ENTHUSIASTIC
INTERNATIONAL JOURNAL OF APPLIED STATISTICS AND DATA SCIENCE
Volume 3, Issue 2, October 2023, pp. 176-188

The Implementation of the Generalized Space-Time Autoregressive (GSTAR) Model for Inflation Prediction

Feby Hestuningtias [a,1], Muhammad Hasan Sidiq Kurniawan [a,2,*]

[a] Statistic Study Program, Universitas Islam Indonesia, Jalan Kaliurang km. 14.5 Yogyakarta, Indonesia
[1] 19611058@students.uii.ac.id; [2] hasan.sidiq@uii.ac.id*
* Corresponding author

ARTICLE INFO

Keywords
Inflation
Time series
GSTAR
Forecasting

ABSTRACT

The macroeconomic indicator used to measure a country's economic balance is inflation. The increase in the price of goods and services causes an increase in inflation, which impacts the decrease in the value of money so that people's purchasing power for goods and services will decrease and result in slow economic growth. One way to determine future inflation is by forecasting. The Generalized Space-Time Autoregressive (GSTAR) model is a time series model involving time and location. This study aims to predict future inflation using the GSTAR model, which uses differencing without uniform location weights, inverse distance, and normalized cross-correlation. The results showed that the models obtained were the GSTAR (2,1) and GSTAR (5,1)I(1) models. The best model to predict inflation is the GSTAR (5,1)I(1) model with the normalized cross-correlation weight, which had Root Mean Square Error (RMSE) value of 0.5743, which was smaller than the GSTAR (2,1) model.

1. Introduction

The economic balance of a country can be measured by the macroeconomic indicator such as inflation. Inflation is a continuous increase in the price of goods and services in general [1]. The increase in the price of goods and services could cause inflation in a country. The impact of inflation includes the decrease in the value of money, resulting in purchasing power for goods and services as the basic needs of society. In other words, the decline in people's purchasing power will result in slow economic growth. Index numbers can be used to measure changes in inflation from time to time which are calculated from the prices of goods and services consumed by the public, and are referred to as the Consumer Price Index (CPI) [1]. CPI is one of the important economic indicators that can provide information about the development of prices for goods and services paid by consumers.

Inflation is an economic problem and a monetary phenomenon that always worries countries, including Indonesia [2]. The discussion about inflation has been increased globally. Sometimes, the inflation in some country is not affected by the domestic factors, but the global factors [3]. Provinces in Indonesia have experienced high inflation. The government explained that the inflation rate in Indonesia in July 2022 reached 4.94%, exceeding the government's target limit of 3%. One of the reasons for the soaring inflation in Indonesia was the high inflation in the food product, which reached 11.47% when it should not have exceeded 6%. One provinces in Indonesia with high inflation in July 2022 were Jambi, West Sumatra, and the Bangka Belitung Islands.

In July 2022, Jambi's inflation was 1.27%, and the calendar year inflation rate and year-on-year inflation were 6.96% and 8.55%, respectively [4]. West Sumatra's inflation was 1.22%, with a calendar year inflation rate of 6.49%, and year-on-year inflation of 8% in July 2022 [5]. In the same month and year, the Bangka Belitung Islands' inflation was 1.05% with calendar year inflation of 5.48%, and year-on-year inflation of 7.77% [6]. The elevated inflation in the three provinces can be attributed to numerous expenditures that were driven by people's needs, resulting in a higher inflation rate compared to other provinces. Hence, it is imperative to closely monitor inflation in order to stabilize the Indonesian economy, particularly in these provinces.

One way to determine future inflation is by forecasting. Inflation in a region is sometimes affected by an imbalance between demand and supply because not all provinces can provide goods and services. Therefore, each region needs its surrounding area to provide goods and services that cause dependence between regions. In another words, inflation has a spatial relationship [7]. The rise and fall of inflation between regions are influenced by spatial aspects so the influence of the region becomes the subject of this study in predicting the value of inflation. In addition to the spatial effect, the inflation rate tends to be influenced by inflation in the previous period. Considering these two factors, it can be said that inflation data can be classified as space-time data. [8]. The decision to take the spatial effects into consideration because the inflation is also affected by the trading decision of investors and the policy by government, which are maybe different between one and other area [9].

Of statistical analysis that can describe time series data and pay attention to location or spatial aspects is Generalized Space-Time Autoregressive (GSTAR). The GSTAR model is a time series model that involves locations with heterogeneous location parameters and stationary data conditions. This model consists of a time order and a spatial order. The temporal order can be determined using the Vector Autoregressive (VAR) model. The VAR model is a classic way to model time series data from nearby locations that tend to be related. Meanwhile, the determination of spatial order is limited to order 1 or λ_1 since higher orders are difficult to interpret [10]. The effect of time on the GSTAR model is indicated by the Autoregressive (AR) parameter and the spatial effect is expressed in the form of a weighting matrix. The location weight matrix is a matrix that expresses the relationship of the observation area measuring $N \times N$ and is symbolized by W. Some weighting matrices that can be applied to the GSTAR model include uniform weights, distance inverse, and cross-correlation normalization [11]. The chosen weighting matrices can impact the accuracy and the prediction results [12]. Some research about GSTAR has discussed the modified weighting matrices in order to get better prediction results. In other words, the decision to choose what weighting matrices that will be used equal to decide the influence of the surrounding area studied [13].

Research conducted by Islamiyah *et al*. used GSTAR to predict data on Tuberculosis (BTA+) sufferers in DKI Jakarta. The results showed that the GSTAR (1; 1) model with normalized cross-correlation location weights was the best GSTAR model compared to distance inverse weighting because it had a small Root Mean Square Error (RMSE) value [14]. Study by Balqis *et al*. modelled and forecasted inflation data using the GSTAR method in three cities in West Java [15]. Meanwhile, several studies have tried to predict sea tides on Java Island using the GSTAR model, the best GSTAR model results for tidal height data was the GSTAR (1;1)I(1) model using cross-correlation normalized weights [16].

This study examines the comparison between differencing and non-differentiated GSTAR models in the application of the GSTAR model to see valid forecasting results on the differences in the GSTAR models obtained. This study aims to determine the general description of inflation in the Provinces of Jambi, West Sumatra, and the Bangka Belitung Islands from January 2017 to December 2022. This involves estimating the best GSTAR Model and forecasting inflation using the best GSTAR Model.

p-ISSN 2798-253X
e-ISSN 2798-3153

2. Material and Method

2.1. Data and Resource

The data used in this study obtained from the websites of Statistics Indonesia of Jambi. The data used were monthly inflation data in Jambi, West Sumatra, and the Bangka Belitung Islands from January 2017 to December 2022.

2.2. Method of Analysis

The analysis method used in this study is descriptive analysis to determine an overview of inflation in Jambi, West Sumatra, and Bangka Belitung Islands. In addition, the GSTAR model was analyzed to predict inflation in those three provinces. The steps carried out in this study are shown in the flowchart of Fig. 1.

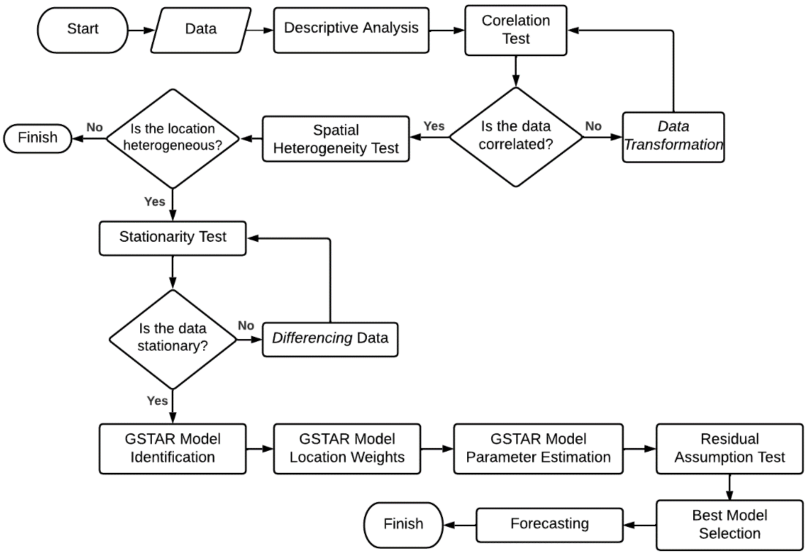


Fig. 1 Research flowchart.

The followings provide detailed explanation of the research steps. Descriptive analysis was used to visualize and to determine the time series components contained in the data. This step conducted to know the general description of the data. Next, the correlation test was used to determine the relationship between variables. If the variables are not related, then data transformation is carried out until the variables are related between each other. The spatial heterogeneity test used the Gini Index test, aiming to determine the presence of spatial diversity in the data. If the condition is not fulfilled, the GSTAR model cannot be estimated. The data stationary test was performed using the Augmented Dickey-Fuller (ADF) test. If the data are not stationary, it is necessary to do differencing until the condition is met. Identification of the GSTAR model consists of two orders: time order and spatial order. The time order can be seen by looking at the smallest AIC value, while the spatial order uses order 1. In this paper, the estimated GSTAR model was formed into two models: without differencing and with differencing. The location weights used in this study were uniform, distance inverse, and normalized cross-correlation. Parameter estimation of the GSTAR model used the Ordinary Least Square (OLS) method. The residual assumption test was conducted to see whether the residual was white noise. If the residual is white noise, then the model is feasible to use. Selection of the best model can be done by looking at the smallest RMSE value of the GSTAR model. The best model obtained was used to forecast the next period.

2.3. Generalized Space-Time Autoregressive (GSTAR)

The GSTAR model is one of the models used to predict time series data that has a relationship between time and location. This model is an extension of the Space-Time Autoregressive (STAR) model and tends to be more flexible than the STAR model.

p-ISSN 2798-253X
e-ISSN 2798-3153

The initial STAR model has a weakness in the assumption of autoregressive parameters that have the same value at all locations. Therefore, the STAR model is more suitable for locations with the same (homogeneous) characteristics [11]. The weakness of the STAR model was then corrected through the GSTAR model. The GSTAR model is more flexible because the assumptions of autoregressive parameters in this model are different for each location. As a result, it can be applied to locations that have heterogeneous characteristics shown in the form of a weighting matrix [9].

The GSTAR model is a model of several $Z_i(t)$ observations found at every N location in space (i = 1,2,..., N) over t periods. The time effect is formulated as a time series model and the spatial effect is formulated as a location weighting matrix [11]. The GSTAR modelled with autoregressive order (p) and spatial order $\lambda_1, \lambda_2, \dots, \lambda_k$ is denoted by GSTAR (p, λ_k). If the data are not stationary within the mean, then differencing is performed so that the GSTAR model formed is GSTAR (p, λ_k)I(d) [13]. The GSTAR model is formulated as follows:

$$Z_i(t) = \sum_{k=1}^{p} \left[\Phi_{k0} + \sum_{l=0}^{\lambda_k} \Phi_{kl} W^{(l)}\right] Z_i(t-k) + e_i(t) \quad (1)$$

where $Z_i(t)$ is the vector observation at time t and location i with size (N × 1). p is an autoregressive order (AR), λ_k is the spatial order to-k. N is the number of locations. $W^{(l)}$ is a location weighting matrix of size (N × N) with weighted values chosen to meet the conditions $W_{ii}^{(l)} = 0$ and $\sum_{i \neq j} W_{ij}^{(l)} = 1$, where $W^{(0)}$ is defined as the identity matrix I. Φ_{k0} is the diagonal of the autoregressive parameter matrix at the time lag to-k and spatial lag to-0(zero) with diagonal elements ($\Phi_{k0}^{1}, \dots, \Phi_{k0}^{N}$). Φ_{kl} is the diagonal of the autoregressive parameter matrix at the time lag to-k and spatial lag to-l with diagonal elements ($\Phi_{kl}^{1}, \dots, \Phi_{kl}^{N}$). $e_i(t)$ is a residual vector of size (N × 1) at the time to-t. The GSTAR Model is better than any other methods in case of point forecasting. Nevertheless, this approach does have limitations. It exhibits reduced accuracy when used for density forecasting and prediction in uncertain forecasting environments [17].

2.4. Spatial Heterogeneity Test

The Gini Index method is commonly used to determine the level of distribution of people's income [15]. It is a very representative analytical ratio for heterogeneous community data [18]. The Gini index is divided into several criteria: G = 0 which means perfect equality and G = 1 which means imperfect equality. If G ≥ 1 then the data are heterogeneous. The following is a spatial heterogeneity hypothesis test.

H_0 : Homogeneous location or perfect even distribution
H_1 : Heterogeneous location or imperfect distribution
Test statistics:

$$G = 1 + \frac{1}{n} - \frac{2}{(n^2 \bar{z}_i)} \sum_{i=1}^{n_i} z_i \quad (2)$$

where z_i is the value of the observed variable, $\bar{z}_i$ is the average value of the observed variables, n is the amount of data, and n_i is the amount of data at the location to-i. H_0 is rejected if the value of G is greater than or equal to 1.

2.5. GSTAR Model Location Weight

The characteristic of the space-time model is the relationship between time and location. Location relatedness in the GSTAR model is expressed in the W weighting matrix. A good location weight is a location weight that forms a model with a small prediction error [19]. Several weights that can be applied to the GSTAR model include uniform weights, distance inverse weights, and normalized cross-correlation weights.

2.5.1. Uniform Weights

Uniform location weights give the same weight value for each location. Thus, this location weight is often used on data that has the same (homogeneous) distance between locations [19]. Uniform weight is defined in the following equation:

$$W_{ij} = \begin{cases} \frac{1}{n_{ij}}, i \neq j \\ 0, i = j \end{cases} ; \sum_{j=1}^{N} W_{ij} = 1 \quad (3)$$

2.5.2. Distance Inverse Weights

Distance inverse weights are calculated based on the actual distance between locations. This weight gives a large weight value for close distances and gives a small weight for long distances. The distance used in this weight takes latitude and longitude into account. Latitude and longitude are then converted to kilometers. The distance inverse location weights are expressed in the following equation:

$$d_{ij} = \sqrt{\left(u_i - u_j\right)^2 + \left(v_i - v_j\right)^2} \quad (4)$$

$$W_{ij} = \frac{\frac{1}{d_{ij}}}{\sum_{j=1}^{n} \frac{1}{d_{ij}}} \quad (5)$$

where $i \neq j$ dan $\Sigma_{i \neq j} W_{ij} = 1$. d_{ij} is the distance from location i to j, $(u_i - u_j)$ is the latitude coordinate and $(v_i - v_j)$ is the longitude coordinate.

2.5.3. The Normalized Cross-Correlation Weights

The normalized weight of cross-correlation is based on the results of normalizing cross-correlation between locations at the appropriate time lag [20]. This weight does not imply certain rules, such as depending on the distance between locations [14]. Estimators of cross-correlation in sample data are as follows:

$$r_{ij}(k) = \frac{\sum_{t=k+1}^{n} [z_i(t) - \bar{z}_i][z_j(t-k) - \bar{z}_j]}{\sqrt{\left(\sum_{t=1}^{n} [z_i(t) - \bar{z}_i]^2\right)\left(\sum_{t=1}^{n} [z_j(t-k) - \bar{z}_j]^2\right)}} \quad (6)$$

Furthermore, determining location weights can be done by normalizing the cross-correlation quantities between locations at the corresponding time. This process generally produces location weights as follows:

$$W_{ij} = \frac{r_{ij}(k)}{\sum_{k \neq 1} |r_{ik}(k)|} \quad (7)$$

where $i \neq j$ and $\Sigma_{i \neq j} |W_{ij}| = 1$.

2.5.4. GSTAR Model Parameter Estimation

The GSTAR model can be expressed as a linear model. The autoregressive parameters of the GSTAR model can be estimated using the least squares method or the OLS method, which minimizes the sum of the squares of the residuals [15]. The estimated model parameters using the OLS method are as follows:

$$\hat{\Phi} = (Z^{*\prime} Z^{*})^{-1} (Z^{*\prime} Z) \quad (8)$$

The purpose of the forecasting model is to predict future values with the smallest possible error, one alternative for model selection based on error values can be done with RMSE for each model. RMSE values range from 0 to ∞. The smaller the RMSE value, the better the model used [11]. RMSE is determined using the following formula:

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^{n} (z_t - \hat{z}_t)^2} \quad (9)$$

where z_t is actual data and $\hat{z}_t$ is forecast data.

2.6. Residual Assumption Test

After obtaining the GSTAR model parameter estimates, the basic assumption that must be met is that the residual error vector is white noise to see the feasibility of the GSTAR model [21]. A model feasibility test is needed for the subsequent step to determine whether the model can be used for forecasting. The detection of white noise residuals can be conducted with the Ljung Box-Pierce test as follows.

$H_0 : \rho_1 = \cdots = \rho_k = 0$ (Residuals fulfilling the assumption of white noise)
$H_1 : \exists \rho_1 \neq 0$ (Residuals do not meet the white noise assumption)
Test statistics:

$$Q = n(n+2)\sum_{k=1}^{K}\frac{\hat{\rho}_k^2}{n-k} \tag{10}$$

$\hat{\rho}_k^2$ is residual lag autocorrelation to-k. H_0 is rejected if $Q > \chi^2_{(\alpha;df)}$.

3. Results and Discussion

3.1. Descriptive Analysis

This descriptive analysis was undertaken to provide an overview of inflation in three provinces on Sumatra Island. The time series plot for inflation data shown in Fig. 2.

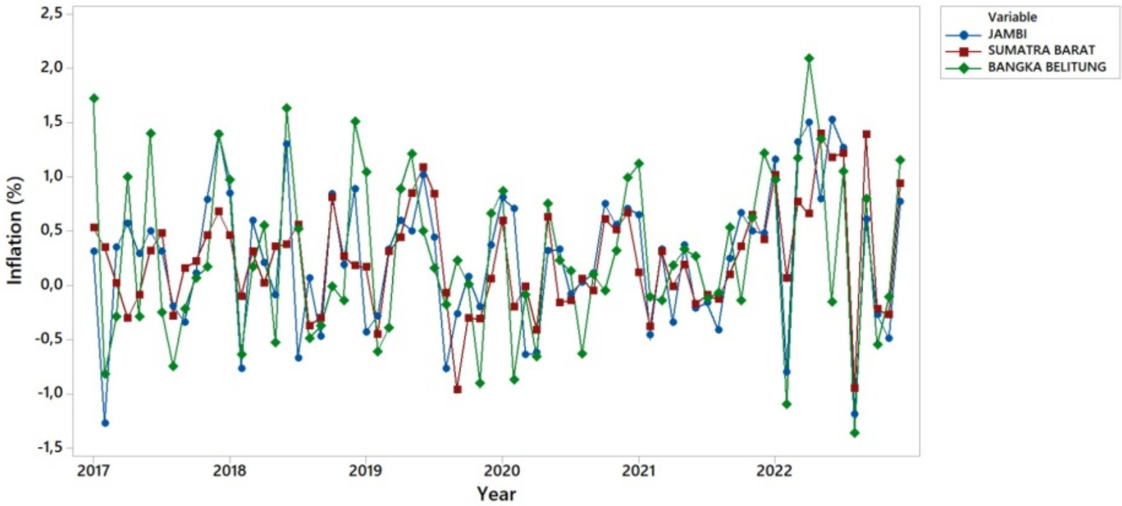


Fig. 2. Inflation Data Plot of Three Provinces on Sumatra Island

Based on Fig. 2, the inflation in the three provinces on Sumatra Island fluctuated but did not show a tendency to increase or decrease. Instead, it tended to be stationary on the mean value. From January 2017 to December 2022, the Bangka Belitung Islands Province saw the highest and lowest levels of inflation among the three provinces on Sumatra Island.

Table 1. Descriptive Statistics of Inflation Data of Three Provinces on Sumatra Island

Province	Mean (%)	Maximum (%)	Minimum (%)
Jambi	0.2507	1.53	-1.27
West Sumatra	0.2429	1.40	-0.96
Bangka Belitung Islands	0.2640	2.09	-1.36

Table 1 shows that the Bangka Belitung Islands Province had the highest average inflation rate. It indicates that the price of goods and services purchased in the Bangka Belitung is more expensive than Jambi and West Sumatra Provinces. Meanwhile, the lowest average inflation rate was in West Sumatra Province, which was 0.2429%. It indicates that the prices of goods and services purchased in West Sumatra Province are cheaper than in Jambi and Bangka Belitung. The highest inflation occurred in the Bangka Belitung Islands Province which was caused by an increase in prices for the commodities of air transportation, cooking oil, and kerisi fish.

3.2. Inflation Data Correlation Between Locations

The correlation between locations can be seen in Table 2.

p-ISSN 2798-253X
e-ISSN 2798-3153

Table 2. Correlation of Inflation between Provinces on Sumatra Island

	Jambi	West Sumatra	Bangka Belitung Islands
Jambi	1	0.6729	0.6562
West Sumatra	0.6729	1	0.5533
Bangka Belitung Islands	0.6562	0.5533	1

Table 2 shows that the correlation of inflation in Jambi, West Sumatra, and Bangka Belitung Islands had a value between 0.5533 to 0.6729. Hence, it can be said that inflation between these provinces is strong. In addition, that the surrounding provinces influence the inflation. It shows the spatial effect of the province.

3.3. Spatial Heterogeneity Test

The application of the GSTAR data model used must meet the assumption of heterogeneous characteristics. The spatial heterogeneity test was conducted by estimating the characteristics of each observation location using the Gini Index test statistic.

Table 3. Spatial Heterogeneity

Province	Gini Index Value
Jambi	1.001543
West Sumatra	1.001543
Bangka Belitung Islands	1.001543

Based on Table 3, the Gini Index value obtained for each province is larger than 1, then the decision is rejected H_0, which means there is heterogeneity between locations in inflation data in the Provinces of Jambi, West Sumatra, and the Bangka Belitung.

3.4. Inflation Data Stationary

To check the stationary of the data in this study, the ADF test was conducted.

Table 4. ADF Test Inflation Data

Province	ADF Test	*p-value*
Jambi	-4.539	0.01
West Barat	-3.4273	0.0583
Bangka Belitung Islands	-4.4216	0.01

According to the findings presented in Table 4, the results showed that inflation data for Jambi and Bangka Belitung were stationary, while inflation data for West Sumatra Province is not stationary using a significance level of $\alpha = 5\%$. Since there is only one province which is nonstationary, it can be said that the data is stationary. Therefore, the differencing process was not necessary. On the other hand, if the differencing process is being conducted to make the data stationary, it should be conducted on, not only the province that is not stationary, but also for all those three provinces. Therefore, this research compares the differencing and nondifferencing GSTAR models. Based on the results of the initial stationary test, the inflation data in West Sumatra Province is not stationary, thus necessitating the implementation of a differencing method. The differencing step is performed once.

Table 5. ADF Test Inflation Data after Differencing Process

Province	ADF Test	*p-value*
Jambi	-8.0150	0.01
West Barat	-7.0977	0.01
Bangka Belitung Islands	-8.9048	0.01

Based on Table 5, the p-value for the three locations is less than $\alpha = 5\%$, so H_0 is rejected. By using a 95% confidence level, it can be concluded that the inflation data for Jambi, West Sumatra, and the Bangka Belitung Islands are already stationary, or the data does not contain a unit root so that it can be proceed to the next analysis to compare the differencing and non-differencing GSTAR models.

p-ISSN 2798-253X
e-ISSN 2798-3153

3.5. GSTAR Model Identification

The GSTAR model has two orders: time order and spatial order. Spatial order is generally limited to order 1 because higher orders are difficult to interpret [9]. Meanwhile, the time order (autoregressive) can be done using the VAR model order, in this case, obtaining the appropriate model is determined by the optimal lag length based on the smallest AIC value.

Table 6. AIC Value in the VAR Model for the GSTAR Model that is not differentiated

Lag	**1**	**2**	**3**	**4**	**5**
AIC	-4.0684	-4.1548	-3.9343	-3.7848	-3.7295
Lag	**6**	**7**	**8**	**9**	**10**
AIC	-3,6381	-3,5672	-3,5333	-3,4096	-3,5423

Based on Table 6, it is found that the time order for the non-differencing GSTAR model was in the lag 2 because lag 2 had the smallest AIC value (-4.1548). Hence, the appropriate GSTAR model is GSTAR (2,1).

Table 7. AIC value in the VAR Model for the GSTAR Model which is differentiated

Lag	**1**	**2**	**3**	**4**	**5**
AIC	-2,6837	-2,9612	-3,0167	-3,1144	-3,2728
Lag	**6**	**7**	**8**	**9**	**10**
AIC	-3,1849	-3,2005	-3,1658	-3,1969	-3,1384

Based on Table 7, it is found that the time order for the differencing GSTAR model was on lag 5 because it has the smallest AIC value of -3.2728. Therefore, the appropriate GSTAR model is GSTAR (5,1)I(1).

3.6. Location Weight

The characteristic of the space-time model is the relationship between time and location. Location-relatedness in the GSTAR model is expressed in a weighting matrix. This study used three weights, uniform location weights, distance inverse location weights, and cross-correlation normalized weights. The results of the location Weight for each method are written.

a. Uniform location weight

$$W = \begin{bmatrix} 0 & 0.5 & 0.5 \\ 0.5 & 0 & 0.5 \\ 0.5 & 0.5 & 0 \end{bmatrix}$$

b. Distance inverse location weights

$$W = \begin{bmatrix} 0 & 0.5093 & 0.4907 \\ 0.6707 & 0 & 0.3293 \\ 0.6624 & 0.3376 & 0 \end{bmatrix}$$

c. Location weights of normalized cross-correlation for non-differencing models

$$W = \begin{bmatrix} 0 & 0.6009 & -0.3991 \\ 0.6222 & 0 & 0.3778 \\ -0.5225 & 0.4775 & 0 \end{bmatrix}$$

d. Location weights of normalized cross-correlation for differencing models

$$W = \begin{bmatrix} 0 & -0.5439 & -0.4561 \\ -0.7085 & 0 & -0.2915 \\ -0.6708 & -0.3292 & 0 \end{bmatrix}$$

3.7. GSTAR Model Parameter Estimation

GSTAR model parameter estimation used OLS method, which is to minimize the sum of the squares of the residuals. There are two GSTAR model parameter estimation results, GSTAR model parameter estimation (2,1) for the model without differencing and GSTAR model parameter estimation (5,1)I(1) for differencing model.

p-ISSN 2798-253X
e-ISSN 2798-3153

Table 8. Parameter Estimation of GSTAR (2,1) model with Uniform Location Weight, Distance Inverse, and Cross-Correlation Normalization

Parameter	Estimate		
	Uniform Location Weight	*Distance Inverse Weight*	*Cross-Correlation Normalization Weight*
Φ_{10}^{1}	0.0312	0.0297	0.1786
Φ_{10}^{2}	0.1176	0.0997	0.1074
Φ_{10}^{3}	-0.1686	-0.1883	0.0855
Φ_{20}^{1}	-0.0733	-0.0741	0.0887
Φ_{20}^{2}	0.1653	0.1562	0.1601
Φ_{20}^{3}	-0.3175	-0.2955	-0.0018
Φ_{11}^{1}	0.1861	0.1414	0.3293
Φ_{11}^{2}	0.1173	0.1596	0.1183
Φ_{11}^{3}	0.3951	0.5152	-0.4667
Φ_{21}^{1}	0.2066	0.1556	-0.0293
Φ_{21}^{2}	0.0777	0.1030	0.0771
Φ_{21}^{3}	0.4215	0.4522	0.7417

The GSTAR model (2,1) with uniform location weights of Jambi Province is

$$Z_1(t) = 0{,}0312Z_1(t-1) + 0{,}0931Z_2(t-1) + 0{,}0931Z_3(t-1) - 0{,}0733Z_1(t-2) + 0{,}1033Z_2(t-2) + 0.1033Z_3(t-2) + e_1(t)$$

Table 9. Parameter Estimation of GSTAR (5,1)I(1) model with Uniform Location Weight, Distance Inverse, and Cross-Correlation Normalization

Parameter	Estimate		
	Uniform Location Weight	*Distance Inverse Weight*	*Cross-Correlation Normalization Weight*
Φ_{10}^{1}	-0.8325	-0.8340	-0.8391
Φ_{10}^{2}	-0.7709	-0.7819	-0.7846
Φ_{10}^{3}	-0.9285	-0.9399	-0.9410
Φ_{20}^{1}	-0.5871	-0.5886	-0.5937
Φ_{20}^{2}	-0.4358	-0.4513	-0.4554
Φ_{20}^{3}	-0.9834	-0.9687	-0.9670
Φ_{30}^{1}	-0.2528	-0.2536	-0.2564
Φ_{30}^{2}	-0.3290	-0.3556	-0.3624
Φ_{30}^{3}	-0.9205	-0.9247	-0.9249
Φ_{40}^{1}	-0.2880	-0.2876	-0.2864
Φ_{40}^{2}	-0.3492	-0.3538	-0.3555
Φ_{40}^{3}	-0.4545	-0.4276	-0.4246
Φ_{50}^{1}	-0.3092	-0.3089	-0.3079
Φ_{50}^{2}	-0.2293	-0.2121	-0.2087
Φ_{50}^{3}	-0.2728	-0.2595	-0.2582
Φ_{11}^{1}	0.0164	0.0138	-0.0184
Φ_{11}^{2}	-0.0014	0.0195	-0.0236
Φ_{11}^{3}	0.0679	0.1378	-0.1575
Φ_{21}^{1}	-0.0445	-0.0320	0.0263
Φ_{21}^{2}	-0.0308	-0.0080	0.0010
Φ_{21}^{3}	0.2988	0.3422	-0.3726
Φ_{31}^{1}	-0.3312	-0.2481	0.2386
Φ_{31}^{2}	-0.0669	-0.0396	0.0290
Φ_{31}^{3}	0.4293	0.5203	-0.5695
Φ_{41}^{1}	-0.3880	-0.2922	0.2862
Φ_{41}^{2}	-0.1581	-0.1699	0.1603
Φ_{41}^{3}	-0.2850	-0.3492	0.3827
Φ_{51}^{1}	-0.2190	-0.1646	0.1602
Φ_{51}^{2}	-0.1979	-0.2485	0.2435

p-ISSN 2798-253X
e-ISSN 2798-3153

Parameter	Estimate		
	Uniform Location Weight	*Distance Inverse Weight*	*Cross-Correlation Normalization Weight*
Φ_{51}^{3}	-0.5889	-0.7061	0.7718

One of the GSTAR model equations (5,1)I(1) with cross-correlation normalization weight for Jambi Province is as follows:

$$Z_1(t) = -0.8391Z_1(t-1) + 0.0100Z_2(t-1) + 0.0084Z_3(t-1) - 0.5937Z_1(t-2) - 0.0143Z_2(t-2)$$

$$- 0.0120Z_3(t-2) - 0.2564Z_1(t-3) - 0.1298Z_2(t-3) - 0.1088Z_3(t-3) - 0.2864Z_1(t-4)$$

$$- 0.1557Z_2(t-4) - 0.1305Z_3(t-4) - 0.3079Z_1(t-5) - 0.0871Z_2(t-5) - 0.0731Z_3(t-5)$$

$$+ e_1(t)$$

3.8. Testing Residual Assumptions

The GSTAR (2,1) model and the GSTAR (5,1)I(1) model are said to be suitable if the residuals are white noise. Testing the assumption of residual white noise uses the Ljung-Box Pierce (LB) test. This test is used to determine whether the model can be used for forecasting. If the assumptions are not white noise, then the model is not suitable for forecasting.

Table 10. Result of White Noise Residual Assumption Test

GSTAR Models	Location Weight	Ljung Box-Pierce Value
GSTAR (2,1)	Uniform	34.356
	Distance Inverse	33.764
	Cross-Correlation Normalization	41.311
GSTAR (5,1)I(1)	Uniform	63.067
	Distance Inverse	62.060
	Cross-Correlation Normalization	61.946

Based on Table 10, uniform location weighting, distance inverse, and normalized cross-correlation in the GSTAR (2,1) and GSTAR (5,1)I(1) models had met the assumption of residual white noise or there is no autocorrelation between residuals because the Ljung Box-Pierce value < $\chi^2_{(0,05;71)} = 91.67024$ so it fails to reject H_0. It means that the GSTAR (2,1) and GSTAR (5,1)I(1) models are suitable for use on uniform location weights, inverse distances, or normalized cross-correlation in forecasting inflation in Jambi, West Sumatra, and the Bangka Belitung.

3.9. Best GSTAR Model Selection

The criteria used in selecting the best GSTAR model can be seen from the smallest RMSE value. RMSE is used as a measure of model goodness and forecast accuracy on inflation data in Jambi, West Sumatra, and the Bangka Belitung. The smaller the RMSE value, the better the model will be.

Table 11. RMSE on GSTAR Model (2,1) and GSTAR Model (5,1)I(1)

GSTAR Models	Location Weight	RMSE
GSTAR (2,1)	Uniform	0.6340
	Distance Inverse	0.6344
	Cross-Correlation Normalization	0.6412
GSTAR (5,1)I(1)	Uniform	0.5762
	Distance Inverse	0.5747
	Cross-Correlation Normalization	0.5743

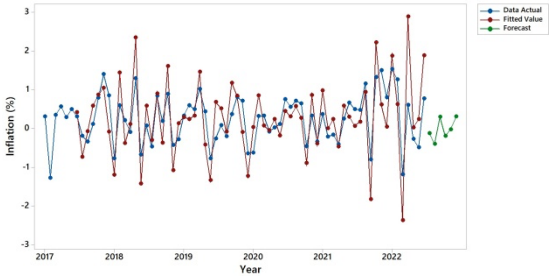
Based on Table 11, the results show that the GSTAR (5,1)I(1) model with normalized cross-correlation weights has the smallest RMSE mean value of 0,5743 compared to the GSTAR (2,1) model with other weights. The GSTAR (5,1)I(1) model with normalized cross-correlation weights is the best model that can be used to predict inflation data in Jambi, West Sumatra, and Bangka Belitung in the next period.

3.10. Inflation Data Forecasting

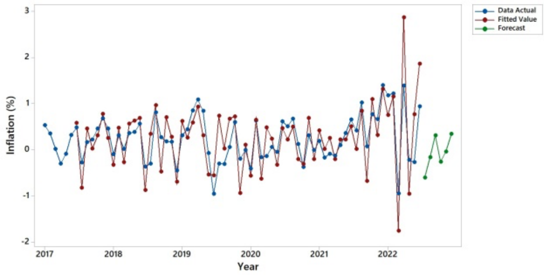
Forecasting the GSTAR (5,1)I(1) model with normalized cross-correlation weights on inflation data in the Provinces of Jambi, West Sumatra, and the Bangka Belitung Islands is carried out for the next six periods. Forecasting results are shown in Table 12.

Table 12. Forecasting using the GSTAR Model (5,1)I(1)

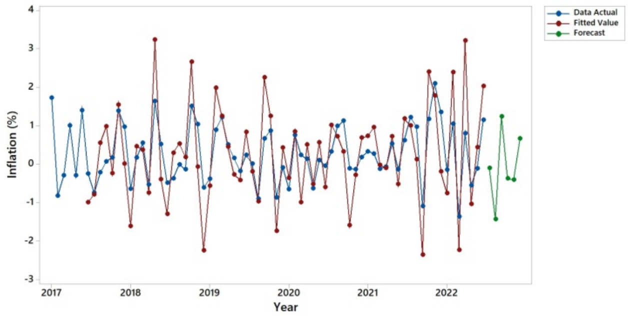
Period	Jambi Province	West Sumatra Province	Bangka Belitung Islands Province
January 2023	-0.12	-0.60	-0.10
February 2023	-0.40	-0.17	-1.44
March 2023	0.30	0.31	1.24
April 2023	-0.19	-0.26	-0.38
May 2023	-0.02	-0.04	-0.41
June 2023	0.31	0.34	0.66



(a)



(b)



(c)

Fig. 3 Inflation forecasting plot of (a) Jambi Province, (b) West Sumatra Province, and (c) Bangka Belitung Islands Province.

Based on Fig. 3, it is found that the forecasting results are quite capable of following the actual data pattern in each province, even though there are fitted values that are quite different from the actual data values. This is because the results of the fitted value data do not only look at the actual data elements but also involve the location weights used. Forecasting results show that inflation in the Provinces of Jambi, West Sumatra, and the Bangka Belitung Islands over the next six months tends to be constant around the mean value.

4. Conclusion

In the Bangka Belitung Province, the largest inflation in January 2017 until December 2022 was 0.2640%. It shows that the price of goods and services purchased in the Bangka Belitung is more expensive than in Jambi and West Sumatra Provinces. We obtained the differencing and non-differencing GSTAR models respectively, GSTAR (2,1) and GSTAR (5,1)I(1) model. The GSTAR (5,1)I(1) model with normalized cross-correlation location weights was better than the GSTAR (2,1) model. It is because the GSTAR (5,1)I(1) model has the smallest RMSE value of 0.5743. Forecast using the best model, the GSTAR (5,1)I(1) model with normalized cross-correlation location weights indicated that it was quite capable of following the actual data pattern. Forecasting results show that

p-ISSN 2798-253X
e-ISSN 2798-3153

inflation in the Provinces of Jambi, West Sumatra, and the Bangka Belitung over the next six months will tend to be constant around the mean value.

References

[1] S. Suseno and S. Astiyah, *Inflasi*. Jakarta: Pusat Pendidikan dan Studi Kebanksentralan (PPSK) Bank Indonesia, 2009.

[2] P.A. Daniel, "Analisis Pengaruh Inflasi terhadap Laju Pertumbuhan Ekonomi di Kota Jambi," *Jurnal of Economics and Business*, vol. 2, no. 1, pp. 131–136, Mar. 2018, doi: 10.33087/ekonomis.v2i1.37.

[3] A. Kabukçuoğlu and E. Martínez-García, "Inflation as a Global Phenomenon—Some Implications for Inflation Modeling and Forecasting," *Journal of Economic Dynamics and Control*, vol. 87, pp. 46–73, Feb. 2018, doi: 10.1016/j.jedc.2017.11.006.

[4] Badan Pusat Statistik Jambi, *Perkembangan Indeks Harga Konsumen Provinsi Jambi Juli 2022*. Jambi, Indonesia, 2022.

[5] Badan Pusat Statistik Sumatera Barat, *Perkembangan Indeks Harga Konsumen Sumatra Barat Juli 2022*. Sumatra Barat, Indonesia, 2022.

[6] Badan Pusat Statistik Kepulauan Bangka Belitung, *Perkembangan Indeks Harga Konsumen Provinsi Kepulauan Bangka Belitung*. Bangka Belitung, Indonesia, 2022.

[7] L. Faizah and S. Setiawan, "Pemodelan Inflasi di Kota Semarang, Yogyakarta, dan Surakarta dengan pendekatan GSTAR," *Jurnal Sains Dan Seni Pomits*, vol. 2, no. 2, pp. 317–322, Apr. 2013, doi: 10.12962/j23373520.v2i2.4866.

[8] N. Nur'Eni, D. Lusiyanti, and I. Gunawan, "Identifikasi Model Generalized Space-Time Autoregressive (Gstar) untuk Nilai Inflasi di Pulau Sulawesi," *Jurnal Ilmiah Matematika dan Terapan*, vol. 18, no. 1, pp. 75–83, Jun. 2021, doi: 10.22487/2540766X.2021.v18.i1.15522.

[9] G. Zhao, M. Xue, and L. Cheng, "A New Hybrid Model for Multi-Step WTI Futures Price Forecasting Based on Self-Attention Mechanism and Spatial–Temporal Graph Neural Network," *Resources Policy*, vol. 85, pp. 1–18, Aug. 2023, doi: 10.1016/j.resourpol.2023.103956.

[10] D.S. Wutsqa, Suhartono, and B. Sutijo, "Generalized Space-Time Autoregressive Modeling," in *Proceedings of the 6th IMT-GT Conference on Mathematics, Statistics and its Applications (ICMSA2010)*, 2010, pp. 752–761.

[11] M.I.T. Mario, Kartiko, and R.D. Bekti, "Pemodelan Generalized Space Time Autoregressive (GSTAR) untuk Peramalan Tingkat Inflasi di Pulau Jawa," *Jurnal Statistika Industri dan Komputasi*, vol. 6, no. 2, pp. 171–184, Jul. 2021.

[12] N.M. Huda and N. Imro'ah, "Determination of the Best Weight Matrix for the Generalized Space Time Autoregressive (GSTAR) Model in the COVID-19 Case on Java Island, Indonesia," *Spatial Statistics*, vol. 54, p. 100734, Apr. 2023, doi: 10.1016/j.spasta.2023.100734.

[13] U.S. Pasaribu, U. Mukhaiyar, N.M. Huda, K.N. Sari, and S.W. Indratno, "Modelling COVID-19 Growth Cases of Provinces in Java Island by Modified Spatial Weight Matrix GSTAR Through Railroad Passenger's Mobility," *Heliyon*, vol. 7, no. 2, Feb. 2021, doi: 10.1016/j.heliyon.2021.e06025.

[14] A.N. Islamiyah, W. Rahayu, and E.D. Wiraningsih, "Pemodelan Generalized Space Time Autoregressive (GSTAR) dan Penerapannya pada Penderita TB Paru (BTA+) di DKI Jakarta," *Jurnal Statistika dan Aplikasinya*, vol. 2, no. 2, pp. 36–48, Dec. 2018, doi: 10.21009/JSA.02205.

[15] V.P. Balqis, E. Kurniati, and O. Rohaeni, "Model Peramalan Data Inflasi dengan Metode Generalized Space Time Autoregressive (Gstar) pada Tiga Kota di Jawa Barat," in *Prosiding Matematika*, 2020, pp. 43–50.

[16] C.A. Widyastuti, A. Hoyyi, and R. Rahmawati, "Peramalan Pasang Surut Air Laut di Pulau Jawa Menggunakan Model Generalized Space Time Autoregressive (GSTAR)," *Jurnal Gaussian*, vol. 5, no. 4, pp. 623–632, Nov. 2016, doi: 10.14710/j.gauss.5.4.623-632.

[17] E.Z. Chini, "Forecasting Dynamically Asymmetric Fluctuations of the U.S. Business Cycle," *International Journal of Forecasting*, vol. 34, no. 4, pp. 711–732, Oct.–Dec. 2018, doi: 10.1016/j.ijforecast.2018.05.003.

[18] H.D. Karlina, R. Cahyandari, and A.S. Awalluddin, "Aplikasi Model Generalized Space Time Autoregressive (GSTAR) pada Data Jumlah TKI Jawa Barat dengan Pemilihan Lokasi Berdasarkan Klaster DBSCAN," *Jurnal Matematika Integratif*, vol. 10, no. 1, pp. 37–48, Apr. 2014, doi: 0.24198/jmi.v10.n1.10183.37-48.

[19] D. Anggraeni, A. Prahutama, and S. Andari, "Aplikasi Generalized Space Time Autoregressive (GSTAR) Pada Pemodelan Volume Kendaraan Masuk Tol Semarang," *Jurnal Media Statistika*, vol. 6, no. 2, pp. 71–80, Dec. 2013, doi: 10.14710/medstat.6.2.61-70.

[20] F.N. Aryani, S.S. Handajani, and E. Zukhronah, "Penerapan Model Generalized Space Time Autoregressive (GSTAR) pada Data Nilai Tukar Petani 3 Provinsi di Pulau Sumatra," in *Prosiding Seminar Nasional Pendidikan Matematika Universitas Pekalongan*, 2020, pp. 209–220.

[21] E. Siswanto, H. Yasin, and Sudarno, "Pemodelan Generalized Space Time Autoregressive (GSTAR) Seasonal pada Data Curah Hujan Empat Kabupaten di Provinsi Jawa Tengah," *Jurnal Gaussian*, vol. 8, no. 4, pp. 418–427, Nov. 2019, doi: 0.14710/j.gauss.8.4.418-427.

p-ISSN 2798-253X
e-ISSN 2798-3153