Journal of Social Research

P-ISSN: 2827-9832
E-ISSN: 2828-335x
http://ijsr.internationaljournallabs.com/index.php/ijsr

IMPLEMENTATION OF THE AUTOREGRESSIVE INTEGRATED (VARI) VECTOR MODEL FOR STRESS-TESTING ANALYSIS OF INDONESIAN BANKING

Lucy Fitria Dewi, Yusep Suparman, I Gede Nyoman Mindra Jaya
Faculty of MIPA, Universitas Padjadjaran
lucy.fitriadewi@gmail.com

ABSTRACT

Stress testing in the banking sector is an important tool for assessing a bank's risk level and provides a basis for assisting financial institutions and regulators in making decisions by taking into account the impact of macroeconomic changes on banking risk, one of which is credit risk, which can be seen from the level of non-performing loans (NPL). bank. NPL ratio data and macroeconomic data are periodic data presented in monthly, quarterly or semi-annual periods and are categorized as time *series data*. At the stationarity test stage it is known that the data is not stationary, so to overcome this a *differencing process is carried out so that it uses the Vector Autoregressive Integrated (VARI)* model in the *stress testing analysis*. Based on the Granger test, there is a causal relationship between macroeconomic variables and NPL variables, so the model can be used to predict NPL values as a *stress-testing analysis*. The model that meets all assumptions and has the maximum lag is the VARI (5.1) model. The results of the model analysis show that the largest and most significant contribution to changes in the NPL variable is from real GDP shocks followed by shocks from Bank Indonesia's policy variables, namely the BI *Rate* and Inflation.

Keywords: *Vector Autoregressive Integrated (VARI), Stress Testing, Stationarity*

INTRODUCTION

Stress testing is test tool for measure strength and endurance system banking with do evaluation the connection between condition macro economy with risk credit from time to time. Variable risk credit used _ is ratio *non-performing loans* (credit problematic), because the more tall NPL ratio then the more big risk credit borne by the Bank. NPL ratio data and macroeconomic data is periodic data presented _ in period time monthly, quarterly nor semiannual and categorized as series data time. In matter this, will done analysis row time multivariate for know the connection between variable *nonperforming loans* (NPL) with variables macro economy (Real GDP, Inflation, and BI Rate) (Waelchli, 2016). *Vector autoregressive* model (VAR) which is a stochastic process model can used for see connection each other dependency from various variable with using string data time (Onwukwe & Nwafor, 2014). No all variable in study This have stationary data, so for cope with data that is not stationary in the VAR model, the *differencing* process is carried out. modelling row time multivariate in the form of *differencing* data use *vector autoregressive integrated* (VARI). Objective from study This is do stress *testing* analysis banking specifically risk commercial bank credit that is affected by the variable macroeconomy (Real GDP, Inflation , and BI Rate) using the *vector autoregressive integrated* (VARI) model and predicting mark ratio credit problem (NPL) of commercial and variable banks macro economy in the future come (Fauzi & Yuliati, 2021). Besides that for know impact shock variable macro economy to ratio credit problematic Commercial Banks as well as impact shock ratio credit problem to Indonesian economy through components of the VARI model viz *Impulse Response Function* (IRF) and *Forecast Error Variance Decomposition* (FEVD).

METHODS

Research data

The data used in this study uses quarterly secondary data from Indonesian Banking Statistics and the Central Bureau of Statistics including the ratio of *non-performing loans* (NPL) to commercial banks, real gross domestic product (GDP), the value of Indonesian inflation and the value of Bank Indonesia's interest rate (B *rate*) for data period 2003 – 2021. Data analysis used Software R and a multivariate time series model. Each observation can be expressed as a random variable $\boldsymbol{Z_t}$with the notation $Z_{t1}, Z_{t2}, \ldots, Z_{tk}$where $k$is the number of variables.

Stationarity

One of the assumptions of *the time series* model is that the data is stationary. There are two types of stationarity (Cryer, 1986), namely strong stationary (*strictly stationarity*) and weak stationarity (*weakness stationarity*). The data is said to be strongly stationary if the distribution is from$Z_{t_1}, Z_{t_2}, \ldots, Z_{t_n}$ is the same as the distribution from $Z_{t_{1-k}}, Z_{t_{2-k}}, \ldots, Z_{t_{n-k}}$for every time to $t_1, t_2, \ldots, t_n$and every time lag $k$while the data is said to be weak *stationarity* if

(i) Average : $E(Z_t) = c$independent with t,$\forall t \in \mathbb{Z}$

(ii) Variance : $E(|Z_t|^2) < \infty$,$\forall t \in \mathbb{Z}$

(iii)Covariance : $\gamma_Z(t,s) = \gamma_Z(t+h, s+h)$,$\forall t,s,h \in \mathbb{Z}$

Stationary test is done to avoid *spurious regression.* Stationary detection can be done by unit root test using the *Augmented Dickey Fuller Test* (ADF Test).

Optimum Lag Determination

In identifying the VARI(p, d) is to determine the order $p$of VAR(p). Determination of the optimum lag (p) will be used in the analysis and parameter estimation for the *vector autoregressive integrated* (VARI) model. The length of the lag represents the degrees of freedom in the VARI model. The best model is the model that has the *smallest Akaike Information Criterion, Hanna Quinn, Schwarz Information Criterion* and Final Prediction Error values (William & Wei, 2006). To obtain the optimum lag, the previous step is to determine the determinant value of the residual covariance.

$$det\left(\sum p\right) = det\ \frac{1}{n-p}\Sigma_t \hat{e}_t \hat{e}_t' \qquad (1)$$

where $p$is the parameter number of each equation in the VARI model.

Then proceed with using the AIC, HQ, SC and FPE values by selecting the smallest value.

$$AIC(p) = \ln\left(det\left(\sum p\right)\right) + \frac{2M^2 \sum p}{n} \qquad (2)$$

$$HQ = \ln\left(det\left(\sum p\right)\right) + \frac{\ln n}{n} \sum p\, M^2 \qquad (3)$$

$$SIC = \ln\left(det\left(\sum p\right)\right) + \frac{M^2 \sum p \ \ln n}{n} \tag{4}$$

$$FPE = \left(\frac{n + M\sum p + 1}{n - M\sum p - 1}\right)^M det\left(\sum p\right) \tag{5}$$

where M is the number of parameters in the model, $p$is the length of the lag, $det(\sum p)$is the determinant of the *variance – covariance matrix* of the residuals for the model with lag p, while $n$is the number of observations.

***Vector Autoregressive* (VAR)**

Sims (1980) proposed the *vector autoregressive* (VAR) method, which is a dynamic linear model that is widely used for forecasting economic variables in the long and short term. This model does not require a distinction between endogenous and exogenous variables because all variables in the VAR model are endogenous. In addition, the VAR model can also be used to determine cause and effect (Peña & Rodríguez, 2002). The estimation method is simple, namely the least squares method and a separate model can be made for each endogenous variable (Dijkstra, 2014). The forecasting results obtained are better than using a more complex simultaneous equation model. The VAR model with $k$variables and orders $p$can be written as follows.

$$\mathbf{Z}_t = \mathbf{\Phi}_0 + \mathbf{\Phi}_1\mathbf{Z}_{t-1} + \mathbf{\Phi}_2\mathbf{Z}_{t-2} + \cdots + \mathbf{\Phi}_p\mathbf{Z}_{t-p} + \mathbf{a}_t \tag{6}$$

d ith

$\mathbf{Z}_t = \begin{bmatrix} Z_{1,t} \\ Z_{2,t} \\ \vdots \\ Z_{k,t} \end{bmatrix}$, is a vector of size $k \times 1$at the t-th time t,

$\mathbf{Z}_{t-i} = \begin{bmatrix} Z_{1,t-i} \\ Z_{2,t-i} \\ \vdots \\ Z_{k,t-i} \end{bmatrix}$, is a vector of size $k \times 1$with$i = 1,2,\dots,p$

$\mathbf{\Phi}_0 = \begin{bmatrix} \phi_{10} \\ \phi_{20} \\ \vdots \\ \phi_{k0} \end{bmatrix}$, is a constant vector of size $k \times 1$.

$\mathbf{\Phi}_1$, $\mathbf{\Phi}_2$, …, $\mathbf{\Phi}_p$is a matrix of sized VAR coefficients$k \times k$

with $\mathbf{\Phi}_1 = \begin{bmatrix} \phi_{11(1)} & \phi_{12(1)} & \cdots & \phi_{1k(1)} \\ \phi_{21(1)} & \phi_{22(1)} & \cdots & \phi_{2k(1)} \\ \vdots & \vdots & \ddots & \vdots \\ \phi_{k1(1)} & \phi_{k2(1)} & \cdots & \phi_{kk(1)} \end{bmatrix}$,$\mathbf{\Phi}_p = \begin{bmatrix} \phi_{11(p)} & \phi_{12(p)} & \cdots & \phi_{1k(p)} \\ \phi_{21(p)} & \phi_{22(p)} & \cdots & \phi_{2k(p)} \\ \vdots & \vdots & \ddots & \vdots \\ \phi_{k1(p)} & \phi_{k2(p)} & \cdots & \phi_{kk(p)} \end{bmatrix}$

$\mathbf{a}_t = \begin{bmatrix} a_{1,t} \\ a_{2,t} \\ \vdots \\ a_{k,t} \end{bmatrix}$, is a residual vector of size $k \times 1$, Equation 6 can be written as follows:

$$\begin{bmatrix} Z_{1,t} \\ Z_{2,t} \\ \vdots \\ Z_{k,t} \end{bmatrix} = \begin{bmatrix} \phi_{10} \\ \phi_{20} \\ \vdots \\ \phi_{k0} \end{bmatrix} + \begin{bmatrix} \phi_{11(1)} & \phi_{12(1)} & \cdots & \phi_{1k(1)} \\ \phi_{21(1)} & \phi_{22(1)} & \cdots & \phi_{2k(1)} \\ \vdots & \vdots & \ddots & \vdots \\ \phi_{k1(1)} & \phi_{k2(1)} & \cdots & \phi_{kk(1)} \end{bmatrix} \begin{bmatrix} Z_{1,t-i} \\ Z_{2,t-i} \\ \vdots \\ Z_{k,t-i} \end{bmatrix} + \ldots + \begin{bmatrix} \phi_{11(p)} & \phi_{12(p)} & \cdots & \phi_{1k(p)} \\ \phi_{21(p)} & \phi_{22(p)} & \cdots & \phi_{2k(p)} \\ \vdots & \vdots & \ddots & \vdots \\ \phi_{k1(p)} & \phi_{k2(p)} & \cdots & \phi_{kk(p)} \end{bmatrix} \begin{bmatrix} Z_{1,t-p} \\ Z_{2,t-p} \\ \vdots \\ Z_{k,t-p} \end{bmatrix} + \begin{bmatrix} a_{1,t} \\ a_{2,t} \\ \vdots \\ a_{k,t} \end{bmatrix}$$

with

$Z_{j,t}$: variable data $j$at time to t, with$j = 1,2,\ldots,k$ And$t = 1,2,\ldots,n$

$Z_{j,t-i}$: variable data $j$at time $t - i$to with$j = 1,2,\ldots,k$; $t = 1,2,\ldots,n$And$i = 1,2,\ldots,p$

ϕ_{j0} : constant variable equation $j$with$j = 1,2,\ldots,k$

$\phi_{jj(i)}$: variable equation parameter coefficient j for variable to $j$lag i; with$i = 1,2,\ldots,p$

$a_{j,t}$: variable residual $j$at time to$t = 1,2,\ldots,n$

VAR model writing can be done using *backwards shift* operator B as follows.

$$\begin{aligned} B\boldsymbol{Z}_t &= \boldsymbol{Z}_{t-1} \\ B^2\boldsymbol{Z}_t &= \boldsymbol{Z}_{t-2} \\ &\ldots \\ B^p\boldsymbol{Z}_t &= \boldsymbol{Z}_{t-p} \end{aligned}$$

then the equation can be written

$$\begin{aligned} \boldsymbol{Z}_t - \boldsymbol{\Phi}_1\boldsymbol{Z}_{t-1} - \cdots - \boldsymbol{\Phi}_p\boldsymbol{Z}_{t-p} &= \boldsymbol{\Phi}_0 + \mathbf{a}_t \\ \boldsymbol{Z}_t - \boldsymbol{\Phi}_1 B\boldsymbol{Z}_t - \cdots - \boldsymbol{\Phi}_p B^p\boldsymbol{Z}_t &= \boldsymbol{\Phi}_0 + \mathbf{a}_t \\ \left(\boldsymbol{I} - \boldsymbol{\Phi}_1 B - \ldots - \boldsymbol{\Phi}_p B^p\right)\boldsymbol{Z}_t &= \boldsymbol{\Phi}_0 + \mathbf{a}_t \end{aligned}$$

or can also be written

$$\boldsymbol{\Phi}_p(\mathrm{B})\boldsymbol{Z}_t = \boldsymbol{\Phi}_0 + \mathbf{a}_t \quad (7)$$

with $\boldsymbol{\Phi}_p\boldsymbol{Z}_t = \mathbf{I} - \boldsymbol{\Phi}_1 B - \cdots - \boldsymbol{\Phi}_1 B^p$, is the $\mathbf{a}_t$ *Vector White Noise* processing circuit dimensions k (VWN ($\mathbf{0}, \sum$))

Integrated Vector Autoregressive Model (VARI)

For example, the VAR model ( ) is not stationary, so the ρ*differencing* process is carried out several $d$times so that the data is stationary, and the VARI model (ρ, d) obtained by

$k$variable formulation. If $\rho = 1$ and $d = 1$, then the VARI model (1,1) can be obtained in matrix form as follows:

$$\begin{bmatrix} Z_{1,t} \\ Z_{2,t} \\ \vdots \\ Z_{k,t} \end{bmatrix} - \begin{bmatrix} Z_{1,t-1} \\ Z_{2,t-1} \\ \vdots \\ Z_{k,t-1} \end{bmatrix} = \begin{bmatrix} \phi_{11(1)} & \phi_{12(1)} & \cdots & \phi_{1k(1)} \\ \phi_{21(1)} & \phi_{22(1)} & \cdots & \phi_{2k(1)} \\ \vdots & \vdots & \ddots & \vdots \\ \phi_{k1(1)} & \phi_{k2(1)} & \cdots & \phi_{kk(1)} \end{bmatrix} \left(\begin{bmatrix} Z_{1,t-1} \\ Z_{2,t-1} \\ \vdots \\ Z_{k,t-1} \end{bmatrix} - \begin{bmatrix} Z_{1,t-2} \\ Z_{2,t-2} \\ \vdots \\ Z_{k,t-2} \end{bmatrix} \right) + \begin{bmatrix} a_{1,t} \\ a_{2,t} \\ \vdots \\ a_{k,t} \end{bmatrix}$$

or equal to the equation

$$\mathbf{Z}_t - \mathbf{Z}_{t-1} = \mathbf{\Phi}_1(\mathbf{Z}_{t-1} - \mathbf{Z}_{t-2}) + \mathbf{a}_t \quad (8)$$

It is assumed that the error values follow a normal distribution, ie $\mathbf{a_t} \sim N(0, \sigma^2 \boldsymbol{I})$and $\mathbf{Y}_t = \mathbf{Z}_t - \mathbf{Z}_{t-1}$. Then the VARI equation with $t = 1,2, \dots, n$is

$$\mathbf{Y}_{t((k\times(n-1))\times 1} = \mathbf{\Phi}_{(k\times k)\times 1} \mathbf{Y}_{t-1((k\times(n-1))\times(k\times k)} + \mathbf{a}_{t(k\times(n-1))\times 1} \quad (9)$$

Parameter Estimation

Parameter estimation used for the *vector autoregressive integrated (VARI) model was obtained by the Ordinary Least Square* method (Ordinary Least Squares Method/MKTB), namely determining the function of the sum of squared errors by changing the matrix $\boldsymbol{Y}, \boldsymbol{\Phi}, \boldsymbol{A}$ into vector form. The following illustrates the VARI estimation (ρ, d) with the formulation of 2 variables.

$$\boldsymbol{Y} = \begin{bmatrix} Y_{1,1} \\ Y_{1,2} \\ \vdots \\ Y_{1,n} \\ Y_{2,1} \\ Y_{2,2} \\ \vdots \\ Y_{2,n} \end{bmatrix}_{\boldsymbol{2nx1}} \quad \boldsymbol{\phi} = \begin{bmatrix} \phi_{10} \\ \phi_{11} \\ \phi_{12} \\ \phi_{20} \\ \phi_{21} \\ \phi_{22} \end{bmatrix}_{\boldsymbol{6x1}} \quad \boldsymbol{a} = \begin{bmatrix} a_{1,1} \\ a_{1,2} \\ \vdots \\ a_{1,n} \\ a_{2,1} \\ a_{2,2} \\ \vdots \\ a_{2,n} \end{bmatrix}_{\boldsymbol{2nx1}} \quad \boldsymbol{W} = \begin{bmatrix} 1 & Y_{1,1-1} & Y_{2,1-1} \\ 1 & Y_{1,2-1} & Y_{2,2-1} \\ \vdots & \vdots & \vdots \\ 1 & Y_{1,n-1} & Y_{2,n-1} \end{bmatrix}$$

with$\boldsymbol{V} = \boldsymbol{I}_{2x2} \otimes \boldsymbol{W}_{nx3} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \otimes \begin{bmatrix} 1 & Y_{1,1-1} & Y_{2,1-1} \\ 1 & Y_{1,2-1} & Y_{2,2-1} \\ \vdots & \vdots & \vdots \\ 1 & Y_{1,n-1} & Y_{2,n-1} \end{bmatrix} =$

$$\begin{bmatrix} 1 & Y_{1,1-1} & Y_{2,1-1} & 0 & 0 & 0 \\ 1 & Y_{1,2-1} & Y_{2,2-1} & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & Y_{1,n-1} & Y_{2,n-1} & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & Y_{1,1-1} & Y_{2,1-1} \\ 0 & 0 & 0 & 1 & Y_{1,2-1} & Y_{2,2-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 1 & Y_{1,n-1} & Y_{2,n-1} \end{bmatrix}_{2nx6} \quad (10)$$

then equation 9 can be written as:

$$\boldsymbol{Y} = \boldsymbol{V\phi} + \boldsymbol{a} \quad (11)$$

The equation for the sum of squared errors is as follows:

$$S = \boldsymbol{Y}^T\boldsymbol{Y} - 2\boldsymbol{Y}^T\boldsymbol{V\phi} + \boldsymbol{\phi}^T\boldsymbol{V}^T\boldsymbol{V\phi} \quad (12)$$

Furthermore obtain derivative First function squared error against $\boldsymbol{\phi}$partially, and set the first derivative equation to zero. So that parameter Φestimation with the OLS method is

$$\hat{\Phi}_{OLS} = (\boldsymbol{V}^T\boldsymbol{V})^{-\mathbf{1}}\boldsymbol{V}^T\boldsymbol{Y} \quad (13)$$

In order to obtain the minimum squared error function, a second derivative of the squared error function is required which has a positive value so that equation 13 is derived Φas follows

$$\frac{d^2S}{d\boldsymbol{\phi}\, d\boldsymbol{\phi}^T} = \frac{d(-2\boldsymbol{Y}^T\boldsymbol{V} + 2\boldsymbol{\phi}^T\boldsymbol{V}^T\boldsymbol{V})}{d\boldsymbol{\phi}^T} = 2\boldsymbol{V}^T\boldsymbol{V} \quad (14)$$

The estimation of the VARI model uses the OLS model in equation 14 with using *time series* data that has been stationary (differencing). Those models Then tested parameter significance more formerly before residual test with objective for determine the parameters that influence against VARI models. Testing is done by test t that is

$$t_{hitung} = \frac{\hat{\phi}_{jj(i)}}{SE(\hat{\phi}_{jj(i)})} \quad (15)$$

where $\hat{\phi}$is the coefficient of each parameter, $SE(\hat{\phi}_{jj(i)})$is the standard error of each parameter. Testing the hypothesis as follows:

$H_0 : \phi_{jj(i)} = 0$(coefficient of parameter is not significant/influence in the model)

$H_1 : \phi_{jj(i)} \neq 0, j = 0,1,2,\ldots,k$ (coefficient of significant/influential parameter in the model)

If $|t_{hitung}| > t_{\alpha/2,(n-k)}$, with $n$is the number of observations, $k$is the number of parameters or the *p-value* < 0.05 then H_0 it is rejected so that the parameters are significant and have an effect on the model.

Outliers

An outlier is an observation or a sub-group that is inconsistent with other observations in a data set (Barnett & Seth, 2014). In this study, the observational data was replaced with residuals, so that after obtaining the residuals from all observations, the upper and lower quartile values of the absolute residual values were determined.

Model Diagnostics

Lagrange Multiplier Test

Engle developed a test to find out whether there is a problem of heteroscedasticity in *time series data* (ARCH effect) using the *Lagrange Multiplier test* (Tsay, 2005). The *Lagrange Multiplier* test is a test by regressing the squared residual values up to the lag $\boldsymbol{p}$, so that it will produce the regression equation as follows:

$$\mathrm{a}_t^2 = \mathrm{a}_0 + \alpha_1\mathrm{a}_{t-1}^2 + \cdots + \alpha_k\mathrm{a}_{t-p}^2, \quad (17)$$

so that the test hypothesis is obtained:

H_0 : $\alpha_1 = \cdots = \alpha_k = 0$(No heteroscedasticity)

H_1 : At least one$\alpha_k \neq 0$, (There is heteroscedasticity)

with test statistics:

$$LM = n \times R^2 \tag{18}$$

d ith LM is the *Lagrange Multiplier* test ,n is the number of observations, $R^2$is the coefficient of determination. The *Lagrange Multiplier* test follows a distribution $\chi^2$with degrees of freedom of $(\alpha, p + q)$. Reject H_0if $LM \geq \chi^2_{(\alpha,p+q)}$. If there is an ARCH *effect* on the residual VARI model selected, then symmetric volatility modeling using ARCH/GARCH or the GJR-GARCH model is used to deal with asymmetric volatility problems.

Autocorrelation Test

Autocorrelation is a condition where there is a correlation between the residuals in a certain period and the residuals in the previous period. Autocorrelation testing on data can use the *Portmanteau Ljung-Box* test which is a modification of the *Portmanteau test* to increase the power of the test on finite samples by replacing the residual autocorrelation coefficient (r_p) with its standard value $(\tilde{r}_p)$

$$\tilde{r}_p^2 = \frac{(n+2)}{(n-p)} r_p^2 \tag{19}$$

so that the test hypothesis is obtained:

H_0 : $\tilde{r}_1 = \tilde{r}_2 = \cdots = \tilde{r}_p = 0$(There is no autocorrelation until the 1st lag p)

H_1 : At least one $\tilde{r}_p \neq 0$(There is autocorrelation until the 1st lag p)

Q test statistic :

$$Q = T(T+2) \sum_{p=1}^{p} \frac{\tilde{r}_p^2}{T-p} \tag{20}$$

where $T$is the number of residuals, $\tilde{r}_p$is the autocorrelation between residuals, $p$is the lag. If p *-value* >α then accept _ H_0which means that there is no significant autocorrelation until the 1st lag p.

Causality

Causality test is a test that measures the strength of the relationship between two or more variables (Granger, 1969a). This starts from the absence of information regarding the influence between variables, so that the causality test can indicate whether the variable has a two-way or one-way relationship. The Granger Causality Test can be used to test the hypothesis whether a time series is useful in predicting another series (Granger, 1969b). Suppose there are two variables x and y, variable x is said to be *Granger-Cause* variable y if the past value of x contains information that helps predict y (Barnett & Seth, 2014). The equation for the Granger test is as follows

$$\boldsymbol{Y_t} = \sum_{i=1}^{n} \boldsymbol{\alpha_i Y_{i-1}} + \sum_{i=1}^{n} \boldsymbol{\beta_i X_{i-1}} + \mathbf{a_{1t}} \tag{7}$$

$$\boldsymbol{X_t} = \sum_{i=1}^{n} \boldsymbol{\gamma_i X_{i-1}} + \sum_{i=1}^{n} \boldsymbol{\tau_i Y_{i-1}} + \mathbf{a_{2t}} \tag{8}$$

According to Granger (1969) to complete the causality model can be done through the F test criteria with the following formula

$$F_{hitung} = \frac{(RSS_R - RSS_{UR})/p}{RSS_{UR}/(n-b)} \quad (9)$$

where $RSS_R$is *the Restricted Residual Sum of Square* , $RSS_{UR}$is *the Unrestricted Residual Sum of Square* , $p$is the lag length, $n$is the amount of observational data, $b$is the number of parameters estimated in the *unrestricted* model. Suppose the variable y is the dependent variable, then the restricted model is obtained by regressing the variable y with all the lag y values without including lag x as the independent variable. The following is the form of the restricted model.

$$\boldsymbol{Y_t = \sum_{i=1}^{n} \alpha_i Y_{i-1} + e_{1t}} \quad (10)$$

while Equation 7 is an unrestricted form of the model.

There are two hypotheses used in the Granger causality test as follows

H_0: X is not the cause of Granger Y

H_1: X is the cause of Granger Y

H_0:Y is not the cause of Granger X

H_1: Y is the cause of Granger X

If the calculated F value > the $F_{(p,n-b)}$ calculated value $\leq \alpha$or *p-value* , then H_0it is rejected so that the causality conditions are met and the best VARI model can be selected.

Model Accuracy

One way to validate the accuracy of the model is to calculate the *Mean Absolute Percentage Error* (MAPE) value with the following formula (William & Wei, 2006).

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{Z_t - \hat{Z}_t}{Z_t}\right| \times 100\% \quad (16)$$

where $\boldsymbol{n}$is the number of observations, $\boldsymbol{Z_t}$is the actual data and $\boldsymbol{\hat{Z}_t}$is the predicted data obtained from the VARI model after the *inverse differencing process* .

Stress Testing Analysis

The focus of bank *stress testing* is on credit risk which represents the most significant banking system risk (Sorge & Virolainen, 2006). The credit risk variable used is the *non-performing loan ratio* , because the higher the NPL ratio, the greater the credit risk borne by the Bank. This NPL ratio is influenced by internal factors and external factors. Macroeconomic Variables are factors that affect NPL externally. The decline in Gross Domestic Product (GDP), the increase in the Consumer Price Index (CPI) and the BI Rate (BI Rate) contributed to driving up NPLs (Indra, 2019). In calculating the Bank's capital adequacy in facing crisis situations, a stress test analysis is carried out to calculate the impact of credit quality in the face of extreme changes that are influenced by macroeconomic aspects. The *stress test* scenario is divided into three, namely the normal scenario, tolerable scenario, and crisis scenario (Agustin & Wiranatakusuma, 2017). The normal scenario is prepared based on forecasts for the development of the Indonesian economy, the tolerable scenario is developed to test bank resilience when economic conditions experience contraction, while the crisis scenario is developed to test bank resilience when economic conditions experience a deep contraction. Scenarios are built by adding (or subtracting) 1,2 or 3 standard deviations to the estimates produced by the VARI model [9]. The analysis component in the VAR model, namely *the impulse response function* (IRF), makes it possible to observe the response of this data and the future of each variable due to changes or *shocks* caused by other variables, while the variance decomposition provides information regarding the contribution (percentage) of changes in the variance of each variable to change in another variable.

Stability

VAR stability testing is carried out before carrying out *the Impulse Response Function* (IRF) and *Forecast Error Variance Decomposition* (FEVD) analysis. The VAR stability test is carried out by calculating the roots of the polynomial function or known as *the roots of characteristic polynomial* . If all the roots of the polynomial function have an absolute value less than one, then the VAR model is stable so that the resulting IRF and FEVD are considered valid. Stable VAR model, VAR can be shown if

$$det\left(\boldsymbol{I}_k - \boldsymbol{A}_1\boldsymbol{z} - \cdots - \boldsymbol{A}_p\boldsymbol{z}^p\right) \neq 0 \textbf{ For } |z| \leq 1 \quad (11)$$

d ith $\boldsymbol{I}_k$is a dimensionless identity matrix$(k \times k)$, $\boldsymbol{A}_1, \dots, \boldsymbol{A}_p$ is a dimensionless coefficient matrix $(k \times k)$that represents the parameter value z.

Impulse Response Function

Impulse Response Function can be used to determine the effect of one variable on the variable itself or other variables. The estimation made in this IRF is the response of a variable to a change of one standard deviation from the variable itself or from other variables contained in the model. IRF analysis uses *the Cholesky Decomposition* which aims to generate *an impulse response* that depends crucially on the ordering *of* the variables in the system with the following calculations

$$IRF(h) = \Gamma^C \quad (12)$$

d ith $\Gamma$is the parameter matrix of the VAR I model, $h$is the forecasting period , meanwhile $C$is the *Cholesky Decomposition* matrix of the *shock covariance variant matrix* . IRF is described in form graph , with vertical axis as mark response (default deviation) a variable to *shock* and horizontal axis as time For a number of time forward after happening *shock.*

Forecast Error Variance Decomposition (FEVD)

Variance Decomposition aims to predict the percentage contribution of the error variance of a variable, namely the magnitude of the difference between before and after the shocks, both originating from the variable itself and from other variables. The FEVD equation can be obtained by the following illustration

$$E_t Z_{t+1} = \phi_0 + \phi_1 Z_1 \quad (13)$$

by value $\phi_0$and $\phi_1$used to estimate the predicted valueZ_{t+1}

$$E_t Z_{t+n} = a_{t+n} + \phi_1^2 a_{t+n-2} + \cdots + \phi_1^{n-1} a_{t+1} \quad (14)$$

with a higher variable FEDV value explaining the higher contribution of endogenous variable variance to other endogenous variables. According to Lutkepohl (2005), the *Variance Decomposition calculation* is as follows:

$$w_{jk,h} = \frac{\sum_{i=0}^{h-1}\left(a_j'\theta_i e_k\right)^2}{\sum_{i=0}^{h-1}\sum_{i=0}^{h-1}\left(a_j'\theta_i e_k\right)^2} \qquad (15)$$

with $\theta_i = \Phi_i P$, with $P$is the lower triangular matrix of the *Cholesky Decomposition matrix* variance covariance , $\Phi_i = JA^i j'$with $J = [I_k \quad 0 \quad \dots \quad 0]$and $A$is the coefficient matrix of the VARI model . FEVD analysis explains the relative importance of the diversity of each variable in the system due to *shock* . This analysis used For know contribution with see percentage variance every variable during a number of future time . _

RESULTS AND DISCUSSION

Descriptive Research Variables

The data used in this study is *time series data* for the period 2003 – 2021 with quarterly data positions. The results of the descriptive analysis are presented in the following table and figure.

Table 1. Variable Descriptive Analysis

Variable	**Minimum**	***Q 1***	**Median**	***Means***	***Q 3***	**Maximum**
Real GDP (%)	-5.32	4.93	5,14	4.95	5.95	7,16
Inflation (%)	1.33	3,295	5.06	5,68	6.95	17,10
BI *Rate* (%)	4.75	6.00	7,12	7,27	7,81	12.75
NPLs (%)	1.77	2.55	3.02	3.73	4,15	8.33

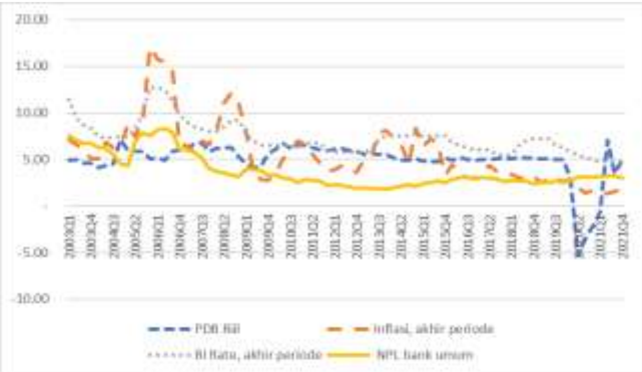

Figure 1: Macroeconomic Variables and NPL

From Table 1 and Figure 1 it can be seen that the macroeconomic variables and the variable NPL ratio of commercial banks show fluctuating data. The Real GDP variable experienced a very sharp decline with the lowest value in the last 16 years of -5.32%. At the end of 2005, the inflation variable reached the highest increase of 17.10% with an average of 5.68 in the last 16 years. The BI Rate variable also achieved a very high increase as did the inflation variable, with a value of 12.75% at the end of 2005 with an average of 7.27% in the last 16 years. The variable ratio of commercial banks' NPLs reached a maximum value of 8.33% in the second quarter of 2006, 6 months after inflation and the BI Rate reached their highest values in the last 16 years (Schechtman & Gaglianone, 2012).

Stationarity Test

Testing is done with the level of significance $\boldsymbol{\alpha = 5}\%$. The following table shows the results of the stationarity test.

Table 2. ADF Test on Macroeconomics and NPL Results of Differentencing

Variable	ADF Partial *Test*			Simultaneous ADF Tests		
	Levels$\boldsymbol{I(1)}$	***p-values***	**Information**	**Levels$\boldsymbol{I(1)}$**	***p-values***	Information
GDP	-5.1287	0.01	It has been stationary	102,066	0.0000	It has been stationary
Inflation	-5.1975	0.01	It has been stationary			
BI *Rate*	-4.1787	0.01	It has been stationary			
NPLs	-4,459	0.01	It has been stationary			
Variable	**ADF Partial *Test***			**Simultaneous ADF Tests**		
	Levels$\boldsymbol{I(0)}$	***p-values***	**Information**	**Levels$\boldsymbol{I(0)}$**	***p-values***	Information
Real GDP	-3.6936	0.03204	stationary	18.0590	0.0208	Not Stationary
Inflation	-3.3304	0.0743	stationary			
BI *Rate*	-3,043	0.1515	Not Stationary			
NPLs	-2.5126	0.3669	Not Stationary			

The ADF test shows that the macroeconomic variables and NPL variables are stationary *differencing* 1, so that the data can be used in the analysis of the vector autoregressive integrated (VARI) model. The data plots that have been stationary are presented in Figure 2 below.

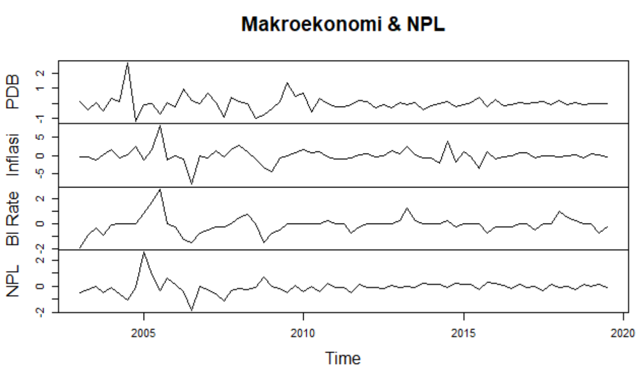

Figure 2: Data Plots

Determination of Lag Parameters

Based on the AIC, HQ, SC and FPE criteria, lag 5 is obtained as the best model because it has the smallest AIC, HQ and FPE values, so the model to be used is the VARI model (5.1).

Table 3. Optimum Lag Selection

lag	AIC	HQ	SC	FPE
1	-3.08030997	-2.81090139	-2.39413759*	0.04600957
2	-2.96442580	-2.47949034	-1.72931551	0.05201806
3	-2.94989152	-2.24942919	-1.16584333	0.05368404
4	-3.90239004	-2.98640084	-1.56940394	0.02139074
5	-4.1673303*	-3.0358143*	-1.2854063	0.0173193*

Parameter Estimation

The next process is to look at the significance level of the VARI model parameters (5.1). The significance of all parameters was partially tested using the t-test statistic so that significant parameter estimates were obtained as follows.

Table 4. Final VAR Model Results (5,1)

Variable	**VARI models (5,1)**
$\Delta\ PDB\ Riil_t$	$dZ_{1,t} = -0{,}49\ dZ_{1,t-1} - 0{,}45\ dZ_{3,t-3} + 0{,}40\ dZ_{4,t-4} + 0{,}39\ dZ_{4,t-5}$
$\Delta\ Inflasi_t$	$dZ_{2,t} =\ 0{,}72 dZ_{1,t-1} + 0{,}92 dZ_{3,t-1} + 1.71 dZ_{1,t-4} - 0{,}43 dZ_{2,t-4}$
$\Delta\ BIRate_t$	$dZ_{3,t} =\ 0{,}39 dZ_{1,t-1} + 0{,}35 dZ_{3,t-1} + 0{,}58 dZ_{4,t-1} + 0.25 dZ_{1,t-2} + 0{,}35 dZ_{1,t-4} - 0.40 dZ_{4,t-5}$
$\Delta\ NPL\ _t$	$dZ_{4,t} =\ 0{,}51 Z_{4,t-1} + 0.56 dZ_{1,t-2} + 0.19 dZ_{3,t-3} + 0{,}33 dZ_{4,t-4} - 0{,}52 dZ_{4,t-5}$

After obtaining the VARI model equation (5.1), further detection of outliers is carried out *which* can be seen through the residual boxplot of the model.

Model Diagnostic Test

From the LM Test it shows H_0 that it is accepted (*p-value* = 0.3682 > 0.05). It can be concluded that for the four equations the VARI model (5.1) fulfills the residual homogeneity requirements. The results of the Portmanteau Ljung-Box test with a significance level of 5% obtained a p-value of 0.3252, it can be concluded that the residual data has the assumption of non-autocorrelation

Causality Test

The results of the Granger causality test show that there are only 3 simultaneous causality relationships with a p-value <0.05 as follows:

a. Real GDP is the cause of Granger Inflation, BI *Rate* and NPL,
b. The BI *Rate* is a Granger cause of Real GDP, Inflation and NPLs, as well as
c. NPL is the cause of Granger Real GDP, Inflation and the BI *Rate*

Table 5. Model Causality Test Results

H_0	F_{hitung}	*p-values*	Conclusion
PDB *do not Granger-cause* BI Inflation .NPL *Rate*	4.2014	1.549e-06	H_0 rejected
Inflation *does not Granger-cause* BI GDP. NPL *rates*	0.56986	0.8947	H_0 accepted
BI. *Rate do not Granger-cause* GDP NPL inflation	2.3841	0.003873	H_0 rejected
NPL *do not Granger-cause* PDB BI Inflation. *Rate*	2.5289	0.002127	H_0 rejected

Stability Test

It is obtained that the value of the modulus range is less than 1 or in other words the VARI model (5.1) is stable so that it can carry out *Impulse Response Function* and *Forecast Error Decomposition Variance* (FEDV) analysis.

Table 6. Stability Test Results

Roots	*Modulus*
-0.091974 + 0.921081i	0.9256612
-0.091974 - 0.921081i	0.9256612
0.621722 + 0.6881182i	0.9222513
0.621722 - 0.6881182i	0.9222513
0.789770 + 0.470365i	0.9192276
0.789770 - 0.470365i	0.9192276
-0.876795 – 0.062235i	0.8790013
-0.876795 + 0.062235i	0.8790013
-0.643057- 0.553253i	0.8482989
-0.643057+ 0.553253i	0.8482989
0.344928-0.727333i	0.8049774
0.344928+0.727333i	0.8049774
0.738165	0.7381645
-0.338656-0.641669i	0.7255526
-0.338656+0.641669i	0.7255526
-0.666902	0.6669023
0.406026+0.409147i	0.5764191
0.406026-0.409147i	0.5764191
-0.051054+0.231554i	0.2371156
-0.051054-0.231554i	0.2371156

Impulse Response Function (IRF) Analysis

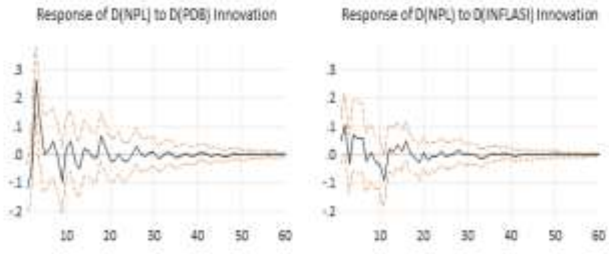

Figure 3: Variable Response of NPL growth to macroeconomic shocks

Figure 3 shows the response of the NPL growth variable to macroeconomic growth variable shocks (Real GDP, Inflation and the BI Rate) throughout the 60 observed periods. The real GDP variable shock of one standard deviation received a response in a negative direction in the first period and the second period of -0.118233 and -0.059429, then turned positive in the third period to the seventh period and tended to reach a balance like the condition before the shock from Real GDP in the 40th period was 0.005 until the end of the period.

Forecast Error Decomposition Variance (FEDV) Analysis

Figure 4 shows the *Variance Decomposition value* of the NPL variable, it can be seen that the dominant NPL variance is influenced by the variance of the variable itself until the 32nd period. The Real GDP variable has a fairly large percentage in influencing changes in the NPL variable compared to the Inflation and BI Rate variables.

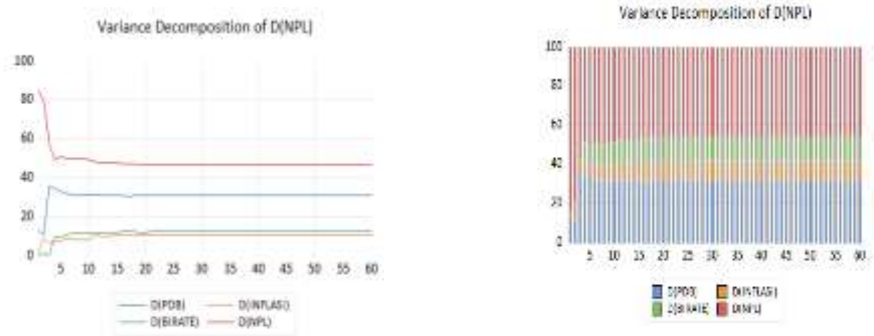

Figure 4: *Variance Decomposition* of NPL Variables

Stress Testing Analysis

The results of backtesting using data for 2003 - 2019 with the VAR model (5.1) obtained the MAPE value in accordance with the Appendix which shows the MAPE value resulting from the prediction of the GDP variable of 9.14%, which means that the model has very good accuracy, the MAPE value of the forecast variable Inflation and the NPL variable is 37.95% and 23.41% respectively, which means the model has sufficient accuracy, while the MAPE value forecasting the BI Rate variable is 11.33%, which means the model has good accuracy.

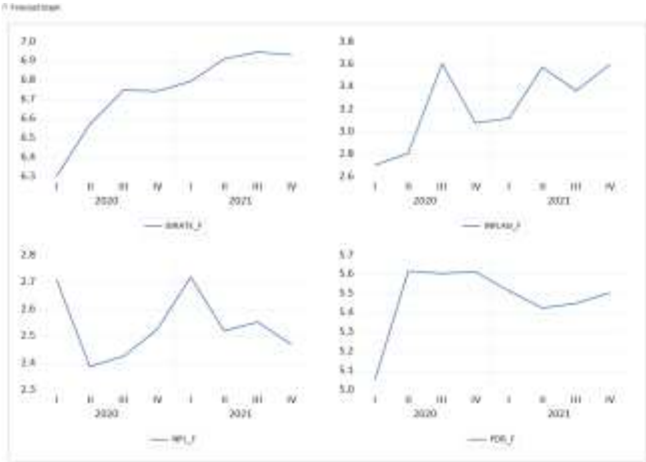

Figure 5: *Forecasting of* Macroeconomic Variables and NPL for 2020 – 2021

Forecasting the NPL variable for the normal scenario uses the VARI model (5.1) with forecasting macroeconomic variables, so that an NPL value of 2.50% for the first quarter, 2.27% for the second quarter, 2.31% for the third quarter is obtained. and by 2.28% for the fourth quarter of 2022. Meanwhile, the tolerable scenario is constructed by adding 1 standard deviation to the estimated NPL value generated by the VARI model for 2021 as well as the crisis scenario which is constructed by adding 2 standard deviations to the estimated NPL value.

CONCLUSION

Based on the results of the analysis, modeling and forecasting of Indonesian banking *stress testing using the integrated autoregressive vector* model, it can be concluded that there is non-stationary in the data so that the *differencing* process is carried out and the ADF test results show that macroeconomic variables and NPL variables are stationary at *differencing* 1. AIC criteria value, HQ and FPE produce the optimal lag at lag 5, so that an integrated autoregressive vector model (VARI, (5,1)) is obtained. This model can be used for modeling the relationship between macroeconomic variables and non-performing loan variables for commercial banks as well as analysis of banking *stress testing* , especially for commercial bank credit risk which is influenced by macroeconomic variables. As for the results of the IRF analysis, the NPL variable for macroeconomic variable shocks of one standard deviation received a negative response to the Real GDP variable and the BI Rate variable in the first period of the shock, while the Inflation variable received a positive response in the first period of the shock. This shows that if there is a fairly high increase in inflation as well as a decrease in Real GDP and the BI Rate, it will cause the NPL variable to increase, in the sense that there is an increase in the ratio of non-performing loans to commercial banks. Meanwhile, the FEDV results state that the largest and most significant contribution to changes in the NPL variable is from real GDP shocks followed by shocks from Bank Indonesia's policy variables, namely the BI *Rate* and Inflation.

REFERENCES

Agustin, L. A., & Wiranatakusuma, D. B. (2017). Analisis Stress test pada Perbankan Syariah di Indonesia. *Journal of Economics Research and Social Sciences, 1*(2), 78–89.

Barnett, L., & Seth, A. K. (2014). The MVGC multivariate Granger causality toolbox: a new approach to Granger-causal inference. *Journal of Neuroscience Methods*, *223*, 50–68.

Dijkstra, T. K. (2014). PLS'Janus face–response to professor Rigdon's 'rethinking partial least squares modeling: in praise of simple methods.' *Long Range Planning*, *47*(3), 146–153.

Fauzi, M. I., & Yuliati, N. (2021). Pemanfaatan Media Sosial Instagram sebagai Upaya Edukasi Pencegahan Penyebaran COVID-19. *Jurnal Riset Public Relations*, 148–155.

Granger, C. W. J. (1969a). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica: Journal of the Econometric Society*, 424–438.

Granger, C. W. J. (1969b). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica: Journal of the Econometric Society*, 424–438.

Indra, I. (2019). Macro Stress Test Model Risiko Kredit: Studi Empiris Perbankan Konvensional dan Syariah di Indonesia [A Macro Stress Test Model of Credit Risk: An Empirical Studies of Conventional and Islamic Banking in Indonesia]. *Jurnal Ekonomi & Kebijakan Publik*, *9*(2), 113–130.

Onwukwe, C. E., & Nwafor, G. O. (2014). A multivariate time series modeling of major economic indicators in Nigeria. *American Journal of Applied Mathematics and Statistics*, *2*(6), 376–385.

Peña, D., & Rodríguez, J. (2002). A powerful portmanteau test of lack of fit for time series. *Journal of the American Statistical Association*, *97*(458), 601–610.

Schechtman, R., & Gaglianone, W. P. (2012). Macro stress testing of credit risk focused on the tails. *Journal of Financial Stability*, *8*(3), 174–192.

Sorge, M., & Virolainen, K. (2006). A comparative analysis of macro stress-testing methodologies with application to Finland. *Journal of Financial Stability*, *2*(2), 113–151.

Tsay, R. S. (2005). *Analysis of financial time series*. John wiley & sons.

Waelchli, B. (2016). A proximity based macro stress testing framework. *Dependence Modeling*, *4*(1).

William, W., & Wei, S. (2006). Time series analysis: univariate and multivariate methods. *USA, Pearson Addison Wesley, Segunda Edicion. Cap*, *10*, 212–235.