MALAYSIAN JOURNAL OF SCIENCE
MJS
ISSN 1394 - 3065

5th International Conference on Applied & Industrial Mathematics and Statistics 2025 (ICoAIMS 2025)

Special Issue

Malaysian Journal of Science 45 (Special Issue 2): 95-104(November 2026) https://mjs.um.edu.my

Forecasting Indonesia's Inflation-at-Risk Using Vector Autoregression with Exogenous Variables Model

Ni Kadek Juliarini[1a], Dedy Dwi Prastyo[2a*], Muhammad Sjahid Akbar[3a]

Abstract: Inflation is a critical economic indicator influencing stability and monetary policy, making accurate forecasting essential. This study employs the Vector Autoregression with Exogenous Variables (VARX) model to simultaneously forecast Indonesia's inflation across various sectors, including food, health, housing, clothing, and education, incorporating the Financial Condition Index (FCI) and Financial Stress Index (FSI). The FSI evaluates financial stress using data from banking, foreign exchange, stock markets, and government securities, while the FCI assesses economic conditions through Principal Component Analysis (PCA). VARX is rarely used in this context and differs from previous studies in that it focuses on the risk of extreme inflation. In addition, this study introduces the use of the Correlation Information Criterion (CIC) as an evaluation tool, which has not been widely applied in similar research. The fluctuation patterns of each sectoral inflation generally display distinct characteristics, enabling a comprehensive evaluation of the VARX model's ability to forecast data with varying fluctuation patterns. This approach differs from previous studies, which typically utilize data with uniform fluctuation patterns. Furthermore, the Seasonal Trend Decomposition Based on Loess (STL) method is employed to mitigate seasonal effects in the data and improve forecasting accuracy. This technique is rarely applied in this context. Analyzing monthly data from February 2010 to December 2024, the VARX(1) model effectively forecasts Indonesia's Inflation-at-Risk (IaR) with high accuracy, as measured using the Symmetric Mean Absolute Percentage Error (SMAPE) criterion. Seasonal inflation in the health sector is addressed using the Seasonal Trend Decomposition Based on Loess (STL) method, which further improves forecasting precision. However, the model struggles with extreme fluctuations, highlighting the need for future research into more advanced predictive models for volatile economic conditions.

Keywords: IaR, FCI, FSI, VARX, Forecasting.

1. Introduction

Economic growth is a key measure of societal welfare and is closely linked to inflation. High inflation often slows economic expansion, while low and stable inflation supports sustained growth. Inflation is defined as a continuous, widespread rise in the prices of goods and services, pushing overall price levels upward. Estimates of inflation-related information are essential for monitoring economic pressures that may impact the financial sector. Inflation-at-Risk (IaR) is a concept used to forecast inflation by considering risks arising from financial variables, specifically the Financial Stability Index (FSI) and the Financial Conditions Index (FCI). A widely used approach in analyzing economic dynamics is the Vector Autoregression (VAR) model, which was first introduced by Christopher Albert Sims in 1972, building upon ideas developed by Granger and Clive in 1969. The VAR model can be used for multivariate analysis, where all variables are considered to be endogenous. Each variable is explained by its past values as well as the past values of other endogenous variables in the system (Gujarati, 2003).

Authors information:
[a]Department of Statistics, Institut Teknologi Sepuluh Nopember, Surabaya 60111, INDONESIA. E-mail: juliarini.kdk@gmail.com[1], dd.prastyo@its.ac.id[2], m_sjahid_a@its.ac.id[3]

*Corresponding Author: dd.prastyo@its.ac.id

The VAR model is commonly utilized for multivariate time series analysis and has several extensions, including VARX for incorporating exogenous variables (Sesay et al., 2020; Farih & Prastyo, 2022), Generalized Space-Time Autoregressive (GSTAR), and GSTARMA for accommodating spatial effects (Akbar, et al., 2020), also with exogenous variables become GSTARX (Suhartono et al., 2018), and hybrid models combining VAR with Machine Learning techniques (Suhartono et al., 2018; Prastyo et al., 2018; Hikmawati et al., 2021; Novianto et al., 2018). In comparison to previous studies, Safuan et al. (2022) conducted a study to investigate the intermittent relationship between financial stress and key macroeconomic variables in Indonesia using the Financial Stress Index (FSI) model. The conclusion drawn from the study is that there is an episodic relationship between financial stress and Indonesia's key macroeconomic variables. Additionally, the FSI developed in the study can capture unusual financial stress events, such as the global financial crisis of 2008 and the early stages of the COVID-19 pandemic in Indonesia. Therefore, for future research related to macroeconomic variables, the FSI can be considered a key variable in the study.

Based on this discussion, given the importance of inflation forecasting, this research aims to forecast IaR using the VARX model, employing inflation data from categories such as food, health, housing, clothing, and education, which reflect the basic needs of the population. Furthermore, economic indicators like

Received: October 28, 2025
Accepted: May 6, 2026
Published: November 15, 2026

DOI: https://doi.org/ 10.22452/mjs.vol45sp2.10
Malaysian Journal of Science 45 (Special Issue 2): 95-104(November 2026)

the FSI and the FCI are also used. The FSI is developed to assess reactions to systemic vulnerabilities within the financial system (Park & Mercado, 2014). Meanwhile, FCI provides greater benefits for examining the relationship between macroeconomic variables (Juhro & Iyke, 2019).

2. Literature Review

This section discusses the literature review that underpins the research. The statistical theories that form the foundation of this study include Vector Autoregression with Exogenous Variables (VARX), granger causality, Seasonal Trend Decomposition Based On Loess (STL), Financial Stress Index (FSI), Financial Condition Index (FCI), Indonesia's Inflation, and Symmetric Mean Absolute Percentage Error (SMAPE).

2.1 Vector Autoregression with Exogenous Variables (VARX) Model

The VAR model expands upon the univariate Autoregression model and is used to analyze multivariate time series data (Zivot & Wang, 2003). The equation for the VAR(*p*) model is expressed in the following form:

$$\mathbf{y}_t = \mathbf{\Phi}_1 \mathbf{y}_{t-1} + \cdots + \mathbf{\Phi}_p \mathbf{y}_{t-p} + \boldsymbol{\varepsilon}_t, \quad p > 0. \tag{1}$$

where $\mathbf{y}_t$ represents the vector of endogenous variables, $\boldsymbol{\phi}_0$ represents the intercept, $\mathbf{\Phi}_1$ represents the parameter of the first-order autoregression (AR), and $\{\boldsymbol{\varepsilon}_t\}$ is assumed to be white noise, with a mean of zero and a variance of σ^2. Meanwhile, VARX is an extension of the VAR model that includes exogenous variables. According to Ocampo & Rodríguez (2012), the VARX(*p*) model is expressed in the following equation form:

$$\mathbf{y}_t = \mathbf{\Phi}_1 \mathbf{y}_{t-1} + \cdots + \mathbf{\Phi}_p \mathbf{y}_{t-p} + \boldsymbol{\gamma}_t \mathbf{y}_t + \boldsymbol{\varepsilon}_t; \ \boldsymbol{p > 0} \tag{2}$$

where $\mathbf{x}_t$ denotes a vector of exogenous variables and $\boldsymbol{\theta}_q$ represents the parameter of exogenous variables at lag *q*.

2.2 Granger Causality

The Granger causality test identifies causal relationships between variables in a model and is usually conducted after selecting the optimal lag for the VAR model. For example, in the context of a bivariate model where $y_t = (y_{1,t}, y_{2,t})'$ and $\varepsilon_t = (\varepsilon_{1,t}, \varepsilon_{2,t})'$, the VAR(1) model is represented by the following two equations (Tsay, 2002):

$$y_{1,t} = \phi_{1,0} + \Phi_{1,11} y_{1,t-1} + \Phi_{1,12} y_{2,t-1} + \varepsilon_{1,t} \tag{3}$$

$$y_{2,t} = \phi_{2,0} + \Phi_{1,21} y_{1,t-1} + \Phi_{1,22} y_{2,t-1} + \varepsilon_{2,t} \tag{4}$$

where $\Phi_{1,ij}$ represents the ij-th element of Φ_1 and $\phi_{i,0}$ denotes the i-th element of ϕ_0. Based on Equation (3), $\Phi_{1,12}$ indicates the linear dependence of $y_{1,t}$ on $y_{2,t-1}$ given $y_{1,t-1}$. Therefore, $\Phi_{1,12}$ represents the conditional effect of $y_{2,t-1}$ on $y_{1,t}$ condition on $y_{1,t-1}$. If $\Phi_{1,12} = 0$, then $y_{1,t}$ does not depend on $y_{2,t-1}$ and the model shows that $y_{1,t}$ only depends on its own past value. Similarly, if $\Phi_{1,21} = 0$, then Equation (4) shows that $y_{2,t}$ does not depend on $y_{1,t-1}$ when $y_{2,t-1}$ is known. Considering Equations (3) and (4) simultaneously, if $\Phi_{1,12} = 0$ and $\Phi_{1,21} \neq 0$ there is a one-way relationship from $y_{1,t}$ to $y_{2,t}$. If $\Phi_{1,12} = \Phi_{1,21} = 0$ then $y_{1,t}$ and $y_{2,t}$ are independent. However, if $\Phi_{1,12} \neq 0$ and $\Phi_{1,21} \neq 0$, there is a two-way relationship between the two variables (Tsay, 2002).

2.3 Seasonal Trend Decomposition Based on LOESS (STL)

Seasonal Trend Decomposition, using Locally Estimated Scatterplot Smoothing (LOESS), also known as STL (Cleveland et al., 1990), is an algorithm developed to break down time series data into three key components: trend, seasonal, and residual. The LOESS is a nonparametric regression method that offers high flexibility, as the curve estimates are automatically formed based on the data, unaffected by outliers. Mathematically, the STL model can be expressed as follows:

$$\mathbf{y}_t = \mathbf{S}_t + \mathbf{T}_t + \boldsymbol{\varepsilon}_t \tag{5}$$

where S_t represents the seasonal component at time *t*, T_t refers to the trend component at time *t*, and ε_t represents the residual at time *t*. When applied to the VAR model, for instance, VAR(1), the equation model becomes as follows:

$$\mathbf{D}_t = \mathbf{\Phi}_1 + \mathbf{D}_{t-1} + \boldsymbol{\varepsilon}_t \tag{6}$$

where $\mathbf{D}_t$ is the vector of seasonal exogenous variables at time *t*, consisting of $D_{1t}, D_{2t}, \ldots, D_{nt}$.

2.4 Financial Stress Index (FSI)

The FSI is an index designed to detect periods of significant stress within the financial system, although not every instance of high stress will result in a crisis (Safuan et al., 2022). FSI measures financial pressure using data from banking, foreign exchange markets, stock markets, and government securities. The FSI is represented by the following equation (Safuan et al., 2022):

$$\text{FSI}_t = \omega_1 \beta_t + \omega_2 \text{EMPI}_t + \omega_3 \vartheta_t^2 + \omega_4 \text{SRET}_t + \omega_5 \text{DMP}_t, \tag{7}$$

where $\omega_j; j = 1,2,3,4,5$ represents the weight of each component based on its standard deviation. The equation for each component is presented as follows:

$$\beta_t = \frac{cov(rf, rj)}{var\,(rj)}, \tag{8}$$

where β_t is the banking stress level at the time t, the rf is the return on IDX Finance, and rj is the return on JKSE. The second component is the foreign exchange market pressure ($EMPI_t$) which can be calculated with the following equation:

$$\text{EMPI}_t = \frac{(\Delta e_t - \mu_{\Delta e})}{\sigma_{\Delta e}} - \frac{(\Delta \text{RES}_t - \mu_{\Delta \text{RES}})}{\sigma_{\Delta \text{RES}}}, \tag{9}$$

where Δe and ΔRES denote the percentage change in the value of the Indonesian rupiah against the US dollar and Indonesia's foreign exchange reserves, respectively. Meanwhile, μ and σ represent the mean and standard deviation. The third component is stock market volatility ϑ_t^2 which is assumed to follow the Generalized Autoregressive Conditional Heteroskedasticity (GARCH(1,1)) model, expressed by the following equation:

$$\vartheta_t^2 = \omega + \theta_1 \varepsilon_{t-1}^2 + \theta_2 \sigma_{t-1}^2 \tag{10}$$

The fourth component is the stock market return ($SRET_t$), which can be calculated using the following equation:

DOI: https://doi.org/ 10.22452/mjs.vol45sp2.10
Malaysian Journal of Science 45 (Special Issue 2): 95-104(November 2026)

$$\mathrm{SRET_t} = log(P_t) - log(P_{t-1}) \quad (11)$$

where P_t is the closing stock price of the JKSE at period t. If the return on the JKSE is lower, it indicates that stress in the stock market is higher, and conversely, if the return is higher, it indicates lower stress. The final component is the government securities market pressure (DMP_t), which can be calculated with the following equation:

$$\mathrm{DMP}_t = \mathrm{YIDGB}_t - \mathrm{YUST}_t \quad (12)$$

where *YIDGB* is the annual yield on Indonesian government bonds with a five-year maturity, and *YUST* is the annual yield on US government bonds with a five-year maturity.

2.5 Financial Condition Index

The FCI is a variable used to examine the connection between economic and financial factors (Juhro & Iyke, 2019). The FCI variable is constructed based on the concept of Principal Component Analysis (PCA). PCA is a multivariate analysis technique that aims to reduce the dimensionality of data by transforming correlated variables into a set of new, uncorrelated variables, while retaining as much variance as possible (Ferrara et al., 2022). These new variables are referred to as Principal Components (PC). In linear algebra, a PC is a linear combination of n random variables (y_1,y_2,…,y_n). Geometrically, this linear combination represents a new coordinate system obtained by rotating the original system to serve as the new coordinate axes. These new axes explain the maximum variance, resulting in a simpler covariance structure. The condition for forming a PC, which is a linear combination of the vector $y = (y_1, y_2, \cdots, y_n)'$ in order to have the most significant variance, it is achieved by selecting eigenvalues $e_i = (e_1,e_2,\ldots,e_n)'$ such that $Var(e_i' y)$ is maximized, and $e_i'e_i = 1$ and $Cov(e_i'y, e_m'y) = 0$ for $m < i$.

The formation of PCs can be done using either the covariance matrix or the correlation matrix. The formation of PCs using the correlation matrix is done when the observed variables do not have the same unit of measurement, requiring standardization. As a result, the variance-covariance matrix of the standardized data will be the same as the correlation matrix of the original data, and the total variance of the principal components will be equal to the sum of the variances of the original variables. The random vector $y = (y_1, y_2, \cdots, y_n)'$ has a variance-covariance matrix Σ with eigenvalues in descending order $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n \geq 0$, The linear combination of the PCs is as follows.

$$\begin{aligned} \mathrm{PC}_1 &= e_1'\mathbf{y} = e_{11}y_1 + e_{21}y_2 + \cdots + e_{n1}y_n \\ \mathrm{PC}_2 &= e_2'\mathbf{y} = e_{12}y_1 + e_{22}y_2 + \cdots + e_{n2}y_n \\ &\vdots \\ \mathrm{PC}_n &= e_k'\mathbf{y} = e_{1n}y_1 + e_{2n}y_2 + \cdots + e_{nn}y_n \end{aligned} \quad (13)$$

where PC_1 is the first PC which has the largest variance, PC_2 is the second PC with the second-largest variance, PC_n is the *n*-th PC with the *n*-th largest variance, and e_n is the eigenvector of the *n*-th variable. The general model of the *j*-th PC can be written as follows:

$$\mathrm{PC}_j = e_j'\mathbf{y}, \quad j = 1,2,\ldots,n. \quad (14)$$

Thus, $Var\,(PC_j) = e_j'\mathbf{\Sigma}e_j = \lambda_i$, $j = 1,2,\ldots,n$ and $Cov\,(PC_j, PC_m) = e_j'\mathbf{\Sigma}e_j = 0$, $j \neq m$. PCs are independent of each other, and their variances correspond to the eigenvalues of the covariance matrix **Σ**, as follows:

$$\begin{aligned} \sigma_{11} + \sigma_{22} + \cdots + \sigma_{nn} &= \sum_{i=1}^{n} Var(y_i) \\ &= \lambda_1 + \lambda_2 + \cdots + \lambda_n \\ &= \sum_{j=1}^{n} Var(\mathrm{PC}_j). \end{aligned} \quad (15)$$

The percentage of total variance explained by PC_j is as follows:

$$\text{Proportion of variance for the } j\text{-th} = \frac{\lambda_j}{\sum_{i=1}^{n}\lambda_i}. \quad (16)$$

If *m* PCs are chosen where $m < n$, then:

$$\text{Proportion of variance for the } m\text{-thPC} = \frac{\lambda_1+\lambda_2+\cdots+\lambda_m}{\lambda_1+\lambda_2+\cdots+\lambda_n} \quad (17)$$

If PCA is linear, then the covariance matrix of the standardized data is used, as the main diagonal of the matrix contains values of one. Thus, the total population variance for the standardized variables is *n* which is the sum of the diagonal elements of the correlation matrix (ρ), so:

$$\text{Proportion of variance for the } j\text{-th} = \frac{\lambda_j}{n} \times 100\%. \quad (18)$$

The number of PCs selected should ideally explain a high cumulative variance, for example, between 70% and 80%.

2.6 Symmetric Mean Absolute Percentage Error (SMAPE)

According to (Chicco et al., 2021), the SMAPE is a forecasting accuracy measure used to address the shortcomings of the Mean Absolute Percentage Error (MAPE), where MAPE becomes undefined when the actual value y_t is zero. The equation for MAPE is as follows:

$$\mathrm{MAPE} = \frac{1}{T}\sum_{t=1}^{T}\left|\frac{y_t - \hat{y}_t}{y_t}\right| \times 100\%, \quad (19)$$

where y_t is the actual data at the time *t* and $\hat{y}_t$ is the predicted data of y_t. This differs from SMAPE, as it can still provide results even when the actual value is zero. The equation for SMAPE is written as follows:

$$SMAPE = \frac{1}{T}\sum_{t=1}^{T}\frac{|\hat{y}_t - y_t|}{(|\hat{y}_t| + |y_t|)} \times 100\%. \quad (20)$$

SMAPE has a lower bound of 0% and an upper bound of 100%,

Table 1. Forecasting Accuracy Levels

SMAPE	Accuracy level
≤ 10%	Highly Precise Forecasting
10% < SMAPE ≤ 20%	Precise Forecasting
20% < SMAPE ≤ 50%	Moderately Precise Forecasting
> 50%	Imprecise Forecasting

DOI: https://doi.org/ 10.22452/mjs.vol45sp2.10
Malaysian Journal of Science 45 (Special Issue 2): 95-104(November 2026)

where SMAPE will automatically reach the upper limit when the actual value is zero. The evaluation values produced by SMAPE have the following criteria as stated in Table 1.

3. Methodology

This study utilizes data associated with macroeconomics, finance, and policy matters in Indonesia. All data used are numerical with a monthly frequency from February 2010 to December 2024, totaling 179 data points for each variable. The variables and data structure used in the study are presented in Tables 2 and 3, respectively.

where T represents the total count of data points utilized. The detailed study analysis is depicted in the flowchart presented in Figure 1.

4. Results and Discussion

The inflation data used in this study include inflation for food, health, housing, clothing, and education. Other variables include FSI and FCI. FSI is calculated using Equation (7), and the results are shown in Figure 2.

The FSI graph illustrates the dynamics of financial stress that occur over a certain period. An increase in FSI reflects significant stress, which can be caused by various factors such as global or domestic economic crises, market uncertainty, tight monetary policies, or external shocks like commodity price fluctuations and natural disasters. This condition is often followed by a decline in economic activity, reduced investment, and diminished market confidence. Conversely, a weakening FSI indicates reduced

Table 2. Research Variables

Variable		Description	
Inflation (y_i); i = 1,…,5	$INFL_1$ (y_1)	Food Inflation	
	$INFL_2$ (y_2)	Health Inflation	
	$INFL_3$ (y_3)	Housing Inflation	
	$INFL_4$ (y_4)	Clothing Inflation	
	$INFL_5$ (y_5)	Education Inflation	
FSI (x_1)		IDX FIN ($x_{1,1}$)	Indonesia Stock Exchange Finance Index
		JKSE ($x_{1,2}$)	Jakarta Stock Exchange
		EXC ($x_{1,3}$)	Exchange Rate (USD IDR)
		CADEV ($x_{1,4}$)	Indonesia Foreign Exchange Reserves
		YIDGB ($x_{1,5}$)	Indonesia Bond Yield
		YUST ($x_{1,6}$)	US Bond Yield
$x_{1,i}$; i =1, …6 represents the components of FSI			
FCI (x_{2j}); $j = 1,2,\dots,s$		RATE ($x_{2j,1}$)	Interest Rate
		JIBOR ($x_{2j,2}$)	Jakarta Interbank Offered Rate
		JUB ($x_{2j,3}$)	Money Supply (M_2)
		CCI ($x_{2j,4}$)	Consumer Confidence Index
$x_{2\,j,i}$; i =1, …4 represents the components of FCI. s: the number of principal components (PCs) formed, such that the cumulative variance reaches between 70% and 80%.			

Table 3. Data Structure

		Endogenous Variables					Exogenous Variables			
t	Date	$y_{1,t}$	$y_{2,t}$	$y_{3,t}$	$y_{4,t}$	$y_{5,t}$	$x_{1,t}$	$x_{21,t}$	⋯	$x_{2s,t}$
1	1/2/2010	$y_{1,1}$	$y_{2,1}$	$y_{3,1}$	$y_{4,1}$	$y_{5,1}$	$x_{1,1}$	$x_{21,1}$	⋯	$x_{2s,1}$
2	1/3/2010	$y_{1,2}$	$y_{2,2}$	$y_{3,2}$	$y_{4,2}$	$y_{5,2}$	$x_{1,2}$	$x_{21,2}$	⋯	$x_{2s,2}$
3	1/4/2010	$y_{1,3}$	$y_{2,3}$	$y_{3,3}$	$y_{4,3}$	$y_{5,3}$	$x_{1,3}$	$x_{21,3}$	⋯	$x_{2s,3}$
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋱	⋮
T	1/12/2024	$y_{1,T}$	$y_{2,T}$	$y_{3,T}$	$y_{4,T}$	$y_{5,T}$	$x_{1,T}$	$x_{21,T}$	⋯	$x_{2s,T}$

Table 4. Results of PCA Financial Condition Index (FCI)

	PC_1	PC_2	PC_3	PC_4
Standard Deviation	1.4726	0.9550	0.8947	0.34479
Proportion of Variance	0.5422	0.2280	0.2001	0.02972
Cumulative Proportion	0.5422	**0.7702**	0.9703	1

DOI: https://doi.org/ 10.22452/mjs.vol45sp2.10
Malaysian Journal of Science 45 (Special Issue 2): 95-104(November 2026)

financial stress, economic recovery, supportive monetary policies, and increased market confidence. This situation typically fosters economic growth through increased investment and consumption. Based on Figure 2, the FSI rose sharply in 2020, reflecting the impact of the COVID-19 pandemic on financial markets, followed by a significant decline during the 2021-2023 period, indicating the success of the economic recovery. However, fluctuations at the end of the graph signal the continued risk of financial stress that should still be monitored. The overall FSI graph effectively reflects the pressure on financial stability in Indonesia. Meanwhile, the PCA process based on Equation (18) to form the FCI variable yields the following results.

The results of the Principal Component Analysis (PCA) show that there are four main components (PC_1 to PC_4) with different standard deviations and variance contributions. Based on the cumulative variance, the selection of principal components was made using the criterion that the component must explain at least 70% of the total variance of the data. According to Table 4, PC_1 contributes 54.22% to the total variance. When combined with PC_2, the cumulative variance explained increases to 77.02%, exceeding the 70% threshold. Therefore, PC_1 and PC_2 were selected as the main components because they represent most of

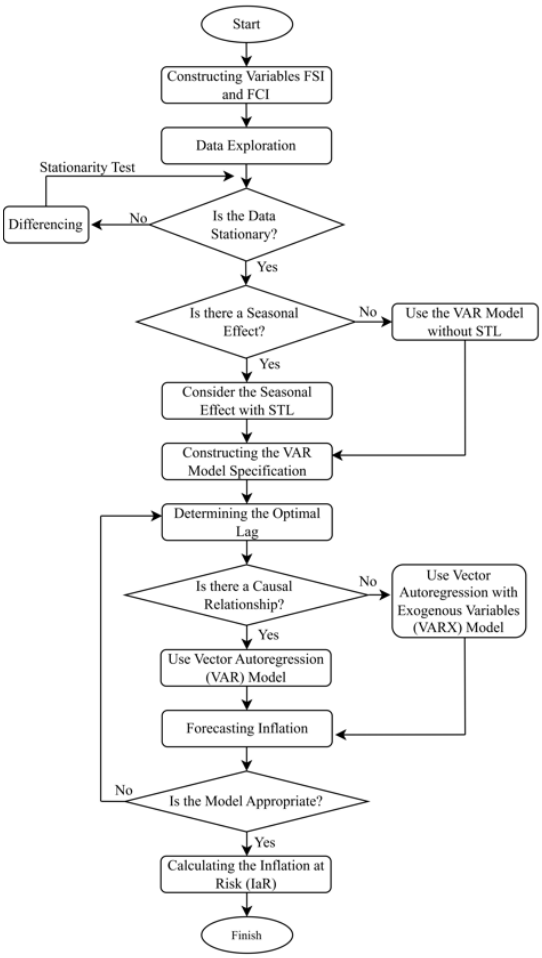


Figure 1. Step Analysis of Research

the information in the data. The combined visualization of the FSI and FCI is shown in Figure 3.

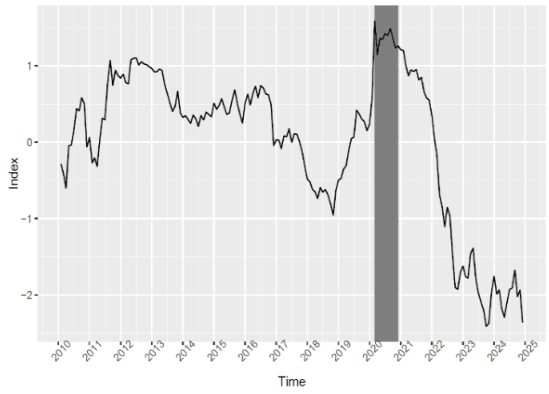


Figure 2. The Financial Stress Index (FSI) shows financial pressures in Indonesia.



Figure 3. Plot of FSI and FCI.

FCI is an index that reflects financial conditions. The visualization of FCI, represented by FCI_1 and FCI_2, illustrates the dynamics of financial conditions from February 2010 to December 2024. FSI and FCI are interrelated, where an increase in stress in financial markets tends to lead to a deterioration in financial conditions, and vice versa. Based on Figure 3, it is observed that during the 2022–2023 period, when the FSI declined in the aftermath of the COVID-19 pandemic, there was an increase in FCI as represented by FCI_2. Meanwhile, the visualization for inflation data is presented in Figure 4.

Based on the visualization of inflation data in Figure 4, it is observed that there is a tendency for seasonal patterns in the health inflation data shown in panel (b). Therefore, the Seasonal Trend Decomposition Based on LOESS (STL) method will be applied to the variable y_2. The next step is to perform a stationarity test on all variables using the Phillips-Perron test, with the following results as stated in Table 5.

With the alternative hypothesis that the data is stationary and a significance level (alpha) of 5%, it was concluded that all inflation variables are stationary, while FSI and FCI are not stationary. Therefore, differencing was performed on x_1, x_{21}, and x_{22}. After differencing, a stationarity test was conducted again, resulting in a p-value of 0.01 for all three variables, indicating that

DOI: https://doi.org/ 10.22452/mjs.vol45sp2.10
Malaysian Journal of Science 45 (Special Issue 2): 95-104(November 2026)

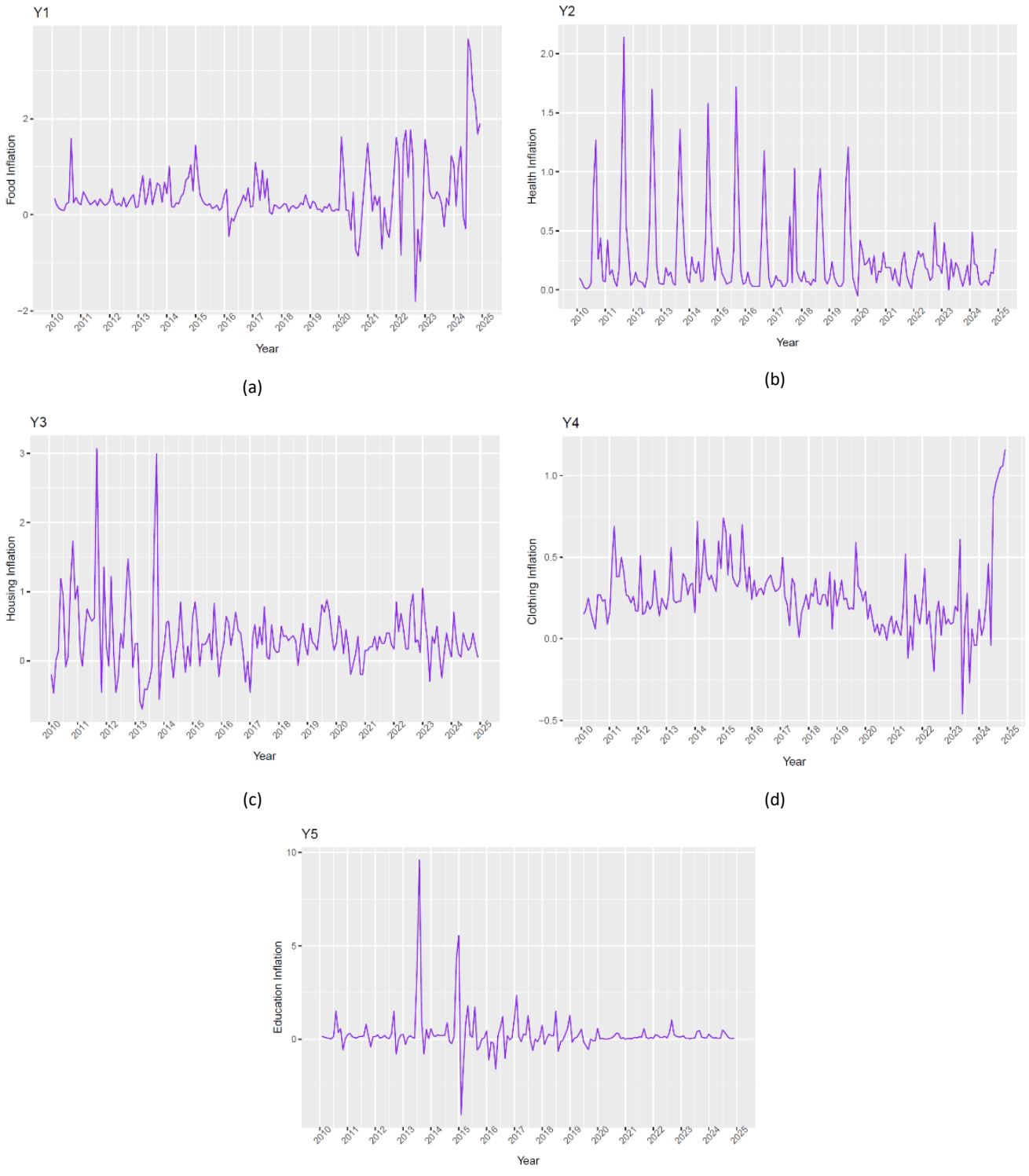


Figure 4. Plot Inflation Data

the data is now stationary. Subsequently, the data was divided into 20% out-of-sample and 80% in-sample portions. The VAR model to be used was then selected based on the lag with the smallest Correlation Information Criterion (CIC) value, as shown in Table 6.

Based on the CIC values in Table 6, it was determined that the optimal lag is in the VAR(1) model, as it has the smallest CIC value. The next step is to perform a Granger causality test to determine the relationships between variables, as presented in Table 7

DOI: https://doi.org/ 10.22452/mjs.vol45sp2.10
Malaysian Journal of Science 45 (Special Issue 2): 95-104(November 2026)

Table 5. Stationarity Test

Variable	p-value
y_1	0.01
y_2	0.01
y_3	0.01
y_4	0.01
y_5	0.01
x_1	0.7789
x_{21}	0.7253
x_{22}	0.654

Table 6. The optimal lag of the VAR model is determined based on the CIC value.

Model	CIC
VAR(1)	**127.3302**
VAR(2)	176.9530
VAR(3)	208.3535
VAR(4)	241.7554

The Granger causality test in Table 7 indicates that the FSI and FCI variables do not have a significant causal relationship with other endogenous variables. Therefore, FSI and FCI can be used as exogenous variables. Additionally, Table 7 shows that several inflation variables do not have a causal relationship with other inflation variables. As a result, a restricted VAR approach is employed to simplify the model or eliminate insignificant variables. Subsequently, the VARX(1) model is estimated as follows:

$$y_1 = 0.421y_{1,t-1} + 0.375d_{2,t-1} + 0.073y_{5,t-1} + 0.078x_{1,t} + 0.028x_{21,t} + 0.125x_{22,t}$$

$$d_2 = 0.368d_{2,t-1} + 0.475y_{4,t-1} + 0.039y_{5,t-1} + 0.226x_{1,t} + 0.036x_{21,t} + 0.13x_{22,t}$$

$$y_3 = 0.483y_{3,t-1} + 0.149y_{5,t-1} + 0.015x_{22,t}$$

$$y_4 = 0.856y_{4,t-1} + 0.078x_{1,t} + 0.004x_{21,t} - 0.042x_{22,t}$$

$$y_5 = 0.561y_{4,t-1} + 0.226y_{5,t-1} + 0.124x_{1,t}$$

To forecast food inflation, there is a positive influence from its past data, as well as health inflation, education inflation, and external factors such as FSI and FCI. Health inflation is positively influenced by its past data, as well as inflation in clothing, education, and external factors such as FSI and FCI. Housing inflation is positively influenced by its past data, education inflation, and FCI. Clothing inflation is positively influenced by its past data, FSI, FCI sub 1, and FCI sub 2 have a negative influence. Education inflation is positively influenced by its past data, clothing inflation, and external factors such as FSI. The model is then used to forecast inflation. The predictions for the in-sample data are presented in a visualized format in Figure 5.

Table 7. Granger Causality Test

	F	df_1	df_2	*p*		*Chisq*	*df*	*p*	
y_1 and y_2	4.95	1	134	0.028	*	4.95	1	0.026	*
$y_1$and y_3	0.51	1	134	0.476		0.51	1	0.475	
$y_1$and y_4	1.83	1	134	0.179		1.83	1	0.177	
$y_1$and y_5	6.07	1	134	0.015	*	6.07	1	0.014	*
$y_1$and x_1	0.01	1	134	0.924		0.01	1	0.924	
$y_1$and x_{21}	0.30	1	134	0.588		0.30	1	0.587	
$y_1$and x_{22}	0.22	1	134	0.637		0.22	1	0.636	
y_1 ALL	3.40	7	134	0.002	**	23.79	7	0.001	**
y_2 and y_1	0.23	1	134	0.629		0.23	1	0.629	
$y_2$and y_3	4.16	1	134	0.043	*	4.16	1	0.041	*
$y_2$and y_4	21.87	1	134	<0.001	***	21.87	1	<0.001	***
$y_2$and y_5	6.69	1	134	0.011	*	6.69	1	0.010	*
$y_2$and x_1	0.82	1	134	0.367		0.82	1	0.365	
$y_2$and x_{21}	1.61	1	134	0.206		1.61	1	0.204	
$y_2$and x_{22}	2.84	1	134	0.094	.	2.84	1	0.092	.
y_2 ALL	6.66	7	134	<0.001	***	46.60	7	<0.001	***
y_3 and y_1	1.17	1	134	0.281		1.17	1	0.279	
$y_3$and y_2	3.58	1	134	0.061	.	3.58	1	0.058	.
$y_3$and y_4	5.18	1	134	0.024	*	5.18	1	0.023	*
$y_3$and y_5	10.36	1	134	0.002	**	10.36	1	0.001	**
$y_3$and x_1	2.17	1	134	0.143	.	2.17	1	0.141	.
$y_3$and x_{21}	1.86	1	134	0.175		1.86	1	0.173	
$y_3$and x_{22}	1.15	1	134	0.285		1.15	1	0.283	
y_3 ALL	6.74	7	134	<0.001	***	47.17	1	<0.001	***
y_4 and y_1	0.76	1	134	0.385		0.76	1	0.383	
$y_4$and y_2	1.32	1	134	0.253		1.32	1	0.251	
$y_4$and y_3	2.26	1	134	0.135		2.26	1	0.132	
$y_4$and y_5	3.48	1	134	0.064	.	3.48	1	0.062	.
$y_4$and x_1	0	1	134	0.953		0	1	0.953	
$y_4$and x_{21}	0.29	1	134	0.588		0.29	1	0.587	
$y_4$and x_{22}	0.79	1	134	0.376		0.79	1	0.374	
y_4 ALL	1.71	7	134	0.111		12	7	0.1	
y_5 and y_1	0.01	1	134	0.931		0.01	1	0.931	
$y_5$and y_2	0	1	134	0.963		0	1	0.963	
$y_5$and y_3	1.55	1	134	0.216		1.55	1	0.213	
$y_5$and y_4	2.84	1	134	0.094	.	2.84	1	0.092	.
$y_5$and x_1	0.65	1	134	0.422		0.65	1	0.420	
$y_5$and x_{21}	1.29	1	134	0.259		1.29	1	0.257	
$y_5$and x_{22}	0.24	1	134	0.622		0.24	1	0.622	
y_5 ALL	1.48	7	134	0.179		10.36	7	0.169	
x_1 and y_1	2.35	1	134	0.128		2.35	1	0.125	
$x_1$and y_2	0.82	1	134	0.367		0.82	1	0.366	
$x_1$and y_3	0.56	1	134	0.457		0.56	1	0.455	
$x_1$and y_4	0.15	1	134	0.701		0.15	1	0.701	
$x_1$and y_5	0.05	1	134	0.828		0.05	1	0.828	
$x_1$and x_{21}	2.80	1	134	0.096	.	2.80	1	0.094	.
$x_1$and x_{22}	3.20	1	134	0.076	.	3.20	1	0.074	.
x_1 ALL	1.02	7	134	0.417		7.17	7	0.411	
x_{21} and y_1	0.20	1	134	0.653		0.20	1	0.652	
$x_{21}$and y_2	0	1	134	0.997		0	1	0.997	
$x_{21}$and y_3	1.29	1	134	0.259		1.29	1	0.257	
$x_{21}$and y_4	0.12	1	134	0.730		0.12	1	0.730	
$x_{21}$and y_5	0.22	1	134	0.643		0.22	1	0.643	
$x_{21}$and x_1	2.97	1	134	0.087	.	2.97	1	0.085	.
$x_{21}$and x_{22}	2.45	1	134	0.120		2.45	1	0.117	
x_{21} ALL	1.48	7	134	0.179		10.38	1	0.168	
x_{22} and y_1	0.15	1	134	0.695		0.15	1	0.718	
$x_{22}$and y_2	2.26	1	134	0.135		2.26	1	0.182	
$x_{22}$and y_3	0.09	1	134	0.769		0.09	1	0.407	
$x_{22}$and y_4	0.88	1	134	0.349		0.88	1	0.339	
$x_{22}$and y_5	0.27	1	134	0.604		0.27	1	0.630	
$x_{22}$and x_1	0	1	134	0.982		0	1	0.960	
$x_{22}$and x_{21}	24.17	1	134	<0.001	***	24.17	1	<0.001	***
x_{22} ALL	4.09	1	134	<0.01	***	28.64	7	<0.001	***

DOI: https://doi.org/ 10.22452/mjs.vol45sp2.10
Malaysian Journal of Science 45 (Special Issue 2): 95-104(November 2026)

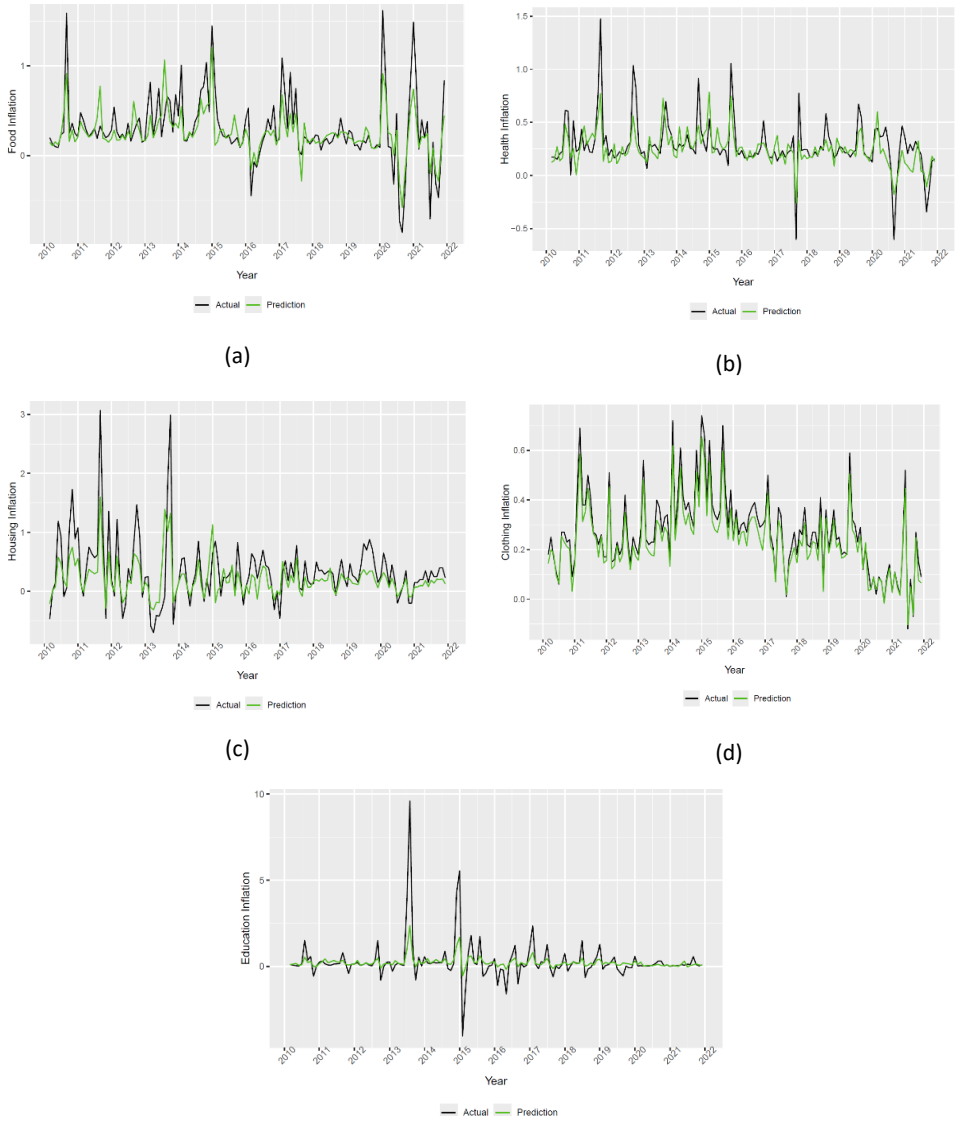


Figure 5. Forecasting In-Sample of Inflation

Figure 5 presents the forecasting results for inflation data during the in-sample period from February 2010 to December 2021, consisting of 143 data points or 80% of the initial dataset. The black line represents the actual inflation data, while the green line represents the model's predictions. Overall, the model accurately follows the actual data patterns. However, it struggles to capture extreme fluctuations. The in-sample forecasting accuracy results are shown in Table 8.

Table 8. Forecasting accuracy based on the SMAPE criterion.

Variable	SMAPE
y_1	21.5%
y_2	21.3%
y_3	38%
y_4	8.8%
y_5	49.7%

Based on the SMAPE criteria in Table 1, it is found that the forecast for clothing inflation is highly accurate, while forecasts for food inflation, health inflation, housing inflation, and education inflation are moderately accurate. Next, the forecast for the out-of-sample data from January 2022 to December 2024, consisting of 35 data points or 20% of the last data, is presented in Figure 6.

Table 9. Forecasting accuracy based on the SMAPE criterion.

Variable	SMAPE
y_1	40%
y_2	44.5%
y_3	27.7%
y_4	19%
y_5	49.6%

DOI: https://doi.org/ 10.22452/mjs.vol45sp2.10
Malaysian Journal of Science 45 (Special Issue 2): 95-104(November 2026)

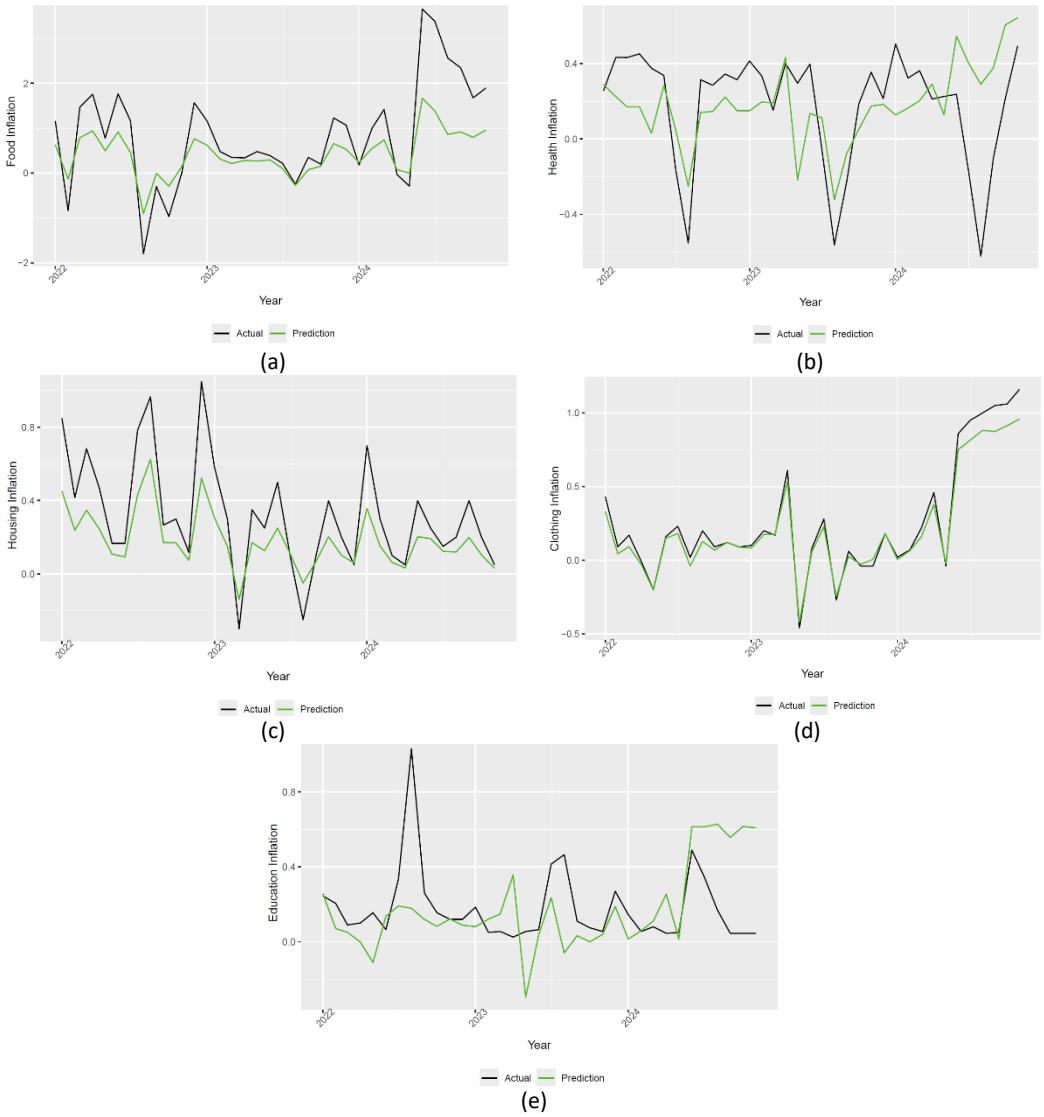


Figure 6. Forecasting the Out-of-Sample of Inflation

In general, the model is able to follow the actual data patterns well; however, it struggles to capture extreme fluctuations. The out-of-sample forecasting accuracy is presented in Table 9.

Based on the SMAPE forecasting criteria in Table 1, it is found that the forecast for clothing inflation is accurate, food inflation, health inflation, housing inflation, and education inflation are moderately accurate.

5. Conclusions

Based on the analysis results, the main finding of this study indicates that the VARX model is effective for predicting IaR in Indonesia. Generally, the forecast results follow the pattern of the actual data. The VARX model performs exceptionally well when the inflation data being analyzed exhibits relatively stable fluctuations, as seen in the clothing inflation component. In such cases, the model successfully captures the dynamic relationships between variables, as reflected in forecast visualizations that closely match the actual data and accurate SMAPE values. However, the model's effectiveness decreases when applied to data with extreme fluctuations, such as the more volatile education inflation component.

6. Acknowledgments

I would like to express my deepest appreciation to *Lembaga Pengelola Dana Pendidikan* (LPDP) for their scholarship support. I would also like to express my gratitude to the reviewers, particularly Wisnowan Hendy Saputra, for his support throughout this research.

7. References

Akbar, M. S., Setiawan, S., Suhartono, S., Ruchjana, B. N., Prastyo, D. D., Muhaimin, A., & Setyowati, E. (2020). A Generalized Space-Time Autoregressive Moving Average (GSTARMA) Model for Forecasting Air Pollutant in Surabaya. J. Phys.: Conf. Ser. 1490 (012022).

DOI: https://doi.org/ 10.22452/mjs.vol45sp2.10
Malaysian Journal of Science 45 (Special Issue 2): 95-104(November 2026)

Chicco, D., Warrens, M. J., & Jurman, G. (2021). The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. PeerJ Computer Science, 18.

Cleveland, R. B., Cleveland, W. S., McRae, J. E., & Terpenning, I. (1990). STL: A Seasonal-Trend Decomposition Procedure Based on Loess. Journal of Official Statistics, 6(1), 3-73.

Farih, I., & Prastyo, D. D. (2022). Forecasting Electricity Consumption Based on Economics and Social Indicators Using VAR Model with Exogenous Variable: Evidence From East Java Province. in 022 6th International Conference on Information Technology, Information Systems and Electrical Engineering (ICITISEE). Yogyakarta.

Ferrara, L., Mogliani, M., & Sahuc, J.-G. (2022). High-frequency monitoring of growth at risk. International Journal of Forecasting, 38, 582-595.

Gujarati, D. N. (2003). Basic Econometrics. New York: Gary Burke.

Hikmawati, F., Suhartono, S., & Prastyo, D. D. (2021). A Hybrid GSTARX-Jordan RNN Model for Forecasting Space-Time Data with Calendar Variation Effect. J. Phys.: Conf. Ser. 1752 (012013).

Juhro, S. M., & Iyke, B. N. (2019). Monetary Policy and Financial Condition in Indonesia. Bulletin of Monetary Economics and Banking, 21, 283-302.

Novianto, M. A., Suhartono, S., Prastyo, D. D., Suharsono, A., & Setiawan, S. (2018). GSTARIX Model for Forecasting Spatio-Temporal Data with Trend, Seasonal and Intervention. J. Phys.: Conf. Ser. 1097 (012076).

Ocampo, S., & Rodríguez, N. (2012). An Introductory Review of a Structural VAR-X Estimation and Applications. Revista Colombiana de Estadística, 35, 479-508.

Park, C.-Y., & Mercado, R. V. (2014). Determinants of financial stress in emerging market economies. Journal of Banking & Finance, 45, 199-224.

Prastyo, D. D., Nabila, F. S., Suhartono, S., Lee, M. H., & Suhermi, N. F.-F. (2018). VAR and GSTAR-Based Feature Selection in Support Vector Regression for Multivariate Spatio-Temporal Forecasting. In Soft Computing in Data Science (SCDS), Communications in Computer and Information Science, vol 937. Springer.

Safuan, S., Sugandi, E. A., Aziz, O. Q., & Triandhari, R. (2022). Indonesia's Financial Stress Events and Macroeconomic Dynamics. Bulletin of Monetary Economics and Banking, 25, 323-370.

Sesay, A., Suhartono, S., & Prastyo, D. D. (2020). Forecasting Exchange Rate Across Countries with Gold Price as Exogenous Variable Using Transfer Function and VARI-X Model. MATEMATIKA, 36(3), 181-196.

Suhartono, S., Gazali, M. M., & Prastyo, D. D. (2018). VARX and GSTARX Models for Forecasting Currency Inflow and Outflow with Multiple Calendar Variations Effect. MATEMATIKA, 34(3), 57–72.

Suhartono, S., Prastyo, D. D., Kuswanto, H., & Lee, M. H. (2018). Comparison between VAR, GSTAR, FFNN-VAR and FFNN-GSTAR Models for Forecasting Oil Production. MATEMATIKA, 34(1), 103–111.

Tsay, R. S. (2002). Analysis of Financial Time Series. Canada: John Wiley & Sons, Inc.

Zivot, E., & Wang, J. (2003). Modeling Financial Time Series with S-Plus. New York: Springer Science+Bu siness Media.

DOI: https://doi.org/ 10.22452/mjs.vol45sp2.10
Malaysian Journal of Science 45 (Special Issue 2): 95-104(November 2026)