	0.0319	0.0275	0.0294	0.0272	0.0277	0.0236	0.0253	0.0233
	0.5	0.7963	0.4822	0.4662	0.5218	0.6920	0.4086	0.4086	0.4462
	1	3.1853	1.7140	1.6449	2.0708	2.7681	1.4620	1.4513	1.8034
0.8	0.1	0.0420	0.0312	0.0360	0.0305	0.0371	0.0272	0.0310	0.0265
	0.5	1.0493	0.6012	0.5751	0.6694	0.9270	0.5159	0.5149	0.5892
	1	4.1972	2.2545	2.1260	2.7192	3.7082	1.9426	1.9210	2.4076
0.9	0.1	0.0743	0.0445	0.0522	0.0461	0.0674	0.0395	0.0472	0.0411
	0.5	1.8587	1.0573	0.9970	1.1904	1.6847	0.9498	0.9503	1.0806
	1	7.4347	4.1782	3.8952	4.7788	6.7387	3.7547	3.7086	4.3385

Table 13. $\phi_1 = 0.1$ and $\phi_2 = 0.7$

γ^2	σ_u	n=30				n=100			
		GLS	RR	Liu	KL	GLS	RR	Liu	KL
0.7	0.1	0.0349	0.0311	0.0327	0.0308	0.0291	0.0253	0.0270	0.0251
	0.5	0.8716	0.5280	0.5127	0.5560	0.7266	0.4400	0.4413	0.4729
	1	3.4863	1.8165	1.7594	2.1701	2.9064	1.5566	1.5468	1.8884
0.8	0.1	0.0447	0.0352	0.0397	0.0344	0.0384	0.0292	0.0330	0.0285
	0.5	1.1172	0.6421	0.6175	0.7051	0.9591	0.5453	0.5448	0.6131
	1	4.4690	2.3426	2.2322	2.8535	3.8364	2.0303	2.0074	2.5107
0.9	0.1	0.0762	0.0482	0.0579	0.0483	0.0684	0.0412	0.0496	0.0421
	0.5	1.9043	1.0851	1.0274	1.2153	1.7096	0.9745	0.9697	1.0970
	1	7.6173	4.2361	3.9368	4.9036	6.8382	3.8348	3.7678	4.4176

Table 14. $\phi_1 = 0.2$ and $\phi_2 = 0.7$

γ^2	σ_u	n=30				n=100			
		GLS	RR	Liu	KL	GLS	RR	Liu	KL
0.7	0.1	0.0442	0.0410	0.0415	0.0410	0.0338	0.0311	0.0320	0.0310
	0.5	1.1049	0.5997	0.5909	0.6022	0.8459	0.5101	0.5129	0.5266
	1	4.4196	1.8850	1.8579	2.0457	3.3836	1.6864	1.6855	1.9705
0.8	0.1	0.0537	0.0471	0.0488	0.0470	0.0429	0.0361	0.0389	0.0357
	0.5	1.3422	0.7100	0.6935	0.7411	1.0728	0.6112	0.6117	0.6593
	1	5.3687	2.4044	2.3366	2.7871	4.2911	2.1509	2.1412	2.6487
0.9	0.1	0.0841	0.0609	0.0694	0.0598	0.0722	0.0487	0.0576	0.0478
	0.5	2.1027	1.1489	1.0904	1.2812	1.8057	1.0297	1.0205	1.1541
	1	8.4106	4.3172	4.0660	5.1806	7.2230	3.9461	3.8599	4.6895

Table 15. $\phi_1 = -0.05$ and $\phi_2 = 0.9$

γ^2	σ_u	n=30				n=100			
		GLS	RR	Liu	KL	GLS	RR	Liu	KL
0.7	0.1	0.0330	0.0304	0.0313	0.0303	0.0260	0.0235	0.0245	0.0234
	0.5	0.8261	0.5022	0.4890	0.5213	0.6497	0.4019	0.4017	0.4241
	1	3.3045	1.6754	1.6277	1.9559	2.5988	1.3984	1.3896	1.6646
0.8	0.1	0.0415	0.0350	0.0377	0.0345	0.0338	0.0273	0.0300	0.0268
	0.5	1.0366	0.6018	0.5754	0.6566	0.8442	0.4842	0.4822	0.5358
	1	4.1463	2.1392	2.0268	2.6397	3.3770	1.7717	1.7522	2.2067
0.9	0.1	0.0684	0.0467	0.0550	0.0464	0.0589	0.0376	0.0446	0.0374
	0.5	1.7101	0.9963	0.9222	1.1392	1.4713	0.8328	0.8247	0.9517
	1	6.8402	3.8390	3.5146	4.6251	5.8853	3.2379	3.1698	3.8635

Table 16. $\phi_1 = -0.025$ and $\phi_2 = 0.9$

γ^2	σ_u	n=30				n=100			
		GLS	RR	Liu	KL	GLS	RR	Liu	KL
0.7	0.1	0.0355	0.0329	0.0336	0.0329	0.0274	0.0251	0.0259	0.0250
	0.5	0.8879	0.5235	0.5106	0.5358	0.6840	0.4228	0.4227	0.4408
	1	3.5516	1.6999	1.6638	1.9319	2.7360	1.4414	1.4327	1.6956
0.8	0.1	0.0439	0.0381	0.0401	0.0378	0.0351	0.0292	0.0317	0.0288
	0.5	1.0981	0.6226	0.5970	0.6688	0.8786	0.5055	0.5031	0.5550
	1	4.3923	2.1631	2.0683	2.6248	3.5143	1.8160	1.7963	2.2567
0.9	0.1	0.0708	0.0500	0.0582	0.0494	0.0602	0.0398	0.0470	0.0393
	0.5	1.7707	1.0146	0.9442	1.1570	1.5059	0.8554	0.8473	0.9726
	1	7.0828	3.8656	3.5559	4.6987	6.0235	3.2971	3.2222	3.9631

Table 17. $\phi_1 = 0.025$ and $\phi_2 = 0.9$

γ^2	σ_u	n=30				n=100			
		GLS	RR	Liu	KL	GLS	RR	Liu	KL
0.7	0.1	0.0465	0.0434	0.0434	0.0434	0.0342	0.0323	0.0325	0.0323
	0.5	1.1615	0.5842	0.5798	0.5880	0.8552	0.4928	0.4936	0.5006
	1	4.6460	1.7447	1.7333	1.8234	3.4207	1.5365	1.5339	1.6602
0.8	0.1	0.0548	0.0496	0.0500	0.0496	0.0420	0.0378	0.0387	0.0377
	0.5	1.3707	0.6813	0.6675	0.6998	1.0494	0.5756	0.5738	0.6050
	1	5.4826	2.2050	2.1585	2.4485	4.1978	1.9141	1.9034	2.2605
0.9	0.1	0.0816	0.0642	0.0689	0.0635	0.0670	0.0505	0.0562	0.0496
	0.5	2.0400	1.0730	1.0199	1.1810	1.6759	0.9264	0.9178	1.0350
	1	8.1600	3.9407	3.7190	4.6695	6.7035	3.4248	3.3496	4.1995

Table 18. $\phi_1 = 0.05$ and $\phi_2 = 0.9$

γ^2	σ_u	n=30				n=100			
		GLS	RR	Liu	KL	GLS	RR	Liu	KL
0.7	0.1	0.0614	0.0563	0.0559	0.0562	0.0450	0.0425	0.0422	0.0424
	0.5	1.5357	0.6245	0.6305	0.6443	1.1259	0.5490	0.5506	0.5627
	1	6.1430	1.7610	1.7634	1.8079	4.5035	1.5716	1.5726	1.5983
0.8	0.1	0.0698	0.0629	0.0621	0.0630	0.0528	0.0487	0.0483	0.0487
	0.5	1.7442	0.7223	0.7243	0.7356	1.3198	0.6320	0.6333	0.6484
	1	6.9766	2.2167	2.2118	2.2865	5.2792	1.9508	1.9495	2.1034
0.9	0.1	0.0964	0.0800	0.0807	0.0801	0.0778	0.0641	0.0663	0.0639
	0.5	2.4110	1.1103	1.0854	1.1626	1.9451	0.9835	0.9814	1.0535
	1	9.6441	3.9398	3.8119	4.3284	7.7806	3.4693	3.4315	4.0780

The simulation tables clearly showed that sMSE value of the RR, Liu, and KL estimators increased as the strength of multicollinearity (γ^2). However, if we compare the biased estimators with the GLS, it is clearly seen that the estimators biased to the strength of the multicollinearity are more robust. In comparison to RR and Liu, the sMSE value of the KL estimator grew as the γ^2 increased. When $\sigma_u = 0.1$, the KL estimator gave much smaller sMSE values than the GLS, while giving very close sMSE values with the RR and Liu biased estimators. However, for $\sigma_u = 1$, the sMSE values of the KL estimator were always smaller than the GLS but gave higher sMSE values compared to RR and Liu. As n increased, sMSE value of the KL estimator decreased. It should be noted, however, that all of these results were interpreted using the stationary AR(2) model. When ϕ_2 constant, the absolute of the ϕ_1 declined in the AR(2) model, sMSE values of the KL estimator and others increased.

5. Application to Real Data

Two real datasets are analyzed to illustrate the sMSE performance of the KL estimator under the GLR model with AR(2) error structure. While performing the analyzes for both datasets, the multicollinearity problem was determined and in the linear regression model, parameter estimation values of OLS, ridge, liu, and kl estimators, sMSE values, k/d values that minimize sMSE for biased estimators, and CPU time is given in seconds. Then, the autocorrelation problem in the datasets was investigated and after the error structures were determined, the similar outputs of GLS, RR, Liu, and KL estimators in the GLR models were given. The examined datasets are available on request from the corresponding author.

5.1. French Economy Data

The French economy data, used by Malinvard [26], consists of one response variable and three regressor variables. The response variable y is imports, x_1 is domestic production, x_2 is stock formation and x_3 is domestic consumption. All variables cover 1946 to 1845 and are measurements per milliards in French frags. The multiple linear regression model to be estimated as

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \varepsilon \tag{8}$$

The dataset is used by previous authors to evaluate the performance of the KL estimator under the linear model (see, [9, 27]). Figure 1 depicts the linear correlation (r) between the three regressions,

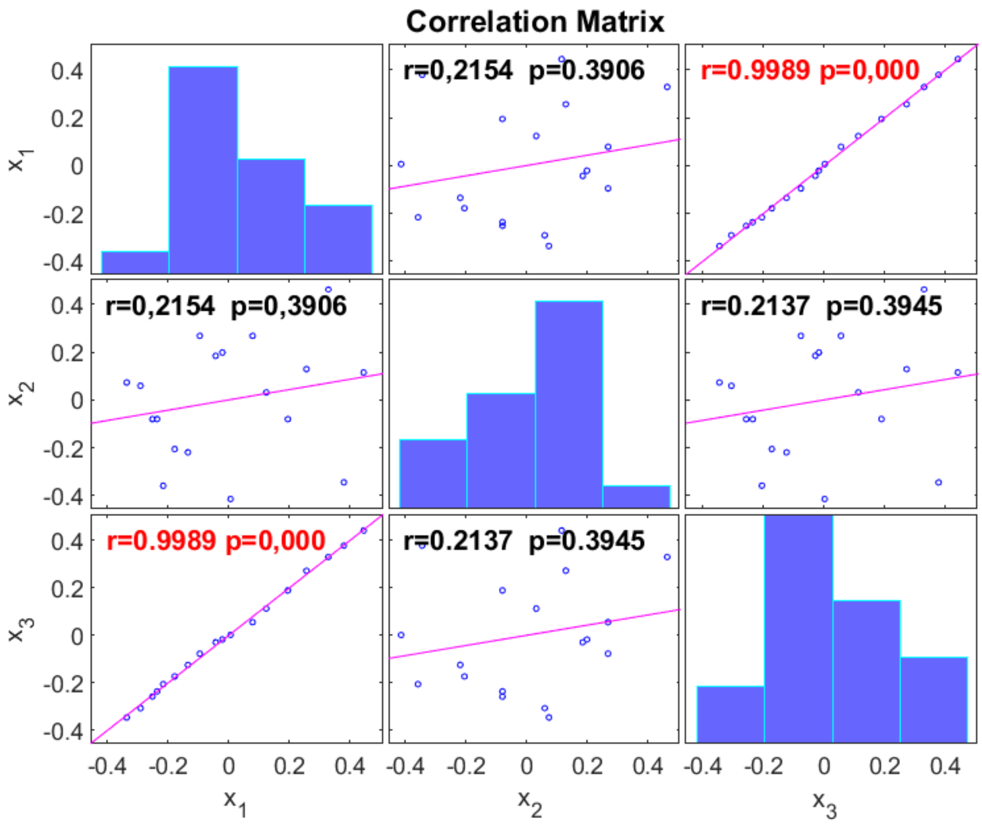


Fig. 1. Correlations between the regressors for French economy data.

as well as their significance at the 0.05 significance level (p). As can be observed, the correlation between the first and third variables is nearly one. In addition, when the condition number is examined, the eigenvalues of the $X'X$ matrix are calculated as 0.2982, 49.6458, 157.9069, 1617795.3772 and $\kappa = 2329.4065$. A strong multicollinearity problem has been detected using these two collinearity diagnostics approaches. As a result, biased estimators were used to estimate the parameters in Equation 1. In the linear regression model with multicollinearity problem, parameter estimates and sMSE values of OLS and alternative biased estimators, and biasing parameters minimizing sMSE for biased estimators are given in Table 19. Also, the times during the minimization of the sMSE values with respect to k/d are given as CPU time.

Table 19. Estimation of model coefficients and sMSE values when autocorrelation is neglected for French economy data

Coeff.	**OLS**	**ridge**	**liu**	**kl**
$\hat{\beta}_0$	-19.7251	-18.8982	-18.8957	-18.8994
$\hat{\beta}_1$	0.03220	0.0628	0.0627	0.0628
$\hat{\beta}_2$	0.4142	0.4008	0.4005	0.4008
$\hat{\beta}_3$	0.2427	0.1947	0.1949	0.1948
sMSE	17.2379	16.515619	16.5154	16.515628
k/d	-	0.0131	0.9452	0.0064
CPU time	-	0.2500	0.3750	0.2656

Table 19 shows that even though the liu estimator fitted to the linear model has the smallest sMSE value, the ridge, liu, and kl estimators are relatively similar in terms of sMSE performance.
Since Kibria and Lukman [9] conducted the paper in a linear regression model, the autocorrelation problem was not included French economy data. However, it has an autocorrelation problem. Autocorrelation function (ACF) and partial autocorrelation function (PACF) are used to determine this status and error structure decision based on ACF and PACF graphs as follows:
Except for the first lag, the other lags fluctuate within the confidence bounds, as shown in the ACF

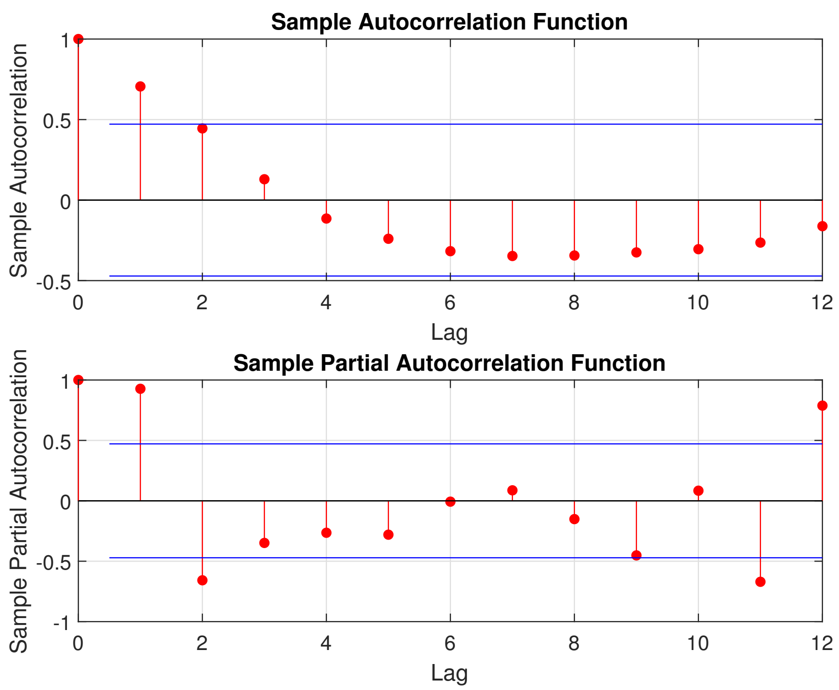


Fig. 2. Correlagram for French economy data

graph of Fig. 2. However, after two lags, the reduction in the PACF graph was abruptly stopped off (Fig. 2). The parameters of the AR(2) model were found to be significant as shown in Table 20.

It can be specify the lag structure, presence of a constant, and innovation distribution of an AR(p) model for this dataset by following Table 20. As shown in Table 20 the constant coefficient can be accepted as statistically zero. Thus, the model to be fit in the AR structure is compatible

Table 20. Choosing appropriate lag in the AR Model for French economy data

p	**T statistic**	**P Value**
Constant	0.6539	0.5132
AR(1)	6.6345	3.2557e-11
AR(2)	-3.5830	0.0003

with Equation 5. The model parameters were estimated in the following order; $\hat{\phi}_1 = 0.7037$ and $\hat{\phi}_2 = 0.0028$.

After the P matrix based on the estimations of the $\hat{\phi}_1$ and $\hat{\phi}_2$ parameters, the response variable and regressor matrix were transformed. The multicollinearity problem still existed in the transformed model ($\kappa = 1271.7991$). Therefore, biased estimators were applied to estimate the regression parameters under the GLR model with AR(2) errors. The GLS, RR, Liu, and KL estimators were used to estimate the parameter estimates in the transformed model, and the sMSE values of the estimators are listed in Table 21.

Table 21. The parameter estimations and sMSE values in GLR models with AR(2) error structure for French economy data

General Linear (AR(2))	**GLS**	**RR**	**Liu**	**KL**
Constant	-21.8389	-21.2531	-21.2582	-21.2452
β_1	-0.0137	-0.0024	-0.0027	-0.0022
β_2	0.4840	0.4805	0.4805	0.4805
β_3	0.3234	0.3040	0.3045	0.3038
$sMSE$	13.0806	12.7340	**12.7339**	12.7342
$opt.k/d$	-	0.0032	0.9703	0.0016
CPU time	-	0.2031	0.5781	0.2812

Table 21 clearly demonstrates that the KL estimator provides an sMSE value that is similar to the well-known RR and Liu estimators when the error structure is AR(2) process. It was also observed that the three-biased estimator had better performance than the unbiased GLS according to the sMSE criterion. Tables 19 and 21 are comparable; if the autocorrelation problem is ignored and a linear regression model is fitted, the sMSE value will differ due to differences in the model parameters.

5.2. Weather Data

Weather data received at the station in the Columbia-Pacific Northwest Region for each 15-minute time period on January 1, 2022, was used (https://www.usbr.gov/pn/agrimet/webagdayread.html). Multiple linear regression model to be estimated as

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \varepsilon \tag{8}$$

where y is wind speed, x_1 is humidity, x_2 is vapor pressure x_3 is dew point temperature.

Examining the correlation matrix graph shown in Figure 3, it can be clearly seen that there is a strong correlation between the vapor pressure and the dew point temperature ($r = 0.9359$ and $p = 1.0393e - 44$).

The fact that this correlation coefficient is very close to 1 indicates a multicollinearity problem in the dataset. In addition, the condition number was calculated. The eigenvalues of the $X'X$ matrix are 0.0092, 0.9884, 1258.8777 and 66381.7020, respectively, and the condition number is $\kappa = 8522.8934$. Both the correlation coefficient and the condition number indicated strong multicollinearity. Therefore, while estimating the regression parameters in the linear regression model given by Equation 1, biased estimators were used as alternatives to OLS, and the optimal k/d values that minimize sMSE values were presented in Table 22. Table 22 shows when fitting the weather data with multicollinearity problem with the linear regression model, biased estimators significantly reduce the sMSE value compared to OLS. In addition, when the CPU times spent in the minimization process of optimal

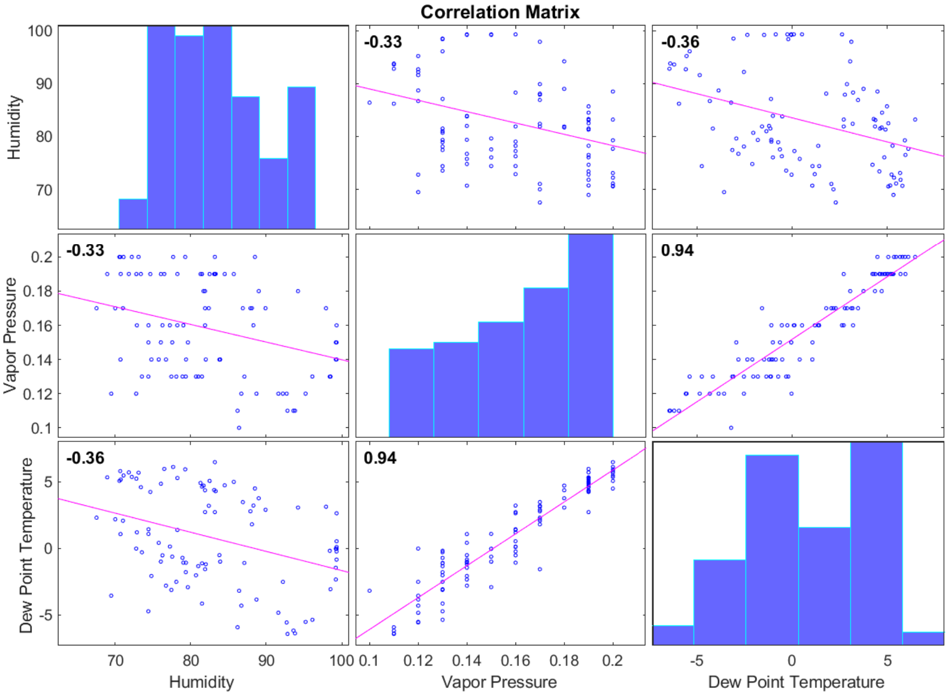


Fig. 3. Correlations between the regressors for weather data

Table 22. Estimation of model coefficients and sMSE values when autocorrelation is neglected for weather data

Coeff.	**OLS**	**ridge**	**liu**	**kl**
$\hat{\beta}_0$	0.2118	0.3051	0.4471	0.8787
$\hat{\beta}_1$	0.01374	0.0214	0.0192	0.0140
$\hat{\beta}_2$	4.9497	0.0683	0.3735	0.4386
$\hat{\beta}_3$	-0.0429	0.0002	-0.0041	-0.0097
sMSE	349.2349	24.1939	23.2219	25.2441
k/d	-	2.0243	0.0536	0.0081
CPU time	-	0.2368	0.2656	0.2543

biasing parameters according to sMSE are examined, it is seen that the algorithms are completed in close seconds. The existence of the autocorrelation problem was investigated after the multicollinearity problem in the weather data was discovered. The ACF and PACF graphs in Figure 4 and the hypothesis tests in Table 23 were used to represent these investigations.

Figure 4 clearly illustrated the properties of ACF/PACF of an AR(2) process: Its ACF decreased sharply and PACF was be nearly zero after lag 2.

Table 23. Choosing appropriate lag in the AR Model for weather data

p	**T statistic**	**P Value**
Constant	0.6669	0.5700
AR(1)	4.2189	2.4541e-05
AR(2)	2.1966	0.0280

Table 23 showed that the two delayed functional associations of errors were statistically significant. In this case, $\hat{\phi}_1$ and $\hat{\phi}_2$ parameters in Equation 5 were estimated as 0.3999 and 0.2561, respectively, in the weather dataset which has autocorrelation problem. After it was determined that the functional associations between the errors were modeled with the AR(2) process, the GLR model was fitted by

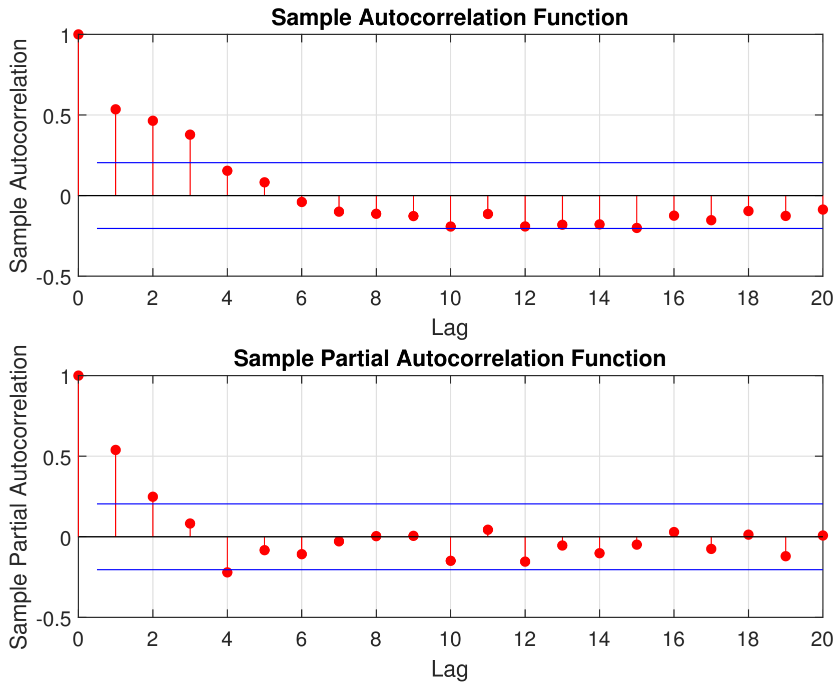


Fig. 4. Correlagram for weather data

the unbiased GLS and the biased RR, Liu, and KL estimators. Estimation of regression parameters, sMSE values, optimal k/d biasing parameters minimizing sMSE, and minimization time as CPU were shown in Table 24.

Table 24. The parameter estimations and sMSE values in GLR models with AR(2) error structure for weather data

General Linear (AR(2))	**GLS**	**RR**	**Liu**	**KL**
Constant	-2.1272	-0.5364	-0.4215	-1.5792
β_1	0.04597	0.0326	0.0310	0.0450
β_2	2.9477	-0.0117	0.1115	-0.0690
β_3	-0.0417	-0.0334	-0.0353	-0.0239
$sMSE$	197.278	12.1687	11.9377	16.8049
$opt.k/d$	-	0.5581	0.0420	0.0100
CPU time	-	0.2012	0.2869	0.2831

It can be obviously seen in Table 24 that applying biased estimators in the weather data with multicollinearity and AR(2) autocorrelation problem significantly reduced the sMSE value compared to the unbiased GLS. The fact that the model variance is higher in the weather data compared to the French economy data supports a situation that is visible in the simulation results: In the French economy data with a small model variance, the sMSE value of the KL estimator was very close to the sMSE values of the RR and Liu estimators, but in the weather data with high model variance, the sMSE value of the KL estimator was slightly higher than the sMSE values of the RR and Liu estimators.

6. Conclusion

The simulation results demonstrated that as the variance of the model increases, the sMSE values of the KL estimator and the others increases. Moreover, the sMSE values of the KL estimator and others appear to increase as the severity of multicollinearity increases. When the model variance is small, sMSE values of the KL estimator under the AR(2) error structure are closely similar to the popular biased estimators' values. The sMSE values of the KL estimator and others decreased when the sample size was increased. Examples of two data sets with autocorrelation problems from both multicollinearity and AR(2) processes are also included in the paper. The results of the two different data sets were generally frugal and the findings support the simulation results. Furthermore, CPU times were discovered to be near to each other while determining the optimal biasing parameter over real datasets. In other words, the Kl estimator was close to the popular estimators in terms of CPU time. It was discovered that KL performed substantially better than GLS with optimum biasing parameters, and its results were extremely near to those of Ridge and Liu estimators. In the statistical literature, new unbiased and biased estimators continue to be proposed. As new estimators are proposed, it is critical to examine the assumptions and to make parameter estimations on the correct model for statistical inference.

Author Contributions

The author read and approved the last version of the paper.

Conflicts of Interest

The author declares no conflict of interest.

References

[1] D. A. Belsley, E. Kuh, R. E. Welsch, Regression Diagnostics: Identifying Influential Data and Sources of Collinearity, John Wiley & Sons, New Jersey, 2005.

[2] A. E. Hoerl, R. W. Kennard, *Ridge Regression: Biased Estimation for Nonorthogonal Problems*, Technometrics 12 (1) (1970) 55–67.

[3] A. E. Hoerl, R. W. Kannard, K. F. Baldwin, *Ridge Regression: Some Simulations*, Communications in Statistics - Theory and Methods 4 (2) (1975) 105–123.

[4] L. JF, *A Simulation Study of Ridge and Other Regression Estimators*, Communications in Statistics - Theory and Methods 5 (4) (1976) 307–323.

[5] K. Liu, *Using Liu-Type Estimator to Combat Collinearity*, Communications in Statistics - Theory and Methods 32 (5) (2003) 1009–1020.

[6] K. Liu, *A New Class of Blased Estimate in Linear Regression*, Communications in Statistics - Theory and Methods 22 (2) (1993) 393–402.

[7] M. R. Özkale, S. Kaçıranlar, *A Prediction-Oriented Criterion for Choosing the Biasing Parameter in Liu Estimation*, Communications in Statistics - Theory and Methods 36 (10) (2007) 1889–1903.

[8] S. Kaçıranlar, S. Sakallıoğlu, *Combining the Liu Estimator and the Principal Component Regression Estimator*, Communications in Statistics - Theory and Methods 30 (12) (2001) 2699–2705.

[9] G. B. Kibria, A. F. Lukman, *A New Ridge-Type Estimator for the Linear Regression Model: Simulations and Applications*, Scientifica Article ID 9758378 (2020) 16 pages.

[10] M. R. Özkale, S. Kaçiranlar, *The Restricted and Unrestricted Two-Parameter Estimators*, Communications in Statistics - Theory and Methods 36 (15) (2007) 2707–2725.

[11] A. F. Lukman, Z. Y. Algamal, B. G. Kibria, K. Ayinde, *The KL Estimator for the Inverse Gaussian Regression Model*, Concurrency and Computation: Practice and Experience 33 (13) (2021) e6222.

[12] A. F. Lukman, G. B. Kibria, *Almon-KL Estimator for the Distributed Lag Model*, Arab Journal of Basic and Applied Sciences 28 (1) (2021) 406–412.

[13] A. F. Lukman, I. Dawoud, B. Kibria, Z. Y. Algamal, B. Aladeitan, *A New Ridge-Type Estimator for the Gamma Regression Model*, Scientifica 2021.

[14] H. E. Oranye, F. I. Ugwuowo, *Modified jackknife Kibria–Lukman estimator for the Poisson regression model*, Concurrency and Computation: Practice and Experience 34 (6) (2022) e6757.

[15] A. A. Hamad, Z. Y. Algamal, *Jackknifing KL Estimator in Poisson Regression Model*, Periodicals of Engineering and Natural Sciences (PEN) 10 (2) (2022) 314–322.

[16] G. Trenkler, *On the Performance of Biased Estimators in the Linear Regression Model with Correlated or Heteroscedastic Errors*, Journal of Econometrics 25 (1-2) (1984) 179–190.

[17] S. Kaçıranlar, *Liu Estimator in the General Linear Regression Model*, Journal of Applied Statistical Science 13 (2003) 229–234.

[18] G. Judge, R. Hill, W. Griffiths, H. Lutkepohl, T. Lee, Introduction to the Theory and Practice of Econometrics, Wiley, New Jersey, 1988.

[19] S. S. Roy, S. Guria, *Regression Diagnostics in an Autocorrelated Model*, Brazilian Journal of Probability and Statistics 18 (2) (2004) 103–112.

[20] H. Vinod, A. Ullah, Recent Advances in Regression Methods, Marcel Dekker, New York, 1981.

[21] G. G. Judge, W. Griffiths, R. Hill, H. Lutkepohl, T. Lee, The Theory and Practice of Econometrics, John Wiley & Sons, New York, 1985.

[22] M. A. Zubair, M. O. Adenomon, *Comparison of Estimators Efficiency for Linear Regressions with Joint Presence of Autocorrelation and Multicollinearity*, Science World Journal 16 (2) (2021) 103–109.

[23] H. Yang, X. Chang, *A New Two-Parameter Estimator in Linear Regression*, Communications in Statistics - Theory and Methods 39 (6) (2010) 923–934.

[24] G. C. McDonald, D. I. Galarneau, *A Monte Carlo Evaluation of Some Ridge-Type Estimators*, Journal of the American Statistical Association 70 (350) (1975) 407–416.

[25] B. G. Kibria, *Performance of Some New Ridge Regression Estimators*, Communications in Statistics - Simulation and Computation 32 (2) (2003) 419–435.

[26] E. Malinvard, Statistical: Methods of Econometrics, North Holland, Amsterdam, 1980.

[27] F. I. Ugwuowo, H. E. Oranye, K. C. Arum, *On the Jackknife Kibria-Lukman Estimator for the Linear Regression Model*, Communications in Statistics - Simulation and Computation (2021) 1–13.