

Performance Evaluation of ARDL Model Stacked with Boosted Ridge Regression on Time Series Data with Multicollinearity*

Evaluasi Kinerja Estimasi Model *ARDL stacked with Boosted Ridge Regression* pada Data Deret Waktu dengan Multikolinearitas

Amir A. Dalimunthe^{1‡}, Agus M. Soleh¹, and Farit M. Afendi¹

¹ Study Program on Statistics and Data Science, IPB University, Indonesia

[‡]corresponding author: amirdalimunthe@apps.ipb.ac.id

Copyright © 2025 Amir A. Dalimunthe, Agus M. Soleh, and Farit M. Afendi. This is an open-access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Time series data plays a vital role in financial and economic study. Two commonly applied models for such data are Vector Autoregression (VAR) and Autoregressive Distributed Lags (ARDL). Nonetheless, interdependence among explanatory variables often leads to multicollinearity, posing challenges for model reliability. This study investigates the effectiveness of the ARDL model integrated with boosted ridge regression as a method to mitigate multicollinearity. Due to limitations in available empirical data, simulation data will be generated to support the analysis. The research consists of two stages: synthetic data generation and analysis on simulated data. Results suggest that ARDL performs well under various multicollinearity conditions, particularly when the training set is sufficiently large and model structure is correctly specified. For smaller training sets, the ARDL Ridge variant demonstrates improved predictive performance.

Keywords: ARDL, multicollinearity, ridge regression, simulation, time series forecast.

* Received: Jun 2025; Reviewed: Jun 2025; Published: Jun 2025

1. Pendahuluan

Data deret waktu merupakan bentuk data yang lazim digunakan dalam studi ekonomi dan keuangan. Dua model statistik yang banyak diterapkan untuk menganalisis data ini adalah *Vector Autoregression* (VAR) dan *Autoregressive Distributed Lags* (ARDL). VAR telah digunakan oleh sejumlah studi, antara lain Alena *et al.* (2017) untuk menganalisis pengaruh peubah makroekonomi terhadap indeks sektoral di BEI; Prabowo *et al.* (2022) dalam menelaah hubungan antara harga ekonomis LPG dan tingkat kemiskinan; serta Qisthina *et al.* (2022) yang meneliti determinan imbal hasil obligasi Pemerintah Indonesia. Zunara *et al.* (2022) menggunakan *Vector Error Correction Model* (VECM), yang merupakan pengembangan dari VAR, untuk mengkaji pengaruh sentimen rasional dan irasional investor terhadap IHSG.

Dibandingkan VAR, model ARDL menawarkan fleksibilitas yang lebih tinggi dalam menangkap dinamika jangka pendek sekaligus hubungan jangka panjang antar peubah. Achsani (2010) membandingkan kinerja VECM dan ARDL dalam studi permintaan uang di Indonesia, dan menemukan bahwa ARDL menghasilkan fungsi yang lebih stabil. Model ARDL juga banyak digunakan dalam konteks internasional, sebagaimana ditunjukkan oleh Dritsakis (2011) di Hungaria, Abdulkheir (2013) di Arab Saudi, Mall (2013) di Pakistan, dan Rishad *et al.* (2018) di India.

Namun, interdependensi antar peubah penjelas sering menimbulkan masalah multikolinearitas, yang mengganggu estimasi dan interpretasi koefisien regresi serta menurunkan reliabilitas model. Lavery *et al.* (2017) menunjukkan bahwa banyaknya peubah penjelas yang saling berkorelasi meningkatkan risiko terjadinya *Type-II error*, meskipun interaksi antar faktor ini sedikit mengurangi kenaikan *error* tersebut. Vatcheva *et al.* (2016) juga menegaskan bahwa bahkan korelasi lemah antar dua atau lebih peubah penjelas dapat memicu multikolinearitas signifikan.

Multikolinearitas umumnya ditangani melalui pendekatan reduksi dimensi, seleksi peubah, atau metode regularisasi. Metode regularisasi seperti *Ridge Regression*, LASSO, dan *Elastic Net* sering diterapkan untuk mengatasi multikolinearitas pada *dataset*. Kajian mengenai kemampuan pendekatan *Ridge Regression* dalam menghadapi berbagai kondisi multikolineritas dapat berkontribusi dalam peningkatan performa model deret waktu. Selain itu, diperlukan juga kajian mengenai potensi metode *boosting* dalam peningkatan kemampuan estimasi model pada berbagai kondisi multikolinearitas.

Penelitian ini berfokus pada performa model dalam melakukan peramalan pada *testing set* atau *dataset* di luar data latih. Model VAR dipilih sebagai *baseline* mengingat VAR sebagai salah satu model umum digunakan dan menawarkan efisiensi komputasi. ARDL yang juga cukup populer dikaji performanya pada berbagai situasi multikolinearitas. Penelitian ini bertujuan mengevaluasi performa model ARDL yang dikombinasikan dengan *boosted ridge regression* sebagai solusi terhadap multikolinearitas, dengan harapan dapat menghasilkan model prediksi yang lebih andal.

2. Metodologi

2.1 Data

Kinerja model dievaluasi menggunakan *Root Mean Square Error* (RMSE) berdasarkan *testing set* atau data di luar *training set*. Untuk menguji performa dalam berbagai kondisi multikolinearitas, penelitian ini menggunakan data simulasi karena keterbatasan data empiris. Sebanyak 27 *dataset* dibangkitkan, masing-masing terdiri dari 20.610 observasi dengan satu peubah respon, dengan pengaturan variasi pada jumlah peubah penjelas (5, 10, dan 15 peubah), jumlah pasangan peubah yang saling berkorelasi (1, 3, dan 5 pasangan), serta tingkat korelasi antar peubah (0,10; 0,50; dan 0,90). Untuk tiap *dataset*, dibentuk lima skenario ukuran *training set* (35, 50, 75, 100, dan 300 observasi), sementara banyaknya *testing set* dijaga tetap, yaitu 1.000 set, masing-masing berukuran 30 observasi. Untuk seluruh *dataset* model diasumsikan menggunakan lag-1.

2.2 Metode Analisis Data

Penelitian ini terbagi menjadi dua tahapan yaitu tahap pembangkitan data simulasi dan tahapan analisis data simulasi. Simulasi, pengolahan data, dan pemodelan dilakukan dengan menggunakan *software* R dengan paket [MASS] (Ripley 2022), paket [vars] (Pfaff 2008), [dynlm](Zeileis 2005), [glmnet] (Friedman et al. 2021), dan [glmnetutils] (Ooi 2017). Untuk mengukur kinerja model akan menggunakan paket [mlr3measures] (Lang et al. 2019).

Prosedur pembangkitan data simulasi adalah sebagai berikut:

1. Koefisien model dibangkitkan secara acak menggunakan distribusi uniform dengan batas bawah -10 dan batas atas 10.
2. Nilai tengah (*mean*) untuk masing-masing peubah penjelas ditentukan secara acak dari sebaran uniform dengan rentang -10 hingga 10.
3. Titik awal (*initial values*) tiap peubah ditetapkan secara acak dari sebaran uniform dengan rentang 1.000 hingga 10.000.
4. Dibangkitkan lima peubah penjelas masing-masing berukuran 20.613 observasi, mengikuti sebaran *multivariate normal*, dengan nilai tengah dari Langkah 2 dan satu pasang peubah memiliki koefisien korelasi sebesar 0,1.
5. Nilai awal untuk peubah respon ΔY dibangkitkan secara acak dari sebaran uniform (-10 hingga 10).
6. Komponen error ε dengan ukuran 20.613 observasi dibangkitkan dari distribusi Normal(0,1).
7. Secara iteratif, peubah respon ΔY dibangkitkan menggunakan struktur model ARDL(1,1,...,1) sebagai berikut:

$$\Delta Y_t = \beta_{00} + \beta_{01}\Delta X01_{t-1} + \beta_{02}\Delta X02_{t-1} + \dots + \beta_{05}\Delta X05_{t-1} + \alpha_1\Delta Y_{t-1} + \varepsilon_t \quad (1)$$
8. Transformasi CUSUM dilakukan pada seluruh peubah dengan titik awal dari Langkah 3 sebagai referensi.
9. Dataset hasil simulasi disimpan sesuai dengan kode skenario.
10. Langkah 4–9 diulang untuk skenario dengan koefisien korelasi sebesar 0,5 dan 0,9.
11. Langkah 4–10 diulang untuk skenario dengan 3-pasangan dan 5-pasangan peubah yang saling berkorelasi.

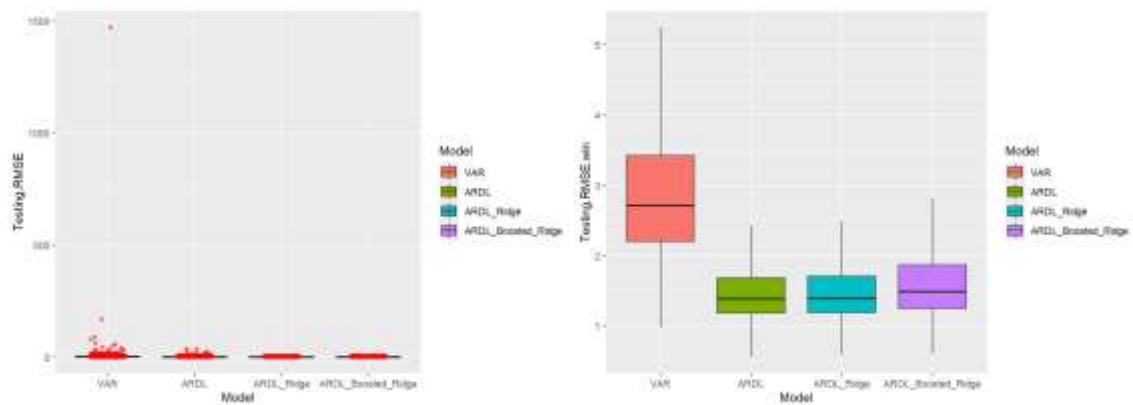
12. Selanjutnya, Langkah 4–11 diulang kembali untuk skenario dengan 10 dan 15 peubah penjelas.

Analisis dilakukan dengan pendekatan *rolling window* dengan langkah sebesar 20 observasi, yang diterapkan pada 135 skenario simulasi (S01–S135). Setiap skenario dievaluasi melalui 1.000 *window*. Prosedur analisis data hasil simulasi adalah sebagai berikut:

1. Tabel performa disiapkan untuk mencatat hasil masing-masing model pada setiap *window* untuk seluruh skenario.
2. Fungsi-fungsi dideklarasikan untuk:
 - a. Menghitung performa model (RMSE),
 - b. Menyusun model ARDL *Ridge*,
 - c. Menyusun model ARDL *stacked with Boosted Ridge Regression* (ARDL_S_BRR).
3. *Testing set* untuk setiap *window* ditetapkan secara bergeser (*rolling*) sebanyak 20 observasi. Contohnya, *window* pertama menggunakan indeks 601–630, *window* kedua menggunakan 621–650, dan seterusnya hingga *window* ke-1.000 (20.581–20.610).
4. *Training set* disusun dengan cara yang sama, bergeser sesuai skenario. Misalnya, pada S01, *window* pertama mengambil data dari indeks 566–600, dan seterusnya hingga *window* terakhir (20.446–20.580).
5. *Dataset* dipraproses sesuai skenario aktif.
6. Untuk setiap *window*:
 - a. Subset *training* dan *testing set* dibentuk sesuai indeks.
 - b. Model VAR, ARDL, dan ARDL_S_BRR dibangun menggunakan *training set*, dan dievaluasi pada *testing set*.
 - c. Performa (RMSE) dari masing-masing model dicatat.
7. Langkah ini diulang untuk seluruh 1.000 *window* dalam setiap skenario.
8. Proses diulang dari awal untuk skenario S02 hingga S135.
9. Hasil akhir digunakan untuk mengevaluasi performa agregat masing-masing model di seluruh kondisi skenario.

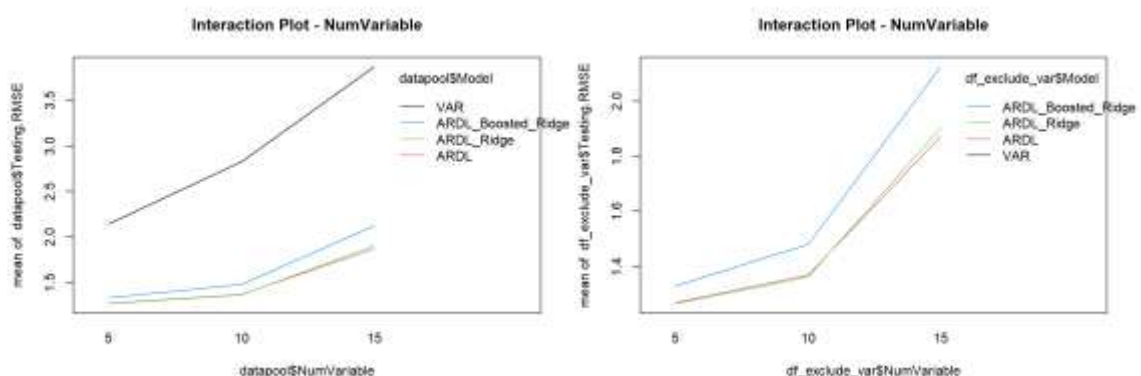
3. Hasil dan Pembahasan

Setelah dilakukan winsorisasi terhadap pencilan yang terdeteksi pada salah satu iterasi model VAR, *boxplot* pada Gambar 1 menunjukkan perbedaan performa antar model. Model VAR mencatatkan RMSE tertinggi, mengindikasikan kinerja yang paling lemah. ARDL dan ARDL *Ridge* menghasilkan performa yang hampir identik dan lebih baik, sementara ARDL *Boosted Ridge* secara konsisten mengungguli VAR, meskipun sedikit tertinggal dari dua model ARDL lainnya.



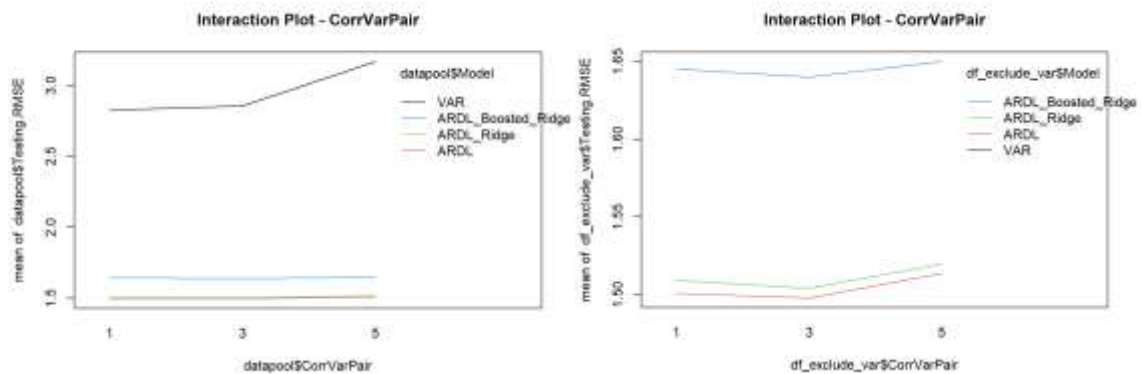
Gambar 1. *Boxplot* Perbandingan RMSE pada *Testing Set* berdasarkan Model

Interaction plot pada Gambar 2 memperlihatkan bahwa semakin banyak peubah penjelas yang memengaruhi peubah dependen, maka RMSE model cenderung meningkat. Temuan ini menunjukkan bahwa kompleksitas model turut berkontribusi terhadap degradasi performa prediktif. Secara umum, model VAR menunjukkan kinerja paling rendah di hampir seluruh kondisi jumlah peubah penjelas. Sementara itu, ARDL dan ARDL *Ridge* konsisten menghasilkan performa yang lebih baik dan relatif sebanding satu sama lain.



Gambar 2. Perbandingan *Interaction Plot* antara Model dan Jumlah Peubah Penjelas

Interaction plot pada Gambar 3 tidak menunjukkan pola yang konsisten dalam performa model terhadap variasi jumlah pasangan peubah yang saling berkorelasi. Namun, model VAR mengalami peningkatan RMSE yang lebih tajam saat jumlah pasangan berkorelasi bertambah dari tiga menjadi lima, dibandingkan peningkatan dari satu ke tiga pasangan. Di sisi lain, model ARDL secara umum menunjukkan performa terbaik dalam berbagai kondisi jumlah pasangan berkorelasi.

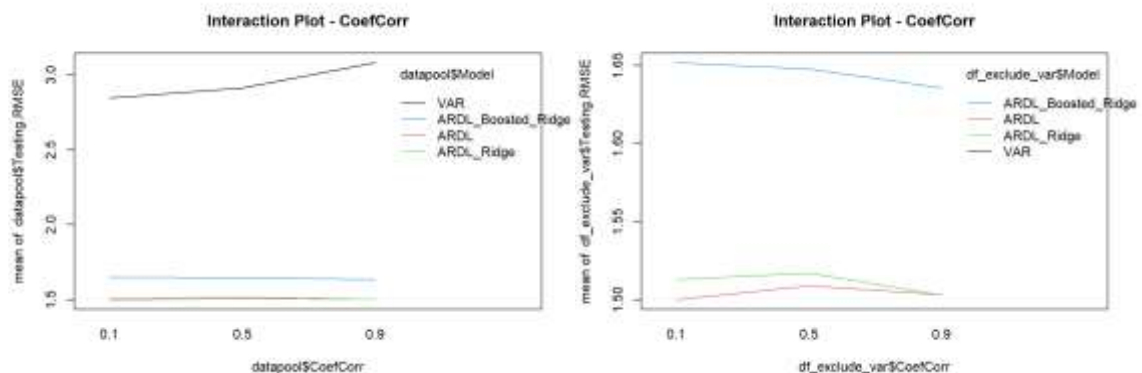


Gambar 3. Perbandingan *Interaction Plot* antara Model dan Banyaknya Pasangan Berkorelasi

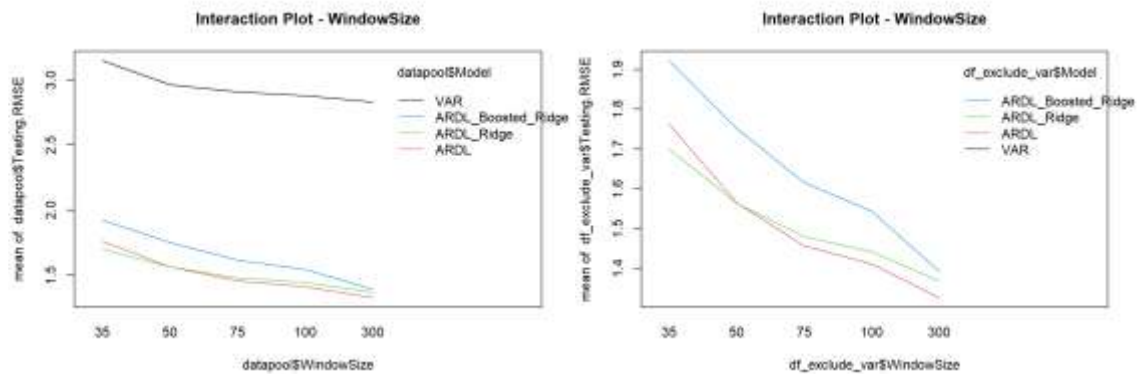
Interaction plot pada Gambar 4 menunjukkan respons yang berbeda antara model VAR dan model lainnya terhadap variasi koefisien korelasi antar peubah penjelas. Kinerja model VAR memburuk seiring meningkatnya korelasi antar peubah penjelas, sedangkan *ARDL Boosted Ridge* justru menunjukkan perbaikan performa dalam kondisi yang sama. Sementara itu, *ARDL* dan *ARDL Ridge* tetap menunjukkan hasil yang relatif unggul dan stabil di berbagai tingkat korelasi.

Interaction plot pada Gambar 5 memperlihatkan bahwa secara umum peningkatan ukuran *training set* menghasilkan RMSE yang lebih rendah. Pada ukuran *training set* kecil (35 observasi), model *ARDL Ridge* menunjukkan performa yang lebih baik dibandingkan *ARDL*. Namun, seiring bertambahnya ukuran data latih, model *ARDL* mencatatkan penurunan RMSE yang signifikan. *ARDL Boosted Ridge* menunjukkan laju penurunan RMSE tercepat, menandakan responsivitas model yang lebih tinggi terhadap ketersediaan data latih.

Penyelidikan lanjutan dilakukan melalui estimasi *Generalized Linear Model* (GLM) dengan RMSE pada *testing set* sebagai peubah dependen, sebagaimana ditampilkan pada Tabel 1. Model referensi (*intercept*) menggunakan VAR dengan konfigurasi dasar: lima peubah penjelas, satu pasangan berkorelasi, koefisien korelasi sebesar 0,1, dan *training set* berukuran 35 observasi.



Gambar 4. Perbandingan *Interaction Plot* antara Model dan Nilai Koefisien Korelasi

Gambar 5. Perbandingan *Interaction Plot* antara Model dan Ukuran *Training Set*

Tabel 1. Hasil Analisis Model Generalized Linear Model (GLM)

Variabel	Estimate	p-value	Significance	Interaction	Estimate	p-value	Significance	Interaction	Estimate	p-value	Significance
(Intercept)	2,120	< 2e-16	***	NumVariable10:ModelARDL	-0,586	< 2e-16	***	WindowSize50:ModelARDL	-0,013	0,613	
NumVariable10	0,686	< 2e-16	***	NumVariable15:ModelARDL	-1,122	< 2e-16	***	WindowSize75:ModelARDL	-0,067	0,008	**
NumVariable15	1,724	< 2e-16	***	NumVariable10:ModelARDL_Ridge	-0,587	< 2e-16	***	WindowSize100:ModelARDL	-0,086	0,001	***
CorrVarPair3	0,029	0,038	*	NumVariable15:ModelARDL_Ridge	-1,087	< 2e-16	***	WindowSize300:ModelARDL	-0,121	0,000	***
CorrVarPair5	0,340	< 2e-16	***	NumVariable10:ModelARDL_Boosted_Ridge	-0,535	< 2e-16	***	WindowSize50:ModelARDL_Ridge	0,050	0,048	*
CoefCorr0.5	0,067	0,000	***	NumVariable15:ModelARDL_Boosted_Ridge	-0,930	< 2e-16	***	WindowSize75:ModelARDL_Ridge	0,020	0,426	
CoefCorr0.9	0,235	< 2e-16	***	CorrVarPair3:ModelARDL	-0,032	0,104		WindowSize100:ModelARDL_Ridge	0,010	0,709	
WindowSize50	-0,185	< 2e-16	***	CorrVarPair5:ModelARDL	-0,328	< 2e-16	***	WindowSize300:ModelARDL_Ridge	-0,015	0,563	
WindowSize75	-0,239	< 2e-16	***	CorrVarPair3:ModelARDL_Ridge	-0,034	0,084		WindowSize50:ModelARDL_Boosted_Ridge	0,015	0,560	
WindowSize100	-0,266	< 2e-16	***	CorrVarPair5:ModelARDL_Ridge	-0,330	< 2e-16	***	WindowSize75:ModelARDL_Boosted_Ridge	-0,067	0,008	**
WindowSize300	-0,315	< 2e-16	***	CorrVarPair3:ModelARDL_Boosted_Ridge	-0,034	0,083		WindowSize100:ModelARDL_Boosted_Ridge	-0,112	0,000	***
ModelARDL	-0,599	< 2e-16	***	CorrVarPair5:ModelARDL_Boosted_Ridge	-0,336	< 2e-16	***	WindowSize300:ModelARDL_Boosted_Ridge	-0,214	< 2e-16	***
ModelARDL_Ridge	-0,666	< 2e-16	***	CoefCorr0.5:ModelARDL	-0,058	0,003	**				
ModelARDL_Boosted_Ridge	-0,506	< 2e-16	***	CoefCorr0.9:ModelARDL	-0,232	< 2e-16	***				
				CoefCorr0.5:ModelARDL_Ridge	-0,063	0,001	**				
				CoefCorr0.9:ModelARDL_Ridge	-0,245	< 2e-16	***				
				CoefCorr0.5:ModelARDL_Boosted_Ridge	-0,071	0,000	***				
				CoefCorr0.9:ModelARDL_Boosted_Ridge	-0,251	< 2e-16	***				

* 5%

** 1%

*** 0,1%

Untuk analisis *post-hoc*, dilakukan pemeriksaan melalui Uji Kontras sebagaimana disajikan pada Tabel 2. Hasil uji kontras mengindikasikan bahwa model VAR secara signifikan memiliki RMSE lebih tinggi dibandingkan ketiga model lainnya. Sementara itu, tidak terdapat perbedaan signifikan antara ARDL dan ARDL *Ridge* dalam hal performa RMSE yang dihasilkan.

Tabel 3 menyajikan perbandingan umum performa antar model. Model VAR mencatatkan waktu komputasi tercepat, namun menghasilkan RMSE tertinggi secara konsisten. ARDL membutuhkan waktu komputasi sedikit lebih lama, namun memberikan RMSE yang lebih rendah.

ARDL *Boosted Ridge* merupakan model dengan waktu komputasi terlama, namun tidak menunjukkan keunggulan performa dibandingkan ARDL dan ARDL *Ridge*. ARDL *Boosted Ridge* tidak secara konsisten mengungguli ARDL *Ridge*, meskipun waktu komputasinya tercatat hampir lima puluh kali lebih lama dibandingkan ARDL *Ridge*.

Tabel 2. Hasil Analisis *post-hoc*

contrast	estimate	SE	df	t.ratio	p.Value
VAR - ARDL	1,442	0,008	539956	179,254	<0,0001
VAR - ARDL_Ridge	1,435	0,008	539956	178,392	<0,0001
VAR - ARDL_Boosted_Ridge	1,301	0,008	539956	161,728	<0,0001
ARDL - ARDL_Ridge	-0,007	0,008	539956	-0,862	0,825
ARDL - ARDL_Boosted_Ridge	-0,141	0,008	539956	-17,526	<0,0001
ARDL_Ridge - ARDL_Boosted_Ridge	-0,134	0,008	539956	-16,664	<0,0001

Tabel 3. Perbandingan Performa Model pada Data Simulasi dan Data Empiris

Model	Median Testing.RMSE	Median Building Duration (detik)
VAR	2,717	0,034
ARDL	1,386	0,045
ARDL_Ridge	1,393	0,088
ARDL_Boosted_Ridge	1,488	4,519

4. Simpulan dan Saran

Secara umum, model ARDL menunjukkan performa yang lebih baik dibandingkan ARDL Ridge maupun ARDL Boosted Ridge dalam menghadapi berbagai kondisi multikolinearitas, selama ukuran *training set* mencukupi dan model terspesifikasi dengan baik. Pada ukuran *training set* yang kecil, pendekatan ARDL Ridge mampu meningkatkan kinerja model secara signifikan.

Penelitian lebih lanjut terkait dampak interaksi antara multikolinearitas dengan pencilaan dan spesifikasi model terhadap performa model dapat memberikan pemahaman yang lebih mendalam terhadap penanganan masalah multikolinearitas. Pemrograman yang lebih efisien juga dibutuhkan untuk membangun model ARDL Ridge dan ARDL Boosted Ridge yang lebih cepat dan *cost effective*.

Daftar Pustaka

- Abdulkheir A.Y. (2013). An Analytical Study of the Demand for Money in Saudi Arabia. *International Journal of Economics and Finance*, 5(4):31-38. doi:10.5539/ijef.v5n4p31.
- Achsani N.A. (2010). Stability of Money Demand in an Emerging Market Economy: An Error Correction and ARDL Model for Indonesia. *Research Journal of International Studies*, 13: 54-62.
- Alena E., Achsani N.A., Andati T. (2017). Dampak Guncangan Variabel Makroekonomi terhadap Beta Indeks Sektoral di BEI. *Jurnal Aplikasi Bisnis dan Manajemen*.
- Dritsakis N. 2011. Demand for Money in Hungary: An ARDL Approach. *Review of Economics and Finance*, 1:1–16.
- Friedman J., Hastie T., Tibshirani R., Narasimhan B., Tay K., Simon N., Yang J. (2021). glmnet: Lasso and Elastic-Net Regularized Generalized Linear Models [R Package]. CRAN.
- Hastie T., Friedman J., Tibshirani R. (2001). *The Elements of Statistical Learning*. Springer.
- Lang M., Becker M., Koers L. (2019). mlr3measures: Performance Measures for mlr3 [R Package]. CRAN.
- Lavery M.R., Acharya P., Sivo S.A., Xu L. (2017). Number of predictors and multicollinearity: What are their effects on error and bias in regression? *Communications in Statistics – Simulation and Computation*, 48(1):27–38. doi:10.1080/03610918.2017.1371750.
- Lütkepohl H. (2005). *New Introduction to Multiple Time Series Analysis*. Springer.
- Mall S. (2013). Estimating a Function of Real Demand for Money in Pakistan: An Application of Bounds Testing Approach to Cointegration. *International Journal of*

- Computer Applications*, 79(5):32–50. doi:10.5120/13740-1548.
- Ooi H. (2017). glmnetUtils: Utilities for Glmnet [R Package]. CRAN.
- Pesaran M.H., Shin Y., Smith R.J. (2001). Bounds testing approaches to the analysis of level relationships. *Journal of Applied Econometrics*. 16(3):289–326. doi:10.1002/jae.616.
- Pfaff B. (2008). VAR, SVAR and SVEC Models: Implementation Within R Package vars. *Journal of Statistical Software*. 27(4). doi:10.18637/jss.v027.i04.
- Prabowo E., Harianto, Juanda B., Indrawan D. (2022). The Economic Price of Liquid Petroleum Gas, Poverty and Subsidy Removal Compensation Scenario in Indonesia. *International Journal of Energy Economics and Policy*. 12(5):169–177. doi:10.32479/ijee.13356.
- Qisthina G.F., Achsani N.A., Novianti T. (2022). Determinants of Indonesian Government Bond Yields. *Jurnal Aplikasi Bisnis dan Manajemen*.
- Ripley B. (2022). MASS: Support functions and datasets for Venables and Ripley's MASS [R package]. CRAN.
- Rishad A., Sharma A., Gupta S. (2018). Demand for Money in India: An ARDL Approach. *International Journal of Economics and Financial Research*. hlm 27–42.
- Vatcheva, K. P., Lee, M., McCormick, J. B., & Rahbar, M. H. (2016). Multicollinearity in Regression Analyses Conducted in Epidemiologic Studies. *Epidemiology: Open Access*, 06(02). doi:10.4172/2161-1165.1000227.
- Zeileis A. (2005). dynlm: Dynamic linear regression [R package]. CRAN.
- Zunara E, Achsani N. A, Hakim D. B, Sembel R. (2022). The Effect of Rational and Irrational Sentiments of Individual and Institutional Investors on Indonesia Stock Market. *Jurnal Aplikasi Bisnis dan Manajemen*.