# FORECASTING FINANCIAL INSTRUMENTS PRICES WITH VECM AND ARIMA MODELS

Michał Budasz 439087, Laura Piontek 429977

2025-07-05

## Contents



### 1 Introduction

The aim of the project is to compare accuracy of forecasts of prices of two cointegrated financial instruments with VECM model and two independent univariate ARIMA models.

### 2 Dataset description and data preparation

The analysis was conducted on a provided dataset that consists of 10 time-series type assets (from *y1* to *y10*), for each variable in total 600 daily observations for time horizon from 2023-09-09 up to 2025-04-30. No NA value was identified in the dataset, therefore after changing 'date' variable into timestamp and splitting the dataset into in-sample (575 observations) and out-of-sample (25 observations) batches, the authors proceeded with further analysis.

### 3 Finding a cointegrated variables

In order to develop VECM model, one needs to find cointegrated pair out of ten provided time series in analyzed dataset. As we can see on the chart provided below, there is more than one cointegrated pair in provided dataset (for instance *y5-y6* pair or *y3-y8* pair seem to be cointegrated as these time-series behave similarly for a given time period). So as to ascertain conclusions drawn from the visual analysis, proper statistical tests were conducted.

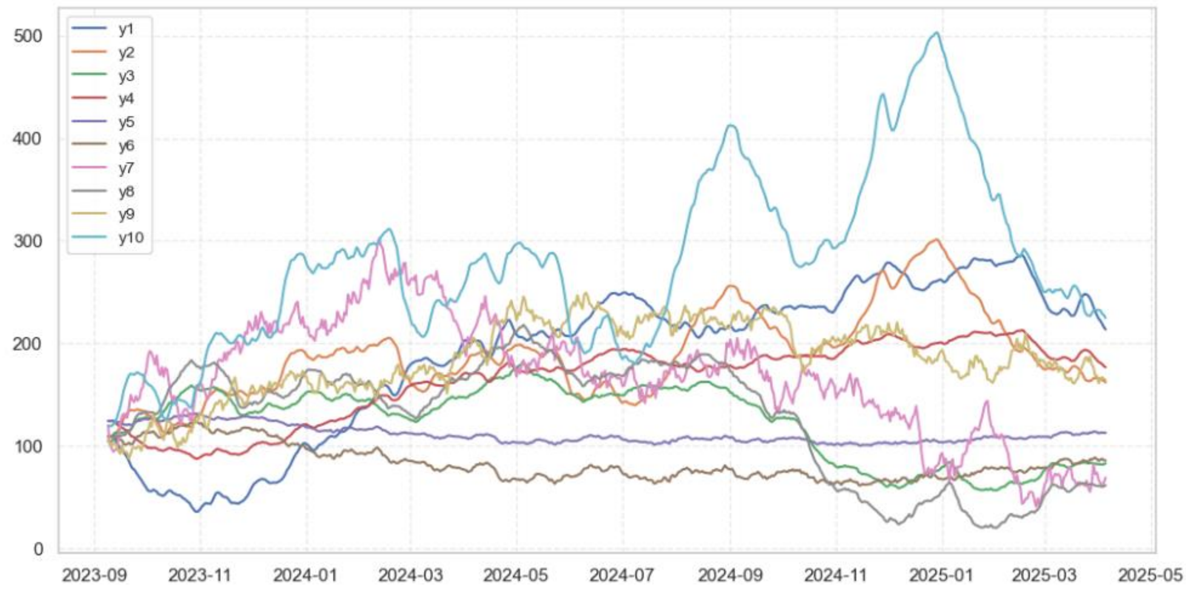


Figure 1: Visualization of each variable values over time

Firstly, order of integration for each variable was checked (as it is necessary that for cointegrated variables, both variables in pair need to have the same order of integration). To do so, an Augmented Dickey-Fuller test was performed- below we present an exemplary procedure for *y1* variable. For the interpretation of ADF statistic (in second column), there is a requirement of no autocorrelation within variable (that is a situation where we fail to reject the null hypothesis in Breush-Godfrey test, by comparing p-values provided in 7$^{th}$ and 9$^{th}$ column with assumed $\alpha$=5% level). As we can see in this case, there is no autocorrelation only when we add 4 augmentations into *y1* time-series: in this situation one can interpret ADF statistic (2$^{nd}$ column) with proper critical value (in our case for $\alpha$=5%, one should compare the values with the ones provided in 4$^{th}$ column). Here we can see that computed test statistic "doesn't fall" into critical value interval- therefore one fails to reject null hypothesis of non-stationarity of time-series and as a result, first differences of given variable need to be analyzed.

| | number of augmentations | ADF test statistic | ADF critival value (1%) | ADF critival value (5%) | ADF critival value (10%) | BG test (1 lag) (statistic) | BG test (1 lag) (p-value) | BG test (5 lags) (statistic) | BG test (5 lags) (p-value) |
|---|---|---|---|---|---|---|---|---|---|
| **0** | 0 | -1.1336934755 | -3.4340000000 | -2.8630000000 | -2.5680000000 | 410.2961000000 | 0.0000000000 | 427.3515000000 | 0.0000000000 |
| **1** | 1 | -1.3639695631 | -3.4340000000 | -2.8630000000 | -2.5680000000 | 34.0159000000 | 0.0000000000 | 63.8021000000 | 0.0000000000 |
| **2** | 2 | -1.1758716648 | -3.4340000000 | -2.8630000000 | -2.5680000000 | 12.4360000000 | 0.0004000000 | 33.9837000000 | 0.0000000000 |
| **3** | 3 | -1.1344141052 | -3.4340000000 | -2.8630000000 | -2.5680000000 | 10.7458000000 | 0.0010000000 | 23.7983000000 | 0.0002000000 |
| **4** | 4 | -1.3004105245 | -3.4340000000 | -2.8630000000 | -2.5680000000 | 0.6391000000 | 0.4240000000 | 8.0554000000 | 0.1532000000 |
| **5** | 5 | -1.2797185681 | -3.4340000000 | -2.8630000000 | -2.5680000000 | 5.4454000000 | 0.0196000000 | 13.2055000000 | 0.0215000000 |
| **6** | 6 | -1.2910704347 | -3.4340000000 | -2.8630000000 | -2.5680000000 | 1.2720000000 | 0.2594000000 | 8.9866000000 | 0.1096000000 |
| **7** | 7 | -1.3219757771 | -3.4340000000 | -2.8630000000 | -2.5680000000 | 1.7460000000 | 0.1864000000 | 6.1195000000 | 0.2948000000 |
| **8** | 8 | -1.2659160421 | -3.4340000000 | -2.8630000000 | -2.5680000000 | 4.6751000000 | 0.0306000000 | 6.9005000000 | 0.2281000000 |
| **9** | 9 | -1.3910229592 | -3.4340000000 | -2.8630000000 | -2.5680000000 | 0.5093000000 | 0.4754000000 | 7.2003000000 | 0.2062000000 |

Table 1: ADF test for *y1* variable

Having in mind the remarks from previous paragraph, the authors conducted similar ADF test for first differences of given variable- conducting analogous analysis one can see that after adding 3 augmentations to time-series (in order to "eliminate" the autocorrelation) the null hypothesis of non-stationarity is rejected, basing on ADF test statistic value. In conclusion, one can finally state that *y1* is a process integrated of order 1 (*y1*~I(1)).

| | number of augmentations | ADF test statistic | ADF critival value (1%) | ADF critival value (5%) | ADF critival value (10%) | BG test (1 lag) (statistic) | BG test (1 lag) (p-value) | BG test (5 lags) (statistic) | BG test (5 lags) (p-value) |
|---|---|---|---|---|---|---|---|---|---|
| **0** | 0 | -6.8612734587 | -3.4340000000 | -2.8630000000 | -2.5680000000 | 34.4703000000 | 0.0000000000 | 64.4254000000 | 0.0000000000 |
| **1** | 1 | -8.4841425877 | -3.4340000000 | -2.8630000000 | -2.5680000000 | 12.5545000000 | 0.0004000000 | 33.2698000000 | 0.0000000000 |
| **2** | 2 | -9.2734079970 | -3.4340000000 | -2.8630000000 | -2.5680000000 | 10.9282000000 | 0.0009000000 | 23.5807000000 | 0.0003000000 |
| **3** | 3 | -7.3817740488 | -3.4340000000 | -2.8630000000 | -2.5680000000 | 0.7455000000 | 0.3879000000 | 7.9635000000 | 0.1583000000 |
| **4** | 4 | -7.3289507288 | -3.4340000000 | -2.8630000000 | -2.5680000000 | 4.3071000000 | 0.0380000000 | 12.9039000000 | 0.0243000000 |
| **5** | 5 | -7.7057963016 | -3.4340000000 | -2.8630000000 | -2.5680000000 | 1.3407000000 | 0.2469000000 | 8.7618000000 | 0.1190000000 |
| **6** | 6 | -6.8786753165 | -3.4340000000 | -2.8630000000 | -2.5680000000 | 1.7860000000 | 0.1814000000 | 5.8018000000 | 0.3260000000 |
| **7** | 7 | -6.9325771173 | -3.4340000000 | -2.8630000000 | -2.5680000000 | 4.6107000000 | 0.0318000000 | 6.7138000000 | 0.2428000000 |
| **8** | 8 | -6.1850202306 | -3.4340000000 | -2.8630000000 | -2.5680000000 | 0.6425000000 | 0.4228000000 | 6.5905000000 | 0.2529000000 |
| **9** | 9 | -6.2306163763 | -3.4340000000 | -2.8630000000 | -2.5680000000 | -0.5681000000 | 1.0000000000 | 7.9145000000 | 0.1610000000 |

Table 2: ADF test for *y1* first differences

Same analysis was conducted for each variable in the dataset- the conclusion is that each variable is integrated of order 1 so theoretically, each variable can be cointegrated with other variable (that is 45 unique pairs)!

| | number of augmentations | ADF test statistic | ADF critival value (1%) | ADF critival value (5%) | ADF critival value (10%) | BG test (1 lag) (statistic) | BG test (1 lag) (p-value) | BG test (5 lags) (statistic) | BG test (5 lags) (p-value) |
|---|---|---|---|---|---|---|---|---|---|
| **0** | 0 | -24.2455006056 | -3.4340000000 | -2.8630000000 | -2.5680000000 | 0.0254000000 | 0.8734000000 | 1.6062000000 | 0.9005000000 |
| **1** | 1 | -17.3776892316 | -3.4340000000 | -2.8630000000 | -2.5680000000 | 0.2651000000 | 0.6067000000 | 2.0621000000 | 0.8405000000 |
| **2** | 2 | -14.0539324208 | -3.4340000000 | -2.8630000000 | -2.5680000000 | 0.0887000000 | 0.7658000000 | 6.3292000000 | 0.2755000000 |
| **3** | 3 | -12.4969244786 | -3.4340000000 | -2.8630000000 | -2.5680000000 | -8.6086000000 | 1.0000000000 | 3.2004000000 | 0.6691000000 |
| **4** | 4 | -11.4382490412 | -3.4340000000 | -2.8630000000 | -2.5680000000 | 2.3281000000 | 0.1271000000 | 6.4451000000 | 0.2653000000 |
| **5** | 5 | -10.5994607765 | -3.4340000000 | -2.8630000000 | -2.5680000000 | 2.7238000000 | 0.0989000000 | 7.6697000000 | 0.1754000000 |
| **6** | 6 | -9.0997121235 | -3.4340000000 | -2.8630000000 | -2.5680000000 | -0.0893000000 | 1.0000000000 | 4.8703000000 | 0.4319000000 |
| **7** | 7 | -8.5761613486 | -3.4340000000 | -2.8630000000 | -2.5680000000 | -0.2051000000 | 1.0000000000 | 5.0347000000 | 0.4117000000 |
| **8** | 8 | -7.9610594935 | -3.4340000000 | -2.8630000000 | -2.5680000000 | 0.6982000000 | 0.4034000000 | 6.8609000000 | 0.2312000000 |
| **9** | 9 | -7.3341944884 | -3.4340000000 | -2.8630000000 | -2.5680000000 | 2.4733000000 | 0.1158000000 | 4.0399000000 | 0.5437000000 |

Table 3: ADF test for residuals in *y3~y8* OLS regression

In order to identify cointegrated pair of variables, the authors performed several OLS regression models and conducted ADF test on its residuals- basing on Engle-Granger methodology, for I(1) processes one needs to include stationary residuals into system (that is $\varepsilon \sim I(0)$). Confirming the assumption from the introduction to this chapter, *y5-y6* and *y3-y8* pairs consist of cointegrated variables- the authors proceeded with *y3-y8* pair in their analysis.

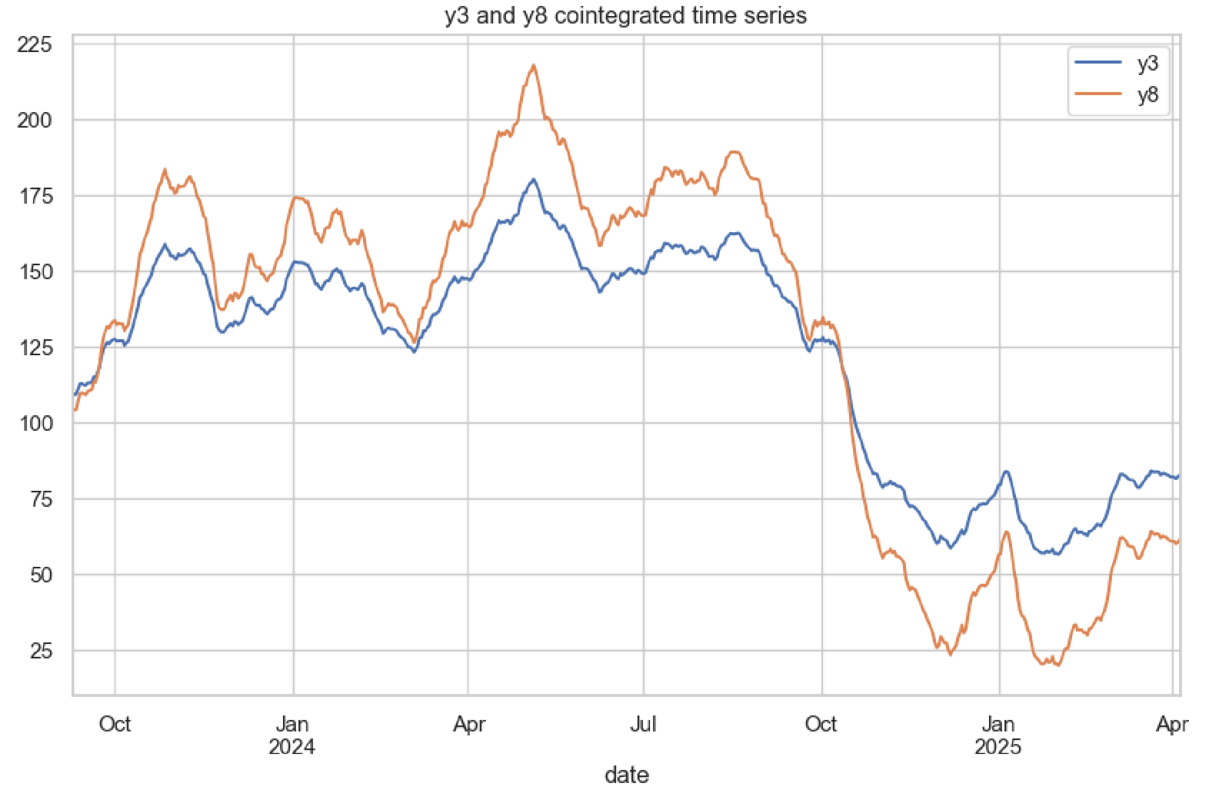


Figure 2: Visualization of variables analyzed within VECM and ARIMA framework

## 4 Development of VECM model

### 4.1 Granger Causality

Before the creation of VECM (Vector Error Correction Mechanism) model, the authors analyzed the relationship between variables, specifically whether there exists a Granger causality among variables. Firstly, we analyzed if *y3* is a Granger cause of *y8* values- the results from the excerpt below show that one fails to reject the null hypothesis of no causality at the 5% level for each lag. Hence, y3 is not a Granger cause of y8.

```
Granger Causality
number of lags (no zero) 1
ssr based F test:         F=0.3533  , p=0.5525  , df_denom=571, df_num=1
ssr based chi2 test:   chi2=0.3551  , p=0.5512  , df=1
likelihood ratio test: chi2=0.3550  , p=0.5513  , df=1
parameter F test:         F=0.3533  , p=0.5525  , df_denom=571, df_num=1

Granger Causality
number of lags (no zero) 2
ssr based F test:         F=0.0775  , p=0.9255  , df_denom=568, df_num=2
ssr based chi2 test:   chi2=0.1563  , p=0.9248  , df=2
likelihood ratio test: chi2=0.1563  , p=0.9248  , df=2
parameter F test:         F=0.0775  , p=0.9255  , df_denom=568, df_num=2

Granger Causality
number of lags (no zero) 3
ssr based F test:         F=0.1783  , p=0.9111  , df_denom=565, df_num=3
ssr based chi2 test:   chi2=0.5416  , p=0.9097  , df=3
likelihood ratio test: chi2=0.5413  , p=0.9097  , df=3
parameter F test:         F=0.1783  , p=0.9111  , df_denom=565, df_num=3

Granger Causality
number of lags (no zero) 4
ssr based F test:         F=0.1958  , p=0.9406  , df_denom=562, df_num=4
ssr based chi2 test:   chi2=0.7956  , p=0.9390  , df=4
likelihood ratio test: chi2=0.7950  , p=0.9391  , df=4
parameter F test:         F=0.1958  , p=0.9406  , df_denom=562, df_num=4

Granger Causality
number of lags (no zero) 5
ssr based F test:         F=0.2336  , p=0.9477  , df_denom=559, df_num=5
ssr based chi2 test:   chi2=1.1912  , p=0.9457  , df=5
likelihood ratio test: chi2=1.1900  , p=0.9458  , df=5
parameter F test:         F=0.2336  , p=0.9477  , df_denom=559, df_num=5
```

Table 4: Granger causality test: *y3* as a Granger cause of *y8*

After that, the authors conducted similar test for opposite direction, that is whether *y8* is a Granger cause of *y3* values- as we can see from results below, one rejects the null hypothesis of no causality at the 5% level for models with 2,3,4 lags. (whereas for 1 lag, we fail to reject the null with 5% level and reject with 10% level). Hence, y8 is a Granger cause of y3 but the relationship is not fully certain. We conclude our thoughts with a statement that there exists unidirectional (but no bidirectional) Granger causality on *y8*-*y3* line.

```
Granger Causality
number of lags (no zero) 1
ssr based F test:         F=2.8503  , p=0.0919  , df_denom=571, df_num=1
ssr based chi2 test:   chi2=2.8653  , p=0.0905  , df=1
likelihood ratio test: chi2=2.8582  , p=0.0909  , df=1
parameter F test:         F=2.8503  , p=0.0919  , df_denom=571, df_num=1

Granger Causality
number of lags (no zero) 2
ssr based F test:         F=14.1616 , p=0.0000  , df_denom=568, df_num=2
ssr based chi2 test:   chi2=28.5726 , p=0.0000  , df=2
likelihood ratio test: chi2=27.8831 , p=0.0000  , df=2
parameter F test:         F=14.1616 , p=0.0000  , df_denom=568, df_num=2

Granger Causality
number of lags (no zero) 3
ssr based F test:         F=8.0415  , p=0.0000  , df_denom=565, df_num=3
ssr based chi2 test:   chi2=24.4233 , p=0.0000  , df=3
likelihood ratio test: chi2=23.9162 , p=0.0000  , df=3
parameter F test:         F=8.0415  , p=0.0000  , df_denom=565, df_num=3

Granger Causality
number of lags (no zero) 4
ssr based F test:         F=5.9689  , p=0.0001  , df_denom=562, df_num=4
ssr based chi2 test:   chi2=24.2578 , p=0.0001  , df=4
likelihood ratio test: chi2=23.7567 , p=0.0001  , df=4
parameter F test:         F=5.9689  , p=0.0001  , df_denom=562, df_num=4

Granger Causality
number of lags (no zero) 5
ssr based F test:         F=4.5080  , p=0.0005  , df_denom=559, df_num=5
ssr based chi2 test:   chi2=22.9836 , p=0.0003  , df=5
likelihood ratio test: chi2=22.5323 , p=0.0004  , df=5
parameter F test:         F=4.5080  , p=0.0005  , df_denom=559, df_num=5
```

Table 5: Granger causality test: *y8* as a Granger cause of *y3*

#### 4.2 VECM order selection

After identifying the existence of Granger causality, the authors proceeded with VECM model development. In order not to overestimate the results nor to obtain spurious regression, the authors analyzed what would be the “optimal” order of such VECM model (which describes number of lags of the differenced variables included in the system). To do so, firstly a theoretical proper VAR model order was investigated, basing on several information criteria values (specifically: AIC, BIC, HQIC, FPE).

Lag selection results:
VAR Order Selection (* highlights the minimums)

| | AIC | BIC | FPE | HQIC |
|---|---|---|---|---|
| 0 | 3.920 | 3.935 | 50.40 | 3.926 |
| 1 | -2.737 | -2.691 | 0.06477 | -2.719 |
| 2 | -3.248 | -3.171* | 0.03886 | -3.218 |
| 3 | -3.278 | -3.170 | 0.03769 | -3.236* |
| 4 | -3.280* | -3.141 | 0.03762* | -3.226 |
| 5 | -3.274 | -3.104 | 0.03785 | -3.208 |
| 6 | -3.268 | -3.067 | 0.03808 | -3.190 |
| 7 | -3.265 | -3.033 | 0.03821 | -3.174 |
| 8 | -3.260 | -2.998 | 0.03837 | -3.158 |
| 9 | -3.252 | -2.958 | 0.03870 | -3.137 |
| 10 | -3.242 | -2.917 | 0.03909 | -3.115 |
| 11 | -3.232 | -2.876 | 0.03950 | -3.093 |
| 12 | -3.219 | -2.832 | 0.04001 | -3.068 |
| 13 | -3.217 | -2.799 | 0.04010 | -3.054 |
| 14 | -3.208 | -2.759 | 0.04046 | -3.033 |
| 15 | -3.210 | -2.731 | 0.04037 | -3.023 |

Table 6: Lag selection for VAR model

As one can see from table above and from the graph below, VAR(5) model would be the most appropriate one for selected pair of cointegrated variables (it may be more intuitive and visible on the graph: after 5th lag one can see the "stabilization" of criteria values).

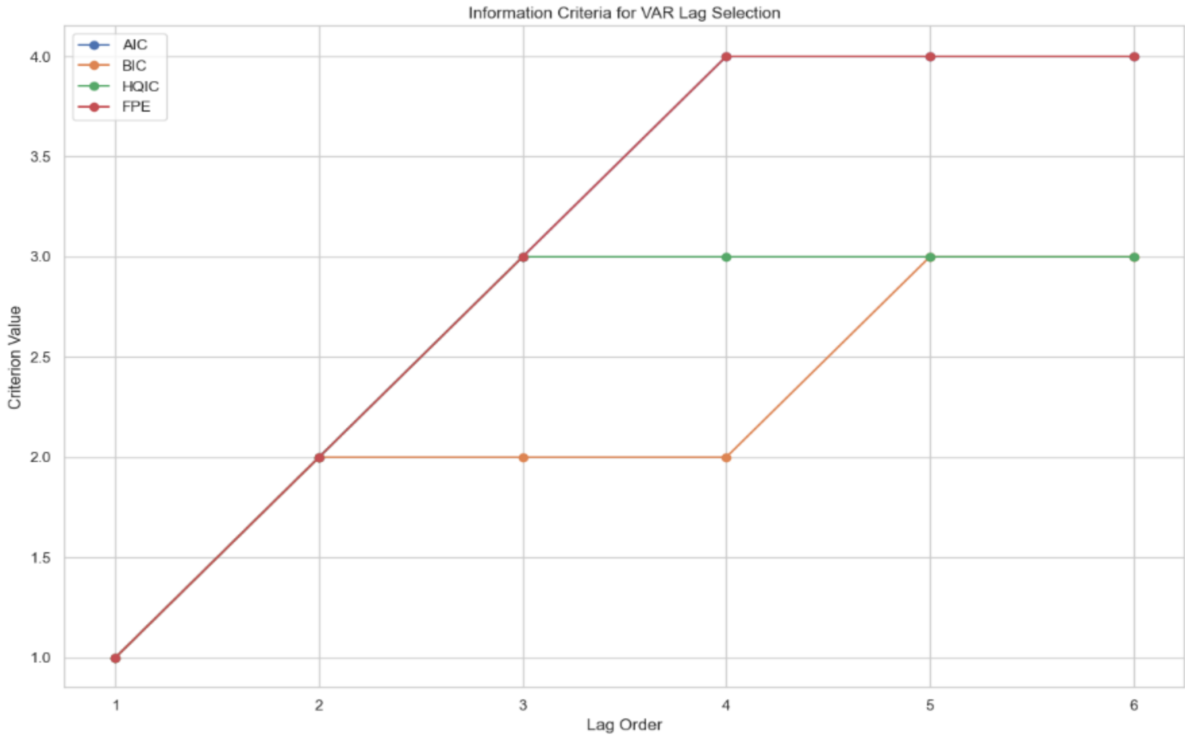


Figure 3: Graphical visualization of lag selection for VAR model

In order to reassure with selected VAR order, the authors conducted Ljung-Box test for residual autocorrelation. As one can see from the results attached below, there is no autocorrelation in residuals for each lag and variable as in each instance one fails to reject the null hypothesis stating that residuals are independently distributed (*ergo:* there is no autocorrelation among residuals).

```
Ljung-Box Test for Residual Autocorrelation:

Y3 series:
Lag 5: p-value = 0.9820
Lag 10: p-value = 0.7432
Lag 15: p-value = 0.7342
Lag 20: p-value = 0.6048

Y8 series:
Lag 5: p-value = 0.9674
Lag 10: p-value = 0.7034
Lag 15: p-value = 0.8060
Lag 20: p-value = 0.6868
```

Table 7: Ljung-Box test for VAR(5) model

As a result, the authors proceed with the VAR(5) model- this implies that potentially VECM(4) model will be most suitable for further analysis.

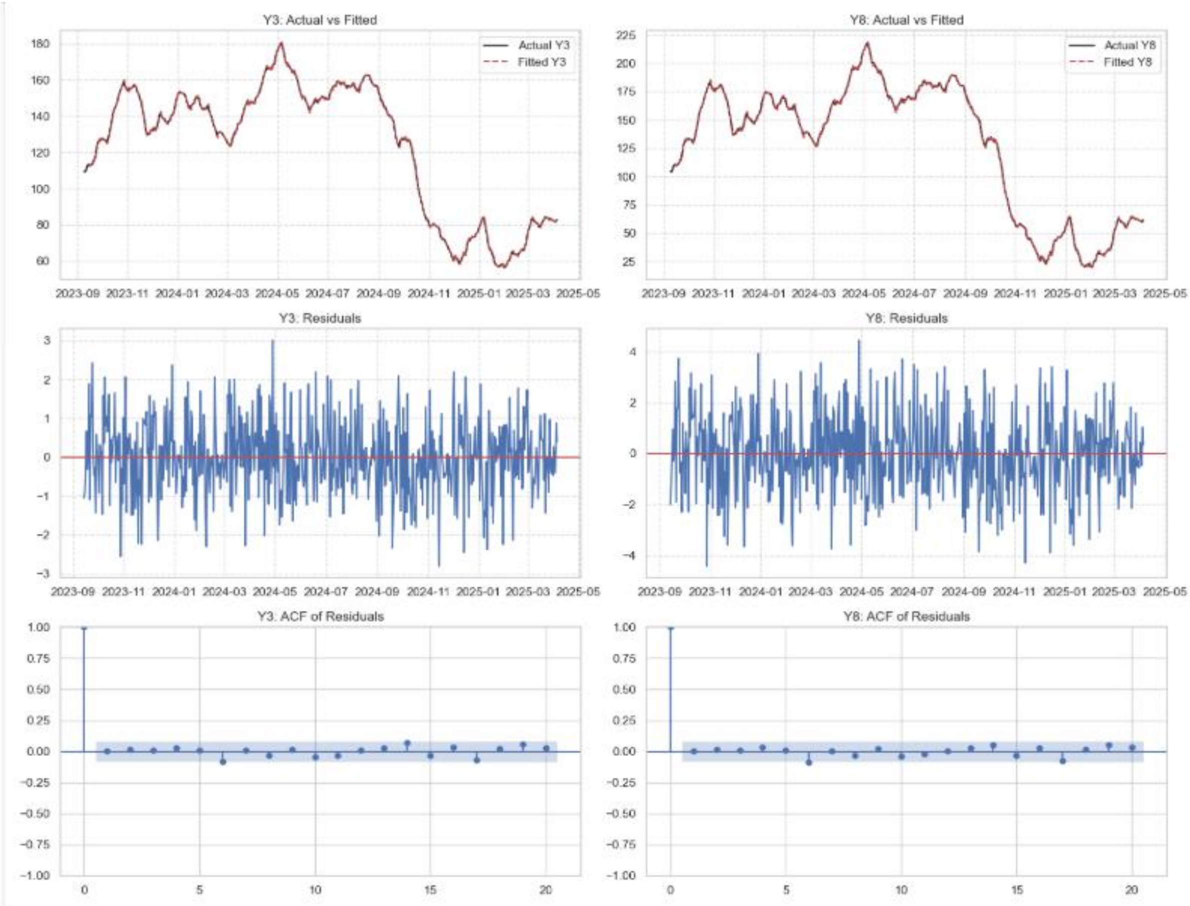


Figure 4: Graphical visualization VAR(5) model components

```
  Summary of Regression Results
==================================
Model:                         VAR
Method:                        OLS
Date:           Thu, 03, Jul, 2025
Time:                     09:49:27
--------------------------------------------------------------------
No. of Equations:         2.00000    BIC:                  -3.10783
Nobs:                     570.000    HQIC:                 -3.21012
Log likelihood:          -662.056    FPE:                 0.0377962
AIC:                     -3.27556    Det(Omega_mle):      0.0363785
--------------------------------------------------------------------
Results for equation y3
========================================================================
           coefficient       std. error           t-stat            prob
------------------------------------------------------------------------
const       -33.730835        34.505105           -0.978           0.328
L1.y3         1.603322         0.340447            4.709           0.000
L1.y8        -0.077842         0.210396           -0.370           0.711
L2.y3        -0.151042         0.347147           -0.435           0.663
L2.y8        -0.115042         0.210389           -0.547           0.585
L3.y3         0.115657         0.347499            0.333           0.739
L3.y8        -0.129069         0.210322           -0.614           0.539
L4.y3         0.062449         0.347565            0.180           0.857
L4.y8        -0.021571         0.210425           -0.103           0.918
L5.y3         0.135432         0.337195            0.402           0.688
L5.y8        -0.135861         0.209804           -0.648           0.517
========================================================================

Results for equation y8
========================================================================
           coefficient       std. error           t-stat            prob
------------------------------------------------------------------------
const      -132.749740        55.851542           -2.377           0.017
L1.y3         2.597006         0.551063            4.713           0.000
L1.y8        -0.140728         0.340557           -0.413           0.679
L2.y3        -0.179078         0.561907           -0.319           0.750
L2.y8        -0.215718         0.340545           -0.633           0.526
L3.y3         0.161797         0.562479            0.288           0.774
L3.y8        -0.211818         0.340437           -0.622           0.534
L4.y3         0.157608         0.562585            0.280           0.779
L4.y8        -0.063253         0.340604           -0.186           0.853
L5.y3         0.272074         0.545799            0.498           0.618
L5.y8        -0.250469         0.339598           -0.738           0.461
========================================================================
```

Table 8: VAR(5) model summary table

As stated in previous paragraph, authors conducted a Johansen test (in both “types”: basing on maximum eigenvalues statistic and on trace statistic) so as to analyze how many cointegrating vectors exist within the system. As we can see from the results below, there exists exactly one cointegrating vector! We obtain normalized *y8* beta coefficient of value -0.62495506: it implies that increase of *y8* value by 1 unit in long-run will result in decrease of *y3* value by calculated beta coefficient.

```
--- Interpretation (Trace Test) ---
H0: r <= 0
  Trace Statistic: 119.997
  Critical Value (95%): 15.494
  Result: Reject H0 at 5% significance level.
H0: r <= 1
  Trace Statistic: 1.222
  Critical Value (95%): 3.841
  Result: Cannot reject H0 at 5% significance level.

--- Interpretation (Max Eigenvalue Test) ---
H0: r = 0
  Max Eigenvalue Statistic: 118.775
  Critical Value (95%): 14.264
  Result: Reject H0 at 5% significance level.
H0: r = 1
  Max Eigenvalue Statistic: 1.222
  Critical Value (95%): 3.841
  Result: Cannot reject H0 at 5% significance level.
```

Table 9: Johansen test for VECM(4) system

Therefore, the authors chose VECM(4) model with cointegration matrix of rank 1 for further analysis and prediction.

```
Det. terms outside the coint. relation & lagged endog. parameters for equation y3
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
L1.y3         -0.1569      0.689     -0.228      0.820      -1.507       1.193
L1.y8          0.3986      0.429      0.929      0.353      -0.442       1.239
L2.y3         -0.3119      0.593     -0.526      0.599      -1.475       0.851
L2.y8          0.2852      0.369      0.773      0.440      -0.438       1.009
L3.y3         -0.1984      0.479     -0.415      0.678      -1.136       0.740
L3.y8          0.1571      0.298      0.527      0.598      -0.427       0.741
L4.y3         -0.1385      0.334     -0.414      0.679      -0.794       0.517
L4.y8          0.1361      0.208      0.654      0.513      -0.272       0.544
Det. terms outside the coint. relation & lagged endog. parameters for equation y8
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
L1.y3         -0.4035      1.115     -0.362      0.718      -2.589       1.782
L1.y8          0.7365      0.694      1.061      0.289      -0.625       2.098
L2.y3         -0.5888      0.960     -0.613      0.540      -2.471       1.293
L2.y8          0.5234      0.597      0.876      0.381      -0.647       1.694
L3.y3         -0.4305      0.775     -0.556      0.578      -1.949       1.088
L3.y8          0.3131      0.482      0.650      0.516      -0.631       1.258
L4.y3         -0.2769      0.541     -0.512      0.609      -1.337       0.784
L4.y8          0.2509      0.337      0.745      0.456      -0.409       0.911
                 Loading coefficients (alpha) for equation y3
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
ec1            0.7627      0.775      0.984      0.325      -0.756       2.281
                 Loading coefficients (alpha) for equation y8
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
ec1            3.0045      1.254      2.396      0.017       0.547       5.462
          Cointegration relations for loading-coefficients-column 1
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
beta.1         1.0000          0          0      0.000       1.000       1.000
beta.2        -0.6250   7.97e-05  -7841.774      0.000      -0.625      -0.625
const        -44.1565      0.011  -3952.458      0.000     -44.178     -44.135
==============================================================================
```

Table 10: VECM(4) model summary table

For further analysis of developed framework, Impulse Response Functions (for shock's impact on system) and Forecast Error Variance Decomposition (for final confirmation of *y8→y3* Granger Causality existence) was created and analyzed.

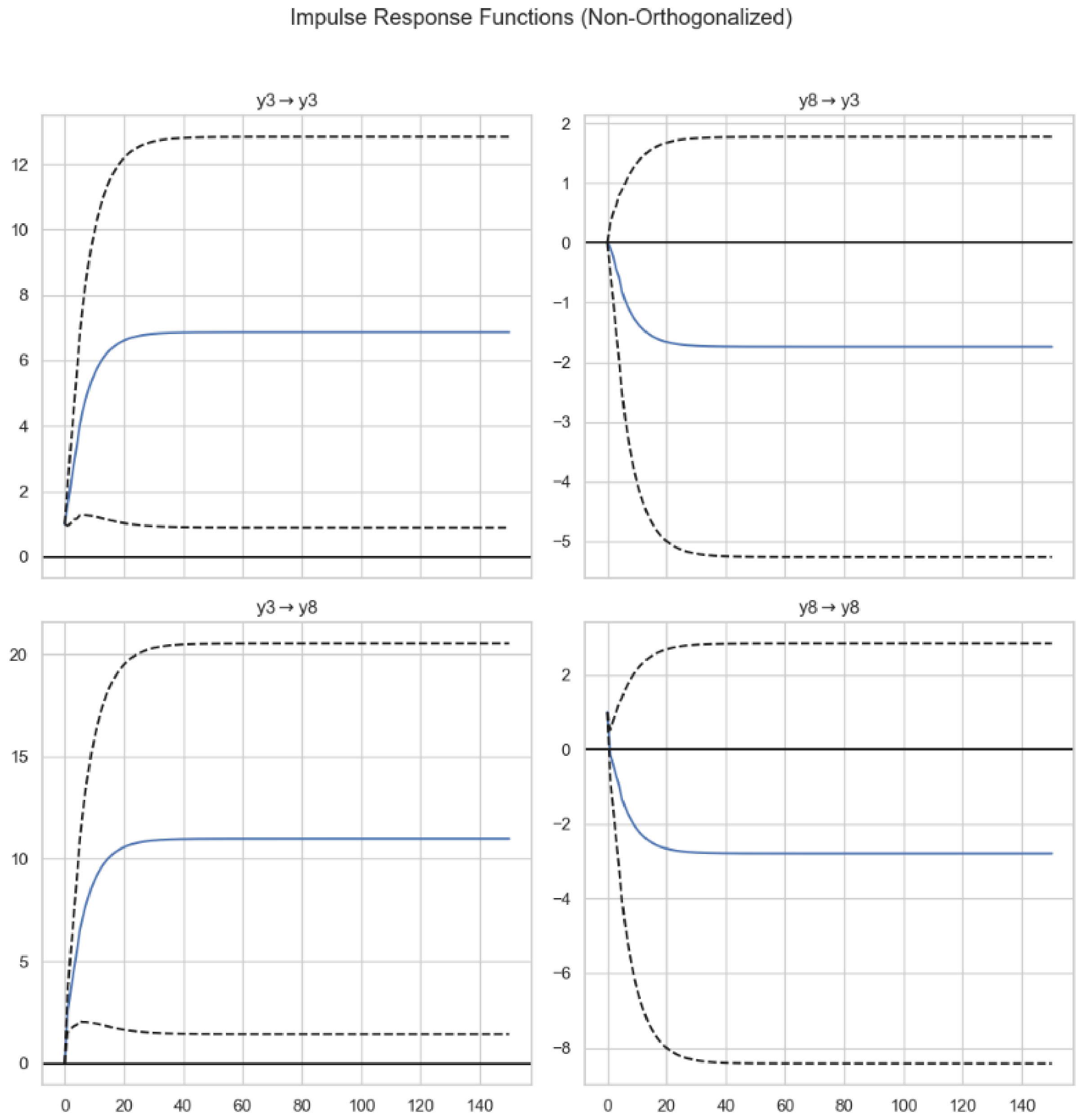


Figure 5: IRF functions (non-orthogonalized)

When it comes to interpretation of Impulse Response Function, one can see that firstly for *y3→y3* graph there is initial increase up to 25 days and afterward slow decrease toward equilibrium, secondly for *y3→y8* graph shock has bigger initial impact on *y8* but as in previous case impacted variable is slowly converging toward 0, next or *y8→y3* graph there is no first period impact of the *y8* on *y3* but we can observe negative impact on *y3* and which converges back to 0 and last but not least for *y8→y8* graph there is quick small jump at the beginning but as in previous case we are observing negative impact on first periods and convergence starting prior 25 period.

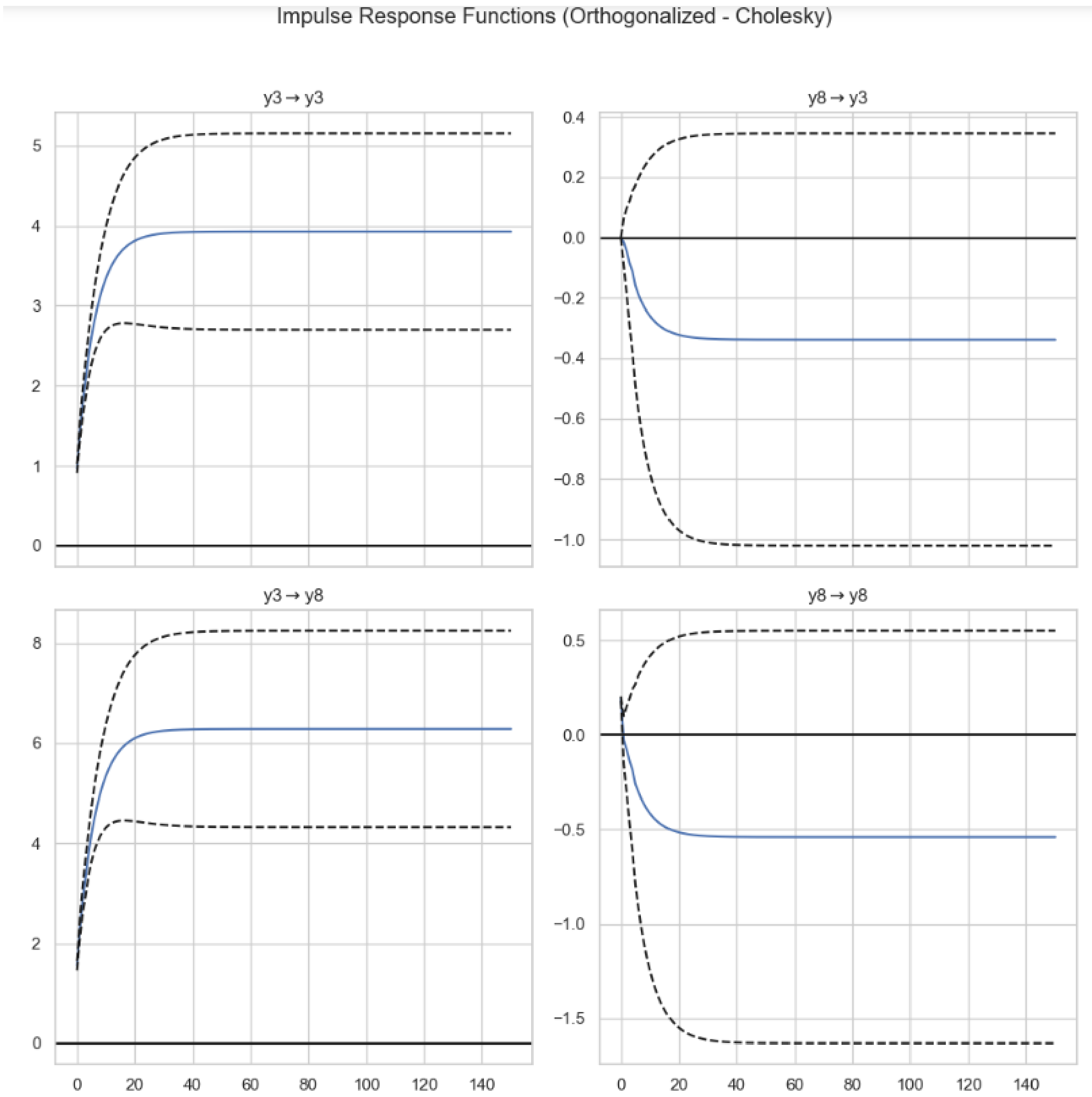


Figure 6: IRF functions (orthogonalized by Cholesky decomposition)

One can also look at Forecast Error Variance Decomposition provided below. We can see that variance of forecast error terms are explained mainly by changes in *y3* for both variables. That confirms our main previous observations that shocks in *y3* have impact on *y8* but no other way around (no bidirectional causality).

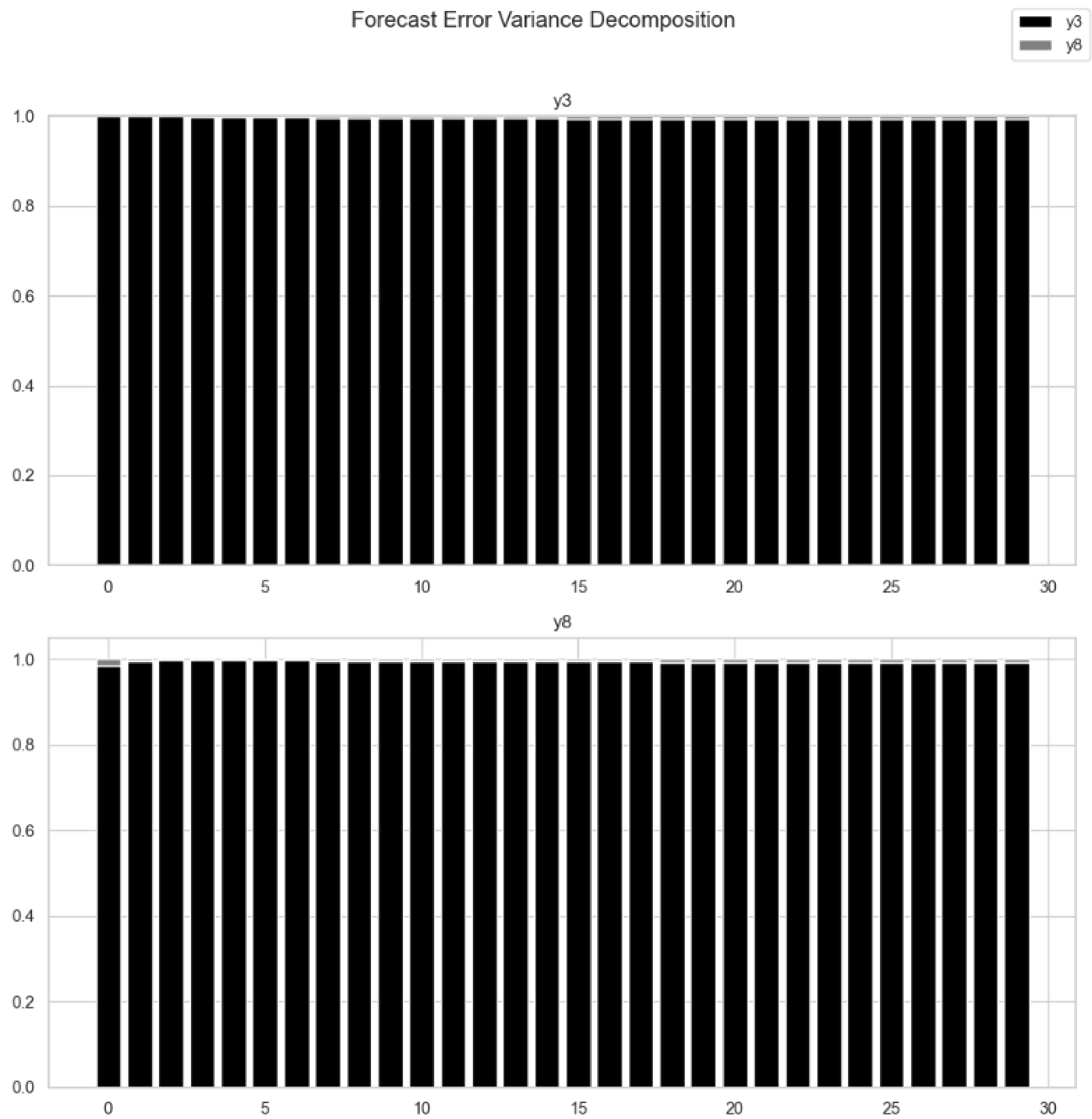

Figure 7: FEVD among y3-y8 variables

At the end of this section of analysis, Ljung-Box test for residuals of model was conducted so as to see whether there is no autocorrelation among residuals. Indeed, basing on the results one fail to reject null hypothesis of no autocorrelation and therefore, selected model is appropriate and the authors used it for the prediction and comparison with out-of-sample values.

```
Residuals (first 5 rows):
                      y3            y8
date
2023-09-14 -1.0597422347 -2.0089833968
2023-09-15 -0.8904504007 -1.2000212456
2023-09-16 -0.1786331779 -0.2323526760
2023-09-17  0.6779125796  0.9693181159
2023-09-18 -0.3187428485 -0.3086562064

Ljung-Box Test for Serial Correlation in Residuals:

y8 Residuals:
          lb_stat    lb_pvalue
5   1.0394042393 0.9593253466
10  7.4171412648 0.6855640975
15 10.3252413891 0.7987979760
20 16.4185595499 0.6903416518

y3 Residuals:
          lb_stat    lb_pvalue
5   0.8105597745 0.9763549240
10  7.0009853942 0.7253519228
15 11.3295182285 0.7289048283
20 17.6170731479 0.6126173317
```

Table 11: Ljung-Box test for VECM(4) model residuals

### 4.3 Forecast with selected VECM model

After selecting VECM(4) model, the authors computed forecast for 25 daily observations ahead and compared it with the OOS values.

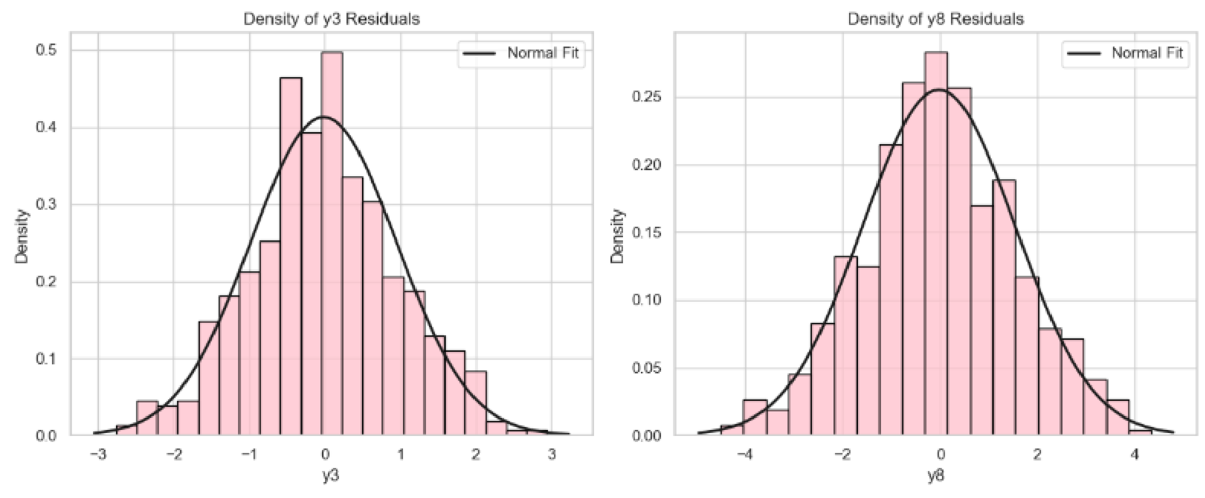


Figure 8: Distribution of residuals within VECM(4) model

Before presenting the results of prediction, here are the plots and Jarque-Bera test results provided respectively- by interpreting both of them one can say that the residuals come from normal standard distribution, a fact that facilitates prediction computation and interpretation.

```
Jarque-Bera Normality Test for Residuals:
y3 Residuals: Statistic=0.188, p-value=0.910
y8 Residuals: Statistic=0.613, p-value=0.736

Conclusion at alpha=0.05:
 - Cannot reject normality for y3 residuals.
 - Cannot reject normality for y8 residuals.
```

Table 12: Jarque-Bera test for VECM(4) model residuals

Below we provided the graphs for bivariate VECM model prediction for the identified pair. As we can see, computed confidence interval “covers” out-of-sample values (fact in favour of stating that the forecast is correctly calculated)- however we can see that the forecasted values (depicted with blue line on graph) seems to alling into “stable” horizontal line, a fact that one should be aware of when assessing the quality of forecast. At the end of the report, proper *ex-post* measures will be provided (so as to compare the results with ARIMA modelling and analogous prediction).

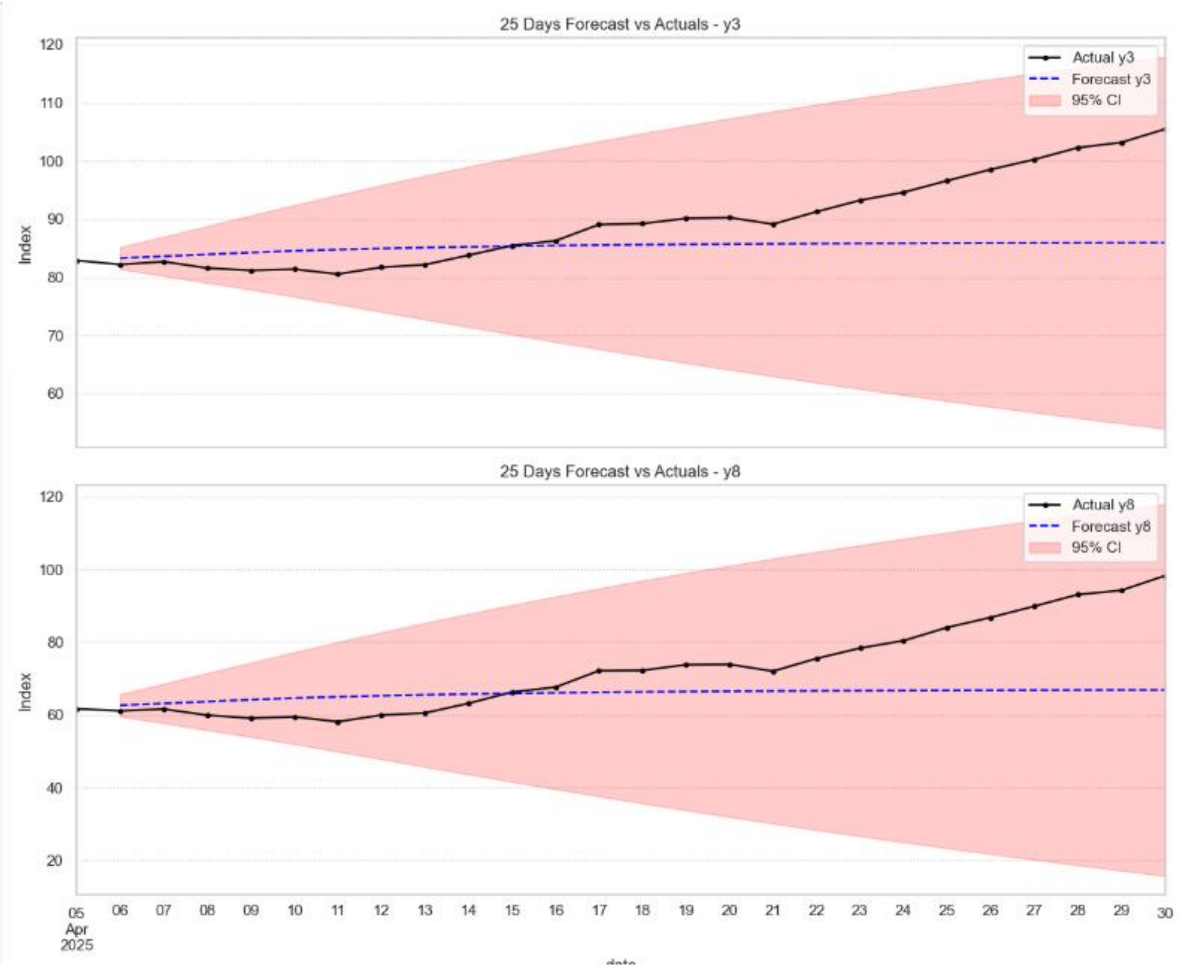


Figure 9: 25-days ahead forecast for y3 and y8 using VECM(4) model

## 5 Development of ARIMA model

### 5.1 Box-Jenkins procedure for *y3* variable

Now we will present our findings concerning ARIMA modelling for selected variables- to do so, Box-Jenkins procedure was performed for selecting optimal 2 univariate ARIMA(p,d,q) models (which consist of: Identification of model's order→ Estimation of model parameters→ Diagnostics with proper statistical tests→ Forecasting for n=25 period ahead).

For first step of described procedure (that is Identification of [p,d,q] values), the authors utilize ACF (Autocorrelation Function) and PACF (Partial Autocorrelation Function) to determine correspondingly right MA(q) order and right AR(p) order. Referring to *"Finding a cointegrated variables"* chapter, we already have order of integration as we have shown that series are both non-stationary but their differences are stationary implying integration of order 1.

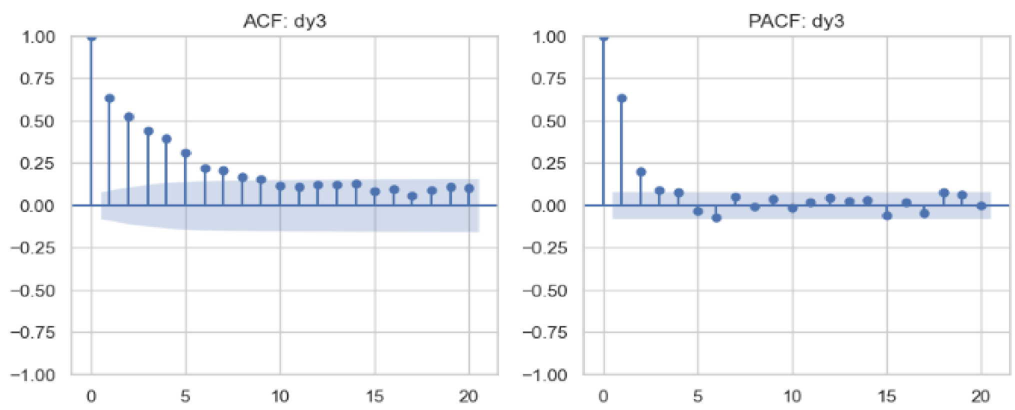


Figure 10: ACF and PACF graphs for y3 variable

We observe quite interesting behavior - for *y3* variable, ACF is decaying exponentially, which would be expected for simple AR(2) models. PACF on the other hand oscillates around 0, but also the greatest spike, which at the same time is significantly greater than other spikes is the first two lags. There is a possibility that the right model will be just AR(2) for differenced series, so ARIMA(2,1,0), but one should also observe that 6th lag could be statistically significant for model prediction. Authors proceeded with ARIMA(2,1,0) model further on.

```
                               SARIMAX Results
==============================================================================
Dep. Variable:                      y   No. Observations:                  575
Model:               ARIMA([2], 1, 0)   Log Likelihood                -868.233
Date:                Tue, 10 Jun 2025   AIC                           1740.466
Time:                        09:44:10   BIC                           1749.171
Sample:                             0   HQIC                          1743.861
                                - 575
Covariance Type:                  opg
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
ar.L2          0.5213      0.035     15.038      0.000       0.453       0.589
sigma2         1.2047      0.074     16.226      0.000       1.059       1.350
===================================================================================
Ljung-Box (L1) (Q):                  66.01   Jarque-Bera (JB):                 1.32
Prob(Q):                              0.00   Prob(JB):                         0.52
Heteroskedasticity (H):               0.94   Skew:                             0.09
Prob(H) (two-sided):                  0.68   Kurtosis:                         2.84
===================================================================================
```

Table 13: ARIMA(2,1,0) model summary table for *y3* variable

In order to confirm whether chosen model specification is correct, the authors conduct Ljung-Box test so as to see whether there is no autocorrelation among residuals. Indeed, for each lag one fails to reject null hypothesis about no autocorrelation of residuals in the model.

Ljung-Box Test for Residual Autocorrelation:

| | lb_stat | lb_pvalue |
|---|---|---|
| 5 | 0.8119677315 | 0.9762637318 |
| 10 | 0.9877312986 | 0.9998373676 |
| 15 | 1.8301737787 | 0.9999835869 |
| 20 | 1.9711284503 | 0.9999999024 |
| 25 | 2.0583105491 | 0.9999999997 |

Table 14: Ljung-Box test for ARIMA(2,1,0) univariate model for *y3* variable

This fact can be also confirmed with analogous ACF and PACF plots for residuals- there is no autocorrelations in model residuals, and as usually we prefer simple models we could just take this one and move forward with the analysis as this model is basically the simplest one we can get. However, we can't really say that this model is certainly the best one without some comparison. It is therefore useful to study also different model specifications and compare them on information criteria basis (lower is better).

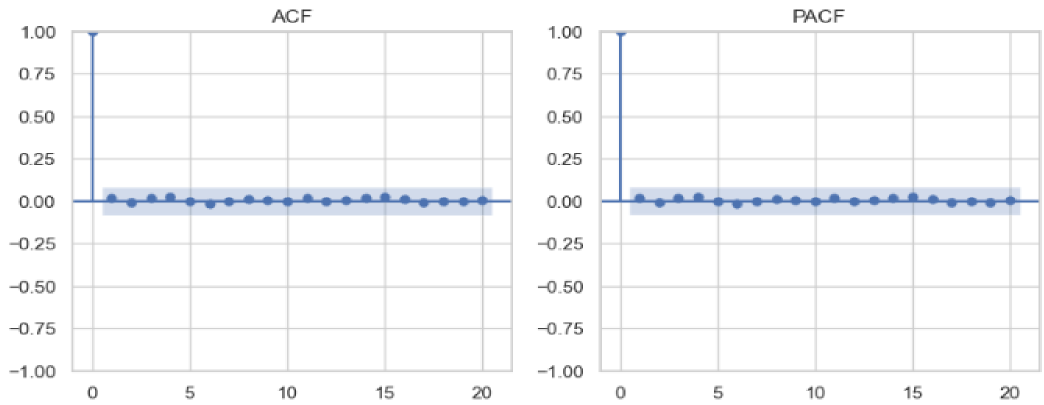


Figure 11: ACF and PACF graphs for ARIMA(2,1,0) univariate model for y3 variable

To finalize Box-Jenkins procedure, below is the graph of obtained predictions and comparison to OOS factual values. The main issue here is the fact that the confidence interval (depicted with red "zone") doesn't "cover" all OOS values, unlike in VECM(4) prediction- this may suggest lower accuracy and this statement will be reevaluated in the *Summary* chapter of this report.

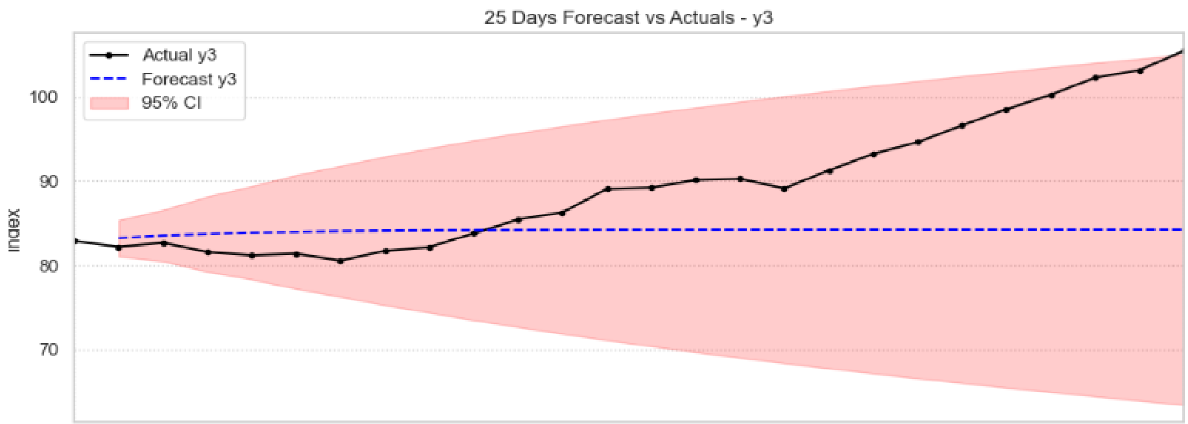


Figure 12: 25-days ahead forecast for y3 using ARIMA(2,1,0) model

**5.2 Box-Jenkins procedure for *y8* variable**

Analogous analysis and Box-Jenkins procedure from previous subchapter was applied to *y8* estimation and prediction.

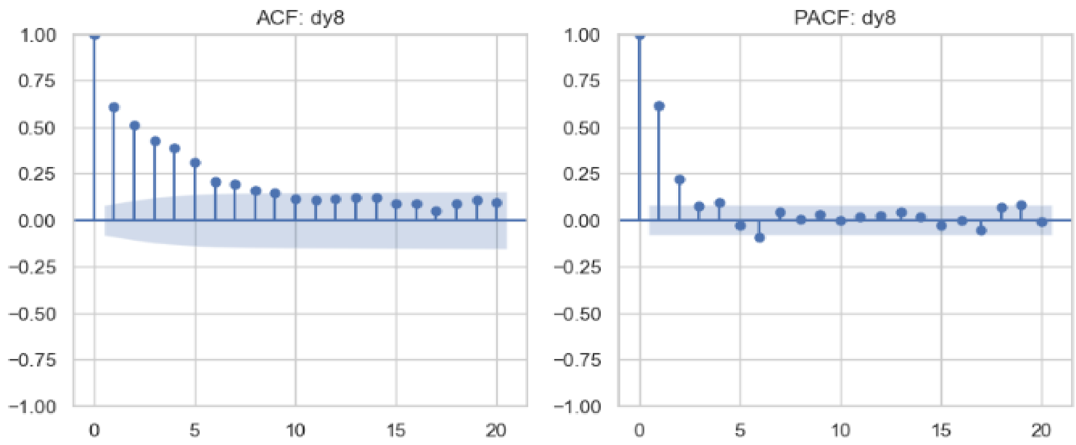


Figure 13: ACF and PACF graphs for y8 variable

The results of analysis are actually the same as for *y3* variable with minor discrepancies- for *y8* also ARIMA(2,1,0) model seems to be most suitable, however for instance on ACF and PACF graph attached above one can see that lag no. 17$^{th}$ is statistically significant. Nevertheless, authors carried on with ARIMA(2,1,0) model estimation and evaluation.

```
                               SARIMAX Results
==============================================================================
Dep. Variable:                      y   No. Observations:                  575
Model:               ARIMA([2], 1, 0)   Log Likelihood               -1148.805
Date:                Tue, 10 Jun 2025   AIC                           2301.611
Time:                        09:44:31   BIC                           2310.316
Sample:                             0   HQIC                          2305.006
                                - 575
Covariance Type:                  opg
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
ar.L2          0.5115      0.035     14.778      0.000       0.444       0.579
sigma2         3.2023      0.197     16.235      0.000       2.816       3.589
===================================================================================
Ljung-Box (L1) (Q):                  62.24   Jarque-Bera (JB):                 0.85
Prob(Q):                              0.00   Prob(JB):                         0.65
Heteroskedasticity (H):               0.94   Skew:                             0.05
Prob(H) (two-sided):                  0.67   Kurtosis:                         2.84
===================================================================================
```

Table 15: ARIMA(2,1,0) model summary table for *y8* variable

Again, by analyzing Ljung-Box test output one can assume that there is no autocorrelation among the residuals in the model as one fails to reject null hypothesis for each subsequent lag period.

```
Ljung-Box Test for Residual Autocorrelation:
        lb_stat     lb_pvalue
5   3.3508856262 0.6460605790
10  3.7183441239 0.9591597938
15  5.8681659083 0.9818781024
20  6.3037660221 0.9984201650
25  6.4730838352 0.9999286190
```

Table 16: Ljung-Box test for ARIMA(2,1,0) univariate model for *y8* variable

The fact is also confirmed with visual interpretation of ACF and PACF graphs for residuals for created ARIMA(2,1,0) univariate model.

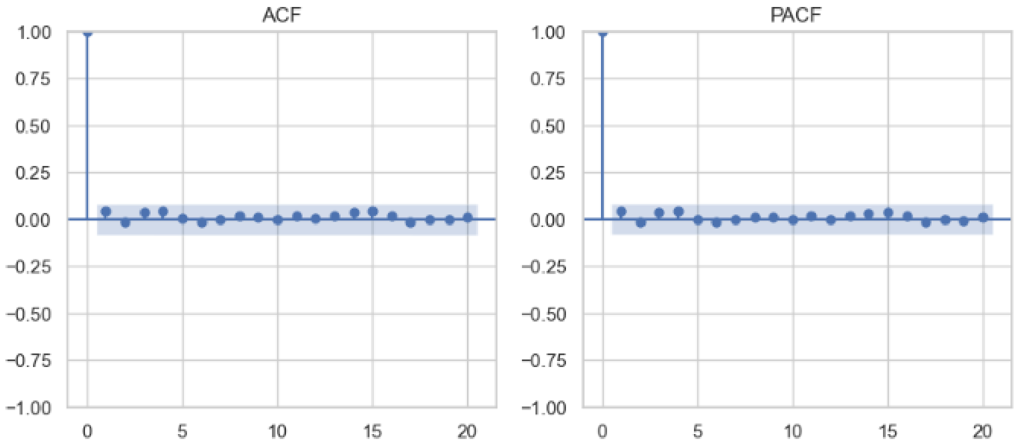


Figure 14: ACF and PACF graphs for ARIMA(2,1,0) univariate model for y3 variable

To finalize Box-Jenkins procedure, similar graph for prediction is provided below. To add up to conclusion already stated in previous chapter, one should be cautious for "stable" prediction that aligns into horizontal line- a fact that was already brought up during VECM modelling and prediction.

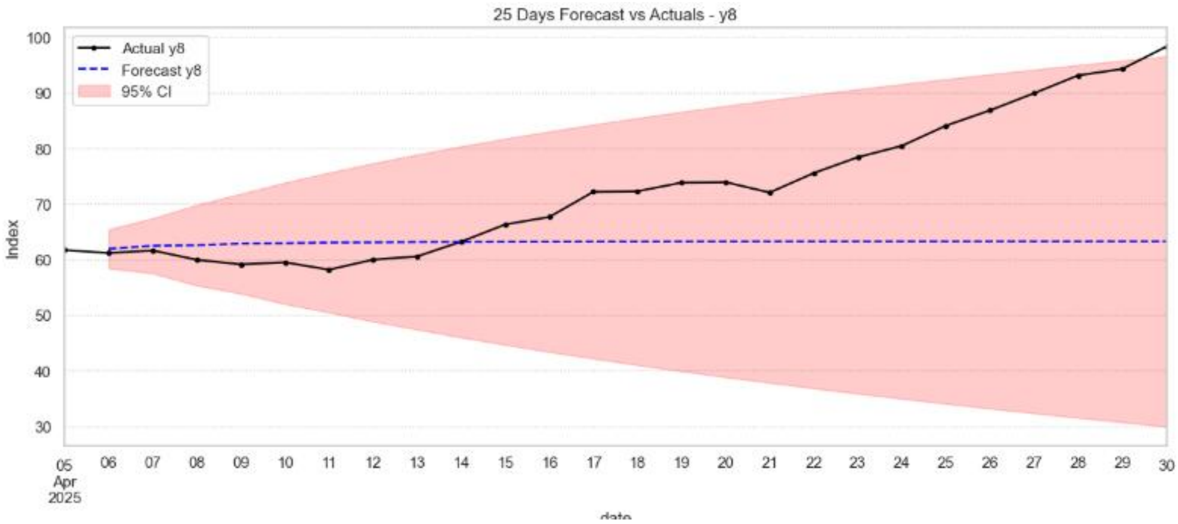


Figure 15: 25-days ahead forecast for y8 using ARIMA(2,1,0) model

## 6 Models comparison and final conclusions

Now it’s the time to compare and evaluate estimated models- below are the prediction plots from both models both for *y3* and *y8* variables. As one can see, the predictions are very close to each other and similar. However the confidence interval for ARIMA(2,1,) have narrowed range but doesn’t “cover” all OOS observations (specifically last 25th observation), therefore the accuracy of ARIMA(2,1,0) should be lower.

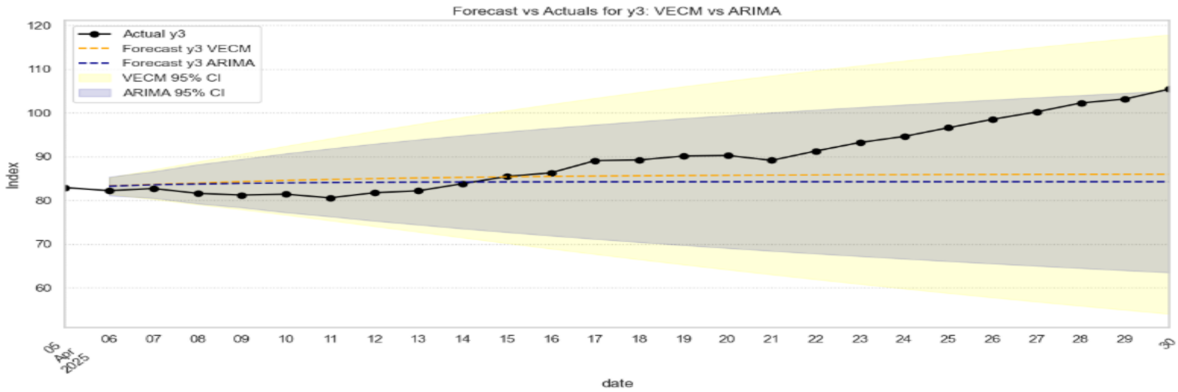


Figure 16: 25-days ahead forecast for y3 obtained by VECM(4) and ARIMA(2,1,0)

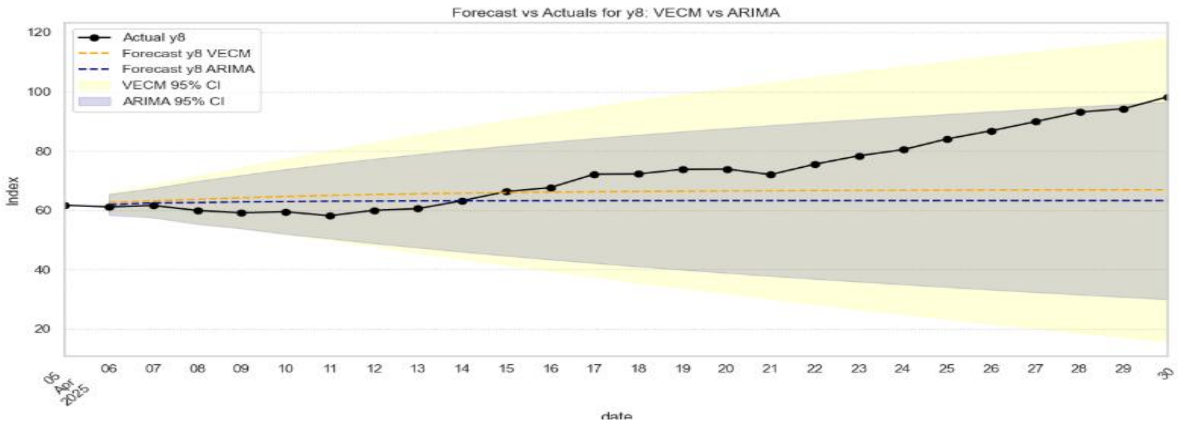


Figure 17: 25-days ahead forecast for y8 obtained by VECM(4) and ARIMA(2,1,0)

Indeed, as we can see from the table below, majority of accuracy metrics (only except Mean Absolute Error metric) suggest that VECM(4) model is a “better one” as the values of error seems to be lower for each variable. As a final conclusion to this report, we would say that VECM(4) model outperformed ARIMA(2,1,0) model both for cointegrated *y3* and *y8* variables.

Combined Forecast Accuracy Metrics:

| | Metric | ARIMA_y3 | VECM_y3 | ARIMA_y8 | VECM_y8 |
|---|---|---|---|---|---|
| 0 | MAE | 1.9868000000 | 6.2102000000 | 2.4830000000 | 10.0166000000 |
| 1 | MSE | 86.9272000000 | 69.1717000000 | 240.7672000000 | 177.9095000000 |
| 2 | RMSE | 9.3235000000 | 8.3170000000 | 15.5167000000 | 13.3383000000 |
| 3 | MAPE (%) | 7.3075000000 | 6.5162000000 | 14.1275000000 | 12.4201000000 |
| 4 | AMAPE (%) | 7.7580000000 | 6.8520000000 | 15.8588000000 | 13.5389000000 |

Table 17: *Ex-post* accuracy metrics for each model

## 7 Reference

All the results and Python code is publicly available on GitHub, under the link:
https://github.com/MrMajszi/TSA_Project_2025

**List of tables**



**List of figures**