Time Series Modelling

Technical Documentation · Module 2

ARIMA, ARIMAX, GARCH, VAR

English

Wolfgang Boehme

wolfgang.boehme@economic-modeling.de

www.economic-modeling.de

25 October 2026

The documentation on this website was developed with the support of the AI Claude (Anthropic).

0 Table of Contents

1 Overview and Usage **2**
1.1 Data Format . . . 2
1.2 Navigation . . . 2
1.3 Recommended Workflow . . . 2

2 Stationarity and Unit Root Tests **3**
2.1 Augmented Dickey-Fuller (ADF) Test . . . 3
2.2 Usage in the Calculator . . . 3

3 Ljung-Box and ARCH-LM Pre-Tests **4**
3.1 Ljung-Box Test . . . 4
3.2 ARCH-LM Test (Engle 1982) . . . 4
3.3 Usage in the Calculator . . . 4

4 ARIMA Models **4**
4.1 Identification via ACF and PACF . . . 5
4.2 Estimation (OLS Approximation) . . . 5
4.3 Residual Diagnostics . . . 5
4.4 Auto-ARIMA . . . 5
4.5 Usage in the Calculator . . . 5

5 ARIMAX – ARIMA with Exogenous Variables **6**
5.1 Model Comparison: ARIMA vs. ARIMAX . . . 6
5.2 Usage in the Calculator . . . 6

6 GARCH Models **6**
6.1 Value at Risk (VaR) . . . 7
6.2 Volatility Half-Life . . . 7
6.3 Usage in the Calculator . . . 7

7 Forecasting and Confidence Intervals **7**
7.1 Point Forecast . . . 7
7.2 Confidence Intervals . . . 8

8 Rolling Forecast and Out-of-Sample Performance **8**
8.1 Methodology . . . 8
8.2 Performance Metrics . . . 8

9 Model Comparison: AIC and BIC **9**

10 Cointegration (Engle-Granger) **9**

10.1 Engle-Granger Two-Step Procedure 9

11 Granger Causality **10**

11.1 F-Test 10

12 VAR Models **10**

12.1 Lag Selection 11

12.2 Usage in the Calculator 11

13 Worked Example **11**

13.1 Data Description 11

13.2 Step 1: Unit Root Test (ADF) 11

13.3 Step 2: Estimating ARIMA(1,1,0) 12

13.4 Step 3: ARIMAX(1,1,0) with ECB Interest Rate 12

13.5 Step 4: GARCH(1,1) on Log-Changes 13

13.6 Step 5: Forecast (3 Steps Ahead) 13

13.7 Step 6: Granger Causality (Lag 2) 14

13.8 Step 7: Cointegration (Engle-Granger) 14

13.9 Step 8: VAR(1) Model 14

1 Overview and Usage

The **Time Series Modelling Calculator** is a self-contained HTML application that runs in the browser without installation. It extends the Time Series Analyser (Module 1) with advanced modelling methods.

1.1 Data Format

The application expects a **CSV file** with the following structure:

```
Date,Variable_1,Variable_2,...,Variable_k
2020-01,1.5,0.0,...
2020-02,1.7,0.0,...
```

- Column 1: Timestamp (date, month, quarter – any format)
- Columns 2+: Numeric time series values (decimal point)
- Delimiter: Comma
- Header row: Required (first row)

1.2 Navigation

The application uses a **sidebar** with four groups:

Group	Modules
Prerequisites	Unit root (ADF), Ljung-Box / ARCH-LM
Univariate	ARIMA, ARIMAX, GARCH
Forecast & Evaluation	Forecast, Rolling Forecast, Model selection (AIC/BIC)
Multivariate	Cointegration, Granger causality, VAR

1.3 Recommended Workflow

1. **Load data**: Upload CSV or use the sample data.
2. **Check unit roots**: ADF test → determine integration order d.
3. **Pre-tests**: Ljung-Box (is ARIMA needed?) and ARCH-LM (is GARCH needed?).
4. **Model selection**: AIC/BIC grid search → choose optimal (p, d, q).
5. **Estimate ARIMA** (optionally with Auto-ARIMA); check residual diagnostics.
6. **ARIMAX**: Add an exogenous variable and check whether AIC improves.
7. **GARCH**: If ARCH effects are present, estimate a volatility model.

8. **Forecast**: Generate a forward forecast with confidence band.
9. **Rolling Forecast**: Evaluate out-of-sample performance via RMSE, MAE, MAPE.
10. **Multivariate**: Cointegration, Granger and VAR for multiple time series.

2 Stationarity and Unit Root Tests

Definition 2.1 (Stationarity). A time series $\{Y_t\}$ is called *weakly stationary* if:

$$\mathrm{E}[Y_t] = \mu \quad \text{(constant mean)}, \tag{1}$$
$$\mathrm{Var}(Y_t) = \sigma^2 < \infty \quad \text{(finite, constant variance)}, \tag{2}$$
$$\mathrm{Cov}(Y_t, Y_{t-h}) = \gamma(h) \quad \text{(depends only on lag } h\text{, not on } t\text{)}. \tag{3}$$

Non-stationary series have a *stochastic trend* (unit root, $I(1)$) or a deterministic trend. ARIMA models require stationarity after d-fold differencing.

2.1 Augmented Dickey-Fuller (ADF) Test

ADF Test Equation

$$\Delta Y_t = \alpha + \beta\, Y_{t-1} + \sum_{i=1}^{k} \delta_i\, \Delta Y_{t-i} + \varepsilon_t \tag{4}$$

- **Null hypothesis** H_0: $\beta = 0$ – unit root, $Y_t \sim I(1)$.
- **Alternative hypothesis** H_1: $\beta < 0$ – stationary.
- Test statistic: $\tau = \hat{\beta}/\mathrm{SE}(\hat{\beta})$ (left-tailed).

Significance level	Critical value	Decision ($\tau <$ crit.)
1 %	−3.43	Stationary (highly significant)
5 %	−2.86	Stationary
10 %	−2.57	Weakly stationary

2.2 Usage in the Calculator

Step 1: Select the time series in the selector bar.
Step 2: Sidebar → *Unit Root* → click "Calculate Tests".
Step 3: The recommended d is automatically transferred to ARIMA, ARIMAX and Rolling Forecast.

3 Ljung-Box and ARCH-LM Pre-Tests

3.1 Ljung-Box Test

Tests whether the data (or residuals) contain significant autocorrelations.

Ljung-Box Statistic

$$Q_{LB} = n(n+2)\sum_{k=1}^{m}\frac{\hat{\rho}_k^2}{n-k} \sim \chi^2(m) \tag{5}$$

where $\hat{\rho}_k$ is the sample autocorrelation at lag k and n is the sample size. H_0: No autocorrelation. Reject when $Q_{LB} > \chi^2_{m;\,0.95}$.

3.2 ARCH-LM Test (Engle 1982)

Tests for conditional heteroskedasticity (ARCH effects). Basis for deciding whether GARCH is appropriate.

ARCH-LM Test Procedure

1. Compute residuals $\hat{\varepsilon}_t$ from an AR model (or from raw data).
2. Auxiliary regression: $\hat{\varepsilon}_t^2 = \gamma_0 + \gamma_1\hat{\varepsilon}_{t-1}^2 + \cdots + \gamma_q\hat{\varepsilon}_{t-q}^2 + u_t$.
3. Test statistic: $LM = n \cdot R^2 \sim \chi^2(q)$.

H_0: No ARCH effects (constant variance). Reject when $LM > \chi^2_{q;\,0.95} = 3.84$ (for $q = 1$).

3.3 Usage in the Calculator

Sidebar → *Ljung-Box / ARCH-LM* → select input (original, differenced, log-returns) → click "Calculate Tests".
Result: Automatic recommendation of whether ARIMA and/or GARCH is needed.

4 ARIMA Models

Definition 4.1 (ARIMA(p,d,q))**.** A time series Y_t follows an ARIMA(p,d,q) process if $W_t = \Delta^d Y_t$ obeys a stationary ARMA(p,q) process:

$$\boxed{W_t = c + \sum_{i=1}^{p}\phi_i\, W_{t-i} + \varepsilon_t + \sum_{j=1}^{q}\theta_j\,\varepsilon_{t-j}} \tag{6}$$

with $\varepsilon_t \sim \mathrm{WN}(0,\sigma^2)$.

Parameter	Name	Meaning
p	AR order	Number of autoregressive lags
d	Integration order	Number of differences needed for stationarity
q	MA order	Number of moving-average terms
ϕ_i	AR coefficients	Weights on lagged values
θ_j	MA coefficients	Weights on lagged error terms
σ^2	Error variance	Variance of the white noise

4.1 Identification via ACF and PACF

Model	ACF	PACF
AR(p)	Exponential decay	Cuts off after lag p
MA(q)	Cuts off after lag q	Exponential decay
ARMA(p, q)	Exponential decay	Exponential decay

4.2 Estimation (OLS Approximation)

The calculator estimates the AR part via OLS on the differenced series. The MA part is approximated in a subsequent auxiliary regression on the residuals:

$$W_t \approx c + \phi_1 W_{t-1} + \cdots + \phi_p W_{t-p} + \theta_1 \hat{\varepsilon}_{t-1} + \cdots + \theta_q \hat{\varepsilon}_{t-q} \tag{7}$$

4.3 Residual Diagnostics

- **ACF of residuals**: All bars should lie within the 95% confidence band $\pm 1.96/\sqrt{n}$.
- **Ljung-Box on residuals**: $Q_{LB} < \chi^2_{m;\,0.95}$ (no remaining autocorrelation).
- **Normality**: Histogram approximately bell-shaped; skewness ≈ 0, excess kurtosis ≈ 0.

4.4 Auto-ARIMA

The calculator performs a grid search over $p \in \{0, \ldots, 3\}$, $q \in \{0, \ldots, 3\}$ and selects the model with the lowest AIC:

$$\text{AIC} = -2\,\hat{\ell} + 2k, \quad k = p + q + d + 1 \tag{8}$$

4.5 Usage in the Calculator

1. Sidebar → *ARIMA* → set p, d, q (or click "Auto-ARIMA").
2. "Estimate" → equation display, goodness-of-fit metrics, fit chart and residual diagnostics appear.
3. Check the residual ACF: red bars outside the band ⇒ increase p or q.

5 ARIMAX – ARIMA with Exogenous Variables

Definition 5.1 (ARIMAX(p, d, q))**.** Extension of the ARIMA model by a regression term $\beta\, X_t$:

$$\boxed{\Delta^d Y_t = c + \sum_{i=1}^{p} \phi_i\, \Delta^d Y_{t-i} + \sum_{j=1}^{q} \theta_j\, \varepsilon_{t-j} + \beta\, \Delta^d X_t + \varepsilon_t} \tag{9}$$

- X_t: Exogenous (explanatory) variable – also differenced d times.
- β: Effect of a unit change in $\Delta^d X_t$ on $\Delta^d Y_t$.
- **Applications**: Inflation with interest rate, GDP with government spending, sales with advertising budget.

5.1 Model Comparison: ARIMA vs. ARIMAX

The calculator automatically computes the AIC improvement:

$$\Delta\text{AIC} = \text{AIC}_{\text{ARIMA}} - \text{AIC}_{\text{ARIMAX}} \tag{10}$$

$\Delta\text{AIC} > 2$: the exogenous variable contributes significantly to explanatory power.

5.2 Usage in the Calculator

1. Sidebar $\rightarrow$ *ARIMAX* $\rightarrow$ select p, d, q and exogenous variable X.
2. "Estimate" $\rightarrow$ equation with β coefficient, ΔAIC comparison and 4 charts appear.
Note: Y and X must have the same length; the shorter series is automatically trimmed.

6 GARCH Models

Financial returns often exhibit *volatility clustering*: periods of high volatility tend to follow high volatility, and calm periods follow calm periods. This violates the assumption of constant variance.

Definition 6.1 (GARCH(p, q))**.**

$$Y_t = \mu + \varepsilon_t, \quad \varepsilon_t = \sigma_t\, z_t, \quad z_t \sim \text{WN}(0, 1) \tag{11}$$

$$\sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i\, \varepsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j\, \sigma_{t-j}^2 \tag{12}$$

where $\omega > 0$, $\alpha_i \geq 0$, $\beta_j \geq 0$.

Parameter	Name	Meaning
ω	Constant	Long-run variance level
α_i	ARCH coefficient	Short-run reaction to shocks (ε^2)
β_j	GARCH coefficient	Volatility persistence
$\alpha + \beta$	Persistence	Close to 1: very long-lived volatility clustering

6.1 Value at Risk (VaR)

$$\text{VaR}_{95\%,t} = -1.645\,\hat{\sigma}_t \tag{13}$$

Interpretation: with 95% probability, the loss in period t is below $\text{VaR}_{95\%}$.

6.2 Volatility Half-Life

$$t_{1/2} = \frac{\ln(0.5)}{\ln(\alpha + \beta)} \tag{14}$$

6.3 Usage in the Calculator

1. Sidebar → *GARCH* → set p, q and input type (log-returns recommended for financial data).
2. "Estimate" → variance equation, conditional volatility, VaR and persistence appear.
Tip: Run the ARCH-LM pre-test first to confirm that GARCH is actually needed.

7 Forecasting and Confidence Intervals

h-step-ahead forecasts are computed from the estimated ARIMA model.

7.1 Point Forecast

Recursive forward forecast on the differenced series:

$$\hat{W}_{T+h|T} = \hat{c} + \hat{\phi}_1\,\hat{W}_{T+h-1|T} + \cdots + \hat{\phi}_p\,\hat{W}_{T+h-p|T} \tag{15}$$

(For $h > 1$, forecasts replace future values; error terms are set to 0.)

Back-transformation (cumulating the differences) yields the forecast on the original scale:

$$\hat{Y}_{T+h} = Y_T + \sum_{s=1}^{h} \hat{W}_{T+s|T} \tag{16}$$

7.2 Confidence Intervals

The h-step forecast error grows with h:

$$\mathrm{SE}_h = \hat{\sigma}\sqrt{h} \tag{17}$$

The confidence band at level $(1-\alpha)$:

$$\hat{Y}_{T+h} \pm z_{1-\alpha/2}\,\mathrm{SE}_h \tag{18}$$

Available levels: 80 % ($z = 1.282$), 90 % ($z = 1.645$), 95 % ($z = 1.960$).

8 Rolling Forecast and Out-of-Sample Performance

8.1 Methodology

In the rolling-window approach, the model is re-estimated at each time point t on a sliding window of size w and a one-step-ahead forecast is generated:

$$\hat{Y}_{t+1} = f\big(Y_{t-w+1}, \ldots, Y_t\big) \tag{19}$$

This avoids *look-ahead bias* and provides a realistic assessment of forecast quality.

8.2 Performance Metrics

Forecast Performance Metrics

$$\mathrm{RMSE} = \sqrt{\frac{1}{T_{\text{test}}}\sum_{t}\big(Y_t - \hat{Y}_t\big)^2} \tag{20}$$

$$\mathrm{MAE} = \frac{1}{T_{\text{test}}}\sum_{t}\big|Y_t - \hat{Y}_t\big| \tag{21}$$

$$\mathrm{MAPE} = \frac{100}{T_{\text{test}}}\sum_{t}\left|\frac{Y_t - \hat{Y}_t}{Y_t}\right| \% \tag{22}$$

$$\mathrm{Bias} = \frac{1}{T_{\text{test}}}\sum_{t}\big(\hat{Y}_t - Y_t\big) \tag{23}$$

- **RMSE**: Penalises large individual errors more heavily (via squaring).
- **MAE**: More robust to outliers; easier to interpret.
- **MAPE**: Relative error in percent; scale-independent.
- **Bias** > 0: Systematic over-prediction; **Bias** < 0: under-prediction.

9 Model Comparison: AIC and BIC

Information Criteria

$$\text{AIC} = -2\,\hat{\ell} + 2k \tag{24}$$

$$\text{BIC} = -2\,\hat{\ell} + k\,\ln(n) \tag{25}$$

where $\hat{\ell}$ is the maximised log-likelihood, k is the number of free parameters, and n is the sample size.

- **Optimal**: Model with the *lowest* AIC or BIC.
- **BIC vs. AIC**: BIC penalises parameter count more heavily ($\ln n > 2$ for $n > 7$) and selects more parsimonious models.
- **Rule of thumb**: $\Delta\text{AIC} > 2$ is considered a significant improvement.

10 Cointegration (Engle-Granger)

Definition 10.1 (Cointegration)**.** Two $I(1)$ time series Y_t and X_t are *cointegrated* if the linear combination $u_t = Y_t - \alpha - \beta X_t$ is stationary ($I(0)$).

10.1 Engle-Granger Two-Step Procedure

Engle-Granger Test Procedure

Step 1: OLS regression

$$Y_t = \hat{\alpha} + \hat{\beta}\,X_t + \hat{u}_t \tag{26}$$

Step 2: ADF test on the residuals $\hat{u}_t$ (without constant):

$$\Delta\hat{u}_t = \gamma\,\hat{u}_{t-1} + \sum_{k=1}^{m} \delta_k\,\Delta\hat{u}_{t-k} + \varepsilon_t \tag{27}$$

Test statistic: $\tau_{EG} = \hat{\gamma}/\text{SE}(\hat{\gamma})$

Significance level	Critical value (MacKinnon 1991, $n \geq 100$)
1 %	−3.90
5 %	−3.34
10 %	−3.04

Economic interpretation: Cointegrated series return to a common equilibrium after deviations. This is the foundation for:

- **VECM** (vector error correction model)
- **Pairs trading** (statistical arbitrage)

11 Granger Causality

Definition 11.1 (Granger Causality). X_t *Granger-causes* Y_t if past values of X significantly improve the forecast of Y beyond what is explained by Y's own lags alone.

Important

Granger causality represents **predictive usefulness**, *not* true causality in the philosophical sense. Third variables can generate spurious causality.

11.1 F-Test

Granger F-Test by Lag

$$\text{Restricted model (AR):} \quad Y_t = c + \sum_{k=1}^{m} \phi_k Y_{t-k} + \varepsilon_t^r \tag{28}$$

$$\text{Unrestricted model (VAR):} \quad Y_t = c + \sum_{k=1}^{m} \phi_k Y_{t-k} + \sum_{k=1}^{m} \gamma_k X_{t-k} + \varepsilon_t^u \tag{29}$$

$$F = \frac{(\text{SSR}_r - \text{SSR}_u)/m}{\text{SSR}_u/(n-2m-1)} \sim F(m,\, n-2m-1) \tag{30}$$

H_0: $\gamma_1 = \cdots = \gamma_m = 0$ (no Granger causality). Reject when $F > F_{\text{crit}}$.

12 VAR Models

Definition 12.1 (VAR(p) Model). A vector autoregressive model of order p for a k-dimensional process $\mathbf{Y}_t = (Y_{1t}, \ldots, Y_{kt})'$:

$$\boxed{\mathbf{Y}_t = \mathbf{c} + \mathbf{A}_1 \mathbf{Y}_{t-1} + \cdots + \mathbf{A}_p \mathbf{Y}_{t-p} + \boldsymbol{\varepsilon}_t} \tag{31}$$

with $\boldsymbol{\varepsilon}_t \sim \text{WN}(\mathbf{0}, \boldsymbol{\Sigma})$.

Each equation is estimated separately by OLS (efficient when all equations share the same regressors). The model contains $k(1 + kp)$ parameters.

12.1 Lag Selection

Optimal lag via AIC or BIC:

$$\mathrm{AIC}(p) = -2\hat{\ell} + 2k^2 p, \qquad \mathrm{BIC}(p) = -2\hat{\ell} + k^2 p \ln(n) \tag{32}$$

12.2 Usage in the Calculator

1. Sidebar → *VAR* → select variables via checkbox (min. 2).
2. Optional: click "Optimal lag (AIC/BIC)" → best p is entered automatically.
3. "Estimate VAR" → coefficient matrices, fit charts and cross-correlation appear.

13 Worked Example: Eurozone Macroeconomic Data

13.1 Data Description

We use the sample data built into the calculator with monthly frequency (2010–2024):

Variable	Mean	Std. dev.	Unit
Inflation (Y)	3.08	1.71	% p.a.
ECB interest rate (X)	1.04	1.52	%

For the worked calculations, 12 observations (2020–2021) are used:

$$\mathbf{Y}_{\text{Infl}} = (1.5;\ 1.7;\ 2.0;\ 2.3;\ 2.8;\ 3.5;\ 4.2;\ 5.1;\ 5.8;\ 6.1;\ 5.5;\ 4.8)$$

$$\mathbf{X}_{\text{Rate}} = (0.0;\ 0.0;\ 0.0;\ 0.0;\ 0.0;\ 0.5;\ 1.5;\ 2.5;\ 3.5;\ 4.0;\ 4.0;\ 3.75)$$

13.2 Step 1: Unit Root Test (ADF)

Example: ADF Test

Test equation: $\Delta Y_t = \alpha + \beta\, Y_{t-1} + \varepsilon_t$

OLS estimation yields:

$$\hat{\beta} = -0.1194, \quad \mathrm{SE}(\hat{\beta}) = 0.0915$$

$$\tau_{ADF} = \frac{-0.1194}{0.0915} = -1.305$$

Result: $\tau_{ADF} = -1.305 > -2.86$ (5 % critical value) $\Rightarrow H_0$ not rejected.
Conclusion: The inflation series is **not stationary** ($I(1)$) $\Rightarrow d = 1$ recommended.

13.3 Step 2: Estimating ARIMA(1,1,0)

Example: ARIMA Estimation

Differenced series: $W_t = \Delta Y_t$

Design matrix (first differences, lag 1):

$$\mathbf{X} = \begin{pmatrix} 1 & W_1 \\ 1 & W_2 \\ \vdots & \vdots \\ 1 & W_{10} \end{pmatrix}, \qquad \mathbf{Y} = \begin{pmatrix} W_2 \\ W_3 \\ \vdots \\ W_{11} \end{pmatrix}$$

OLS estimate $\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$:

$$\hat{c} = -0.0925, \quad \hat{\phi}_1 = 1.0063$$

Model equation: $\Delta Y_t = -0.0925 + 1.0063\,\Delta Y_{t-1} + \varepsilon_t$

Goodness-of-fit:

RMSE	R^2	AIC	BIC
0.3655	0.6025	14.25	15.16

13.4 Step 3: ARIMAX(1,1,0) with ECB Interest Rate

Example: ARIMAX Estimation

Extended design matrix with ΔX_t (change in interest rate):

$$\hat{c} = -0.0877, \quad \hat{\phi}_1 = 0.6322, \quad \hat{\beta}_Z = 0.3862$$

Model equation:

$$\Delta Y_t = -0.0877 + 0.6322\,\Delta Y_{t-1} + 0.3862\,\Delta X_t + \varepsilon_t$$

Interpretation: A 1 percentage point increase in the interest rate is associated with a 0.39 percentage point rise in inflation (cet. par., contemporaneously).

R^2	AIC	ΔAIC vs. ARIMA
0.6388	16.63	-2.38

$\Delta\text{AIC} = 14.25 - 16.63 = -2.38 < 0$: In this small example ($n = 12$) the interest rate does not improve the AIC. With larger samples ($n \geq 60$) the effect is typically positive.

13.5 Step 4: GARCH(1,1) on Log-Changes

Example: GARCH Estimation

Log-changes of inflation: $r_t = \ln(Y_t/Y_{t-1})$

Initialisation: $\hat{\sigma}_0^2 = \mathrm{Var}(r_t)$

Parameters (iterative method-of-moments):

$$\hat{\omega} = 0.001331, \quad \hat{\alpha} = 0.15, \quad \hat{\beta} = 0.75$$

Variance equation:

$$\hat{\sigma}_t^2 = 0.001331 + 0.15\, r_{t-1}^2 + 0.75\, \hat{\sigma}_{t-1}^2$$

Key metrics:

$$\text{Persistence: } \hat{\alpha} + \hat{\beta} = 0.90$$

$$\text{Half-life: } t_{1/2} = \frac{\ln 0.5}{\ln 0.90} \approx 6.6 \text{ periods}$$

$$\text{Current volatility: } \hat{\sigma}_{12} = 0.1296$$

$$\text{VaR}_{95\%} : -1.645 \times 0.1296 = -0.2131$$

13.6 Step 5: Forecast (3 Steps Ahead)

Example: ARIMA Forecast

Last differenced value: $W_{11} = Y_{11} - Y_{10} = 4.8 - 5.5 = -0.7$

Recursive computation:

$$\hat{W}_{13} = -0.0925 + 1.0063 \times (-0.7) = -0.7969$$

$$\hat{Y}_{13} = 4.8 + (-0.7969) = 4.003$$

Step	Forecast	LB (95 %)	UB (95 %)
$t+1$	4.0031	3.2867	4.7196
$t+2$	3.1088	2.0956	4.1220
$t+3$	2.1163	0.8754	3.3573

The confidence band widens with $\hat{\sigma}\sqrt{h} = 0.3655\sqrt{h}$.

13.7 Step 6: Granger Causality (Lag 2)

Example: Granger Test – Interest Rate → Inflation

Restricted model (AR(2) on inflation): $\text{SSR}_r = 0.4772$
Unrestricted model (AR(2) + 2 lags of interest rate): $\text{SSR}_u = 0.1491$

$$F = \frac{(0.4772 - 0.1491)/2}{0.1491/(10 - 2 \cdot 2 - 1)} = \frac{0.1641}{0.0298} = 5.50$$

Critical value: $F_{2,5;\,0.05} \approx 5.79$
$F = 5.50 < 5.79$ (5 %) – just not significant at the 5 % level, significant at 10 %.
Caution: With $n = 12$ the test power is low. With a larger sample, the interest rate would typically be a significant Granger cause of inflation.

13.8 Step 7: Cointegration (Engle-Granger)

Example: Cointegration – Inflation ~ Interest Rate

Step 1: OLS regression $Y_t = \alpha + \beta X_t + u_t$:

$$\hat{\alpha} = 2.3013, \quad \hat{\beta} = 0.8954, \quad R^2 = 0.8852$$

Long-run relationship:

$$\hat{Y}_t = 2.30 + 0.90\, X_t$$

Step 2: ADF on residuals $\hat{u}_t$:

$$\tau_{EG} = -0.9504$$

Critical value (5 %): −3.34.
$\tau_{EG} = -0.95 > -3.34 \Rightarrow H_0$ not rejected $\Rightarrow$ **no cointegration** at the 5 % level.
Note: Despite the high $R^2 = 0.885$, this is a *spurious regression* – a classic illustration of why the cointegration test is indispensable.

13.9 Step 8: VAR(1) Model

Example: VAR Model

Equation system ($k = 2$, $p = 1$):

$$\hat{Y}_{1t} = -0.882 + 1.636\, Y_{1,t-1} - 0.796\, Y_{2,t-1}$$
$$\hat{Y}_{2t} = -1.465 + 0.796\, Y_{1,t-1} + 0.228\, Y_{2,t-1}$$

where Y_1 = inflation and Y_2 = interest rate.

Interpretation – Equation 1 (Inflation):
$\hat{\phi}_{11} = 1.636$: Strong inflation persistence.
$\hat{\phi}_{12} = -0.796$: A higher interest rate dampens inflation in the following period.

Interpretation – Equation 2 (Interest Rate):

$\hat{\phi}_{21} = 0.796$: Rising inflation pushes up the interest rate (Taylor-rule effect).

13 Quick Reference: All Modules

Module	Purpose	Sidebar
Unit Root	Check stationarity, determine d	Prerequisites
Ljung-Box/ARCH	Is ARIMA/GARCH needed?	Prerequisites
ARIMA	Univariate modelling and forecasting	Univariate
ARIMAX	ARIMA + exogenous variable	Univariate
GARCH	Volatility modelling, VaR	Univariate
Forecast	h-step forecast + confidence band	Forecast & Eval.
Rolling Forecast	Out-of-sample RMSE/MAE/MAPE	Forecast & Eval.
Model Selection	AIC/BIC grid search over ARIMA orders	Forecast & Eval.
Cointegration	Long-run equilibrium relationship	Multivariate
Granger	Predictive causality $X \to Y$	Multivariate
VAR	Simultaneous modelling of k series	Multivariate