JTAM (Jurnal Teori dan Aplikasi Matematika)
http://journal.ummat.ac.id/index.php/jtam
p-ISSN 2597-7512 | e-ISSN 2614-1175
Vol. 8, No. 3, July 2024, pp. 983-993

Robust Continuum Regression Study of LASSO Selection and WLAD LASSO on High-Dimensional Data Containing Outliers

Nurmai Syaroh Daulay[1], Erfiani[1], Agus M Soleh[1]
[1]Departement of Statistics, IPB University, Indonesia
nurmaisyarohdaulay@apps.ipb.ac.id

Article History:
Received : 30-04-2024
Revised : 12-06-2024
Accepted : 20-06-2024
Online : 19-07-2024

Keywords:
Robust Continuum Regression;
LASSO;
WLAD LASSO.

ABSTRACT

In research, we often encounter problems of multicollinearity and outliers, which can cause coefficients to become unstable and reduce model performance. Robust Continuum Regression (RCR) overcomes the problem of multicollinearity by reducing the number of independent variables, namely compressing the data into new variables (latent variables) that are independent of each other and whose dimensions are much smaller and applying robust regression techniques so that the complexity of the regression model can be reduced without losing essential information from data and provide more stable parameter estimates. However, it is hampered in the computational aspect if the data has very high dimensions ($p >> n$). In the initial stage, it is necessary to reduce dimensions by selecting variables. The Least Absolute Shrinkage and Selection Operator (LASSO) can overcome this but is sensitive to the presence of outliers, which can result in errors in selecting significant variables. Therefore, we need a method that is robust to outliers in selecting explanatory variables such as Weighted Least Absolute Deviations with LASSO penalty (WLAD LASSO) in selecting variables by considering the absolute deviation of the residuals. This method aims to overcome the problem of multicollinearity and model instability in high-dimensional data by paying attention to resistance to outliers. Leverages the outlier resistant RCR and variable selection capabilities of LASSO and WLAD LASSO to provide a more reliable and efficient solution for complex data analysis. Measure the performance of RKR-LASSO and RKR-WLAD LASSO; simulations were carried out using low-dimensional data and high-dimensional data with two scenarios, namely without outliers (δ= 0%) and with outliers (δ= 10%, 20%, 30%) with a level of correlation (ρ = 0.1,0.5,0.9). The analysis stage uses RStudio version 4.1.3 software using the "MASS" package to generate data that has a multivariate normal distribution, the "glmnet" package for LASSO variable selection, the "MTE" package for WLAD LASSO variable selection. The simulation results show the performance of RKR-LASSO tends to be superior in terms of model goodness of fit compared to RKR-WLAD LASSO. However, the performance of RKR-LASSO tends to decrease as outliers and correlations increase. RKR-LASSO tends to be looser in selecting relevant variables, resulting in a simpler model, but the variables chosen by LASSO are only marginally significant. RKR-WLAD LASSO is stricter in variable selection and only selects significant variables but ignores several variables that have a small but significant impact on the model.

Crossref
https://doi.org/10.31764/jtam.v8i3.23123

A. INTRODUCTION

In regression analysis, the Ordinary Least Square (OLS) method is one of the approaches commonly used to model the relationship between independent variables and response variables. However, we often need to work on multicollinearity and outliers when applying OLS (Lakshmi et al., 2021). Multicollinearity occurs when two or more independent variables in the model have a high correlation, leading to unstable and challenging to interpret estimates of the regression coefficients (Tsagris & Pandis, 2021). Outliers refer to observations that are far from the general pattern in the data and can significantly affect the analysis results if not appropriately handled (Leys et al., 2018). Both of these phenomena can cause distortions in estimation and interpretation and reduce the reliability of the regression model. Therefore, paying attention to the estimation method used is necessary so that the analysis results are more reliable and accurate.

Continuum Regression (CR) is a development of the Least Squares Regression (LSR), Partial Least Squares Regression (PLSR) and Principal Component Regression (PCR) methods to overcome multicollinearity problems by reducing the number of independent variables, namely compressing the data into new variables (latent variables), which are independent of each other and have much smaller dimensions (Chen & Zhu, 2015). New variables in CR by maximizing the variance of independent variables and the covariance between the independent variable and the response variable. CR is introduced and used to complete the calibration model in several case examples using cross-validation index criteria compared with various levels of adjustment parameters δ; the conclusion is that CR is better compared to the results of LSR, PLSR, and PCR (Stone & Brooks, 1990). Setiawan et al. (2007) researched calibration models using the CR approach, concluding that CR has advantages over PCR and PLSR in solving multicollinearity problems in various independent variable matrix structures (setiawan & Notodiputro, 2007). Sometimes, huge independent variables $(p \gg n)$ cause the matrix structure of the independent variables to experience singularities, which will cause problems in the computational aspect (Ajeel & Hashem, 2020) . So, at the initial stage, it is necessary to reduce the dimensions from the original high dimension $(nxp)$to a lower dimensional space, for example (nxh) where $h < (n-1) < p$ but still retains most of the relevant information from the original data, this process is called preprocessing (Velliangiri et al., 2019).

One of the preprocessing methods with variable selection is the Least Absolute Shrinkage and Selection Operator (LASSO) (Lima et al., 2020). The LASSO method selection reduces the regression coefficient of variables that have a high correlation with error. This reduction aims to make the regression coefficient close to zero or even equal to zero (Cui & Wang, 2016). LASSO parameter estimation is known to have stable regression coefficients, reduces the number of parameters and has good consistency in parameter convergence (ChunRong et al., 2017). The results of Arwini's (2020) study on modelling using CR with variable selection pre-processing using the LASSO method can increase precision and provide reasonably accurate prediction results compared to the CR model with Principal Component Analyst (PCA) preprocessing (Arisandi et al., 2020). The results of LASSO selection will be new variables modelled with CR. Still, LASSO selection is sensitive to outliers, which can result in errors when selecting significant variables.

The WLAD LASSO method was introduced by Arslan (2012) as a combination of WLAD and LASSO penalties. This method is robust against outliers in the response variable and against outliers in the independent variable. It is better at dealing with outliers and variable selection than LASSO and LAD LASSO in selecting variables by considering the absolute deviation of the residuals (Arslan, 2012; Yang & Li, 2018). Septa (2022) studied the performance of the LAD LASSO and WLAD LASSO methods on high-dimensional data containing outliers. Based on the study, WLAD LASSO can overcome the weaknesses of LAD LASSO, which selects few significant variables and LASSO, which selects many variables which are not significant in variable selection, with a higher prediction error compared to LASSO but lower than LAD LASSO (Cahya et al., 2022).

In CR analysis, calculations are based on variance and covariance matrices, which may contain outlier data. To deal with data that contains outliers the Robust Continuum Regression (RCR) method is an effective choice. This method combines the advantages of CR with resistance to outliers so it can provide more stable and consistent parameter estimates in the presence of outliers (Serneels et al., 2005). Sereneels et al. (2005) research examined the performance of the RCR estimates of the Hubert estimator on simulated and observational data with RCR results superior to Classical Continuum Regression (CCR). Khotimah et al (2020) In a robust regression simulation study of MM and Least Median Square (LMS) estimation with data containing outliers $(s = 0\%, 5\%, 10\%, 15\%, 20\%, \text{and } 30\%)$ and dimensions $(n = 50, 200, 1000)$. The Root Mean Square Error (RMSE) results in MM and LMS estimation are superior to using LSR (Khotimah et al., 2020).

In this research, we compare the Robust Continuum Regression method with LASSO selection (RCR-LASSO) with the Robust Continuum Regression method with WLAD LASSO selection (RCR-WLAD LASSO) on low-dimensional data and high-dimensional data containing several levels of outliers (δ) and levels of correlation (ρ). The performance comparison of these two methods aims to assess how efficient the technique is in overcoming multicollinearity and outlier problems and selecting the most relevant variables in forming the regression model. Variable selection with LASSO and WLAD LASSO selection can help simplify the model by identifying the variables that contribute most to the research, especially in high-dimensional data. More efficient methods help improve the quality of analysis and interpretation of results. Further analysis to see the method's performance will be discussed by comparing the Root Mean Squared Error of Prediction (RMSEP) and coefficient of determination (R^2) values to measure how well the regression model fits the research data.

B. METHODS

This research uses simulation data analyzed by RStudio 4.1.3 software. The simulation generates multivariate normal distribution data with several outlier levels (δ) and correlation levels (ρ). This simulation refers to research by Serneel et al. (2005), but there is an additional modification of the correlation level ($\rho = 0.1, 0.5, 0.9$) and additional outlier levels in the response variable ($\delta = 10\%, 20\%, 30\%$). This study's optimal δ* value in Robust Continuum Regression between (0<δ*<1) is δ* =0.3. A small δ*, such as 0.3, reaches the optimal value faster than a larger δ* value and can extract more relevant information from the independent variables (the results of the experiment δ* between the values 0 to 1) (Xie et al., 2020). when δ*=0 Robust Continuum Regression includes Least Squares Regression, δ*=0.5 includes Partial Least Squares Regression, and δ=1 includes Principal Component Regression. This simulation was conducted on low dimensional data ($n = 75, p = 25$) and high dimensional data ($n = 75, p = 100$). The following are the stages and studies in the simulation:

1. Determining the number of observations in the low dimension $n = 75$, the number of independent variables $p = 25$ and high dimension observations $n = 75$, the number of independent variables $p = 100$;
2. Generating a vector of p-dimensional independent variables at the ith observation$-i$ $x_i = (x_{1i}, x_{2i}, \dots, x_{pi})^T$ through the Multivariate Normal distribution $x_i \sim N_p\ (0, \Sigma)$ where the covariance matrix $\Sigma = {r_{ij}}_{pxp}$; $r_j = \rho^{|j-i|}$; $\rho = 0{,}1\ ; 0{,}5\ ; 0{,}9$ dan $i = 1, 2, 3, \dots, n\ ; j, i = 1, 2, 3, \dots, p$;
3. Determine the model to be used, namely:

$$y_i = x_{1i} + x_{2\,i} + x_{3i} + x_{4i} + \cdots + x_{25i} + \varepsilon_i\ ; \text{for } p = 25;$$
$$y_i = x_{1i} + x_{2\,i} + x_{3i} + x_{4i} + \cdots + x_{100i} + \varepsilon_i\ ; \text{for } p = 100.$$

4. Generate response variables with two scenarios as follows:
   a. Scenario 1: Response variable without outliers based on model $y_i = x_i^T\beta + \varepsilon_i\ ;\ \varepsilon_i \sim N(0{,}1)$;
   b. Scenario 2: Response variable contains outliers ($\delta = 10\%, 20\%, 30\%$) in the model $y_i = x_i^T\beta + \varepsilon_i\ ;\ \varepsilon_i \sim N(1-\delta)\ N(0{,}1) + \delta N\ (25{,}1)$;

5. Perform preprocessing with LASSO and WLAD LASSO variable selection.

$$\hat{\beta}_{LASSO} = arg\,min\{\textstyle\sum_{i=1}^{n}(y_i - x_i^T\beta)^2 + \lambda\sum_{j=1}^{p}|\beta_j|\} \quad (1)$$

$$\hat{\beta}_{WLAD\ LASSO} = arg\,min\{\textstyle\sum_{i=1}^{n} w_i(y_i - x_i^T\beta) + \lambda\sum_{j=1}^{p}\omega_j|\beta_j|\} \quad (2)$$

Where y_i is the observed value of the response variable at the i-th observation, $x_i = (x_{1i}, x_{2i}, \dots, x_{pi})^T$, $x_i^T\beta$ is the predicted value for the i-th observation, obtained by multiplying the vectors feature x_i with regression coefficient vector β (Robert Tibshirani, 1996). The weights $w_i = \min\{1, \frac{p}{RD(x_i)}\}$, for $i = 1, 2, 3, \dots, n$ are obtained using Robust Distances (RD) (Wahid et al., 2017). $RD = (x_i - \hat{\mu})^T\, \hat{\Sigma}^{-1}(x_i - \hat{\mu})$; $\hat{\mu}$ dan $\hat{\Sigma}$ are obtained from the *Minimum Regularized Covariance Determinant* (MRCD) estimator (Boudt et al., 2020; Bulut, 2020).The selection of independent variables in LASSO and WLAD LASSO selection is based on the optimum lambda (λ) with a minimum value of Mean Squared Error cross-validation (MSE CV) (Lee et al., 2016). The procedure for k-fold *cross-validation* is as follows (Izenman, 2008):

a. Randomly divides data into k parts or folds into k subsamples.
b. For each k subsamples, one subsample will be used as *testing* data and $(k-1)$ subsamples as *training* data.
c. The *cross-validation* process is repeated k times, and each subsample is used only once as *testing* data.
d. The optimum λ is obtained based on the minimum MSE CV value (Emmert-Streib & Dehmer, 2019).

$$MSE\ CV = \frac{1}{k}\sum_{k=1}^{K}\sum_{(x_i, y_i)\epsilon T}(y_i - \hat{y}_{-k}(x_i))^2 \tag{3}$$

$\hat{y}_{-k}(x_i)$ is the predicted response value for x_i when the model is obtained from data without involving the k-th subsample, andy_i is the ith response variable in the *testing* data.

e. Repeat steps (b) to (d) k times to obtain the minimum CV. Selection of predictor variables is based on selecting the optimum lambda value with the smallest cross-validation value.

6. Carry out robust continuum regression modelling using selected variables. Independent variables used are the selection results from LASSO and WLAD LASSO.

$$y = T_h \xi + \varepsilon \tag{4}$$

y is a vector of response variables of size $nx1$, $T_h = XW_h$ with $W_h = (w_1, w_2, \dots, w_h)$ is a weighting matrix of size pxh. So, the matrix T_h contains h columns of latent variables (Ismah et al., 2024; Zhou, 2019). Where p is the number of independent variables, and h is the number of new variables from LASSO and WLAD LASSO selection. Weighting vector in continuum regression with formula:

$$w_i = arg\,max\{Cov\,(Xw, y)^2\ Var\,(Xw)^{[\delta*/(1-\delta*)]-1}\} \tag{5}$$

with constraints $\|w_i\| = 1$ and $Cov\,(Xw_i, yw_j) = 0$ for $i < j = 1, 2, \dots, h$. In the RCR concept, calculations are based on robust covariance and variance values. Robust estimator by trimming the covariance values Xα, y and trimming the variance value Xα (Zhang et al., 2011).

$$Cov_\alpha(x, y) = \frac{1}{n-2l}\sum_{i=l+1}^{n-1} z_i \tag{6}$$

With $z_i = (x_i - \bar{x}_\alpha)(y_i - \bar{y}_\alpha)$, and $l = [n\alpha] + 1$.

7. Calculate the estimated value of RCR with MM estimation as follows:

$$\hat{\beta}_{MM} = (X^T W_i X)^{-1}(X^T W_i Y) \tag{7}$$

$$\hat{\beta}_{\delta,h}^{RKR} = W\hat{\xi} = W(T_h^T T_h)^{-1} T_h^T \hat{\beta}_{MM} \tag{8}$$

8. The simulation in steps 2 to 6 is repeated 1000 times;
9. Validate the model by looking at the Root Mean Squared Error of prediction (RMSEP) value:

$$RMSEP = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n}} \tag{9}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y}_i)^2} \tag{10}$$

y_i is the observed value for the i-th data, and $\hat{y}_i$ is the predicted value for the ith data obtained from the prediction model. $\bar{y}_i$ is the average of the actual observation values, $\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$ is the sum of squares of residuals, and $\sum_{i=1}^{n}(y_i - \bar{y}_i)^2$ is the total sum of squares.

10. Comparing the evaluation results of RCR-LASSO and RCR-WLAD LASSO modelling via RMSEP values.

C. RESULT AND DISCUSSION

1. Results of Comparative Study of The RCR-LASSO and RCR-WLAD LASSO Methods

The following are the simulation results of the RCR-LASSO selection and RCR-WLAD LASSO selection methods on low-dimensional and high-dimensional data with several levels of outliers (δ) and correlation (ρ). Figure 1 is the simulation result in the low dimensional case (n>p) with the remaining normal distribution. It can be seen in the picture that when the outlier is 0% at several levels of correlation, the RMSEP value of the RCR- LASSO is lower than the RCR-WLAD LASSO. A lower RMSEP value indicates that the average model prediction error is lower. Overall, this means the closer the model predictions are to the actual values. The higher the outlier level provided by RCR, LASSO selection and WLAD LASSO selection, it tends to increase, but LASSO tends to be lower and more stable. When the correlation value becomes greater for outliers (δ=10%, 20%, 30%), RCR- WLAD LASSO selection decreases, and the variance becomes smaller. If the correlation between variables is higher, more variables can be related to each other, eliminating more variables from the model. This can reduce model complexity and increase generalization, reducing the RMSEP value of WLAD LASSO selection, as shown in Figure 1.

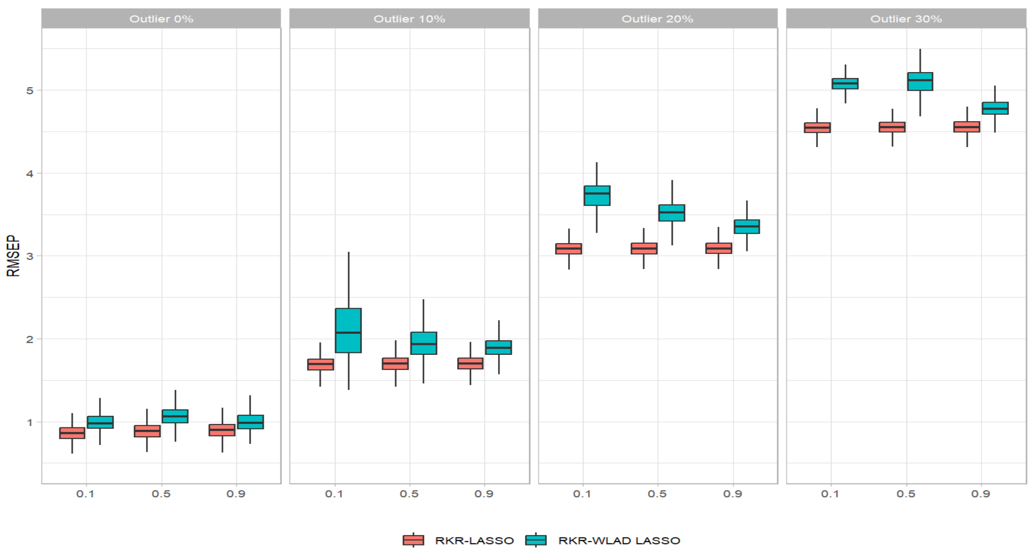

Figure 1. Low-dimensional data simulation results($n = 75 > p = 25$)

Figure 2 is a simulation result of high-dimensional data (n<p) with a residual normal distribution. The RMSEP value of the RCR-LASSO selection tends to be lower than the RCR-WLAD LASSO selection. It can be seen in the picture that when the outlier level is higher, the results of the RCR-LASSO selection and the RCR-WLAD LASSO selection tend to have increasingly similar RMSEP values. Both approaches, RCR-LASSO and RCR-WLAD LASSO selection, use robust estimation techniques for outliers. This allows both models to handle the data better and adapt to the presence of outliers. The two models tend not to be too influenced by outliers in the modeling process, which can produce almost similar RMSEP values. Increasing the correlation value in the RCR-LASSO selection and RCR-WLAD LASSO selection influences the decrease in the RMSEP value. Increasing correlation can increase the stability of regression models, especially when used in RCR methods. As a result, the model can provide more consistent estimates, which have the potential to reduce RMSEP because prediction errors are more controlled, as shown in Figure 2.

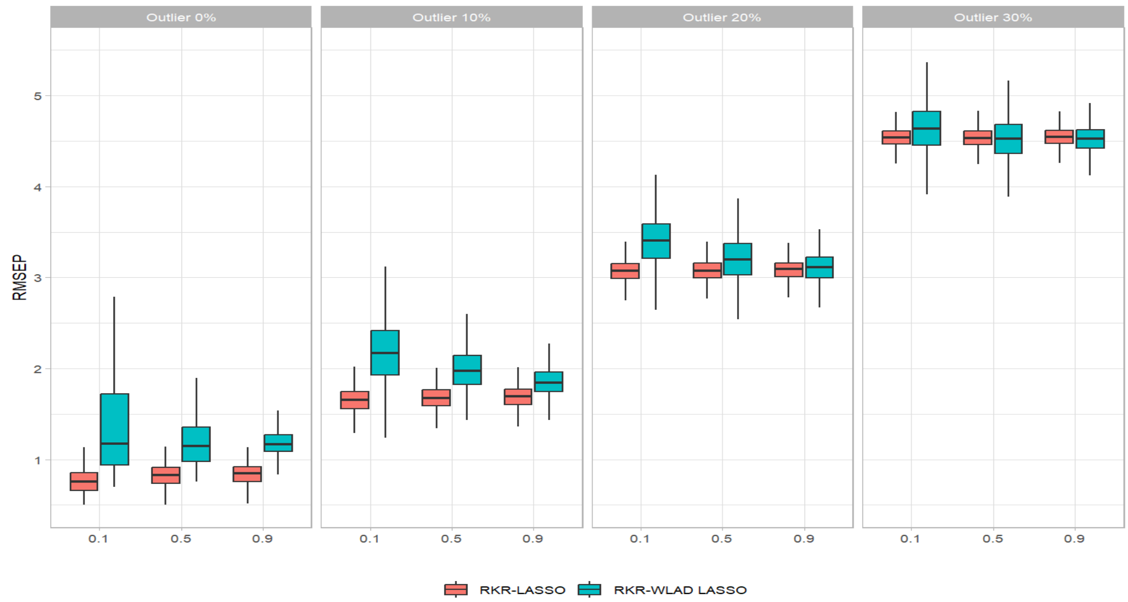

Figure 2. High-dimensional data simulation results($n = 75 < p = 100$)

Table 1. RMSEP and Coefficient of Determination (R^2) Values from Simulation Results of Low-Dimensional Data RCR-LASSO and RCR-WLAD LASSO Selection

δ		**RMSEP**			**R^2**		
		ρ = 0,1	**ρ = 0,5**	**ρ =0,9**	**ρ = 0,1**	**ρ = 0,5**	**ρ =0,9**
0%	RCR-LASSO	0,861	0,883	0,898	0,877	0,951	0,964
	RCR-WLAD LASSO	1,068	1,011	1,002	0,892	0,938	0,949
10%	RCR-LASSO	1,688	1,698	1,699	0,769	0,862	0,884
	RCR-WLAD LASSO	2,101	1,948	1,896	0,756	0,795	0,896
20%	RCR-LASSO	3,085	3,087	3,091	0,643	0,716	0,767
	RCR-WLAD LASSO	3,327	3,538	3,352	0,547	0,639	0,734
30%	RCR-LASSO	4,546	4,553	4,554	0,563	0,632	0,676
	RCR-WLAD LASSO	3,411	5,093	4,779	0,463	0,517	0,631

Table 2. RMSEP and R^2 Values from Simulation Results of High-Dimensional Data RCR-LASSO and RCR-WLAD LASSO Selection

		RMSEP			**R^2**		
		ρ = 0,1	**ρ = 0,5**	**ρ =0,9**	**ρ = 0,1**	**ρ = 0,5**	**ρ =0,9**
0%	RCR-LASSO	1,999	1,014	0,834	0,916	0,969	0,971
	RCR-WLAD LASSO	1,316	1,188	1,181	0,855	0,919	0,939
10%	RCR-LASSO	1,916	1,803	1,681	0,793	0,809	0,894
	RCR-WLAD LASSO	2,123	1,995	1,855	0,567	0,766	0,899
20%	RCR-LASSO	3,266	3,204	3,079	0,632	0,706	0,752
	RCR-WLAD LASSO	3,277	3,203	3,111	0,375	0,634	0,705
30%	RCR-LASSO	4,786	4,725	4,544	0,551	0,607	0,664
	RCR-WLAD LASSO	4,562	4,525	4,524	0,547	0,609	0,673

Next, a significant test was carried out using the t-test to see whether there were fundamental average differences in each RMSEP between the resulting RKR-LASSO and RKR-WLAD LASSO methods. The results show that in low-dimensional data with several correlations and outliers, the p-value is 2.2e-16, smaller than the significance level of 0.05. Therefore, the null hypothesis is rejected, indicating a significant difference between the RMSEP of the two methods. Meanwhile, in high-dimensional data at outlier levels of 10%, 20%, and 30% with several levels of correlation, the results show a p-value of 0.1142, which is greater than the significance level of 0.05. Therefore, the null hypothesis is accepted, indicating there is insufficient evidence to show that there is a significant difference between the two methods.

Figure 3 average simulation results of selected variables using the LASSO and WLAD LASSO methods. The averages of selected variables tend to be almost the same at different levels of correlation. Fewer variables were selected using the WLAD LASSO method on low-dimensional and high-dimensional data than LASSO. This is in line with research by Septa (2022) comparing variable selection between LASSO, LAD LASSO and WLAD-LASSO with the results that WLAD LASSO can overcome the weakness of LASSO which selects many variables that are not significant in the research (Cahya et al., 2022).Variables selected using the LASSO method in low-dimensional data tend to be fewer than in high-dimensional data. In low-dimensional data, LASSO tends to select fewer variables due to various factors, such as the ratio of variables to sample, so that LASSO has fewer variables to consider for adjustment and limits the number of variables that can be selected, as shown in Figure 3.

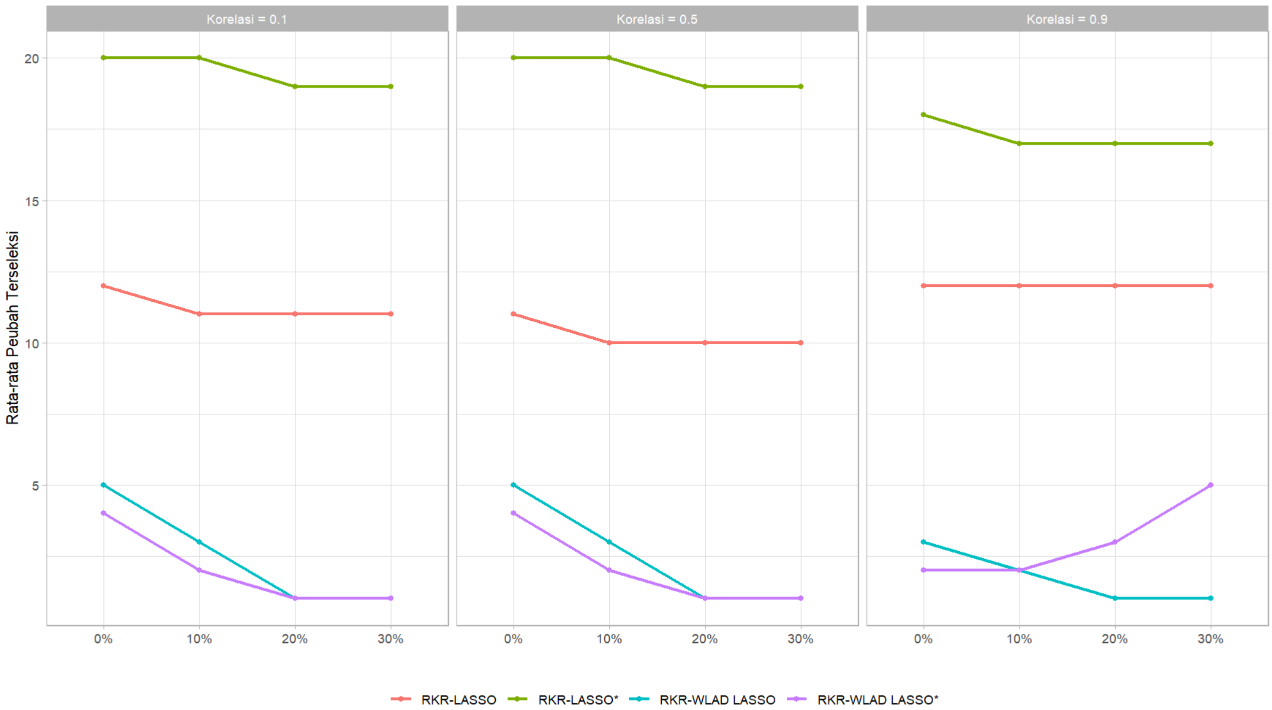

Figure 3. Average simulation results of selected variables

Description: * High dimensional data

D. CONCLUSION AND SUGGESTIONS

Based on the research carried out on low-dimensional data, the performance of RKR-LASSO tends to be superior in terms of model goodness of fit compared to RKR-WLAD LASSO. The combination of LASSO, which tends to be looser in selecting relevant variables, and RKR, which provides resistance to outliers and multicollinearity problems, produces a simpler and more stable model. However, the performance of RKR-LASSO tends to decrease as outliers and correlations increase. In high-dimensional data, the performance of the two methods in handling data complexity is the same. RKR-LASSO tends to be looser in selecting relevant variables, resulting in a simpler model, but the variables chosen by LASSO are only marginally significant. RKR-WLAD LASSO is stricter in variable selection and only selects significant variables but ignores several variables that have a small but significant impact on the model. For further research, it is necessary to apply other methods, such as combining other variable selection methods, such as Least Trimmed Squares (LTS), by minimizing the effect of outliers on parameter estimates with Elastic Net, which combines the LASSO method and the Ridge method to overcome the weaknesses of OLS and WLAD in dealing with outliers. A genetic algorithm is also necessary to find the shrinkage factor for the RKR method. Genetic algorithms can help find the optimal combination of factors to increase resistance to overfitting or underfitting and can improve model performance.

REFERENCES

Ajeel, S. M., & Hashem, H. A. (2020). Comparison Some Robust Regularization Methods in Linear Regression via Simulation Study. *Academic Journal of Nawroz University*, *9*(2), 244–252. https://doi.org/10.25007/ajnu.v9n2a818

Arisandi, A., Wigena, A. H., & Mohamad Soleh, A. (2020). Continuum Regression Modeling with LASSO to Estimate Rainfall. *International Journal of Scientific and Research Publications (IJSRP)*, *10*(10), 380–385. https://doi.org/10.29322/ijsrp.10.10.2020.p10651

Arslan, O. (2012). Weighted LAD-LASSO method for robust parameter estimation and variable selection in regression. *Computational Statistics and Data Analysis*, *56*(6), 1952–1965. https://doi.org/10.1016/j.csda.2011.11.022

Boudt, K., Rousseeuw, P. J., Vanduffel, S., & Verdonck, T. (2020). The minimum regularized covariance determinant estimator. *Statistics and Computing*, *30*(1), 113–128. https://doi.org/10.1007/s11222-019-09869-x

Bulut, H. (2020). Mahalanobis distance based on minimum regularized covariance determinant estimators for high dimensional data. *Communications in Statistics - Theory and Methods*, *49*(24), 5897–5907. https://doi.org/10.1080/03610926.2020.1719420

Cahya, S. D., Sartono, B., Indahwati, I., & Purnaningrum, E. (2022). Performance of LAD-LASSO and WLAD-LASSO on High Dimensional Regression in Handling Data Containing Outliers. *JTAM (Jurnal Teori Dan Aplikasi Matematika)*, *6*(4), 844–856. https://doi.org/10.31764/jtam.v6i4.8968

Chen, X., & Zhu, L. P. (2015). Connecting continuum regression with sufficient dimension reduction. *Statistics and Probability Letters*, *98*(2015), 44–49. https://doi.org/10.1016/j.spl.2014.12.007

ChunRong, C., ShanXiong, C., Lin, C., & YuChen, Z. (2017). Method for Solving LASSO Problem Based on Multidimensional Weight. *Advances in Artificial Intelligence*, *2017*(1), 1–9. https://doi.org/10.1155/2017/1736389

Cui, C., & Wang, D. (2016). High dimensional data regression using Lasso model and neural networks with random weights. *Information Sciences*, *372*(2016), 505–517. https://doi.org/10.1016/j.ins.2016.08.060

Emmert-Streib, F., & Dehmer, M. (2019). High-Dimensional LASSO-Based Computational Regression Models: Regularization, Shrinkage, and Selection. *Machine Learning and Knowledge Extraction*, *1*(1), 359–383. https://doi.org/10.3390/make1010021

Ismah, I., Erfiani, Wigena, A. H., & Sartono, B. (2024). Functional Continuum Regression Approach to Wavelet Transformation Data in a Non-Invasive Glucose Measurement Calibration Model. *Mathematics and Statistics*, *12*(1), 69–79. https://doi.org/10.13189/ms.2024.120110

Izenman, A. J. (2008). *Modern Multivariate Statistical Techniques* (G. Casella, S. Fienbarg, & I. Olkin, Eds.; Vol. 2008). Springer New York. https://doi.org/10.1007/978-0-387-78189-1

Khotimah, K., Sadik, K., & Rizki, A. (2020). Study of Robust Regression Modeling Using MM-Estimator and Least Median Squares. *Proceedings of the 1st International Conference on Statistics and Analytics*, 1–20. https://doi.org/10.4108/eai.2-8-2019.2290533

Lakshmi, K., Mahaboob, B., Rajaiah, M., & Narayana, C. (2021). Ordinary least squares estimation of parameters of linear model. *Journal of Mathematical and Computational Science*, *11*(2), 2015–2030. https://doi.org/10.28919/jmcs/5454

Lee, S., Seo, M. H., & Shin, Y. (2016). The lasso for high dimensional regression with a possible change point. *J. R. Statist. Soc. B*, *78*(1), 193–210. https://doi.org/https://doi.org/10.1111/rssb.12108

Leys, C., Klein, O., Dominicy, Y., & Ley, C. (2018). Detecting multivariate outliers: Use a robust variant of the Mahalanobis distance. *Journal of Experimental Social Psychology*, *74*(2018), 150–156. https://doi.org/10.1016/j.jesp.2017.09.011

Lima, E., Davies, P., Kaler, J., Lovatt, F., & Green, M. (2020). Variable selection for inferential models with relatively high-dimensional data: Between method heterogeneity and covariate stability as adjuncts to robust selection. *Scientific Reports*, *10*(1), 1–11. https://doi.org/10.1038/s41598-020-64829-0

Robert Tibshirani. (1996). Regression shrinkage and selection via the lasso: a retrospective. *Journal of the Royal Statistical Society B*, *58*(1), 267–288. https://doi.org/https://doi.org/10.1111/j.2517-6161.1996.tb02080.x

Serneels, S., Filzmoser, P., Croux, C., & Van Espen, P. J. (2005). Robust continuum regression. *Chemometrics and Intelligent Laboratory Systems*, *76*(2), 197–204. https://doi.org/10.1016/j.chemolab.2004.11.002

setiawan, & Notodiputro, K. A. (2007). Regresi Kontinum dengan Prapemrosesan Transformasi Wavelet Diskret (Continum Regression with Discrete Wavelet Transformation Preprocessing). *Jurnal ILMU DASAR*, *8*(2), 103–109.

Stone, M., & Brooks, R. J. (1990). Continuum Regression: Cross-Validated Sequentially Constructed Prediction Embracing Ordinary Least Squares, Partial Least Squares and Principal Components Regression. *Journal of the Royal Statistical Society: Series B*, *52*(2), 237–269. https://doi.org/10.1111/j.2517-6161.1990.tb01786.x

Tsagris, M., & Pandis, N. (2021). Multicollinearity. *Statistics and Research Design*, *159*(5), 695–696. https://doi.org/10.1016/j.ajodo.2021.02.005

Velliangiri, S., Alagumuthukrishnan, S., & Thankumar Joseph, S. I. (2019). A Review of Dimensionality Reduction Techniques for Efficient Computation. *Procedia Computer Science*, *165*(2019), 104–111. https://doi.org/10.1016/j.procs.2020.01.079

Wahid, A., Khan, D. M., & Hussain, I. (2017). Robust Adaptive Lasso method for parameter's estimation and variable selection in high-dimensional sparse models. *PLOS ONE*, *12*(8), 1–17. https://doi.org/10.1371/journal.pone.0183518

Xie, Z., Feng, X., Chen, X., & Huang, G. (2020). Optimizing a vector of shrinkage factors for continuum regression. *Chemometrics and Intelligent Laboratory Systems*, *206*(2020), 104–111. https://doi.org/10.1016/j.chemolab.2020.104141

Yang, H., & Li, N. (2018). WLAD-LASSO method for robust estimation and variable selection in partially linear models. *Communications in Statistics - Theory and Methods*, *47*(20), 4958–4976. https://doi.org/10.1080/03610926.2017.1383427

Zhang, X. Y., Li, Q. B., & Zhang, G. J. (2011). Modified robust continuum regression by net analyte signal to improve prediction performance for data with outliers. *Chemometrics and Intelligent Laboratory Systems*, *107*(2), 333–342. https://doi.org/10.1016/j.chemolab.2011.05.003

Zhou, Z. (2019). Functional continuum regression. *Journal of Multivariate Analysis*, *173*(20), 1–22. https://doi.org/10.1016/j.jmva.2019.03.006