Remittances Review
April 2024,
Volume: 9, No: 2, pp.4974-4985
ISSN: 2059-6588(Print) | ISSN 2059-6596(Online)
Received : 20 Januanry 2024, Accepted: 30 April 2024
DOI: https://doi.org/10.33282/rr.vx9i2.262

OPTIMIZING FEATURE SELECTION FOR CLIMATE CARBON EMISSION INDICATORS: A MULTICOLLINEARITY-RESILIENT APPROACH WITH L1 AND L2 REGULARIZATION

Zainab Javed[1], Dr. Anam Javaid[2,] Iqra Gulshan[3], Dr. Shahbaz Nawaz[4]

[1,3]*M.Phil Scholar, Department of Statistics, The Women University Multan*

[2]*School of Mathematical Sciences, Universiti Sains Malaysia*
11800 USM, Penang, Malaysia

[2]*Assistant Professor; Department of statistics, The Women University Multan, Pakistan*

[4]*School of Quantitative Sciences, University Utara Malaysia, 06010 Kedah, Malaysia*

[4]*Punjab Bureau of Statistics; Planning and Development Department, Pakistan*

Corresponding Author Email: anamjavaid7860@yahoo.com

ABSTRACT

Carbon emissions, predominantly in the form of carbon dioxide CO_2, plays a vital role in driving climate change and its associated environmental consequences. This research go through into the origins, impacts, and potential remedies concerning carbon emissions, emphasizing their crucial role in global aim to combat climate change. In the context of Pakistan, carbon emissions primarily stem from energy production, industrial activities, transportation, and agricultural practices. This study specifically concentrates on the optimal selection of models for predicting carbon emissions in Pakistan. To achieve this objective statistical analyses perform based on various regression techniques. The response variable is taken as the consumption of CO_2 in solid, liquid, and gaseous forms, while predictors includes population, total value, and agricultural value added as a percentage of GDP, agricultural land area, urban population, gross fixed capital formation, industry value added, fertilizer consumption as a percentage of production, and GDP. The model selection process comprises five stages. In the initial stage, multicollinearity diagnosis is implemented through a correlation matrix. The second stage involves outlier identification using various statistical measures and graphical analyses, such as Box Plots. Subsequently, the third stage focuses on optimal model selection by using Ridge and LASSO regression analyses. The fourth stage. The efficient model selection based on criteria such as the sum of squared error (SSE), mean square error (MSE), root mean square error (RMSE), and forecasting efficiency assessed through the mean absolute percentage error (MAPE). The study result shows that on the basis of maximum R^2 , minimum MSE, MAPE, LASSO regression is considered as the best technique according to the define selection criteria's.

Keywords: Multicollinearity, Outliers, Boxplot, Ridge, LASSO

Introduction

Climate change, characterized by long-term shifts in temperatures and weather patterns, is driven by a combination of natural and human-induced factors. While natural phenomena like solar cycle variations plays a role in shaping climate patterns (Beer and Mende et al., 2000). The significant impact of human activities on climate change cannot be unnoticed. Since the 1800s, the burning of fossil fuels, including coal, oil, and gas, has emerged as a primary source of anthropogenic climate change (Woolway et al., 2020). The combustion of fossil fuels releases greenhouse gases, which act as an insulating layer enveloping the Earth. This process effectively traps heat from the sun within the atmosphere, leading to a gradual rise in global temperatures. Among the various greenhouse gases, carbon dioxide and methane are particularly noteworthy contributors to climate change (Voigt et al., 2017). Carbon dioxide is released through activities such as vehicle operation and the use of coal for heating, while deforestation and land clearing also release significant amounts of this gas. Methane emissions, on the other hand, are largely associated with waste disposal and decomposition processes (Younger et al., 2008).
The sources of greenhouse gas emissions are diverse and encompass multiple sectors of human activity. Energy production, industrial processes, transportation, residential and commercial buildings, agriculture, and land use all contribute substantially to the release of these gases into the atmosphere (Omer et al., 2009). The cumulative impact of these emissions has been instrumental in driving changes in global climate patterns, with far-reaching implications for ecosystems, weather extremes, and sea levels (Handmer and Honda et al., 2012)

Climate change is widely recognized as one of the most pressing issues facing Pakistan today. The country is experiencing significant shifts in its climate, with a notable increase in extreme weather events such as heatwaves, droughts, and severe flooding (Abid and Scheffran et al., 2015). These changes have had a profound impact on both the environment and the population of Pakistan. Over the past two decades, more than 150 extreme weather events have been recorded, highlighting the scale of the challenge (Elahi et al., 2022). In 2022, the country experienced devastating floods that submerged a third of its territory, causing widespread destruction and displacement. The diverse geography of Pakistan means that its climate varies widely, presenting unique challenges across different regions (Ali et al., 2017) Addressing the impact of climate change is crucial for the long-term sustainability and resilience of Pakistan, requiring coordinated efforts at both national and international levels. It is imperative for policymakers and stakeholders to prioritize climate adaptation and mitigation strategies to safeguard the well-being of the country and its people (Lamb and Steinberger 2017).

Reason of Climate Change

Details of various anthropogenic and natural causes associated with human activities. The equilibrium of the full of atmosphere heat reasonable have been distributed by anthropogenic factors (Rasul and Ahmad 2012). Reasons of climate change are observed in Figure 1 as follows.

Remittances Review
April 2024,
Volume: 9, No: 2, pp.4974-4985
ISSN: 2059-6588(Print) | ISSN 2059-6596(Online)

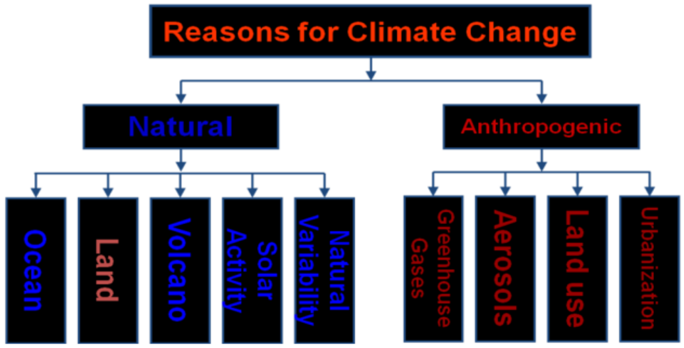


Figure 1: Reason of climate variation natural as well as anthropogenic related to human actions.

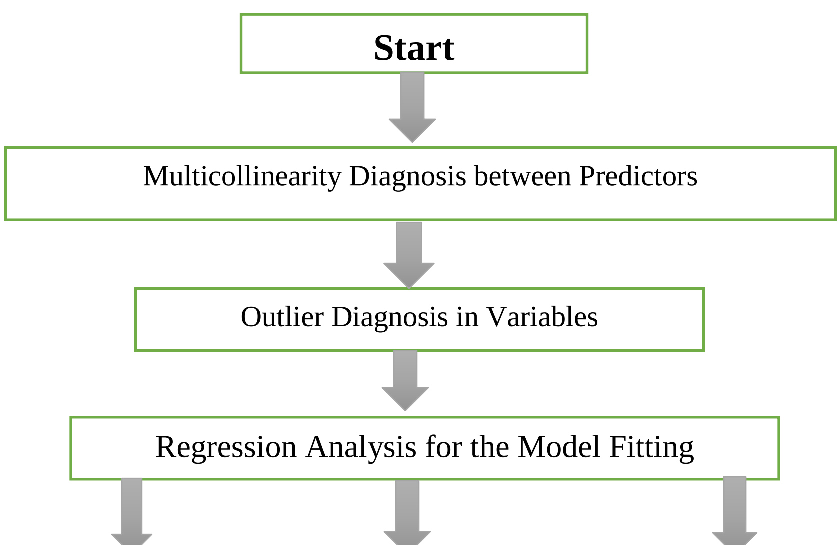
Climate Behavior in Pakistan

In Pakistan, it is expected that heat wave experience will become more frequent and last longer. During a heat wave in June 2015, more than 12,000 people die in Karachi alone (Durham 2016). In Sindh province more than 200 people die in other cities. In Balochistan (Turbat-2017 May) 53.7℃ was recorded the second highest temperature in Pakistan. Heat wave occur in the plains of the country in just May and the June before the monsoon season (Zahid and Rasul 2012). Pakistan being a developing country does not have a high impact of increasing atmospheric greenhouse gasses. It is share of global greenhouse gas emissions ranked 153th is less than 0.8% (Yoro and Daramola 2020). After this, Pakistan is one of the 10 countries most vulnerable to climate change (Ullah and Takaaki 2016). According to the Global Climate Risk Index for 2020 annual report, Pakistan have suffered an economic loss of USD 3792.52 million as a result of 152 extreme weather events between 1999 to 2018 (Eckstein et al., 2021).

METHODOLOGY

The target population for this study is to control the effect of CO_2 emission from which the sample might be drawn.

The following phases are performed are each of dataset both for "Interaction and Without Interaction" the application of the flow chart given in Figure 2



Remittances Review
April 2024,
Volume: 9, No: 2, pp.4974-4985
ISSN: 2059-6588(Print) | ISSN 2059-6596(Online)

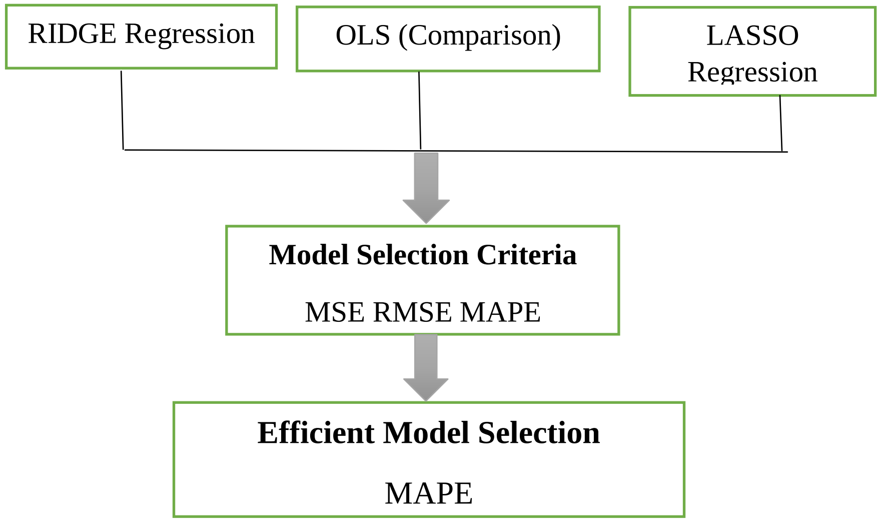


Figure 2: Flow Chart of the Methodology

Ordinary Least Square Method

The concept of the least square estimation process was independently developed by Gauss (1795), Legendre (1805), and Dutka (1990) in the first decade of the 19th century (Dutka 1990). OLS regression is a commonly used method for estimating the coefficients of a linear regression equation, which describes the relationship between one or more independent quantitative variables and a dependent variable (whether it's a simple or multiple linear regression) by (Tyzhnenko and Ryeznik 2019).

Ridge Regression

Ridge regression, also known as Tikhonov regularization or L2 regularization, is a linear regression technique that is used to address the problem of multicollinearity in multiple linear regression (Sajwani et al., 2019). Multicollinearity occurs when there is high correlation among independent variables, leading to instability and unreliable estimates of the regression coefficients Khalaf and Iguernane (2016). In ridge regression, a penalty term is added to the linear regression objective function, and this penalty term is proportional to the square of the magnitude of the coefficients. The objective function in ridge regression is:

$$\text{Minimize } (||Y-X\beta||_2^2+\lambda||\beta||_2^2)$$

- Y is the vector of the dependent variable.
- X is the matrix of independent variables.
- β is the vector of regression coefficients.
- $||Y-X\beta||_2^2$ is the ordinary least squares OLS loss term.
- λ is the regularization parameter, a non-negative hyper parameter that controls the strength of the penalty term.
- $||\beta||_2^2$ is the L2 norm (Euclidean norm) of the coefficient vector.

Remittances Review
April 2024,
Volume: 9, No: 2, pp.4974-4985
ISSN: 2059-6588(Print) | ISSN 2059-6596(Online)

The inclusion of the regularization term helps to shrink the coefficients toward zero, mitigating the problem of multicollinearity and preventing overfitting. The regularization parameter λ determines the trade-off between fitting the data well and keeping the coefficients small. A larger λ results in more shrinkage of coefficients. The ridge regression solution can be found by minimizing the modified objective function using techniques like gradient descent or closed-form solutions (Allerbo and Jonasson, et al., 2023). Ridge regression is particularly useful when dealing with datasets with high-dimensional feature spaces or when multicollinearity is present. It provides a more stable solution compared to ordinary least squares in such cases.

Least Absolute Shrinkage and Selection Operator Regression

Least Absolute Shrinkage and Selection Operator (LASSO) is another form of linear regression that includes a regularization term. Like ridge regression, LASSO aims to address the issue of multicollinearity and prevent overfitting by adding a penalty to the linear regression objective function (Enwere and Nduka et al., 2023).

The objective function for LASSO regression is given by:

$$\text{Minimize } (||Y-X\beta||_2^2+\lambda||\beta||_1)$$

- Y is the vector of the dependent variable.
- X is the matrix of independent variables.
- β is the vector of regression coefficients.
- $||Y-X\beta||_2^2$ is the ordinary least squares OLS loss term.
- λ is the regularization parameter, a non-negative hyper parameter that controls the strength of the penalty term.
- $||\beta||_1$ is the L1 norm (Euclidean norm) of the coefficient vector.

FINDINGS AND DATA ANALYSIS

The dataset of variables used in this study is taken from https://www.kaggle.com/. Total of 56 observations are used for the purpose of analysis. There are 8 variables related to the carbon emission are analyzed in the dataset, CO_2 is taken as dependent variable and all other 7 factors such as Population, Industrialization, GDP, Fertilizer consumption, urban area, Agriculture land and forestry, Gross fixed capital formation are considered as predictors. The codes are given to all the included variables in the analysis. The list of the codes with their respective names are mentioned in Table 1

Table 1: *Variables Codes and Descriptions*

Sr No.	Variable Names	Variable Codes
1	CO_2 emissions from solid fuel consumption (kt) Value, CO_2 emissions from liquid fuel consumption (kt) Value, CO_2 emissions from gaseous fuel consumption (kt) Value	**Y**
2	Population, total Value	$\mathbf{X_1}$
3	Agriculture, forestry, and fishing, value added (% of GDP) Value	$\mathbf{X_2}$
4	Agricultural land (sq. km) Value	$\mathbf{X_3}$

Volume: 9, No: 2, pp.4974-4985
ISSN: 2059-6588(Print) | ISSN 2059-6596(Online)

5	Urban population Value	**X_4**
6	Gross fixed capital formation (current US$)	**X_5**
7	Industry (including construction), value added (constant 2010 US$)	**X_6**
8	Fertilizer consumption (% of fertilizer production)	**X_7**
9	GDP	**X_8**

Table 1 represents that in the climate change CO_2 emission the main factors are Population, Industrialization, GDP, Fertilizer consumption, urban area, Agriculture land and forestry, Gross fixed capital formation. CO_2 emission has interacted with all factors.
The unit of the CO_2 emission is taken as in kiloton (KT) while the population total value, Agricultural land, GDP, and urban population is in per 1000 unit. The Agriculture, forestry, Fertilizer consumption, and fishing is in terms of per 100 unit. The Gross fixed capital formation, Industry per 1 USD. In table 1 the CO_2 emissions from solid fuel consumption (kt) Value, CO_2 emissions from liquid fuel consumption (kt) Value and CO_2 emissions from gaseous fuel consumption (kt) Value are represents the dependent variable Y and the independent variables are coded by X_1, X_2, X_3, X_4, X_5, X_6, X_7 and X_8.

Multicollinearity Analysis in the Data

Correlation matrix is calculated to diagnose the multicollinearity issue in the dataset that is used in the analysis. The result for the matrix are obtained in Table 2

For the CO_2 emission

Correlation matrix for the main factors of climate change is calculated by using Excel software with dependent and independent variables with interaction effect, Table 2 shows the information

Table 2: *Correlation matrix for the main factors of climate change*

	Y	**X1**	**X2**	**X3**	**X4**	**X5**	**X6**	**X7**	**X8**
Y	1								
X1	0.98	1							
X2	-0.71	-0.792	1						
X3	-0.14	-0.128	0.127	1					
X4	0.99	0.998	-0.773	-0.126	1				
X5	0.96	0.943	-0.631	-0.147	0.952	1			
X6	0.99	0.988	-0.720	-0.157	0.992	0.978	1		
X7	-0.42	-0.471	0.508	0.152	-0.463	-0.38578	-0.437	1	
X8	0.94	0.920	-0.575	-0.114	0.933	0.987736	0.960	-0.362	1

The matrix in Table 2 shows that correlation matrix among variables. The correlation matrix shows there is no multicollinearity issue as ($r_{xy} < 0.90$) (Javaid et al.. 2021).

Remittances Review
April 2024,
Volume: 9, No: 2, pp.4974-4985
ISSN: 2059-6588(Print) | ISSN 2059-6596(Online)

Diagnosis of Outliers in Dataset

After diagnosis of Multicollinearity, outliers are detected by using the Boxplots in the dataset. The values outside the boxplot will be considered as outliers. The boxplots are calculated for the predictors and response variables by using Minitab software.

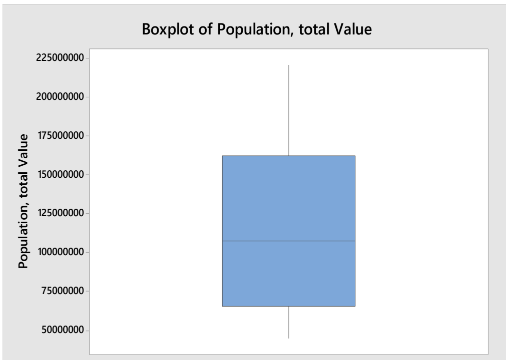


Figure 3

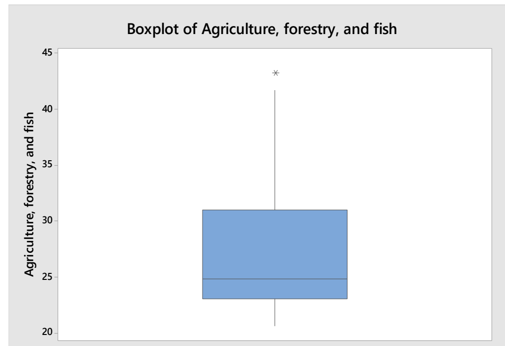


Figure 4

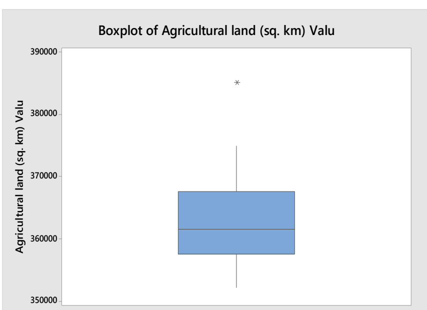


Figure 5

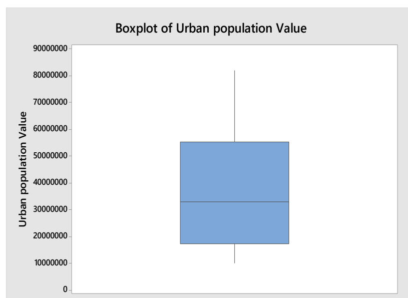


Figure 6

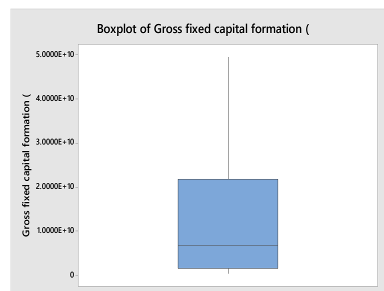


Figure 7

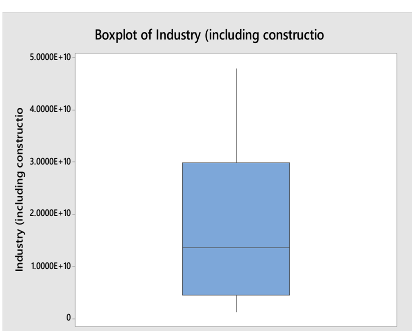


Figure 8

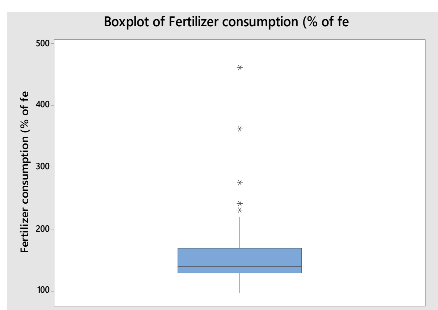


Figure 9

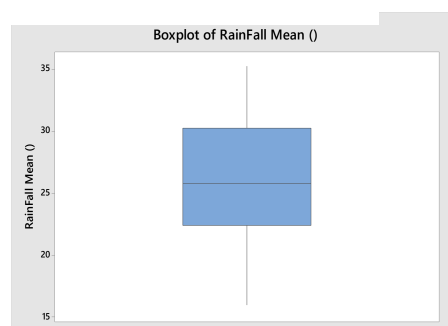


Figure 10

In Figure 3 to 10, outliers can be observed in various variables through boxplot.

Best Model Selection

After the analysis of multicollinearity and outliers, the regression analysis is carried out of the best selection model due to presence of multicollinearity and outliers.

Remittances Review
April 2024,
Volume: 9, No: 2, pp.4974-4985
ISSN: 2059-6588(Print) | ISSN 2059-6596(Online)

Ridge Regression Analysis

To calculate the significant factors in the presence of outliers of climate change Ridge Regression is used. The results are provided in Table 3 by using R-studio software as follows

Table 3: *Results for Ridge Regression*

Factors	Coefficients	*P*-value	Significance
Y	$1.450671e^{+04}$	0.00031	Significant
$\mathbf{X_1}$	$2.599466e^{-04}$	0.00025	Significant
$\mathbf{X_2}$	$-1.610846e^{+02}$	0.0018	Significant
$\mathbf{X_3}$	$-5.127963e^{-02}$	0.006	Significant
$\mathbf{X_4}$	$6.455676e^{-04}$	0.00041	Significant
$\mathbf{X_5}$	$7.167409e^{-07}$	0.000007	Significant
$\mathbf{X_6}$	$9.574423e^{-07}$	0.0000067	Significant
$\mathbf{X_7}$	$3.301657e^{+00}$	0.670	Non-Significant
$\mathbf{X_8}$	$7.738211e^{+01}$	0.820	Non-Significant

Table 3 displays the regression output, indicating that the variables X_1, X_2, X_3, X_5, and X_6 are statistically significant at 5% level of significance. However, the variables X_7 and X_8 are considered non-significant as their p-values are higher than 0.05.

LASSO Regression Analysis

To calculate the significant factors in the presence of outliers of climate change LASSO Regression is used. The significant Factors are provided as in Table 4 are as follows

Table 4: *Results for LASSO Regression*

Factors	Coefficients	*P*-value	Significance
Y	$-4.639530e^{+04}$	0.00014	Significant
$\mathbf{X_1}$	$3.840433e^{-04}$	0.00033	Significant
$\mathbf{X_2}$	$7.847559e^{+02}$	0.0027	Significant
$\mathbf{X_5}$	$4.613493e^{-04}$	0.00039	Significant
$\mathbf{X_6}$	$2.390382e^{-06}$	0.0000028	Significant
$\mathbf{X_7}$	$1.167094e^{+01}$	0.01	Significant
$\mathbf{X_8}$	$-2.435266e^{+01}$	0.0024	Significant

Table 4 presents the output of lasso regression, indicating that the variables X_1, X_2, X_4, X_6, X_7, and X_8 are statistically significant at 5% level of significance (Javaid et al., 2019). On the other hand, the coefficients for X_3 and X_8 are considered non-significant.

Ordinary Least Square Analysis

The results for OLS and MC is observed in Table 5 by using R-studio software are as follows

Remittances Review
April 2024,
Volume: 9, No: 2, pp.4974-4985
ISSN: 2059-6588(Print) | ISSN 2059-6596(Online)

Table 5: *Results for Ordinary Least Square*

Factors	Coefficients	*P*-value	Significance
Y	$-5.958e^{+04}$	0.2494	Significant
X_1	$-3.951e^{-04}$	0.7846	Non-Significant
X_2	$9.033e^{+02}$	0.0250 *	Significant
X_3	$7.513e^{-02}$	0.2990	Non-Significant
X_4	$2.042e^{-03}$	0.5876	Non-Significant
X_5	$-2.639e^{-07}$	0.5632	Non-Significant
X_6	$3.956e^{-06}$	$1.52e^{-06}$ ***	Significant
X_7	$1.453e^{+0}$	0.0883	Non-Significant
X_8	$-1.483e^{+02}$	0.0301 *	Significant

In Table 5 the regression output reveals that the variables X_2, X_6, and X_8 are statistically significant as their p-values are less than 0.05 (Lodhi et al., 2023). On the other hand, the variables X_1, X_3, X_4, X_5, and X_7 are found as non-significant since their p-values exceed 0.05.

By removing multicollinearity variables

Removing the multicollinearity of the variables by using the R-Studio Software without interaction is provided in Table 6

Table 6 : *Results for after removing multicollinearity variables*

Factors	Coefficients	*P*-value	Significance
Intercept	$-2.449e^{+04}$	0.564539	Non-Significant
Population	$3.193e^{-04}$	0.646277	Non-Significant
Agriculture, forestry, and fishing	$1.385e^{+03}$	0.000262***	Significant
Agricultural land	$-1.607e^{-01}$	0.074604.	Non-Significant
Urban Population value	$2.230e^{-03}$	0.187220	Non-Significant
Fertilizer consumption	$1.901e^{+01}$	0.113405	Non-Significant

In Table 6 after addressing multicollinearity, the notable factor is X2: Agriculture, forestry, and fishing (Javaid et al., 2024). This suggests that Agriculture, forestry, and fishing have a significant impact on CO2 emission.

Results and Comparison

The comparison are observed through various methods in Table 7 as follows.

Table 7: *Comparison among results of SSE, SST, R^2, MSE, RMSE and MAPE*

	SSE	SST	R^2	MSE	RMSE	MAPE
Ridge Regression	1836810259	153398709961	0.9880259	321856943	**17940.37187**	**27.78194**
LASSO Regression	**556627306**	153398709961	**0.9963714**	661386440.8	25717.43457	55.14817
OLS				**13552488086**	116415.154	81.33669
After Removing Multicollinearity				346147051.2	18605.02758	55.90684

In Table 7, SSE is minimum for LASSO regression while R^2 is high for LASSO regression, RMSE is lower for ridge regression than other regression techniques. MAPE is less than for ridge regression and other factors. So, it is considered as good model to forecast. Since other majority selection criteria are in favor of ridge regression. So, ridge regression as an efficient model selection to forecast the carbon emission based on various factors.

References

Abid, M., J. Scheffran, et al. (2015). "Farmers' perceptions of and adaptation strategies to climate change and their determinants: the case of Punjab province, Pakistan." Earth System Dynamics **6**(1): 225-243.

Ali, S., Y. Liu, et al. (2017). "Climate change and its impact on the yield of major food Ullah, W. and N. Takaaki (2016). "Climate change vulnerability of Pakistan towards natural disasters: a review." International Journal of Environmental Protection and Policy **4**(5): 126.

Allerbo, O., J. Jonasson, et al. (2023). "Elastic gradient descent, an iterative optimization method approximating the solution paths of the elastic net." Journal of Machine Learning Research **24**(277): 1-53.

Beer, J., Mende, W., & Stellmacher, R. (2000). The role of the sun in climate forcing. *Quaternary Science Reviews*, *19*(1-5), 403-415.

Durham, G. B. (2016). Controlled climates and human variation: Ubiquitous air conditioning and lowering heat thresholds in a hotter world, University of Houston.

Dutka, J. (1990). "Robert Adrain and the method of least squares." Archive for history of exact sciences: 171-184.

Eckstein, D., V. Künzel, et al. (2021). The global climate risk index 2021, Bonn: Germanwatch.
Elahi, E., Z. Khalid, et al. (2022). "Extreme weather events risk to crop-production and the adaptation of innovative management strategies to mitigate the risk: A retrospective survey of rural Punjab, Pakistan." Technovation **117**: 102255.

Enwere, K., E. Nduka, et al. (2023). "Comparative Analysis of Ridge, Bridge and Lasso Regression Models In the Presence of Multicollinearity." IPS Intelligentsia Multidisciplinary Journal **3**(1): 1-8.

Kashif, M., ur Rehman, A., Qadir, M. I., Javed, M. K., Javaid, A., Javaid, S., & Javaid, A. (2022). Impact of the COVID-19 Pandemic on Student's Education. *STATISTICS, COMPUTING AND INTERDISCIPLINARY RESEARCH*, *4*(2), 9-16.

Javaid, A., Muthuvalu, M. S., Sulaiman, J., Ismail, M., & Ali, M. K. M. (2019, December). Forecast the moisture ratio removal during seaweed drying process using solar drier. In *AIP Conference Proceedings* (Vol. 2184, No. 1). AIP Publishing.

Javaid, A., & Akbar, A. (2020). Factors Identification for Human Development Index. *STATISTICS, COMPUTING AND INTERDISCIPLINARY RESEARCH*, *2*(1), 1-8.

Javaid, A., Ismail, M., & Ali, M. K. M. (2020). Efficient Model Selection of Collector Efficiency in Solar Dryer using Hybrid of LASSO and Robust Regression. *Pertanika Journal of Science & Technology*, *28*(1).

Javaid, A., Ismail, M. T., & Ali, M. K. M. (2021). Efficient Model Selection For Moisture Ratio Removal Of Seaweed Using Hybrid Of Sparse And Robust Regression Analysis. *Pakistan Journal of Statistics and Operation Research*, 669-681.

Javed, Z., Javaid, A., Javaid, S., & Javed, A. (2022). Analysis of Child Mortality on Geographical Basis Over Different Centuries. *STATISTICS, COMPUTING AND INTERDISCIPLINARY RESEARCH*, *4*(2), 1-7.

Javaid, et al., (2024). STANDARD OPERATING PROCEDURE FOR EFFICIENT MODEL SELECTION THROUGH HYBRID OF STEPWISE AND ROBUST REGRESSION ANALYSIS. *Remittances Review*, *9*(2), 3599-3615.

Khalaf, G. and M. Iguernane (2016). "Multicollinearity and a ridge parameter estimation approach." Journal of Modern Applied Statistical Methods **15**(2): 25.

Lamb, W. F. and J. K. Steinberger (2017). "Human well-being and climate change mitigation." Wiley Interdisciplinary Reviews: Climate Change **8**(6): e485.

Lim, H. Y., Fam, P. S., Javaid, A., Ali, M., & Khan, M. (2020). Ridge Regression as Efficient Model Selection and Forecasting of Fish Drying Using V-Groove Hybrid Solar Drier. *Pertanika Journal of Science & Technology*, *28*(4).

Lodhi, I., Nawaz, S., Javaid, A., Javaid, S., & Javaid, A. (2023). Geographically Analysis of Wheat Production on Annual Basis. *STATISTICS, COMPUTING AND INTERDISCIPLINARY RESEARCH*, *5*(1), 29-36.

Omer, A. M. (2009). "Energy use and environmental impacts: A general review." Journal of renewable and Sustainable Energy **1**(5).

Rasul, G., & Ahmad, B. (2012). Climate change in Pakistan. *Pakistan Meteorological Department*.

Sajwani, S. A. A. H. (2019). "Using Stagewise (L1) as a Feature Selection Tool to Enhance Kernelized Ridge Regression (L2) Outputs for Regression Problems."

Suhaeri, M. E., Alimudin, A., Javaid, A., Ismail, M., & Ali, M. K. M. (2021, November). Evaluation of clustering approach with euclidean and Manhattan distance for outlier detection. In *AIP Conference Proceedings* (Vol. 2423, No. 1). AIP Publishing.

Tyzhnenko, A. G. and Y. V. Ryeznik (2019). "Ordinary least squares: the adequacy of linear regression solutions under multicollinearity and without it."

Voigt, C., Lamprecht, R. E., Marushchak, M. E., Lind, S. E., Novakovskiy, A., Aurela, M., ... & Biasi, C. (2017). Warming of subarctic tundra increases emissions of all three important greenhouse gases–carbon dioxide, methane, and nitrous oxide. *Global Change Biology*, *23*(8), 3121-3138.

Woolway, R. I., Kraemer, B. M., Lenters, J. D., Merchant, C. J., O'Reilly, C. M., & Sharma, S. (2020). Global lake responses to climate change. *Nature Reviews Earth & Environment*, *1*(8), 388-403.

Yoro, K. O. and M. O. Daramola (2020). CO2 emission sources, greenhouse gases, and the global warming effect. Advances in carbon capture, Elsevier**:** 3-28.

Younger, M., H. R. Morrow-Almeida, et al. (2008). "The built environment, climate change, and health: opportunities for co-benefits." American journal of preventive medicine **35**(5): 517-526.

Zahid, M. and G. Rasul (2012). "Changing trends of thermal extremes in Pakistan." Climatic change **113**(3): 883-896.