

Article

Addition of Subset and Dummy Variables in the Threshold Spatial Vector Autoregressive with Exogenous Variables Model to Forecast Inflation and Money Outflow

Setiawan Setiawan ^{1,*}, Gama Putra Danu Sohibien ², Dedy Dwi Prastyo ¹, Muhammad Sjahid Akbar ¹ and Anton Abdulbasah Kamil ³

¹ Department of Statistics, Faculty of Science and Data Analytics, Institut Teknologi Sepuluh Nopember, Surabaya 60119, Indonesia; dedy-dp@statistika.its.ac.id (D.D.P.); m_sjahid_a@its.ac.id (M.S.A.)

² Department of Applied Statistics, Politeknik Statistika STIS, Jakarta 13320, Indonesia; gamaputra@stis.ac.id

³ Faculty of Economics, Istanbul Gelisim University, Istanbul 34315, Turkey; akamil@gelisim.edu.tr

* Correspondence: setiawan@statistika.its.ac.id

Abstract: The TSpVARX model can be used in inflation and money outflow forecasting by accommodating the reciprocal relationship among endogenous variables, the influence of exogenous variables, inter-regional linkages, and the nonlinearity of the relationship between endogenous and predetermined variables. However, the impact of some events, such as Eid al-Fitr and fuel price adjustments, still cannot be accommodated in the TSpVARX model. This condition causes inflation and money outflow forecasting using TSpVARX to be unsatisfactory. Our study is to improve the forecasting performance of the TSpVARX model by adding subset and dummy variables. We use a 12th lag subset variable to capture seasonal effects and a dummy variable to represent fuel price changes. These additions enhance the model's accuracy in forecasting inflation and money outflow by accounting for recurring patterns and specific events, like fuel price changes. Based on the RMSE values of the training and testing data, we can conclude that forecasting inflation and money outflow using TSpVARX with the addition of subset and dummy variables is better than the regular TSpVARX. The inflation and money outflow forecasting generated after the addition of subset and dummy variables are also more fluctuating as in the movement of the actual data.

Keywords: nonlinear time series; threshold spatial; SpVAR; inflation; subset variable; TSpVARX



Citation: Setiawan, Setiawan, Gama Putra Danu Sohibien, Dedy Dwi Prastyo, Muhammad Sjahid Akbar, and Anton Abdulbasah Kamil. 2024. Addition of Subset and Dummy Variables in the Threshold Spatial Vector Autoregressive with Exogenous Variables Model to Forecast Inflation and Money Outflow. *Economics* 12: 352. <https://doi.org/10.3390/economics12120352>

Academic Editor: Stefan Collignon

Received: 18 October 2024

Revised: 29 November 2024

Accepted: 4 December 2024

Published: 19 December 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Time series models used to forecast inflation and money outflow data include Autoregressive Integrated Moving Average with exogenous variables (ARIMAX), Vector Autoregressive with exogenous variables (VARX), Generalized Space-Time Autoregressive (GSTAR), Threshold Vector Autoregressive (TVAR), Spatial Vector Autoregressive (SpVAR), and Threshold Spatial Vector Autoregressive with exogenous variables (TSpVARX). ARIMAX is a univariate model that can accommodate the influence of predetermined variables in the form of lag endogenous variables and exogenous variables (Wei 2006; Lestari and Dini 2024). The use of the ARIMAX model in inflation modeling is performed to accommodate the influence of several variables, such as fuel increases, outliers, and interest rates.

VARX is a multivariate model that can accommodate the impact of exogenous variables and the reciprocal relationship between endogenous variables (Lütkepohl 2005; Tsay 2014). Two variables that have a mutual relationship but are modeled with a single model will result in biased coefficient estimation (Stock and Watson 2020). VARX models can be used to model inflation and money outflow because inflation and money outflow have a reciprocal relationship (Adelowokan et al. 2019; Bello and Saulawa 2013). However, the VARX model does not include spatial weights that serve to distinguish the size of

the inter-regional relationship. A model that can accommodate the linkage of time series variables between regions is the GSTAR model (Sohibien 2017; Fadlurrohman 2020). Sohibien (2017) found that GSTAR with normalized cross-correlation weights is well used in forecasting the inflation of Dumai and Pekanbaru. Fadlurrohman (2020) found that GSTARX with uniform weights is better at forecasting than GSTARX with inverse distance weights. A limitation of the GSTARX model is its inability to accommodate the relationship between endogenous variables for more than one type of endogenous variable. The SpVAR model can accommodate inter-regional linkages for more than one variable (Beenstock and Felsenstein 2007; Di Giacinto 2010; Sohibien et al. 2024a).

The models described above are less suitable for nonlinear patterns of relationships between endogenous and predetermined variables. The TVAR model can capture the nonlinearity of the relationship between time series variables by forming several VAR model regimes based on a predetermined threshold value (Tsagkanos et al. 2018). Although TVAR has been able to capture the nonlinearity of the relationship between endogenous and predetermined variables, the TVAR model has not been able to accommodate the interrelationship of time series variables between regions. The TSpVARX model captures complex interactions by accounting for relationships among variables, external factors, past data points, regional connections, and nonlinear patterns (Sohibien et al. 2024b). Models, such as ARIMAX, VARX, GSTARX, and SpVARX, can only be used to analyze linear relationship patterns between endogenous variables and predetermined variables. Meanwhile, the TSpVARX model can already analyze non-linear relationship patterns between endogenous variables and predetermined variables. Sohibien et al. (2024b) found that when there is a nonlinear relationship between endogenous variables and predetermined variables, the forecasting performance of the TSpVARX model is better than SpVARX.

Another approach that can be used in handling nonlinear relationships is to use machine learning approaches, such as Neural Networks (NNs) and Support Vector Machine (SVM). With the flexibility of machine learning, it can approach any function so that nonlinear patterns between variables can be learned (Bharadiya 2023; Shoko and Sigauke 2023). However, the TSpVARX model has better advantages over machine learning approaches. TSpVARX provides clear rules about when the relationship between variables changes, for example when a variable crosses a certain threshold. This makes it easier for users to understand than machine learning models. In addition, the structure of the TSpVARX model is also simpler than the machine learning approach. In TSpVARX, the nonlinearity of the relationship is shown by a different model between models below the threshold and above the threshold. It causes the TSpVARX model to also be used for interpretation in understanding the relationship between variables. Meanwhile, the complexity of the machine learning approach means that it cannot be used to interpret the relationship between variables.

Although the TSpVARX model has been able to accommodate many things, it still has shortcomings. The TSpVARX model of Sohibien et al. (2024b) cannot accommodate the influence of events with recurring patterns and events that include nonmetric variables. This is still a weakness if the TSpVARX model is used to forecast inflation and money outflow because the inflation and money outflow are influenced by events with recurring patterns (in this case the Eid al-Fitr event) and events that include nonmetric variables (in this case the increase and decrease in fuel prices). In addition to producing more accurate forecasts, the TSpVARX model with dummy and subset variables also has the potential to be applied in the real world, such as analyzing the impact of fuel price increases and decreases on inflation, the effect of holidays on inflation, and future economic planning based on the threshold value of the TSpVARX model.

Therefore, in this study, we propose a TSpVARX with the addition of the 12th subset and dummy variables. Subset variables are lags of endogenous variables included in the model to capture recurring patterns for any given period. Dummy variables are categorical variables that are coded into numerical form (usually 0 and 1) to represent groups in the

data so that differences in patterns between data groups can be captured by the model. We include the 12th lag of the endogenous variable as the 12th subset variable to capture the effect of the Eid al-Fitr event on inflation and money outflow. Meanwhile, we include fuel price increases and decreases as dummy variables to accommodate the impact of fuel price adjustments on inflation and money outflow. The developed model will be applied to the forecasting of inflation and money outflow of Yogyakarta, Solo, and Semarang.

2. Methodology

In this section, we will explain three things, namely the first subsection regarding the dataset, the second subsection regarding the SpVARX model, the third subsection regarding the TSpVARX model, and the fourth subsection regarding the development of the TSpVARX model with the addition of the 12th subset and dummy variables. In the first subsection, we describe the data and variables used in this study. We also explain the categories of dummy variables used in this study. In the second subsection, we explain the general form of the SpVARX model and how to estimate the coefficients of the SpVARX model. In the third subsection, we explain how the TSpVARX model is formed based on the selected threshold values and variables along with the steps to estimate the TSpVARX model coefficients. In the fourth subsection, we explain how the general form of the TSpVARX model with the addition of the 12th subset and dummy variables to accommodate the existence of a pattern that repeats every 12 months and the effect of increasing and decreasing fuel prices. We will also explain the steps to estimate the model coefficients.

2.1. Dataset

The data used in this study are monthly data from January 2006 to September 2024 for the BI benchmark interest rate, the exchange rate of the rupiah against the US dollar, and Inflation of Semarang, Solo, and Yogyakarta. Besides that, we also use the money outflow data of Semarang, Solo, and Yogyakarta from January 2006 to May 2023. Inflation of Semarang, Solo, and Yogyakarta are from Statistics Indonesia. Meanwhile, the money outflow of Semarang, Solo, and Yogyakarta, the BI benchmark interest rate, and the exchange rate of the rupiah against the US dollar are from Bank Indonesia. The model used for forecasting is built using data from January 2006 to May 2023. Further forecasting will be carried out from June 2023 to December 2024.

The variables used in this study are inflation of Semarang (Y_{1t}^1), inflation of Solo (Y_{1t}^2), inflation of Yogyakarta (Y_{1t}^3), money outflow of Semarang (Y_{2t}^1), money outflow of Solo (Y_{2t}^2), money outflow of Yogyakarta (Y_{2t}^3), depreciation of the rupiah against the United States dollar at $t - i$ ($X_{1,t-i}$), BI benchmark interest rate at $t - i$ ($X_{2,t-i}$), the 12th subset variable of Semarang inflation ($Y_{1,t-12}^1$), the 12th subset variable of Solo inflation ($Y_{1,t-12}^2$), the 12th subset variable of Yogyakarta inflation ($Y_{1,t-12}^3$), the 12th subset variable of Semarang money outflow ($Y_{2,t-12}^1$), the 12th subset variable of Solo money outflow ($Y_{2,t-12}^2$), and the 12th subset variable of Yogyakarta money outflow ($Y_{2,t-12}^3$).

It is important to add dummy variables to the TSpVARX model, to capture the effect of rising and falling fuel prices on inflation. When there is an increase in fuel prices, the prices of other goods also increase. This causes an increase in inflation. Conversely, when there is a decrease in fuel prices, inflation decreases. Meanwhile, the addition of the 12th subset variable is aimed at capturing inflation patterns that recur every 12 months caused by events, such as Eid al-Fitr, Christmas, and New Year holidays.

The subset variable in this study is an endogenous lag variable used in the model. This variable is different from the autoregressive variable, where if using autoregressive order 3 (AR (3)), then the endogenous lag variable used is the 1st, 2nd, and 3rd lags of the endogenous variable. In the subset variable, we only use one lag of the endogenous variable, without including the other lags of the endogenous variable. In this study, we use the 12th lag of the endogenous variable for our subset variable. The addition of the

12th lag of the endogenous variable aims to capture the fluctuation pattern of inflation and money outflow caused by events that recur every year (12 months), such as Eid, Christmas, and New Year.

There are two types of dummy variables that we input in the modeling, namely the dummy variable of whether or not the fuel price increases and the dummy variable of whether or not the fuel price decreases. With these dummy variables, fluctuations in inflation and money outflow caused by increases and decreases in fuel prices can be accommodated by the model. The determination of dummy variable data of the fuel price adjustment is as follows:

$$R_t = \begin{cases} 1, & \text{if the fuel price increase occurs before the 15th day of month } t \text{ or after the 15th day of month } t-1 \\ 0, & \text{for others} \end{cases}$$

$$S_t = \begin{cases} 1, & \text{if the fuel price decrease occurs before the 15th day of month } t \text{ or after the 15th day of month } t-1 \\ 0, & \text{for others} \end{cases}$$

Semarang, Solo, and Yogyakarta are the locus of our research because they are geographically close, have easy access to each other, and are included in the same Bank Indonesia region. In addition, the three cities are known as the Golden Triangle area because they are the center of economic development in Central Java both from the tourism and industrial sectors. This allows for linkages in terms of economic activity between these three cities.

2.2. The SpVARX Model

The SpVARX model with spatial order one, temporal lag order p , and exogenous variable lag order q can be written as SpVARX(1, p , q). The general form of the SpVARX model is as follows:

$$\begin{aligned} Y_{1,t}^1 = & \alpha_{10}^1 + \gamma_{11}^{1,(1)} X_{1,t-1}^1 + \cdots + \gamma_{11}^{1,(q)} X_{1,t-q}^1 + \cdots + \gamma_{1m}^{1,(i)} X_{m,t-1}^1 + \cdots + \gamma_{1M}^{1,(1)} X_{M,t-1}^1 + \cdots + \gamma_{1M}^{1,(q)} X_{M,t-q}^1 \\ & + \phi_{11}^{1,(1,0)} Y_{1,t-1}^1 + \phi_{11}^{1,(1,1)} \left(w_{11(1,2)} Y_{1,t-1}^2 + \cdots + w_{11(1,u)} Y_{1,t-1}^u + \cdots + w_{11(1,N)} Y_{1,t-1}^N \right) + \cdots + \phi_{1r}^{1,(j,0)} Y_{r,t-1}^1 \\ & + \phi_{1r}^{1,(j,1)} \left(w_{1r(1,2)} Y_{r,t-1}^2 + \cdots + w_{1r(1,u)} Y_{r,t-1}^u + \cdots + w_{1r(1,N)} Y_{r,t-1}^N \right) + \cdots + \phi_{1K}^{1,(p,0)} Y_{K,t-p}^1 \\ & + \phi_{1K}^{1,(p,1)} \left(w_{1K(1,2)} Y_{K,t-p}^2 + \cdots + w_{1K(1,u)} Y_{K,t-p}^u + \cdots + w_{1K(1,N)} Y_{K,t-p}^N \right) + \varepsilon_{1t}^1 \end{aligned} \quad (1a)$$

$$\begin{aligned} Y_{1,t}^2 = & \alpha_{20}^2 + \gamma_{21}^{2,(1)} X_{1,t-1}^2 + \cdots + \gamma_{21}^{2,(q)} X_{1,t-q}^2 + \cdots + \gamma_{2m}^{2,(i)} X_{m,t-1}^2 + \cdots + \gamma_{2M}^{2,(1)} X_{M,t-1}^2 + \cdots + \gamma_{2M}^{2,(q)} X_{M,t-q}^2 \\ & + \phi_{21}^{2,(1,0)} Y_{1,t-1}^2 + \phi_{21}^{2,(1,1)} \left(w_{21(2,1)} Y_{1,t-1}^1 + w_{21(2,3)} Y_{1,t-1}^3 + \cdots + w_{21(2,u)} Y_{1,t-1}^u + \cdots + w_{21(2,N)} Y_{1,t-1}^N \right) + \cdots + \phi_{2r}^{2,(j,0)} Y_{r,t-1}^2 \\ & + \phi_{2r}^{2,(j,1)} \left(w_{2r(2,1)} Y_{r,t-1}^1 + w_{2r(2,3)} Y_{r,t-1}^3 + \cdots + w_{2r(2,u)} Y_{r,t-1}^u + \cdots + w_{2r(2,N)} Y_{r,t-1}^N \right) + \cdots + \phi_{2K}^{2,(p,0)} Y_{K,t-p}^2 \\ & + \phi_{2K}^{2,(p,1)} \left(w_{2K(2,1)} Y_{K,t-p}^1 + w_{2K(2,3)} Y_{K,t-p}^3 + \cdots + w_{2K(2,u)} Y_{K,t-p}^u + \cdots + w_{2K(2,N)} Y_{K,t-p}^N \right) + \varepsilon_{2t}^2 \end{aligned} \quad (1b)$$

⋮

$$\begin{aligned} Y_{k,t}^n = & \alpha_{k0}^n + \gamma_{k1}^{n,(1)} X_{1,t-1}^n + \cdots + \gamma_{k1}^{n,(q)} X_{1,t-q}^n + \cdots + \gamma_{km}^{n,(i)} X_{m,t-1}^n + \cdots + \gamma_{kM}^{n,(1)} X_{M,t-1}^n + \cdots + \gamma_{kM}^{n,(q)} X_{M,t-q}^n \\ & + \phi_{k1}^{n,(1,0)} Y_{1,t-1}^n + \phi_{k1}^{n,(1,1)} \left(w_{k1(n,1)} Y_{1,t-1}^1 + \cdots + w_{k1(n,n-1)} Y_{1,t-1}^{n-1} + w_{k1(n,n+1)} Y_{1,t-1}^{n+1} + \cdots + w_{k1(n,N)} Y_{1,t-1}^N \right) + \cdots + \phi_{kr}^{n,(j,0)} Y_{r,t-1}^n \\ & + \phi_{kr}^{n,(j,1)} \left(w_{kr(n,1)} Y_{r,t-1}^1 + \cdots + w_{kr(n,n-1)} Y_{r,t-1}^{n-1} + \cdots + w_{kr(n,n+1)} Y_{r,t-1}^{n+1} + \cdots + w_{kr(n,N)} Y_{r,t-1}^N \right) + \cdots + \phi_{kK}^{n,(p,0)} Y_{K,t-p}^n \\ & + \phi_{kK}^{n,(p,1)} \left(w_{kK(n,1)} Y_{K,t-p}^1 + \cdots + w_{kK(n,n-1)} Y_{K,t-p}^{n-1} + w_{kK(n,n+1)} Y_{K,t-p}^{n+1} + \cdots + w_{kK(n,N)} Y_{K,t-p}^N \right) + \varepsilon_{kt}^n \end{aligned} \quad (1c)$$

⋮

$$\begin{aligned} Y_{K,t}^N = & \alpha_{K0}^N + \gamma_{K1}^{N,(1)} X_{1,t-1}^N + \cdots + \gamma_{K1}^{N,(q)} X_{1,t-q}^N + \cdots + \gamma_{Km}^{N,(i)} X_{m,t-1}^N + \cdots + \gamma_{KM}^{N,(1)} X_{M,t-1}^N + \cdots + \gamma_{KM}^{N,(q)} X_{M,t-q}^N \\ & + \phi_{K1}^{N,(1,0)} Y_{1,t-1}^N + \phi_{K1}^{N,(1,1)} \left(w_{K1(N,1)} Y_{1,t-1}^1 + \cdots + w_{K1(N,u)} Y_{1,t-1}^u + \cdots + w_{K1(N,N-1)} Y_{1,t-1}^{N-1} \right) + \cdots + \phi_{Kr}^{N,(j,0)} Y_{r,t-1}^N \\ & + \phi_{Kr}^{N,(j,1)} \left(w_{Kr(N,1)} Y_{r,t-1}^1 + \cdots + w_{Kr(N,u)} Y_{r,t-1}^u + \cdots + w_{Kr(N,N-1)} Y_{r,t-1}^{N-1} \right) + \cdots + \phi_{KK}^{N,(p,0)} Y_{K,t-p}^N \\ & + \phi_{KK}^{N,(p,1)} \left(w_{KK(N,1)} Y_{K,t-p}^1 + \cdots + w_{KK(N,u)} Y_{K,t-p}^u + \cdots + w_{KK(N,N-1)} Y_{K,t-p}^{N-1} \right) + \varepsilon_{Kt}^N \end{aligned} \quad (1d)$$

where

$X_{m,t-i}^n$ is the m -th metric exogenous variable at the n -th location in the period $t-i$,

$\gamma_{km}^{n,(i)}$ is the coefficient of variable $X_{m,t-i}^n$ in equation Y_{kt}^n ,

M is the number of exogenous variables,

j is the lag order of the autoregressive,

and i is the lag order of the metric exogenous variable from 1 to q .

The SpVARX model in Equations (1a) to (1d) can be made into vector and matrix operations as follows:

$$\begin{bmatrix} Y_{1,t}^1 \\ Y_{1,t}^2 \\ \vdots \\ Y_{k,t}^n \\ \vdots \\ Y_{K,t}^N \end{bmatrix}_{NK \times 1} = \begin{bmatrix} (\mathbf{x}_{1,t}^1)' & (\mathbf{v}_{1,t}^1)' & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & (\mathbf{x}_{1,t}^2)' & (\mathbf{v}_{1,t}^2)' & \cdots & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & (\mathbf{x}_{k,t}^n)' & (\mathbf{v}_{k,t}^n)' & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & (\mathbf{x}_{K,t}^N)' & (\mathbf{v}_{K,t}^N)' \end{bmatrix}_{NK \times NK(1+Mq+2Kp)} \begin{bmatrix} \gamma_1^1 \\ \Phi_1^1 \\ \gamma_1^2 \\ \Phi_1^2 \\ \vdots \\ \gamma_k^n \\ \Phi_k^n \\ \vdots \\ \gamma_K^N \\ \Phi_K^N \end{bmatrix} + \begin{bmatrix} e_{1,t}^1 \\ e_{1,t}^2 \\ \vdots \\ e_{k,t}^n \\ \vdots \\ e_{K,t}^N \end{bmatrix} \quad (2)$$

where:

$$(\mathbf{x}_{k,t}^n)' = \begin{bmatrix} 1 & X_{1,t-1}^n & \cdots & X_{1,t-q}^n & \cdots & X_{m,t-i}^n & \cdots & X_{M,t-1}^n & \cdots & X_{M,t-q}^n \end{bmatrix}_{1 \times (1+Mq)}$$

$$(\mathbf{v}_{k,t}^n)' = \begin{bmatrix} Y_{1,t-1}^n & Y_{k1,t-1}^{n*} & \cdots & Y_{r,t-j}^n & Y_{kr,t-j}^{n*} & \cdots & Y_{K,t-p}^n & Y_{kK,t-p}^{n*} \end{bmatrix}_{1 \times 2Kp}$$

$$Y_{kr,t-j}^{n*} = \sum_{\substack{u=1 \\ u \neq n}}^N w_{kr}(n, u) Y_{r,t-j}^u$$

$$\gamma_k^n = \begin{bmatrix} \alpha_{k0}^n & \gamma_{k1}^{n(1)} & \cdots & \gamma_{k1}^{n(q)} & \cdots & \gamma_{km}^{n(i)} & \cdots & \gamma_{kM}^{n(1)} & \cdots & \gamma_{kM}^{n(q)} \end{bmatrix}'_{(1+Mq) \times 1}$$

$$\Phi_k^n = \begin{bmatrix} \phi_{k1}^{n,(1,0)} & \phi_{k1}^{n,(1,1)} & \cdots & \phi_{kr}^{n,(j,0)} & \phi_{kr}^{n,(j,1)} & \cdots & \phi_{kK}^{n,(p,0)} & \phi_{kK}^{n,(p,1)} \end{bmatrix}'_{2Kp \times 1}$$

$$\text{Let } \mathbf{Y} = \begin{bmatrix} Y_{1,h+1}^1 & Y_{1,h+1}^2 & \cdots & Y_{k,h+1}^n & \cdots & Y_{K,h+1}^N \\ Y_{1,h+2}^1 & Y_{1,h+2}^2 & \cdots & Y_{k,h+2}^n & \cdots & Y_{K,h+2}^N \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ Y_{1,T}^1 & Y_{1,T}^2 & \cdots & Y_{k,T}^n & \cdots & Y_{K,T}^N \end{bmatrix} \text{ with size } (T-h) \times NK,$$

$$\mathbf{Z}_k^n = \begin{bmatrix} (\mathbf{x}_{k,h+1}^n)' & (\mathbf{v}_{k,h+1}^n)' \\ (\mathbf{x}_{k,h+2}^n)' & (\mathbf{v}_{k,h+2}^n)' \\ \vdots & \vdots \\ (\mathbf{x}_{k,T}^n)' & (\mathbf{v}_{k,T}^n)' \end{bmatrix}, \mathbf{Z} = \begin{bmatrix} \mathbf{Z}_1^1 & \mathbf{0} & \cdots & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{Z}_1^2 & \cdots & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{Z}_k^n & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \cdots & \mathbf{Z}_K^N \end{bmatrix}, \mathbf{b}_k^n = \begin{bmatrix} \gamma_k^n \\ \Phi_k^n \end{bmatrix},$$

$$\boldsymbol{\beta} = \begin{bmatrix} \mathbf{b}_1^1 \\ \mathbf{b}_1^2 \\ \vdots \\ \mathbf{b}_k^n \\ \vdots \\ \mathbf{b}_K^N \end{bmatrix}, \text{ and } \boldsymbol{\xi} = \begin{bmatrix} e_{1,h+1}^1 & e_{1,h+1}^2 & \cdots & e_{k,h+1}^n & \cdots & e_{K,h+1}^N \\ e_{1,h+2}^1 & e_{1,h+2}^2 & \cdots & e_{k,h+2}^n & \cdots & e_{K,h+2}^N \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ e_{1,T}^1 & e_{1,T}^2 & \cdots & e_{k,T}^n & \cdots & e_{K,T}^N \end{bmatrix} \text{ then if the number of}$$

observations used for the model is T and $h = \max(p, q)$, then the form of the SpVARX model in Equation (2) can be described as follows:

$$\text{Vec}(\mathbf{Y}) = \mathbf{Z}\boldsymbol{\beta} + \text{Vec}(\boldsymbol{\xi}), \quad (3)$$

where

Vec is an operator that stacks a matrix as a column vector,

$\text{Vec}(\mathbf{Y})$ is a vector of size $NK(T-h) \times 1$ obtained from stacking $\mathbf{Y}$,

$\text{Vec}(\boldsymbol{\xi})$ is a vector of size $NK(T-h) \times 1$ obtained from stacking $\boldsymbol{\xi}$,

$\mathbf{Z}$ is a matrix of size $NK(T-h) \times NK(1 + Mq + 2Kp)$,

$\boldsymbol{\beta}$ is a vector of size $NK(1 + Mq + 2Kp) \times 1$.

There are several steps to estimate the coefficient parameters of the SpVARX model with the addition of the 12th subset and dummy variables using the MLE method (Sohibien et al. 2024a, 2024b).

1. Form vectors and matrices using lag orders p and q that can produce the smallest AIC value of the SpVARX model.
2. Estimate the coefficients of the SpVARX model with OLS using the following formula:

$$\hat{\boldsymbol{\beta}}_{OLS, SpVARX} = (\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\text{Vec}(\mathbf{Y})$$

3. Find the residuals of the SpVARX model obtained by OLS in step 2 with the following formula:

$$\text{Vec}(\boldsymbol{\xi})_{OLS} = \text{Vec}(\mathbf{Y}) - \mathbf{Z}\hat{\boldsymbol{\beta}}_{OLS, SpVARX}$$

4. Obtain the error covariance matrix estimator of the OLS-derived SpVARX with the following formula:

$$\hat{\boldsymbol{\Sigma}}_{OLS, SpVARX} = \begin{bmatrix} \text{Var}(e_{1,t}^1) & \text{Cov}(e_{1,t}^1, e_{1,t}^2) & \cdots & \text{Cov}(e_{1,t}^1, e_{k,t}^n) & \cdots & \text{Cov}(e_{1,t}^1, e_{K,t}^N) \\ \text{Cov}(e_{1,t}^2, e_{1,t}^1) & \text{Var}(e_{1,t}^2) & \cdots & \text{Cov}(e_{1,t}^2, e_{k,t}^n) & \cdots & \text{Cov}(e_{1,t}^2, e_{K,t}^N) \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \text{Cov}(e_{1,t}^n, e_{1,t}^1) & \text{Cov}(e_{1,t}^n, e_{1,t}^2) & \cdots & \text{Var}(e_{k,t}^n) & \cdots & \text{Cov}(e_{1,t}^n, e_{K,t}^N) \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \text{Cov}(e_{k,t}^n, e_{1,t}^1) & \text{Cov}(e_{k,t}^n, e_{1,t}^2) & \cdots & \text{Cov}(e_{k,t}^n, e_{k,t}^1) & \cdots & \text{Var}(e_{K,t}^N) \end{bmatrix}$$

where

$$\mathbf{e}_k^n = [e_{k,h+1}^n \quad e_{k,h+2}^n \quad \cdots \quad e_{k,T}^n], \quad \text{Var}(\mathbf{e}_k^n) = \frac{1}{(T-h)}(\mathbf{e}_k^n)' \cdot (\mathbf{e}_k^n),$$

$$\text{Cov}(\mathbf{e}_k^n, \mathbf{e}_K^N) = \frac{1}{(T-h)}(\mathbf{e}_k^n)' \cdot (\mathbf{e}_K^N).$$

5. Estimate the model coefficient of the SpVARX with MLE using the following formula:

$$\hat{\boldsymbol{\beta}}_{MLE, SpVARX} = (\mathbf{Z}'(\hat{\boldsymbol{\Sigma}}_{OLS, SpVARX} \otimes \mathbf{I})\mathbf{Z})^{-1}\mathbf{Z}'(\hat{\boldsymbol{\Sigma}}_{OLS, SpVARX} \otimes \mathbf{I})^{-1}\text{Vec}(\mathbf{Y}). \quad (4)$$

2.3. Threshold Spatial Vector Autoregressive with Metric Exogenous Variables (TSpVARX)

The TSpVARX model is formed by dividing the observation data into several regimes based on the threshold variable $(Y_{k,t-d}^u)$ and the threshold value (ς) . In each regime, the SpVARX model coefficients will be estimated so that each regime will contain SpVARX models that have different model coefficients. The TSpVARX model consisting of G regimes with delay order d , spatial order 1, temporal lag order p , and exogenous variable lag order q can be written as TSpVARX($G, 1, p, q, d$). The general form of the TSpVARX model is as follows:

$$\mathbf{y} = \begin{cases} \mathbf{Z}^{(1)}\boldsymbol{\beta}^{(1)} + \boldsymbol{\varepsilon}^{(1)}, & \text{when } Y_{k,t-d}^u \leq \varsigma_1 \\ \mathbf{Z}^{(2)}\boldsymbol{\beta}^{(2)} + \boldsymbol{\varepsilon}^{(2)}, & \text{when } \varsigma_1 < Y_{k,t-d}^u \leq \varsigma_2 \\ \vdots \\ \mathbf{Z}^{(g)}\boldsymbol{\beta}^{(g)} + \boldsymbol{\varepsilon}^{(g)}, & \text{when } \varsigma_{g-1} < Y_{k,t-d}^u \leq \varsigma_g \\ \vdots \\ \mathbf{Z}^{(G)}\boldsymbol{\beta}^{(G)} + \boldsymbol{\varepsilon}^{(G)}, & \text{when } \varsigma_{G-1} < Y_{k,t-d}^u \leq \varsigma_G \end{cases} \quad (5)$$

where

$Y_{k,t-d}^u$ is the selected threshold variable,

ς_{g-1} is the selected threshold value for the lower bound of the g -th regime, and

ς_g is a selected threshold value for the upper bound of the g -th regime.

There are several steps in generating TSpVARX coefficient estimates (Sohibien et al. 2024b).

1. Set a lag of the endogenous variable, which will be the threshold variable.
2. Determine the temporal lag order (p) based on the smallest AIC of the SpVAR model.
3. Determine the exogenous variable lag order (q) based on the smallest AIC of the SpVARX model.
4. Determine the delay limit (d) equal to the selected order p so that the threshold variable candidates are $Y_{k,t-1}^u, Y_{k,t-2}^u, \dots, Y_{k,t-p}^u$.
5. For each threshold variable candidate, determine the lowest threshold value; (ς_{DL}) is the 10th percentile of the threshold variable candidate and the highest threshold value, and (ς_{UL}) is the 90th percentile of the threshold variable candidate so that we obtain the threshold value candidate interval as follows $\varsigma_{DL} \leq \varsigma \leq \varsigma_{UL}$.
6. Divide the data into two parts based on all possibilities ς and d ; when $Y_{k,t-d}^u > \varsigma$, the data will fall into the first regime, and when $Y_{k,t-d}^u \leq \varsigma$, the data will fall into the second regime.
7. Estimate the coefficients of the SpVARX model with the addition of the 12th subset and dummy variables in the first regime and the second regime for all possible data splits by using the estimation steps described in Section 2.2.
8. Calculate the ln-likelihood function values in the first regime $l(\hat{\boldsymbol{\beta}}_{MLE,SpVARX}^{(1)}(d, \varsigma) | \boldsymbol{\Omega}^{(1)})$ and the second regime $l(\hat{\boldsymbol{\beta}}_{MLE,SpVARX}^{(2)}(d, \varsigma) | \boldsymbol{\Omega}^{(2)})$ for each possible division of the data. The formula is as follows:

$$l(\hat{\boldsymbol{\beta}}_{MLE,SpVARX}^{(g)}(d, \varsigma) | \boldsymbol{\Omega}^{(g)}) = -\frac{NKT^{(g)}}{2} \ln 2\pi - \frac{1}{2} \ln |\boldsymbol{\Omega}^{(g)}| - \frac{1}{2} \left[(\mathbf{y}^{(g)} - \mathbf{Z}^{(g)}\boldsymbol{\beta}^{(g)})' (\boldsymbol{\Omega}^{(g)})^{-1} (\mathbf{y}^{(g)} - \mathbf{Z}^{(g)}\boldsymbol{\beta}^{(g)}) \right].$$

9. Calculate the total ln-likelihood with the following formula:

$$l(\hat{\boldsymbol{\beta}}_{MLE,SpVARX}(d, \varsigma)) = l(\hat{\boldsymbol{\beta}}_{MLE,SpVARX}^{(1)}(d, \varsigma) | \boldsymbol{\Omega}^{(1)}) + l(\hat{\boldsymbol{\beta}}_{MLE,SpVARX}^{(2)}(d, \varsigma) | \boldsymbol{\Omega}^{(2)}).$$

10. Obtain the estimated delay ($\hat{d}$) and threshold values ($\hat{\varsigma}$) by finding the pair (d, ς) that maximizes $l(\hat{\boldsymbol{\beta}}_{MLE,SpVARX}(d, \varsigma))$ or it can be written down:

$$(\hat{d}, \hat{\varsigma}) = \left\{ (d, \varsigma), \max l(\hat{\boldsymbol{\beta}}_{MLE,SpVARX}(d, \varsigma)) \right\}.$$
11. The coefficient estimator is the estimator that is used $\hat{d}, \hat{\varsigma}$ as the basis for regime division. We can write it as follows:

$$\hat{\boldsymbol{\beta}}_{MLE,TSpVARX \text{ with 2 Regimes}}^{(g)} = \hat{\boldsymbol{\beta}}_{MLE,SpVARX}^{(g)}(\hat{d}, \hat{\varsigma})$$

12. If you want to perform TSpVARX modeling up to $G - 1$ regimes, there will be $G - 1$ threshold variables. The 1st, 2nd, ..., $G - 2$ th threshold and delay estimators use the values obtained from the TSpVARX with the $G - 1$ regime model.
13. The search for the $G - 1$ th threshold estimator is performed by searching from the threshold value candidates in each TSpVARX model regime with $G - 1$ regimes so that the threshold value candidates are in the following range:

$$\zeta_L^{1[G-1]} \leq \zeta_U^{1[G-1]} \leq \zeta_L^{2[G-1]} \leq \zeta_U^{2[G-1]} \leq \zeta_L^{g[G-1]} \leq \zeta_U^{g[G-1]} \leq \zeta_L^{G-1[G-1]} \leq \zeta_U^{G-1[G-1]}$$

where

$\zeta_L^{g[G-1]}$ is the lowest threshold value candidate derived from the 10th percentile of data in the g -th regime of the TSpVARX model with $G - 1$ regimes, and $\zeta_U^{g[G-1]}$ is the highest threshold value candidate derived from the 90th percentile of data in the g -th regime of the TSpVARX model with $G - 1$ regimes.

14. Calculate the total ln likelihood for all possible threshold value candidates as explained in step 13 with the following formula:

$$l(\hat{\beta}_{MLE,SpVARX}(d, \hat{\zeta}_1, \hat{\zeta}_2, \dots, \zeta_{G-1})) = l(\hat{\beta}_{MLE,SpVARX}^{(1)}(d, \hat{\zeta}_1, \hat{\zeta}_2, \dots, \zeta_{G-1}) | \Omega^{(1)}) + l(\hat{\beta}_{MLE,SpVARX}^{(2)}(d, \hat{\zeta}_1, \hat{\zeta}_2, \dots, \zeta_{G-1}) | \Omega^{(2)}) + \dots + l(\hat{\beta}_{MLE,SpVARX}^{(G)}(d, \hat{\zeta}_1, \hat{\zeta}_2, \dots, \zeta_{G-1}) | \Omega^{(G)})$$

15. The $G - 1$ th threshold value estimator is the threshold value that maximizes the total ln likelihood. We can also write it as follows:

$$\hat{\zeta}_{G-1} = \left\{ \zeta_{G-1}, \max l(\hat{\beta}_{MLE,TSpVARX}(\hat{d}, \hat{\zeta}_1, \hat{\zeta}_2, \dots, \hat{\zeta}_{G-1})) \right\}.$$

16. The estimator coefficients of the TSpVARX with G regime in the g -th regime with the addition of the 12th subset and dummy variables are the coefficients obtained by dividing the regime based on the estimation of delay $\hat{d}$ and threshold value $\hat{\zeta}_1, \hat{\zeta}_2, \dots, \hat{\zeta}_{G-1}$. We can write it as follows:

$$\hat{\beta}_{MLE,TSpVARXGRzim}^{(g)} = \hat{\beta}_{MLE,SpVARX}^{(g)}(\hat{d}, \hat{\zeta}_1, \hat{\zeta}_2, \dots, \hat{\zeta}_{G-1}).$$

We include a flow chart in Figure 1 to improve readability and understanding of the TSpVARX model.

2.4. The Development of TSpVARX with the Addition of the 12th Subset and Dummy Variables

The development of the TSpVARX model with the 12th subset and dummy variables is performed by adding the 12th lag of endogenous and dummy variables as predetermined variables in the TSpVARX model. The form of the SpVARX model with the addition of the 12th lag of endogenous and dummy variables is as follows:

$$Y_{1,t}^1 = \alpha_{10}^1 + \gamma_{11}^{1(1)} X_{1,t-1}^1 + \dots + \gamma_{11}^{1(q)} X_{1,t-q}^1 + \dots + \gamma_{1m}^{1(i)} X_{m,t-1}^1 + \dots + \gamma_{1m}^{1(1)} X_{m,t-1}^1 + \dots + \gamma_{1m}^{1(q)} X_{m,t-q}^1 + \lambda_{11}^1 D_{11,t}^1 + \dots + \lambda_{1l}^1 D_{1l,t}^1 + \dots + \lambda_{1L}^1 D_{1L,t}^1 + \phi_{11}^{1(j,0)} Y_{1,t-1}^1 + \phi_{11}^{1(1,1)} (w_{11(1,2)} Y_{1,t-1}^2 + \dots + w_{11(1,u)} Y_{1,t-1}^u + \dots + w_{11(1,N)} Y_{1,t-1}^N) + \dots + \phi_{1r}^{1(j,0)} Y_{r,t-j}^1 + \phi_{1r}^{1(j,1)} (w_{1r(1,2)} Y_{r,t-j}^2 + \dots + w_{1r(1,u)} Y_{r,t-j}^u + \dots + w_{1r(1,N)} Y_{r,t-j}^N) + \dots + \phi_{1K}^{1(p,0)} Y_{K,t-p}^1 + \phi_{1K}^{1(p,1)} (w_{1K(1,2)} Y_{K,t-p}^2 + \dots + w_{1K(1,u)} Y_{K,t-p}^u + \dots + w_{1K(1,N)} Y_{K,t-p}^N) + \phi_{11}^{1(12,0)} Y_{1,t-12}^1 + \phi_{12}^{1(12,0)} Y_{2,t-12}^1 + \dots + \phi_{1r}^{1(12,0)} Y_{r,t-12}^1 + \dots + \phi_{1K}^{1(12,0)} Y_{K,t-12}^1 + e_{1t}^1 \quad (6a)$$

$$Y_{1,t}^2 = \alpha_{10}^2 + \gamma_{11}^{2(1)} X_{1,t-1}^2 + \dots + \gamma_{11}^{2(q)} X_{1,t-q}^2 + \dots + \gamma_{1m}^{2(i)} X_{m,t-1}^2 + \dots + \gamma_{1m}^{2(1)} X_{m,t-1}^2 + \dots + \gamma_{1m}^{2(q)} X_{m,t-q}^2 + \lambda_{11}^2 D_{11,t}^2 + \dots + \lambda_{1l}^2 D_{1l,t}^2 + \dots + \lambda_{1L}^2 D_{1L,t}^2 + \phi_{11}^{2(j,0)} Y_{1,t-1}^2 + \phi_{11}^{2(1,1)} (w_{11(2,1)} Y_{1,t-1}^1 + \dots + w_{11(2,u)} Y_{1,t-1}^u + \dots + w_{11(2,N)} Y_{1,t-1}^N) + \dots + \phi_{1r}^{2(j,0)} Y_{r,t-j}^2 + \phi_{1r}^{2(j,1)} (w_{1r2} Y_{r,t-j}^1 + \dots + w_{1r(2,u)} Y_{r,t-j}^u + \dots + w_{1r(2,N)} Y_{r,t-j}^N) + \dots + \phi_{1K}^{2(p,0)} Y_{K,t-p}^2 + \phi_{1K}^{2(p,1)} (w_{1K(2,1)} Y_{K,t-p}^1 + \dots + w_{1K(2,u)} Y_{K,t-p}^u + \dots + w_{1K(2,N)} Y_{K,t-p}^N) + \phi_{11}^{2(12,0)} Y_{1,t-12}^2 + \phi_{12}^{2(12,0)} Y_{2,t-12}^2 + \dots + \phi_{1r}^{2(12,0)} Y_{r,t-12}^2 + \dots + \phi_{1K}^{2(12,0)} Y_{K,t-12}^2 + e_{1t}^2 \quad (6b)$$

$$\begin{aligned}
& \vdots \\
Y_{k,t}^n = & \alpha_{k0}^n + \gamma_{k1}^{n,(1)} X_{1,t-1}^n + \cdots + \gamma_{k1}^{n,(q)} X_{1,t-q}^n + \cdots + \gamma_{km}^{n,(i)} X_{m,t-1}^n + \cdots + \gamma_{km}^{n,(1)} X_{M,t-1}^n + \cdots + \gamma_{km}^{n,(q)} X_{M,t-q}^n + \lambda_{k1}^n D_{k1,t}^n + \cdots \\
& + \lambda_{kl}^n D_{kl,t}^n + \cdots + \lambda_{kL}^n D_{kL,t}^n + \phi_{k1}^{n,(j,0)} Y_{1,t-1}^n + \phi_{k1}^{n,(1,1)} \left(w_{k1(2,1)} Y_{1,t-1}^n + \cdots + w_{k1(2,u)} Y_{1,t-1}^u + \cdots + w_{k1(2,N)} Y_{1,t-1}^N \right) \\
& + \cdots + \phi_{kr}^{n,(j,0)} Y_{r,t-j}^n + \phi_{kr}^{n,(j,1)} \left(w_{kr2} Y_{r,t-j}^1 + \cdots + w_{kr(2,u)} Y_{r,t-j}^u + \cdots + w_{kr(2,N)} Y_{r,t-j}^N \right) \\
& + \cdots + \phi_{kk}^{n,(p,0)} Y_{K,t-p}^n + \phi_{kk}^{n,(p,1)} \left(w_{kk(2,1)} Y_{K,t-p}^1 + \cdots + w_{kk(2,u)} Y_{K,t-p}^u + \cdots + w_{kk(2,N)} Y_{K,t-p}^N \right) + \phi_{k1}^{n,(12,0)} Y_{1,t-12}^n + \phi_{k2}^{n,(12,0)} Y_{2,t-12}^n \\
& + \cdots + \phi_{kr}^{n,(12,0)} Y_{r,t-12}^n + \cdots + \phi_{kk}^{n,(12,0)} Y_{K,t-12}^n + e_{kt}^n, \\
& \vdots \\
Y_{K,t}^N = & \alpha_{K0}^N + \gamma_{K1}^{N,(1)} X_{1,t-1}^N + \cdots + \gamma_{K1}^{N,(q)} X_{1,t-q}^N + \cdots + \gamma_{Km}^{N,(i)} X_{m,t-1}^N + \cdots + \gamma_{Km}^{N,(1)} X_{M,t-1}^N + \cdots + \gamma_{Km}^{N,(q)} X_{M,t-q}^N + \lambda_{K1}^N D_{K1,t}^N + \cdots \\
& + \lambda_{K2}^N D_{K2,t}^N + \cdots + \lambda_{KL}^N D_{KL,t}^N + \phi_{K1}^{N,(j,0)} Y_{1,t-1}^N + \phi_{K1}^{N,(1,1)} \left(w_{K1(N,1)} Y_{1,t-1}^N + \cdots + w_{K1(N,u)} Y_{1,t-1}^u + \cdots + w_{K1(N,N-1)} Y_{1,t-1}^{N-1} \right) \\
& + \cdots + \phi_{Kr}^{N,(j,0)} Y_{r,t-j}^N + \phi_{Kr}^{N,(j,1)} \left(w_{Kr(N,1)} Y_{r,t-j}^1 + \cdots + w_{Kr(N,u)} Y_{r,t-j}^u + \cdots + w_{Kr(N,N-1)} Y_{r,t-j}^{N-1} \right) \\
& + \cdots + \phi_{KK}^{N,(p,0)} Y_{K,t-p}^N + \phi_{KK}^{N,(p,1)} \left(w_{KK(2,1)} Y_{K,t-p}^1 + \cdots + w_{KK(2,u)} Y_{K,t-p}^u + \cdots + w_{KK(N,N-1)} Y_{K,t-p}^N \right) + \phi_{K1}^{N,(12,0)} Y_{1,t-12}^N + \phi_{K2}^{N,(12,0)} Y_{2,t-12}^N \\
& + \cdots + \phi_{Kr}^{N,(12,0)} Y_{r,t-12}^N + \cdots + \phi_{KK}^{N,(12,0)} Y_{K,t-12}^N + e_{Kt}^N,
\end{aligned} \quad (6c)$$

$$\begin{aligned}
& \vdots \\
& \vdots \\
& \vdots
\end{aligned} \quad (6d)$$

where:

$D_{kl,t}^n$ is the l -th dummy variable in the equation $Y_{k,t}^n$,

λ_{kl}^n is the coefficient of $D_{kl,t}^n$,

L is the number of dummy variables,

$Y_{r,t-12}^n$ is the 12th lag of the r -th endogenous variable at the n -th location, and

$\phi_{kr}^{n,(12,0)}$ is the coefficient of the 12th lag of the r -th endogenous variable at the n -th location in the equation $Y_{k,t}^n$.

The SpVARX model with the 12th subset and dummy variables can be written in vector and matrix form as follows:

$$\begin{bmatrix} Y_{1,t}^1 \\ Y_{1,t}^2 \\ \vdots \\ Y_{k,t}^n \\ \vdots \\ Y_{K,t}^N \end{bmatrix} = \begin{bmatrix} (\mathbf{x}_{1,t}^1)' & (\mathbf{d}_{1,t}^1)' & (\mathbf{f}_{1,t}^1)' & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & 0 & (\mathbf{x}_{1,t}^2)' & (\mathbf{d}_{1,t}^2)' & (\mathbf{f}_{1,t}^2)' & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & (\mathbf{x}_{k,t}^n)' & (\mathbf{d}_{k,t}^n)' & (\mathbf{f}_{k,t}^n)' & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & (\mathbf{x}_{K,t}^N)' & (\mathbf{d}_{K,t}^N)' & (\mathbf{f}_{K,t}^N)' \end{bmatrix} \begin{bmatrix} \gamma_1^1 \\ \lambda_1^1 \\ \Phi_1^1 \\ \gamma_1^2 \\ \lambda_1^2 \\ \Phi_1^2 \\ \vdots \\ \gamma_k^n \\ \lambda_k^n \\ \Phi_k^n \\ \vdots \\ \gamma_K^N \\ \lambda_K^N \\ \Phi_K^N \end{bmatrix}, \quad (7)$$

$$+ \begin{bmatrix} e_{1,t}^1 \\ e_{1,t}^2 \\ \vdots \\ e_{k,t}^n \\ \vdots \\ e_{K,t}^N \end{bmatrix}$$

where

$$(\mathbf{x}_{k,t}^n)' = \begin{bmatrix} 1 & \cdots & X_{1,t}^n & \cdots & X_{1,t-q}^n & \cdots & X_{m,t-i}^n & \cdots & X_{M,t-1}^n & \cdots & X_{M,t-q}^n \end{bmatrix}_{1 \times (1+Mq)},$$

$$(\mathbf{d}_{k,t}^n)' = \begin{bmatrix} D_{k1,t}^n & \cdots & D_{kl,t}^n & \cdots & D_{kL,t}^n \end{bmatrix}_{1 \times L},$$

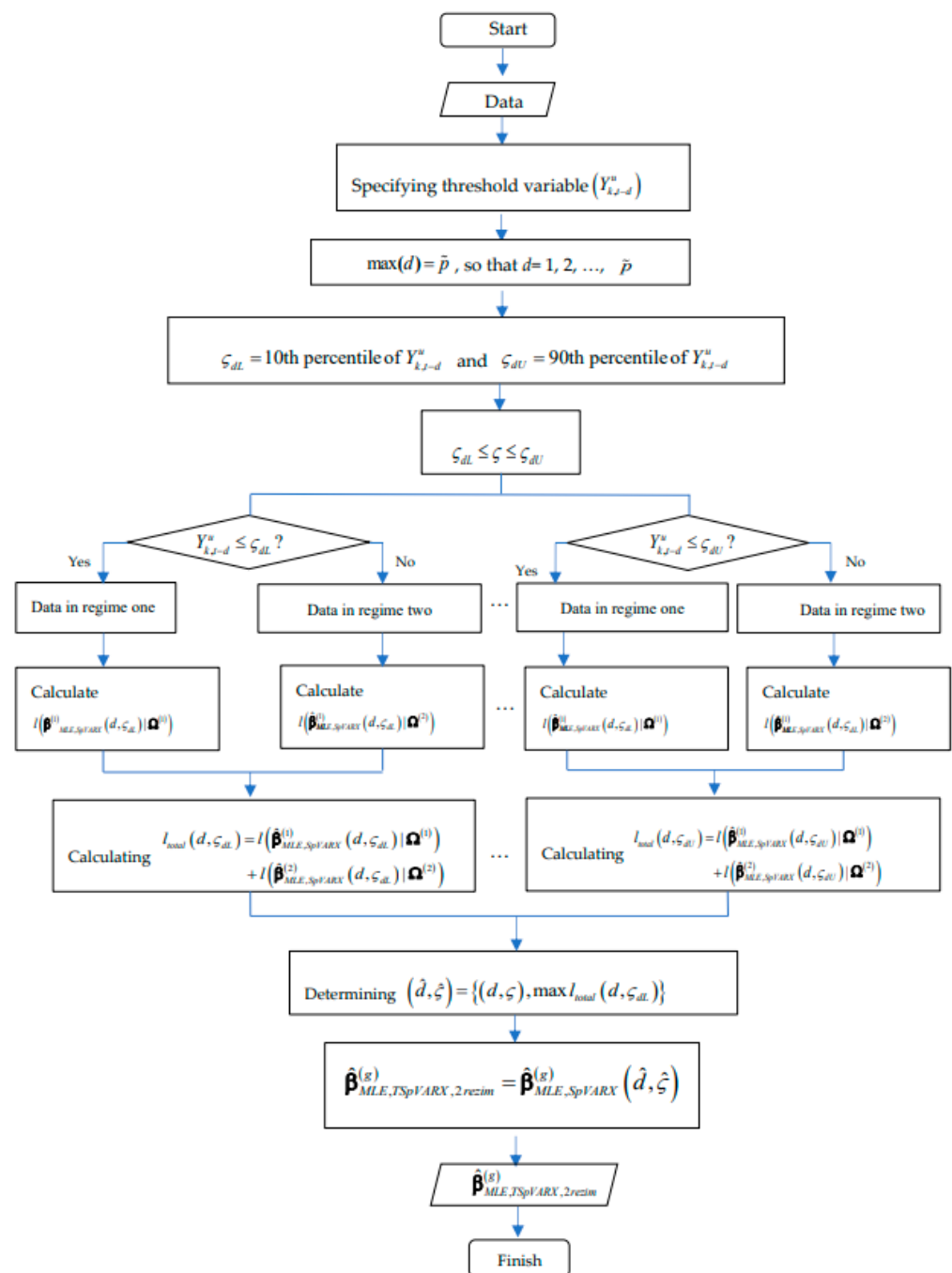


Figure 1. The flow chart of the coefficient estimation of TSsVARX.

$$(\mathbf{f}_{k,t}^n)' = [Y_{1,t-1}^n \quad Y_{k1,t-1}^{n*} \quad \cdots \quad Y_{r,t-j}^n \quad Y_{kr,t-j}^{n*} \quad \cdots \quad Y_{K,t-p}^n \quad Y_{kK,t-p}^{n*} \quad \cdots \quad Y_{1,t-12}^n \quad Y_{2,t-12}^n \quad \cdots \quad Y_{r,t-12}^n \quad \cdots \quad Y_{K,t-12}^n]_{1 \times K(2p+1)},$$

$$Y_{kr,t-j}^{n*} = \sum_{\substack{u=1 \\ u \neq n}}^N w_{kr}(n, u) Y_{r,t-j}^u,$$

$$\gamma_k^n = [\alpha_{k0}^n \quad \gamma_{k1}^{n(1)} \quad \cdots \quad \gamma_{k1}^{n(q)} \quad \cdots \quad \gamma_{km}^{n(i)} \quad \cdots \quad \gamma_{kM}^{n(1)} \quad \cdots \quad \gamma_{kM}^{n(q)}]_{(1+Mq) \times 1}'$$

$$\lambda_k^n = [\lambda_{k1}^n \quad \cdots \quad \lambda_{kl}^n \quad \cdots \quad \lambda_{kL}^n]_{L \times 1}'$$

$$\phi_k^n = \begin{bmatrix} \phi_{k1}^{n,(1,0)} & \phi_{k1}^{n,(1,1)} & \dots & \phi_{kr}^{n,(j,0)} & \phi_{kr}^{n,(j,1)} & \dots & \phi_{kK}^{n,(p,0)} & \phi_{kK}^{n,(p,1)} & \dots & \phi_{k1}^{n,(12,1)} & \dots & \phi_{kK}^{n,(12,1)} \end{bmatrix}'_{K(2p+1) \times 1}.$$

$$\text{Let } \mathbf{F}_k^n = \begin{bmatrix} (\mathbf{x}_{k,h+1}^n)' & (\mathbf{d}_{k,h+1}^n)' & (\mathbf{f}_{k,h+1}^n)' \\ (\mathbf{x}_{k,h+2}^n)' & (\mathbf{d}_{k,h+2}^n)' & (\mathbf{f}_{k,h+2}^n)' \\ \vdots & \vdots & \vdots \\ (\mathbf{x}_{k,T}^n)' & (\mathbf{d}_{k,T}^n)' & (\mathbf{f}_{k,T}^n)' \end{bmatrix}, \mathbf{F} = \begin{bmatrix} \mathbf{F}_1^1 & \mathbf{0} & \dots & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{F}_1^2 & \dots & \mathbf{0} & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{F}_k^n & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \dots & \mathbf{F}_K^N \end{bmatrix},$$

$$\varphi_k^n = \begin{bmatrix} \gamma_k^n \\ \lambda_k^n \\ \Phi_k^n \end{bmatrix}, \text{ and } \boldsymbol{\theta} = \begin{bmatrix} \varphi_1^1 \\ \varphi_1^2 \\ \vdots \\ \varphi_k^n \\ \vdots \\ \varphi_K^N \end{bmatrix}; \text{ then, if the number of observations used for the model}$$

is T and $h = \max(12, p, q, L)$, then the form of the SpVARX model in equation (7) can be described as follows:

$$\text{Vec}(\mathbf{Y}) = \mathbf{F}\boldsymbol{\theta} + \text{Vec}(\boldsymbol{\xi}), \quad (8)$$

where

$\mathbf{F}$ is a matrix of size $NK(T-h) \times NK(1 + Mq + 2Kp + K + L)$, and

$\boldsymbol{\theta}$ is a vector of size $NK(1 + Mq + 2Kp + K + L) \times 1$.

TSpVARX model with the addition of the 12th subset and the dummy variables, formed by dividing the data into several regimes, is based on the threshold value and selected threshold variables. After that, the coefficient estimation of the SpVARX model with the addition of the 12th subset and the dummy variables is carried out.

2.5. Steps of TSpVARX Modeling with the Addition of the 12th Subset and Dummy Variables in the Application

The analysis method used in this research is Threshold Spatial Vector Autoregressive with Exogenous Variables (TSpVARX) by adding the 12th Subset and Dummy Variables. There are some steps to achieve the research objectives.

1. We perform data stationarity tests with the Augmented Dickey–Fuller and Philips Perron Test.
2. We identify spatial relationships between endogenous variables using cross-correlation.
3. We determine spatial weights using uniform and cross-correlation normalization.
4. We determine the lag length of endogenous variables (p) using the AIC of the SpVAR model.
5. We determine the lag length of the exogenous variable (q) using the AIC of the SpVARX model.
6. We identify the nonlinearity of the relationship between endogenous variables and predetermined variables using the RESET Test method. The RESET Test method specifically checks whether the relationship between the endogenous variables and the nonlinear nature has or has not been accommodated in the model. How the testing steps of nonlinearity by using the RESET Test can be seen in Tsay (2010).
7. We estimate the coefficients of TSpVARX with two, three, and four regimes for all possible threshold variables.
8. We select the best of each TSpVARX by looking at the smallest AIC value.
9. We compare the forecasting performance of SpVARX and TSpVARX using the RMSE of the testing data.
10. We insert the 12th subset and fuel price adjustment variables.

11. We compare the forecasting performance of TSpVARX and TSpVARX *with the Addition of the 12th Subset and Dummy Variables*.
12. We perform forecasting for several periods using the best model obtained in step 11.

2.6. Relationship Among SpVARX, TSpVARX, and TSpVARX Models with Subset Variables and Dummy Variable

The SpVARX model is a model that can accommodate reciprocal relationships among variables, space–time relationships of more than one type of endogenous variable, and the influence of exogenous metric variables. However, the SPVARX model cannot accommodate the nonlinearity between endogenous and predetermined variables.

The TSpVARX model accommodates the nonlinearity of the relationship between endogenous and predetermined variables by dividing the data into several regimes. In each regime, a SpVARX model will be created so that each regime will have a different SpVARX model. Regime division is performed based on the threshold value and the selected threshold variable that can maximize the total likelihood. Although it can accommodate the nonlinearity of the relationship, the TSpVARX model has not been able to capture some events that can affect the value of inflation and outflow, such as Eid, Christmas, New Year, and the increase and decrease in fuel prices. The TSpVARX model with the addition of the 12th subset variable and dummy variables can overcome the problem. The formation of the TSpVARX model with the addition of the 12th subset variable and dummy variables is performed by performing SpVARX modeling by adding exogenous variables in the form of the 12th lag of endogenous variables and dummy variables of fuel price increases and decreases in each regime for which the distribution has been determined in TSpVARX modeling.

3. Results and Discussion

3.1. Data Exploration

In Figure 2, we can see that during the period from May 2006 to May 2024, inflation in these three cities increased in July 2013. It could be due to the Eid al-Fitr event in early August 2013. Inflation in June 2008 and December 2014 for Semarang and Yogyakarta were also quite high. The increase in June 2008 could be due to the impact of the fuel price hike. The increase in inflation in December 2014 could be due to the fuel price hike in mid-November 2014.

In Figure 3, we can see that there are some high increases in money outflow of the three cities in some periods, such as in June 2017, June 2018, May 2020, etc. We suspect these conditions are due to the influence of the Eid al-Fitr event, which caused many people to withdraw money for Eid's needs. The high increase in money outflow data causes the money outflow variable in the three cities to vary greatly, ranging from units to thousands of billions. Therefore, in this study, the money outflow data are transformed using the logarithm form so that the value of the money outflow variable is not much different.

The first step taken in modeling time series data is to test the data stationarity. If the data are not stationary, then the differencing process is carried out. The data stationarity test of this study uses the Phillips Perron (PP) and Augmented Dickey-Fuller (ADF) tests. If the p -value is less than the significance level, the data are concluded to be stationary. The results of data stationarity testing can be seen in Table 1.

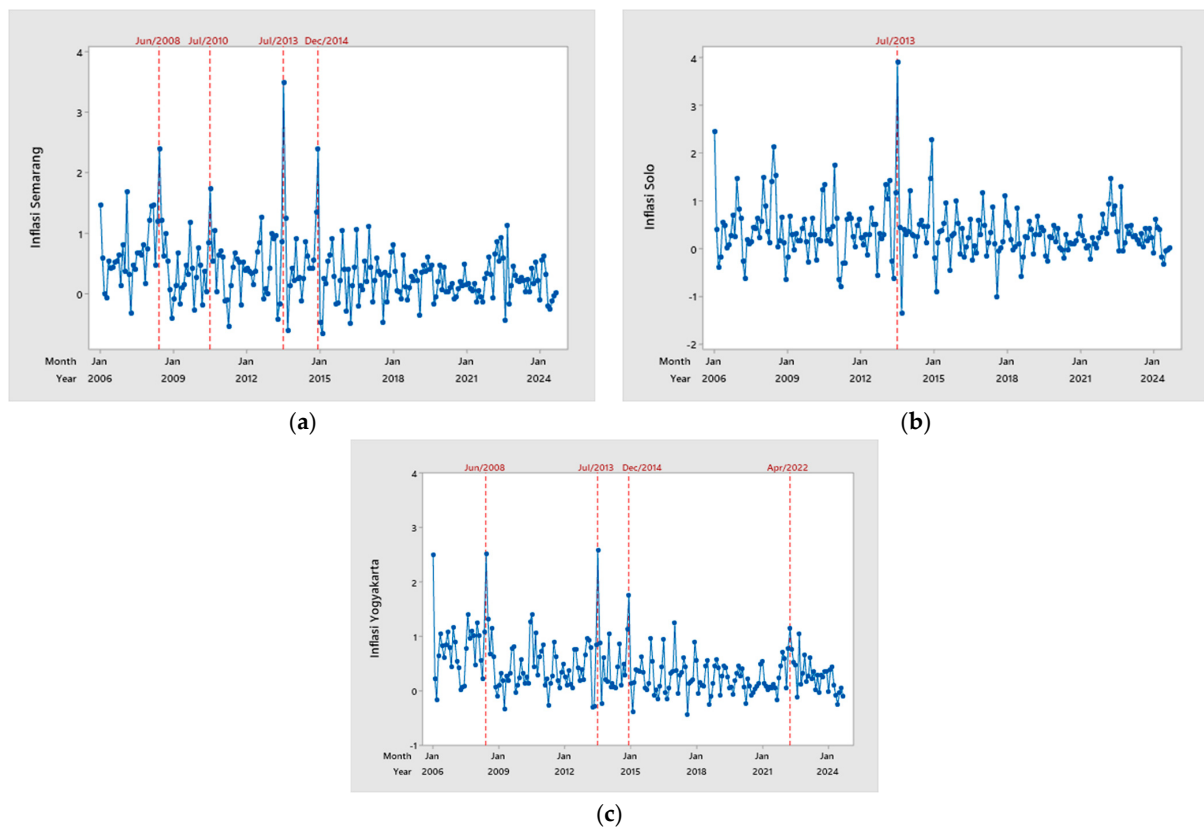


Figure 2. Time series plot of inflation of Semarang, Solo, and Yogyakarta Cities. (a) Semarang, (b) Solo, and (c) Yogyakarta.

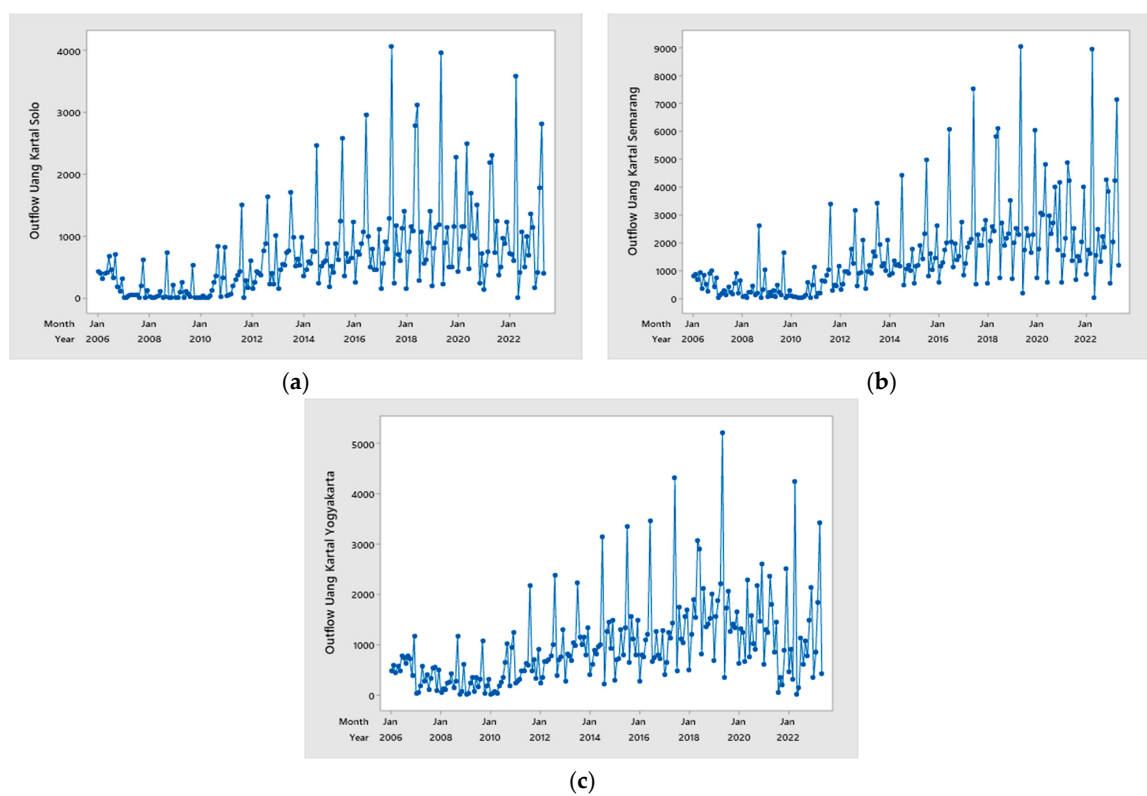


Figure 3. Time series plot of money outflow in Semarang, Solo, and Yogyakarta. (a) Semarang, (b) Solo, and (c) Yogyakarta.

Table 1. Stationarity test results.

Variables	Philips Perron Test		Augmented Dickey-Fuller (ADF)	
	<i>p</i> -Value	Conclusion	<i>p</i> -Value	Conclusion
Inflation of Semarang	0.000	Stationary	0.000	Stationary
Inflation of Solo	0.000	Stationary	0.000	Stationary
Inflation of Yogyakarta	0.000	Stationary	0.000	Stationary
Log Money Outflow of Semarang	0.000	Stationary	0.007	Stationary
Log Money Outflow of Solo	0.000	Stationary	0.002	Stationary
Log Money Outflow of Yogyakarta	0.000	Stationary	0.000	Stationary

The next step is to identify the relationship between endogenous variables in one city and another. This identification is performed by looking at the significance of the cross-correlation value between endogenous and predetermined variables. Based on Table 2, it can be seen that there is a significant cross-correlation between the inflation of Semarang one month earlier and all endogenous variables at the five-percent significance level. Significant cross-correlation also exists between the inflation of Solo one month earlier and the inflation of Yogyakarta and Semarang at the five-percent significance level. Besides that, the inflation of Semarang one month earlier also has a significant cross-correlation with the money outflow of Yogyakarta and Semarang at the 10-percent significance level. The inflation of Yogyakarta one month earlier has a significant cross-correlation with all endogenous variables at the five-percent significance level. What we get here is in line with the previous research from [Sohibien et al. \(2024a\)](#), which found that inflation in these three cities is spatially related. In addition, [Sumarminingsih et al. \(2021\)](#) also show that inflation between cities can be related.

Table 2. The significance test of cross-correlation between the lag of the inflation variable and all endogenous variables.

Endogenous Variables	Inflation of Semarang		Inflation of Solo		Inflation of Yogyakarta	
	Cross-Correlation	PV	Cross-Correlation	PV	Cross-Correlation	PV
Inflation of Semarang	0.290	0.000 **	0.348	0.000 **	0.345	0.000 **
Inflation of Solo	0.180	0.009 **	0.282	0.000 **	0.251	0.000 **
Inflation of Yogyakarta	0.295	0.000 **	0.320	0.000 **	0.372	0.000 **
Log Money Outflow of Semarang	−0.235	0.000 **	−0.151	0.030 **	−0.271	0.000 **
Log Money Outflow of Solo	−0.206	0.003 **	−0.120	0.085 *	−0.235	0.000 **
Log Money Outflow of Yogyakarta	−0.171	0.014 **	−0.132	0.057 *	−0.215	0.002 **

Notes: (*) significant at a 10-percent significance level. (**) significant at a 5-percent significance level.

Table 3 shows that there is a significant cross-correlation between the money outflow in a city and other cities at the five percent significance level. The existence of an interrelationship between the money outflow of these three cities can be caused by the fact that the BI representative offices in these three cities belong to the same group. As a result, if a city experiences a shortage of money stock, the other cities can supply the shortage. Thus, there is a linkage in the money outflow between these three cities.

Table 3. The significance test of cross-correlation between the lag of log money outflow and all endogenous variables.

Endogenous Variable	Log Outflow of Semarang		Log Outflow of Solo		Log Outflow of Yogyakarta	
	Cross-Correlation	PV	Cross-Correlation	PV	Cross-Correlation	PV
Inflation of Semarang	−0.136	0.050 **	−0.109	0.118	−0.064	0.355
Inflation of Solo	−0.011	0.877	0.022	0.754	0.032	0.651
Inflation of Yogyakarta	−0.111	0.115	−0.071	0.310	−0.027	0.702
Log outflow of Semarang	0.512	0.000 **	0.491	0.000 **	0.408	0.000 **
Log outflow of Solo	0.500	0.000 **	0.568	0.000 **	0.470	0.000 **
Log outflow of Yogyakarta	0.389	0.000 **	0.450	0.000 **	0.414	0.000 **

Notes: (**) significant at a 5-percent significance level.

3.2. Temporal Lag Selection (p)

The selection of temporal lag (p) is performed by looking at the smallest Akaike Information Criterion (AIC) value generated from several SpVAR models formed. There are two weights used, namely uniform and normalized cross-correlation weight. The AIC of the SpVAR models can be seen in Table 4. The smallest AIC is obtained from the SpVAR (1, 1) with a constant.

Table 4. Akaike Information Criterion (AIC) values based on SpVAR(1, p) model (for $p = 1, 2, \dots, 4$) and spatial weights.

SpVAR Model		AIC with Uniform Weighting	AIC with Cross-Correlation Normalization Weighting
Without Constant	SpVAR(1, 1)	−1487.846	−1486.08
	SpVAR(1, 2)	−1494.386	−1497.785
	SpVAR(1, 3)	−1422.474	−1425.865
	SpVAR(1, 4)	−1308.21	−1310.738
With Constant	SpVAR(1, 1)	−1586.181	−1587.617
	SpVAR(1, 2)	−1546.97	−1553.562
	SpVAR(1, 3)	−1453.157	−1457.315
	SpVAR(1, 4)	−1331.915	−1335.972

3.3. Order Selection q

The next step is to select the order lag of exogenous variables (q) that will be used in SpVARX and TSpVARX modeling. The selection of order q is performed by looking at the smallest AIC generated from SpVARX models with a temporal lag order (p) of 1. We can write the general form as SpVARX (1, 1, q). The results of the AIC values formed from SpVARX (1, 1, q) models can be seen in Table 5. SpVARX models that are successfully formed using uniform weight are SpVARX models with a lag order q of one to four. Meanwhile, the SpVARX models successfully formed using the normalized cross-correlation weight are SpVARX models with a lag order q of 1, 2, and 3. The SpVARX model with a lag order q of 4 with a normalized cross-correlation weight is not successfully formed due to singularity problems in estimating the SpVARX model coefficients. A coefficient estimation that cannot be generated due to the singularity, causes us to be unable to obtain the AIC value. Based on Table 5, the smallest AIC obtained from the SpVARX model is when the order q is equal to one or both when using the uniform and normalized cross-correlation weight.

Table 5. Akaike Information Criterion (AIC) based on SpVAR (1, 1, q) model (for $q = 1, 2, \dots, 4$) and spatial weight.

SpVARX Model		AIC with Uniform Weight	AIC with Normalized Cross-Correlation Weight
With Constant	SpVARX(1, 1, 1)	−1584.989	−1586.072
	SpVARX(1, 1, 2)	−1567.85	−1569.152
	SpVARX(1, 1, 3)	−1545.433	−1546.449
	SpVARX(1, 1, 4)	−1520.337	−

3.4. Nonlinearity Test Between Endogenous and Lag of Endogenous Variables

The next step is to test whether the endogenous variables are suitable to be modeled linearly with the predetermined variables. This research uses the reset test method. The hypothesis used in the linearity test with the reset test is as follows:

H₀. *The endogenous variable fits a linearly model with predetermined variables.*

H₁. *Endogenous variables are not linearly modeled with predetermined variables.*

The results of the nonlinearity testing between endogenous variables and predetermined variables can be seen in Table 6. If the p -value is less than the significant level, the decision is to reject H_0 , which means that the relationship between the endogenous variables and the predetermined variables is not linear. In Table 6, it can be seen that of the six endogenous variables used, three of them (inflation of Solo, inflation of Yogyakarta, and outflow of Solo) have p -values less than the five- or ten-percent significance level. The inflation of Solo and money outflow of Solo are significant at the ten-percent significance level, while the inflation of Yogyakarta is significant at the five-percent significance level. Thus, these three endogenous variables are not fit to be modeled linearly against the predetermined variables used.

Table 6. Linearity test results using reset test between endogenous variables and all predetermined variables.

Endogenous Variable	p -Value
Inflation of Semarang	0.104
Inflation of Solo	0.054 *
Inflation of Yogyakarta	0.004 **
Ln outflow of Semarang	0.155
Ln outflow of Solo	0.084 *
Ln outflow of Yogyakarta	0.124

Notes: (*) significant at a 10-percent significance level. (**) significant at a 5-percent significance level.

The non-linear relationship between inflation and predetermined variables, such as the lag of money outflow, exchange rate depreciation, and BI benchmark interest rate is in line with the findings of several studies. Çitçi and Kaya (2023), using a sample of 149 countries, found that the exchange rate has a non-linear relationship with inflation. Gök and Bulut (2021) found that interest rates have a non-linear relationship with Turkish inflation.

3.5. Selection of Threshold Variables

After we obtain the temporal lag order (p), the exogenous variable lag order (q), and the indications of non-linear relationships, the next step is to model the TSpVARX. The number of regimes (s) tried in this study are 2, 3, and 4. The TSpVARX ($s, 1, 1, 1$) model formed in this study uses the uniform and normalized cross-correlation weight. The value of the uniform spatial weight and the normalized cross-correlation weight can be seen in Appendix A. Because the order p is one, the threshold variables used in this study are $Y_{1,t-1}^1$, $Y_{1,t-1}^2$, $Y_{1,t-1}^3$, $\log Y_{2,t-1}^1$, $\log Y_{2,t-1}^2$, and $\log Y_{2,t-1}^3$. Based on Figure 4, we can see

that 15 of 18 TSspVARX models formed have smaller AIC values than SpVARX and 14 of 18 TSspVARX models have smaller AIC values than SpVAR. This indicates that most of the TSspVARX formed are better than SpVARX and SpVAR in terms of modeling. Based on the smallest AIC, we select the best three TSspVARX models with a uniform weight, namely the two-regime TSspVARX with the threshold variable $\log Y_{2,t-1}^2$, the three-regime TSspVARX with the threshold variable $Y_{1,t-1}^3$, and the four-regime TSspVARX with the threshold variable $Y_{1,t-1}^2$.

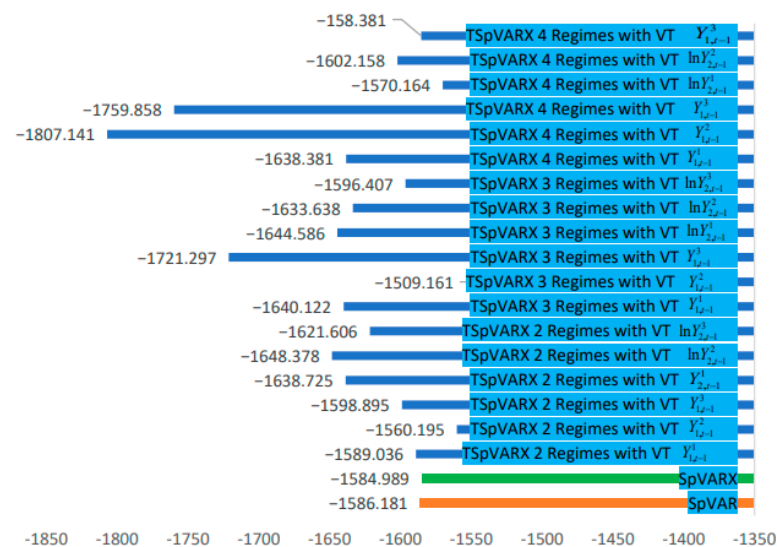


Figure 4. AIC values of SpVAR, SpVARX, and TSspVARX models of two, three, and four regimes with uniform weights. Notes: VT is the threshold variable.

In Figure 5, it can be seen that 13 of 16 TSspVARX models with normalized cross-correlation weights have smaller AIC values than SpVARX and SpVAR. This shows that most of the TSspVARX models with normalized cross-correlation weights are better than SpVAR and SpVARX in terms of modeling. Based on the smallest AIC value, the best three TSspVARX models with normalized cross-correlation weights are the two-regime TSspVARX with the threshold variable $\log Y_{1,t-1}^2$, the three-regime TSspVARX with the threshold variable $\log Y_{1,t-1}^2$, and the four-regime TSspVARX with the threshold variable $\log Y_{1,t-1}^2$.

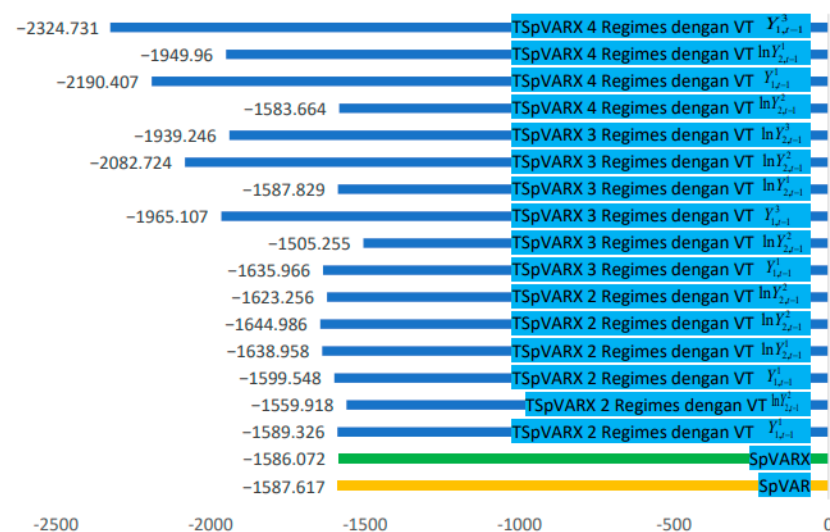


Figure 5. AIC values of SpVAR, SpVARX, and TSspVARX models of two, three, and four regimes with normalized cross-correlation weights. Notes: VT is the threshold variable.

3.6. Evaluation of TSpVARX Model Compared to SpVARX in Modeling Inflation and Money Outflow of Semarang, Solo, and Yogyakarta

After we obtain the best model of TSpVARX with two, three, and four regimes, the next step is to evaluate the performance of TSpVARX and SpVARX models in forecasting the inflation and money outflow of Semarang, Solo, and Yogyakarta. An evaluation of the forecasting performance is performed by looking at the Root Mean Square Error (RMSE) for each endogenous variable. The RMSE is a value used to measure the accuracy of a model in forecasting. A lower RMSE indicates a more precise forecast result. Specifically, for the log money outflow variable, the RMSE is calculated using the residuals of the original forecasting money outflow, not in logarithm form. It is performed because the forecasting results used in reality are the original value of money outflow. According to Table 7, five of six endogenous variables obtain the smallest RMSE in the testing data when using the three-regime TSpVARX model with a uniform weight. Those five endogenous variables are Semarang Inflation ($Y_{1,t-1}^1$), Solo Inflation ($Y_{1,t-1}^2$), money outflow of Semarang ($\log Y_{2,t-1}^1$), money outflow of Solo ($\log Y_{2,t-1}^2$), and money outflow of Yogyakarta ($\log Y_{2,t-1}^3$). Meanwhile, in the training data, four endogenous variables obtain the smallest RMSE when we use a normalized cross-correlation weight.

Table 7. Root Mean Square Error (RMSE) of SpVARX and TSpVARX.

Data	Weights	Model	Inflation of Semarang ($Y_{1,t}^1$)	Inflation of Solo ($Y_{1,t}^2$)	Inflation of Yogyakarta ($Y_{1,t}^3$)	Money Outflow of Semarang ($Y_{2,t}^1$)	Money Outflow of Solo ($Y_{2,t}^2$)	Money Outflow of Yogyakarta ($Y_{2,t}^3$)
Testing	Uniform	SpVAR	1.949	0.366	0.332	3059.76	1164.92	1229.44
		SpVARX	0.391	0.366	0.3	2450.76	940.89	990.31
		TSpVARX with 2 Regimes	0.46	0.405	0.319	2661.00	932.80	952.01
		TSpVARX with 3 Regimes	0.366	0.339	0.299	2130.05	827.40	866.93
		TSpVARX with 4 Regimes	0.402	0.372	0.301	2577.12	958.13	1021.66
	Cross-Correlation Normalization	SpVAR	0.399	0.361	0.292	2728.032	1046.463	1098.674
		SpVARX	0.389	0.359	0.297	2457.53	943.28	991.76
		TSpVARX 2 with 2 Regimes	0.429	0.394	0.303	2661.60	929.26	945.37
		TSpVARX 3 with 3 Regimes	0.542	0.47	0.486	2997.96	1028.87	1175.92
		TSpVARX 4 with 4 Regimes	0.452	0.528	0.383	2604.83	963	892.41
Training	Uniform	SpVAR	0.487	0.552	0.421	1496.206	705.238	820.559
		SpVARX	0.552	0.482	0.418	1442.384	687.528	806.829
		TSpVARX with 2 Regimes	0.467	0.547	0.406	1412.614	677.882	797.165
		TSpVARX with 3 Regimes	0.445	0.535	0.398	1426.346	698.190	795.610
		TSpVARX with 4 Regimes	0.477	0.537	0.398	1426.458	695.333	795.303
	Cross-Correlation Normalization	SpVAR	0.488	0.553	0.421	1492.539	704.125	818.975
		SpVARX	0.488	0.553	0.418	1441.095	687.130	806.098
		TSpVARX with 2 Regimes	0.468	0.548	0.408	1412.025	677.884	796.828
		TSpVARX with 3 Regimes	0.542	0.470	0.486	2997.964	1028.873	1175.918
		TSpVARX with 4 Regimes	0.458	0.540	0.390	1405.814	675.928	790.389

In Figures 6 and 7, the blue line illustrates the movement of actual data. The red line illustrates the movement of the forecasting data of the TSpVARX model, while the green line illustrates the movement of forecast data of the SpVARX model. Based on Figure 6, the up-and-down pattern of actual data can already be followed by the up-and-down pattern of forecast data generated from the three-regime TSpVARX model. However, there are

still some points of actual data that are still far from the forecasting inflation data. The events that cause some of the actual inflation data to be far from the forecasting data results are Christmas and New Year events (December 2006, December 2010), Eid al-Fitr events (19 August 2012, 8 August 2013), fuel price increases (24 May 2008, November 2014, 1 September 2022), and fuel price decreases (1 December 2008, 1 February 2015, 1 April 2016, 10 February 2019).

In Figure 7, it can be seen that there are still actual data at several points of the training and testing periods that are far from the forecasting money outflow data when we use the TSpVARX model. The events that cause this situation are Christmas and New Year events (December 2006, December 2019) and Eid al-Fitr events (October 2007, October 2008, September 2009, September 2010, August 2011, August 2012, August 2013, July 2014, July 2015, July 2016, June 2017, June 2018, June 2019, May 2022, April 2023). The most events that cause this condition are the Eid al-Fitr events.

In Table 8, we can see that the forecasting results of inflation and money outflow of Semarang, Solo, and Yogyakarta using the three-regime TSpVARX model are relatively stable after the next sixteen periods (October 2024, November 2024, and December 2024). It could be due to the use of an autoregressive lag order (p) of one that can only accommodate the data pattern of one previous period. In addition, the other influences of the Eid al-Fitr event and the adjustment of fuel prices have not been captured in the model.

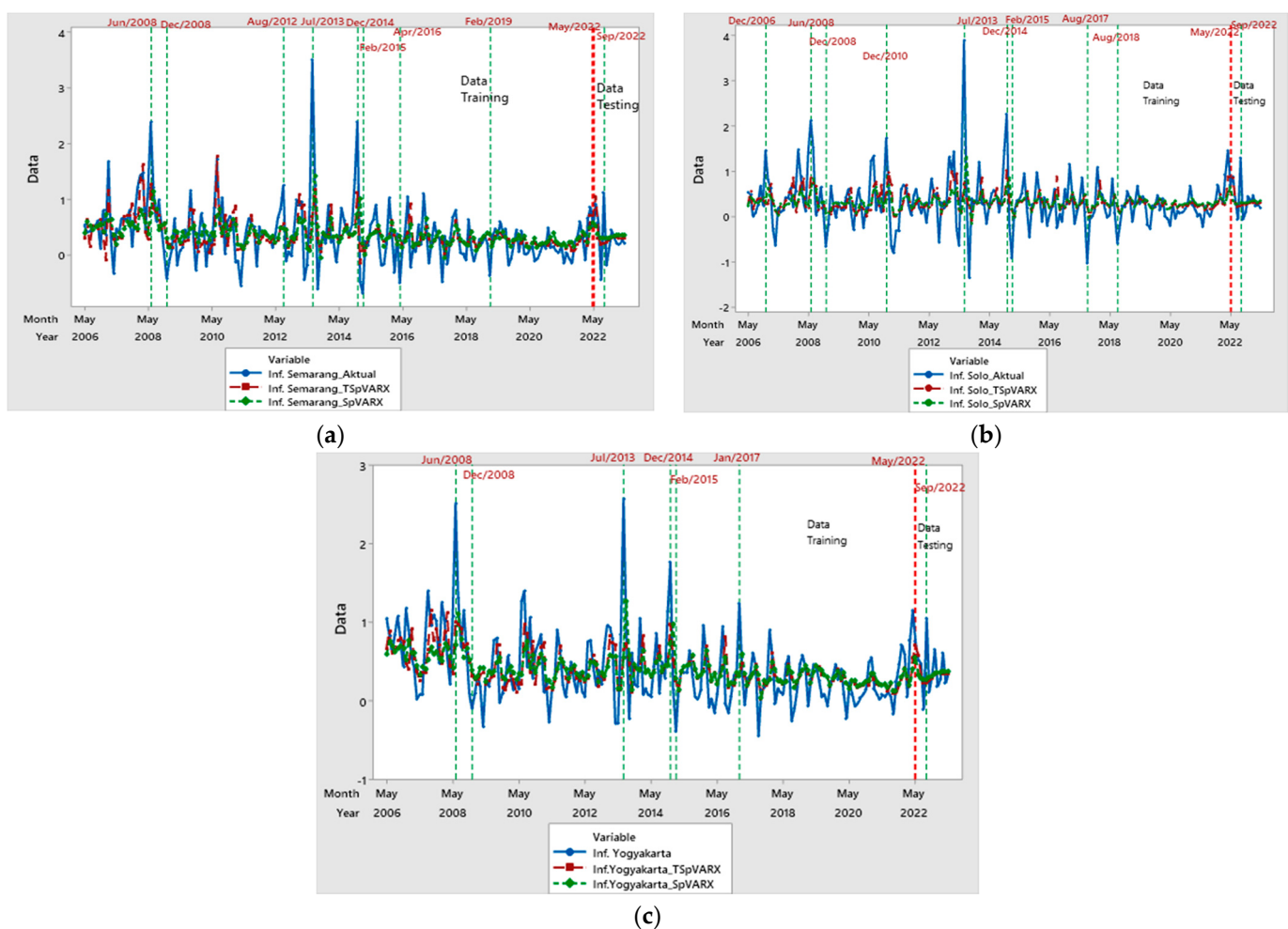


Figure 6. Plot of actual and forecast data using TSpVARX three regimes for (a) Semarang inflation, (b) Solo inflation, and (c) Yogyakarta inflation.

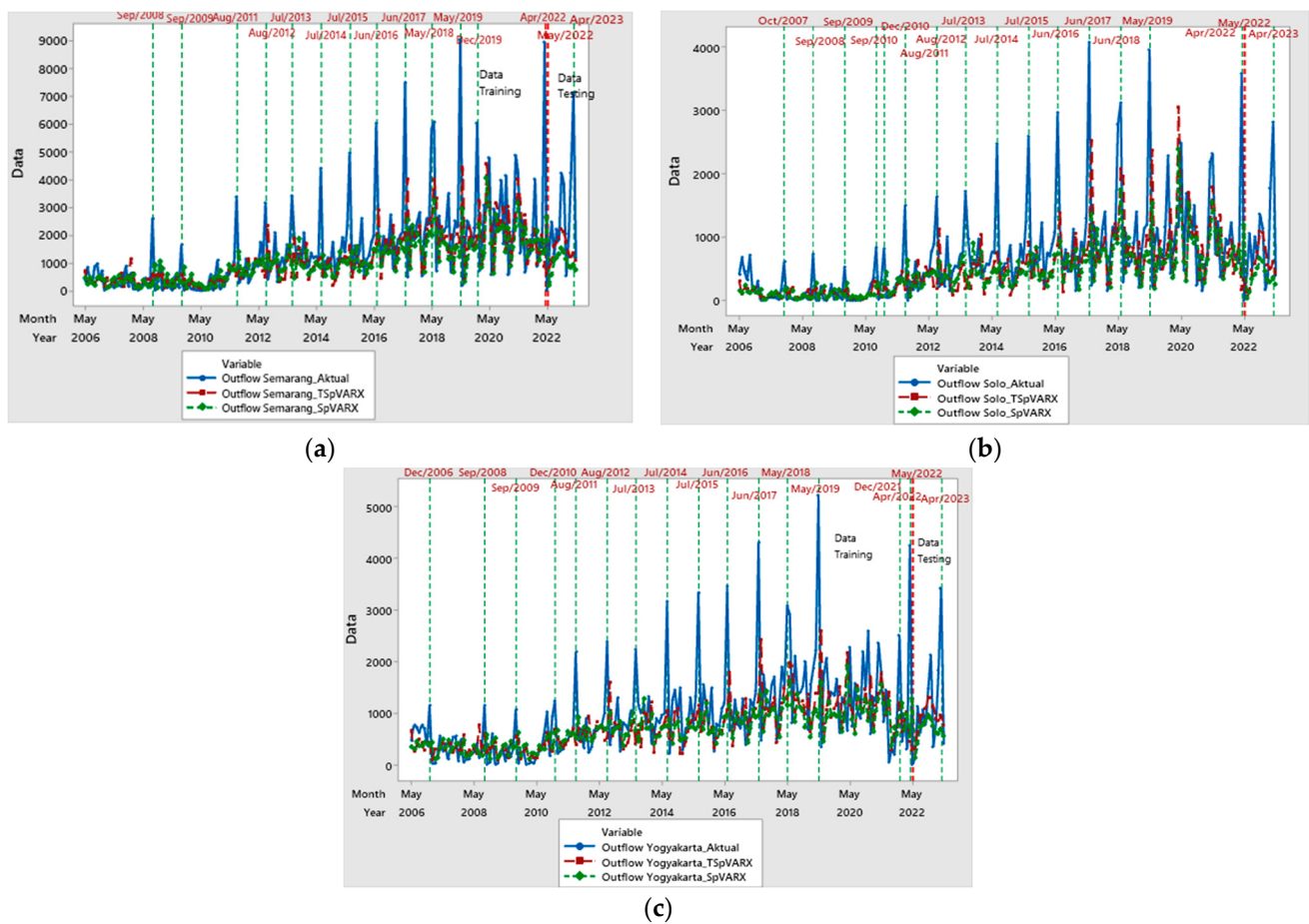


Figure 7. Plot of actual and forecasting data in training and testing data periods using TSpVARX with three regimes for money outflow of (a) Semarang, (b) Solo, and (c) Yogyakarta.

Table 8. Forecasting of inflation (in percent) and money outflow (in billion rupiahs) of Semarang, Solo, and Yogyakarta using TSpVARX three regimes.

Years	Months	Inflation of Semarang	Inflation of Solo	Inflation of Yogyakarta	Money Outflow of Semarang	Money Outflow of Solo	Money Outflow of Yogyakarta
2023	Jun	0.27	0.26	0.32	1170.09	412.69	673.65
2023	Jul	0.23	0.23	0.25	1303.44	527.27	784.50
2023	Ags	0.29	0.28	0.32	1384.05	613.21	910.39
2023	Sep	0.23	0.23	0.25	1523.43	698.45	927.87
2023	Oct	0.32	0.31	0.35	1523.95	733.44	1004.58
2023	Nov	0.28	0.25	0.32	1688.13	866.91	1137.90
2023	Dec	0.36	0.34	0.39	1431.25	659.18	992.05
2024	Jan	0.31	0.29	0.34	1311.64	551.14	917.04
2024	Feb	0.21	0.21	0.24	1368.85	569.88	861.08
2024	Mar	0.36	0.35	0.39	1342.39	640.51	946.17
2024	Apr	0.36	0.33	0.40	1333.95	619.51	985.76
2024	Mei	0.31	0.29	0.33	1519.46	855.06	1050.96
2024	Jun	0.21	0.21	0.24	1401.90	629.53	864.87
2024	Jul	0.17	0.16	0.19	1476.63	638.73	844.42
2024	Ags	0.23	0.23	0.26	1313.11	533.11	777.51
2024	Sep	0.24	0.25	0.28	1025.16	343.08	628.31
2024	Oct	0.29	0.30	0.33	904.17	290.27	608.09
2024	Nov	0.30	0.31	0.33	805.45	241.46	565.92
2024	Dec	0.29	0.30	0.33	766.44	221.82	546.40

3.7. Addition of the 12th Subset and Dummy Variables in the Form of Fuel Price Adjustments

Based on the previous section, it can be seen that the results of forecasting inflation and money outflow using the TSpVARX model are not so good because several forecasting values are still far from the actual data. In addition, the forecasting data of inflation and money outflow using the TSpVARX model are also still stable after the next three periods. This indicates that the influence of the increase and decrease in fuel prices and the Eid al-Fitr event have not been captured by the model. Therefore, in this section, a subset variable of the 12th lag of the endogenous variable and dummy variables of increasing and decreasing fuel prices will be added to the TSpVARX modeling. This addition is expected to help in capturing the effect of the Eid al-Fitr and fuel price adjustment event. Based on Table 9, it can be seen that there are five smallest RMSE values based on the residuals of the testing data obtained from the TSpVARX model with the addition of the 12th subset and the dummy variables. Based on the RMSE of the training data residuals, all the smallest RMSEs are obtained using the TSPVARX model with the addition of the 12th subset and dummy variables. Therefore, we can conclude that the addition of the 12th subset and the dummy variables can improve the performance of the TSpVARX model in forecasting the inflation and money outflow of Semarang, Solo, and Yogyakarta.

Table 9. Root Mean Square Error (RMSE) using testing and training data based on three-regime TSpVARX and three-regime TSPVARX models with the addition of the 12th subset and the dummy variables.

Data	Model	Inflation of Semarang ($\gamma_{1,t}^1$)	Inflation of Solo ($\gamma_{1,t}^2$)	Inflation of Yogyakarta ($\gamma_{1,t}^3$)	Money Outflow of Semarang ($\gamma_{2,t}^1$)	Money Outflow of Solo ($\gamma_{2,t}^2$)	Money Outflow of Yogyakarta ($\gamma_{2,t}^3$)
Testing	TSpVARX 3 Rezim	0.366	0.339	0.299	2130.05	827.40	866.93
	TSpVARX 3 Rezim dengan Subset 12 dan Dummy Pt dan St	0.340	0.318	0.319	1756.249	687.680	769.774
Training	TSpVARX 3 Rezim	0.445	0.535	0.398	1426.346	698.190	795.610
	TSpVARX 3 Rezim dengan Subset 12 dan Dummy Pt dan St	0.335	0.419	0.315	1204.770	634.635	686.866

The addition of the 12th subset variable causes some recurring patterns of inflation and money outflow increases due to the Eid al-Fitr event to be captured in the model. Thus, the resulting forecasting in the period of the Eid al-Fitr event is better than before the subset variable is added. The addition of dummy variables in the form of increases and decreases in fuel prices causes the pattern of increase or decrease in inflation and money outflow due to fuel price adjustments to be captured by the model. Therefore, the forecasting of inflation and money outflow when fuel price adjustments occur is closer to the actual data than before the dummy variables are included.

The performance comparison of TSpVARX with and without subset and dummy variables based on RMSE can be seen in Figure 8. Based on Figure 8, the RMSE value of the testing data and training data for inflation and outflow is smaller than the TSpVARX model with subset and dummy variables. This indicates that the TSpVARX model with subset and dummy variables has better performance compared to the TSpVARX model. The improvement in inflation and money outflow forecasting accuracy due to the addition of subset and dummy variables is useful for the government to take the right policy, especially in price control.

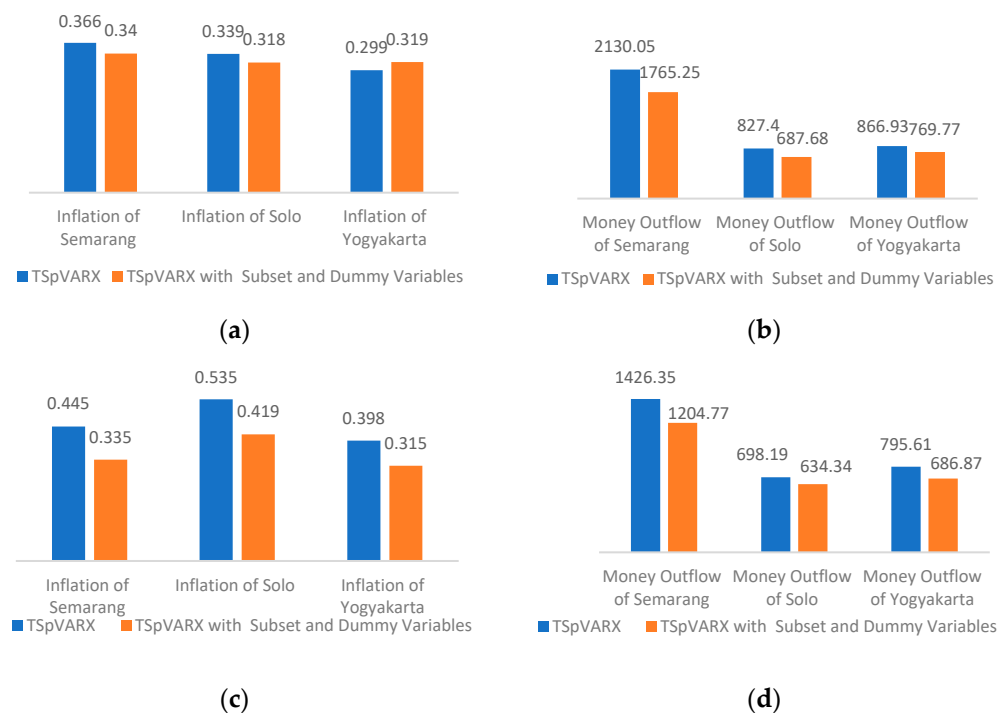


Figure 8. Root Mean Square Error (RMSE) of three-regime TSspVARX and three-regime TSspVARX models with the addition of the 12th subset and the dummy variables of (a) inflation testing data, (b) money outflow training data, (c) inflation testing data, and (d) money outflow testing data.

The comparison between the plots of actual and forecasting data can be seen in Figures 9 and 10. The blue line illustrates the movement of actual data, and the green line illustrates the movement of forecasting data from the TSspVARX model. The red line illustrates the movement of the forecast data of the TSspVARX model with the addition of the 12th subset and the dummy variables. Based on Figure 8, it can be seen that some points of increase and decrease in inflation can be captured by the TSspVARX model with the addition of the 12th subset and the dummy variables. The effects of some events on inflation captured in the model are as follows:

1. the effect of Christmas and New Year on inflation of Semarang, Solo, and Yogyakarta in December 2014;
2. the effect of Eid al-Fitr on the inflation of Semarang in August 2012;
3. the effect of Eid al-Fitr on inflation of Semarang, Solo, and Yogyakarta in July 2013;
4. the effect of a fuel price increase on inflation of Semarang, Solo, and Yogyakarta in June 2008;
5. the effect of a fuel price increase on inflation of Semarang, Solo, and Yogyakarta in November 2014;
6. the effect of a fuel price decrease on inflation of Yogyakarta in December 2008;
7. the effect of a fuel price decrease on inflation of Semarang, Solo, and Yogyakarta in February 2015;
8. the effect of a fuel price decrease on inflation of Semarang and Solo in February 2019.

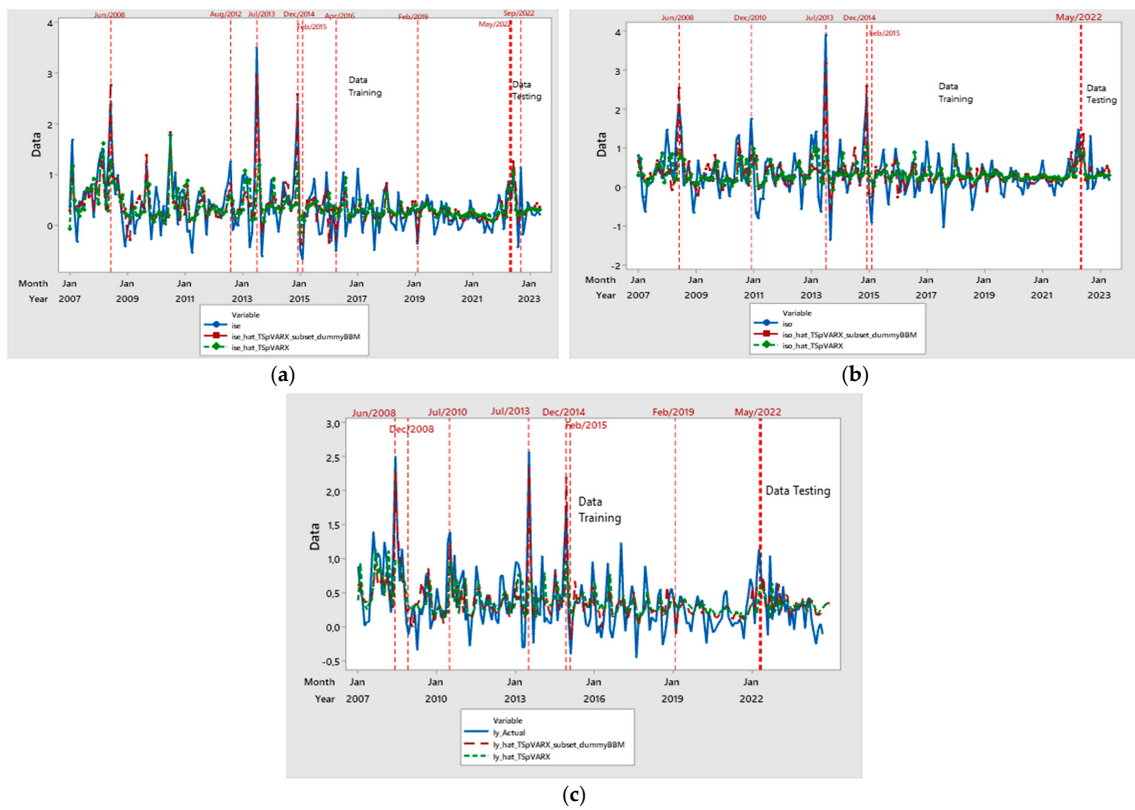


Figure 9. Plot of actual and forecasting data for (a) Semarang inflation, (b) Solo inflation, and (c) Yogyakarta inflation.

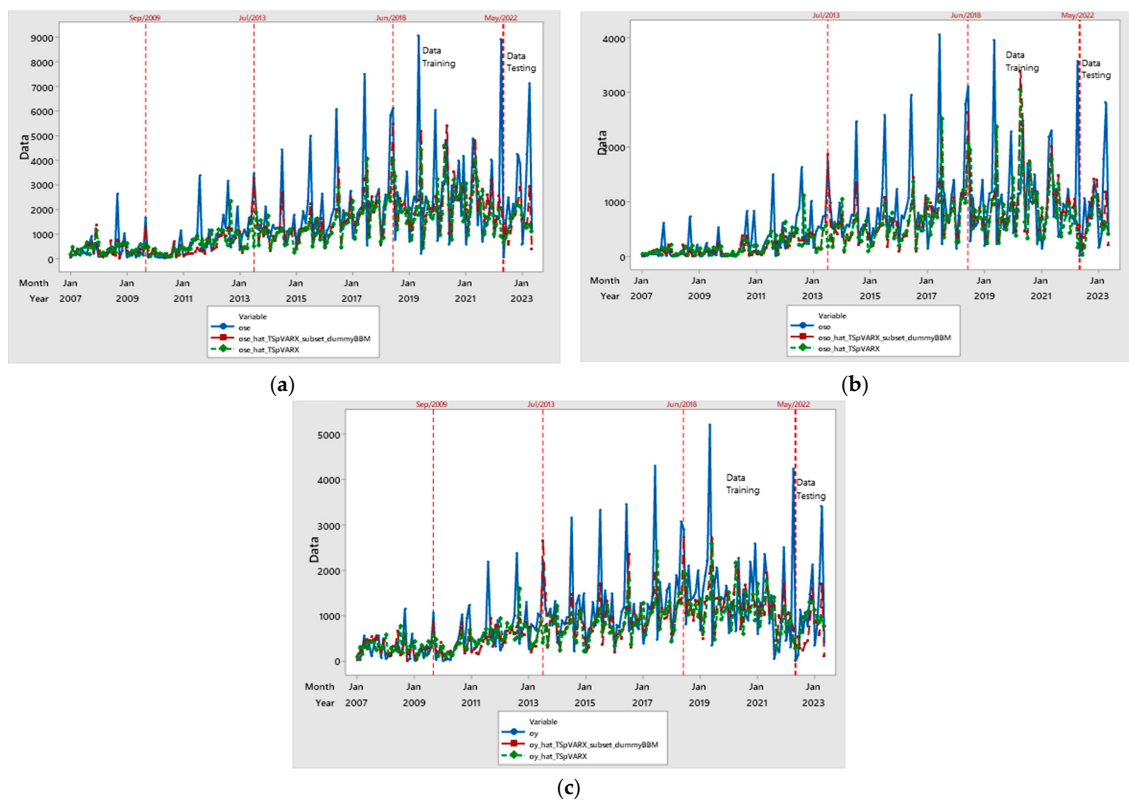


Figure 10. Plot of actual and forecasting data for (a) money outflow of Semarang, (b) money outflow of Solo, and (c) money outflow of Yogyakarta.

Based on Figure 10, it can be seen that some money outflow data points, which previously could not be captured by the three-regime TSpVARX, have been captured by the three-regime TSpVARX model with the addition of the 12th subset and the dummy variables. The effects of some events on money outflow captured in the model are as follows:

1. the effect of Eid al-Fitr on the money outflow of Solo and Yogyakarta in September 2009;
2. the effect of Eid al-Fitr on money outflow of Semarang, Solo, and Yogyakarta in July 2013;
3. the effect of Eid al-Fitr on the money outflow of Semarang, Solo, and Yogyakarta in July 2018;
4. the effect of Eid al-Fitr on money outflow of Semarang, Solo, and Yogyakarta in May 2022.

3.8. Forecasting Results Using the Selected Model

The results of forecasting using TSPVARX with the addition of the 12th subset and the dummy variables for June 2023 to May 2024 can be seen in Table 10. If we compare the forecasting result in Table 10 with Table 8, the forecasting result in Table 10 fluctuates more than in the Table 8. This indicates that the addition of the 12th subset and dummy variables can already capture the data pattern due to the Eid al-Fitr event and fuel price adjustment. By using the forecasting result generated from the TSpVARX model with the addition of subset variables and dummy variables, the government can use the right policy in controlling the inflation rate. For example, if the highest inflation forecast for the October–December 2024 period is in November 2024, then the government can take precautions, such as raising interest rates and conducting open market operations in the previous month. In addition, the government can also find out the estimated increase in inflation and money outflow during Eid al-Fitr. It can be anticipated for the government to ensure the availability of food and non-food needs for the community so that there is no scarcity.

Table 10. Forecasting result of inflation (in percent) and money outflow (in billion Rupiah) of Semarang, Solo, and Yogyakarta using TSpVARX with the addition of the 12th subset and the dummy variables.

Years	Months	Inflation of Semarang	Inflation of Solo	Inflation of Yogyakarta	Money Outflow of Semarang	Money Outflow of Solo	Money Outflow of Yogyakarta
2023	Jun	0.03	0.14	0.16	1496.36	392.09	319.76
2023	Jul	0.30	0.24	0.27	1875.41	622.51	770.40
2023	Ags	0.21	0.20	0.23	1433.36	634.06	832.20
2023	Sep	0.34	0.39	0.37	2069.90	897.37	982.46
2023	Oct	0.27	0.23	0.30	1735.58	809.57	972.46
2023	Nov	0.27	0.21	0.29	2359.80	1007.74	1253.77
2023	Dec	0.42	0.37	0.42	2293.82	797.98	1410.76
2024	Jan	0.32	0.28	0.31	1217.20	469.92	785.45
2024	Feb	0.24	0.23	0.22	1538.97	512.83	835.56
2024	Mar	0.38	0.33	0.42	1976.75	763.40	1207.94
2024	Apr	0.40	0.34	0.35	2633.24	918.13	1669.01
2024	Mei	0.29	0.23	0.33	1838.78	898.66	899.12
2024	Jun	0.24	0.17	0.19	1511.27	568.49	618.91
2024	Jul	0.21	0.16	0.18	1686.85	615.86	752.38
2024	Ags	0.26	0.19	0.19	1391.16	522.05	794.18
2024	Sep	0.37	0.31	0.27	1357.30	400.60	786.30
2024	Oct	0.37	0.31	0.33	1234.65	370.66	819.22
2024	Nov	0.43	0.38	0.35	1344.91	353.58	888.40
2024	Dec	0.39	0.34	0.35	1338.65	334.36	954.85

4. Discussion

The model formed in this study still has limitations. One of the limitations of this model is that it has not been able to overcome the singularity problem. Singularity problems can occur due to a high correlation between predetermined variables. The singularity problem can result in several things, including the model coefficients cannot be estimated and a measure of model goodness cannot be obtained. In this study, the AIC of the SpVARX (1, 1, 4) model with normalized cross-correlation weights could not be obtained due to the singularity problem in estimating the error covariance matrix used when estimating the SpVARX model coefficients with MLE. Future research can develop the TSpVARX model by using estimation methods that can overcome the singularity problem, such as the Least Absolute Shrinkage and Selection Operator (LASSO) and Ridge Regression.

Economic variables move together. This leads to the need for knowledge of how economic variables are interconnected. The TSpVARX model with subset and dummy variables developed in this study still has limitations in terms of the difficulty of interpreting the relationship between variables. Therefore, future research can still develop impulse responses from the TSpVARX model. Impulse responses can be used to help explore how a shock to one economic variable affects other variables in the system over time.

Improving forecasting accuracy based on inflation variables can help policymakers in making decisions. If central banks have more accurate inflation forecasts, they can set interest rates with greater confidence to maintain price stability. This reduces the risk of over-tightening or under-tightening policies, which can negatively impact economic growth. In future research, the TSpVARX model with the addition of subset and dummy variables can be applied to accommodate the influence of other events on inflation and money supply, such as crises, natural disasters, pandemics, tax increases, and electricity tariff increases.

5. Conclusions

The addition of the 12th subset and dummy variable can improve the forecasting performance of inflation and money outflow of Semarang, Solo, and Yogyakarta. We prove this based on the lower RMSE obtained from the TSpVARX model with the addition of subset and dummy variables compared to the regular TSpVARX model. The best model for forecasting inflation and money outflow of Semarang, Solo, and Yogyakarta in this study is the three-regime TSpVARX model with the addition of the 12th lag of endogenous variables as subset variables and a fuel price increase and decrease as dummy variables. The addition of the 12th subset variables can accommodate the movement pattern of inflation and money outflow that occurs due to the events of Eid al-Fitr, Christmas, and New Year. Meanwhile, the addition of dummy variables can accommodate the movement of inflation and money outflow that occurs due to the increase and decrease in fuel prices. By accommodating these events in the model, inflation and money outflow forecasting will be more accurate.

By using the TSpVARX model with the addition of subset and dummy variables, the government can take the right policy in controlling inflation during the Eid, Christmas, and New Year holidays. For example, if the inflation forecasting results are high in the event of Eid in the future, then the government can use a policy by increasing interest rates in that period. In addition, by using the TSpVARX model with the addition of subset and dummy variables, the government can also carefully consider the timing and magnitude of future fuel price increases. The TSpVARX model can also be applied not only to the case of inflation and money supply. For example, we can use the model to create an economic growth forecasting model by accommodating spatial aspects and catastrophic events, such as the COVID-19 pandemic.

At the stage of selecting the exogenous variable lag order (q) using the normalized cross-correlation weight, the AIC can only be calculated up to an order q of 4. This is due to the singularity problem in estimating the SpVARX model coefficients. Therefore, future research is recommended to develop the TSpVARX model with the Least Absolute Shrinkage and Selection Operator (LASSO) that can overcome the singularity problem. In

addition, the TS_pVARX model with the addition of the 12th subset and dummy variables still has limitations in interpretation. The TS_pVARX model may result in different signs of the coefficient estimation on each endogenous lag variable. Thus, it is difficult to obtain a clear conclusion about the relationship between predetermined variables and endogenous variables. Future research is recommended to develop impulse response calculations derived from the model.

Author Contributions: Conceptualization, S.S. and G.P.D.S.; methodology, G.P.D.S.; software, G.P.D.S.; validation, S.S., D.D.P. and M.S.A.; formal analysis, G.P.D.S. and A.A.K.; investigation, G.P.D.S.; resources, S.S.; data curation, G.P.D.S.; writing—original draft preparation, S.S. and G.P.D.S.; writing—review and editing, D.D.P., M.S.A. and A.A.K.; visualization, G.P.D.S.; supervision, D.D.P., M.S.A. and A.A.K.; project administration, S.S.; funding acquisition, S.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by ITS fund partnership research grant 2024.

Informed Consent Statement: Not applicable.

Data Availability Statement: Publicly available datasets were analyzed in this study. This data can be found here: www.bps.go.id (accessed on 1 October 2024).

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A

The value of the uniform spatial weight.

	Inflation of Semarang	Inflation of Solo	Inflation of Yogyakarta	Money Outflow of Semarang	Money Outflow of Solo	Money Outflow of Yogyakarta
Inflation of Semarang	0	0.5	0.5	0	0.5	0.5
Inflation of Solo	0.5	0	0.5	0.5	0	0.5
Inflation of Yogyakarta	0.5	0.5	0	0.5	0.5	0
Money Outflow of Semarang	0	0.5	0.5	0	0.5	0.5
Money Outflow of Solo	0.5	0	0.5	0.5	0	0.5
Money Outflow of Yogyakarta	0.5	0.5	0	0.5	0.5	0

The value of the normalized cross-correlation weight.

	Inflation of Semarang	Inflation of Solo	Inflation of Yogyakarta	Money Outflow of Semarang	Money Outflow of Solo	Money Outflow of Yogyakarta
Inflation of Semarang	0	0.502258	0.497742	0	−0.62789	−0.37211
Inflation of Solo	0.416833	0	0.583167	−0.25419	0	0.745807
Inflation of Yogyakarta	0.479102	0.520898	0	−0.60775	−0.39225	0
Money Outflow of Semarang	0	−0.3571	−0.6429	0	0.545937	0.454063
Money Outflow of Solo	−0.4671	0	−0.5329	0.515899	0	0.484101
Money Outflow of Yogyakarta	−0.56333	−0.43667	0	0.463612	0.536388	0

References

- Adelowokan, Adedayo Oluwaseyi, Bolaji Adesola Adesoye, and Titilope Maria Ogunmuyiwa. 2019. Impact of Open Market Operations and Money Supply on Inflation in Nigeria. *Acta Universitatis Danubius. Economica* 15: 230–44.
- Beenstock, Michael, and Daniel Felsenstein. 2007. Spatial vector autoregressions. *Spatial Economic Analysis* 2: 167–96. [CrossRef]
- Bello, Muhammad Zayyanu, and Mubarak Abdullahi Saulawa. 2013. Relationship between Inflation, Money Supply, Interest Rate and Income Growth (Rgdp) in Nigeria 1980–2010. An Empirical Investigation. *Journal of Economics and Sustainable Development* 4: 7–13.
- Bharadiya, Jasmin Praful. 2023. Exploring the Use of Recurrent Neural Networks for Time Series Forecasting. *International Journal of Innovative Science and Research Technology* 8: 2023–27. [CrossRef]
- Çitçi, Sadettin Haluk, and Hüseyin Kaya. 2023. Exchange rate uncertainty and the connectedness of inflation. *Borsa Istanbul Review* 23: 723–35. [CrossRef]
- Di Giacinto, Valter. 2010. On vector autoregressive modeling in space and time. In *Tem di Discussione*. no. 746. L'Aquila: Bank of Italy, Regional Economic Research Unit, L'Aquila Branch. Available online: <http://onlinelibrary.wiley.com/doi/10.1111/j.1538-4616.2010.00331.x/full> (accessed on 1 May 2023).
- Fadlurrohman, Alwan. 2020. Integration of GSTAR-X and Uniform Location Weights Methods for Forecasting Inflation Survey of Living Costs in Central Java. *Journal of Intelligent Computing and Health Informatics* 1: 23–29. [CrossRef]
- Gök, Burçin Çalışkan, and Ümit Bulut. 2021. The Interest Rate Pass-Through Process in Turkey: Empirical Evidence from Linear and Nonlinear Estimation Techniques. *Politics, Economics and Administrative Sciences Journal of Kırsehir Ahi Evran University* 5: 114–26.
- Lestari, Dian Widya Lestari, and Sekti Kartika Dini. 2024. Forecasting the Price of Shallots and Red Chilies Using the Autoregressive Integrated Moving Average with Exogenous Variable (ARIMAX) Model. *EKSAKTA: Journal of Sciences and Data Analysis* 5: 42–49.
- Lütkepohl, Helmut. 2005. *New Introduction to Multiple Time Series Analysis*. New York: Springer. [CrossRef]
- Shoko, Claris, and Caston Sigauke. 2023. American Journal of Infection Control Short-term forecasting of COVID-19 using support vector regression: An application using Zimbabwean data. *American Journal of Infection Control* 51: 1095–107. [CrossRef]
- Sohibien, Gama Putra Danu. 2017. Perbandingan Model STAR dan GSTAR untuk Peramalan Inflasi Dumai, Pekanbaru, dan Batam. *Statistika* 5: 14–26.
- Sohibien, Gama Putra Danu, Setiawan Setiawan, and Dedy Dwi Prastyo. 2024a. Spatial Vector Autoregressive with Metric Exogenous Variable (SpVARX) for Inflation and Outflow Forecasting. *Procedia Computer Science* 234: 140–47. [CrossRef]
- Sohibien, Gama Putra Danu, Setiawan Setiawan, and Dedy Dwi Prastyo. 2024b. Data forecasting performance evaluation of threshold spatial vector autoregressive with exogenous variables. *International Journal of Data and Network Science* 8: 523–36. [CrossRef]
- Stock, James H., and Mark W. Watson. 2020. *Introduction to Econometrics*, 4th ed. London: Pearson Education.
- Sumarminingsih, E., Setiawan, A. Suharsono, and B. N. Ruchjana. 2021. Comparison of Spatial Weights in Modeling Inflation and Money Supply Using the Spatial Vector Autoregressive Model with Calendar Variations. *IOP Conference Series: Materials Science and Engineering* 1115: 012088. [CrossRef]
- Tsagkanos, Athanasios, Anastasios Evgenidis, and Konstantina Vartholommatou. 2018. Financial and monetary stability across Euro-zone and BRICS: An exogenous threshold VAR approach. *Research in International Business and Finance* 44: 386–93. [CrossRef]
- Tsay, Ruey S. 2010. Analysis of financial time series. In *Analysis of Financial Time Series*, 3rd ed. Hoboken: Wiley. [CrossRef]
- Tsay, Ruey S. 2014. *Multivariate Time Series Analysis*. Chicago: Wiley.
- Wei, William W. S. 2006. *Time Series Analysis Univariate and Multivariate Method*. Boston: Addison-Wesley.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.