Zero : Jurnal Sains, Matematika, dan Terapan
E-ISSN : 2580-5754; P-ISSN : 2580-569X
Volume 9, Number 2, 2025
DOI: 10.30829/zero.v9i2.26119
Page: 511-523

GSTAR (1;1) Transfer Function Model for Forecasting Chili Prices with Rainfall Effect

[1] **Yundari**
Department of Mathematics, Universitas Tanjungpura, Pontianak, Indonesia

[2] **Asri Rahmawati**
Department of Mathematics, Universitas Tanjungpura, Pontianak, Indonesia

[3] **Yuyun Eka Pratiwi**
Department of Mathematics, Universitas Tanjungpura, Pontianak, Indonesia

Article Info

Article history:

Accepted, 30 October 2025

Keywords:

Chili price forecasting;
Exogenous variable;
GSTAR-TF Model;
Rainfall Effect;
Space-Time Analysis.

ABSTRACT

Chili price fluctuations are strongly influenced by climate variability, distribution inefficiencies, and spatial. The ARIMA–TF and GSTARX models can address these issues. However, GSTARX is generally limited to linear responses of exogenous variables, while ARIMA–TF does not capture spatial heterogeneity. This study proposes an integration of the GSTAR and Transfer Function approaches. This model combines spatial-temporal dependencies with the external factors such as rainfall. The modeling process begins with GSTAR modeling, followed by parameter estimation of the transfer function using the Marquardt algorithm. The data used are rainfall and chili prices in three locations in West Kalimantan. The model data estimation results show very good accuracy, with MAPE values of 15%, 8%, and 14%. These results confirm that rainfall is a crucial factor influencing chili price variations and demonstrate the reliability of GSTAR–TF in forecasting agricultural commodities in the face of climate variability.

Corresponding Author:

Yundari,
Department of Mathematics,
Universitas Tanjungpura, Pontianak, Indonesia
Email: yundari@math.untan.ac.id

1. INTRODUCTION

Chili peppers are one of the agricultural commodities with high price volatility, directly affecting regional inflation and community welfare [1]. In West Kalimantan, chili production centers are located in three regencies: Pontianak, Kubu Raya, and Sintang, which supply local markets alongside chili imported from Java. Climatic factors, distribution systems, and market structures are the primary determinants of chili price formation across regions [2]. Three regencies in West Kalimantan are Sambas, Ketapang, and Kapuas Hulu, representing distinct environmental and distributional conditions [3] in Figure 1. Sambas, a coastal region with a humid tropical climate, is characterized by high rainfall. Its proximity to international trade routes at the Malaysia border makes local chili prices influenced by domestic factors and external supply flows [4]. Ketapang, which encompasses coastal and

inland areas, is a major agricultural center in southern West Kalimantan. Its variable rainfall patterns present challenges for price stability. At the same time, the district's reliance on sea and land transportation makes chili prices vulnerable to distribution disruptions caused by weather or infrastructure constraints [5]. Meanwhile, Kapuas Hulu, located upstream and bordering Betung Kerihun National Park, experiences high rainfall [6]. Due to limited accessibility and isolated markets, agricultural products face distribution constraints, making local chili prices more volatile and highly sensitive to local supply shocks. Differences in rainfall patterns and distribution capacity make Kapuas Hulu a critical case for examining the influence of climatic variables on price dynamics.

Despite differing climatic and distributional characteristics, these regions are interconnected through regional trade flows. Price fluctuations in one regency may influence prices in neighboring areas through direct or intermediary market linkages, highlighting spatial interdependence in chili price formation. Given these conditions, forecasting methods that account for spatial and temporal linkages are highly relevant. The Generalized Space-Time Autoregressive (GSTAR) model is designed to model interregional dependencies, which is widely applied to several application data such as chili price [7], Covid case [8], GDP[9], tea production[10], Gamma ray log data[11], copper and gold grade[12], GDP[13], and Rice price[14]. The transfer Function approach allows the integration of exogenous variables.

Previously, the ARIMA–Transfer Function (ARIMA-TF) model [15],[16] was applied to single location modeling, in which ARIMA was employed for the transfer function and residual analysis. However, this approach is less efficient when extended to multiple locations. Another method that incorporates exogenous variables across locations is the GSTARX model [17]. GSTARX extends the GSTAR framework by including explanatory variables (X) that directly influence spatio-temporal dependent variables [18]. Structurally similar to regression within a space–time framework, GSTARX incorporates external variables linearly into the model. This model is particularly suitable when exogenous variables directly influence the primary process and aim to enhance predictive power by explicitly including external factors in the spatio-temporal autoregressive structure.

In this study, the GSTAR–TF model is developed to describe the dynamic relationships between input (exogenous) and output variables, including lagged effects, transient responses, and long-term impacts. Unlike GSTARX, which treats explanatory variables as covariates, the GSTAR–TF model establishes transfer functions that characterize how changes in exogenous variables influence the dependent variable across time and space. This approach offers greater flexibility in capturing complex input–output dynamics, such as the delayed effects of rainfall on chili prices that may emerge only after several weeks. This research expects rainfall, which is time- and location-dependent, to influence chili prices, which are also time- and location-dependent. This model is expected to be a new alternative in spatial and temporal modeling that considers external variables/factors.

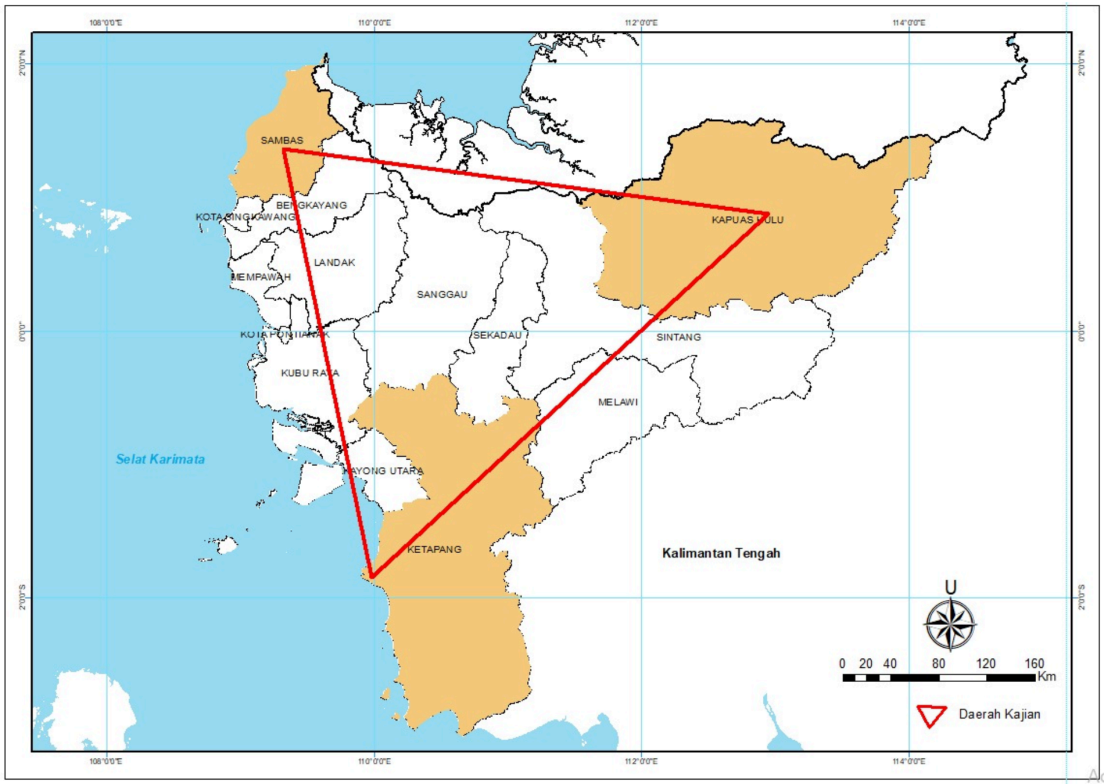


Figure 1. Research locations in three regencies of West Kalimantan

2. RESEARCH METHOD

The paper begins with a literature review necessary for understanding the models, namely the GSTAR model and the Transfer Function model. The GSTAR modelling process follows the three iterations of the Box–Jenkins's methodology, which include the identification stage starting with stationarity testing, parameter estimation, and model diagnostic testing. The Transfer Function modelling involves prewhitening the input and output series, computing the cross-correlation function to determine the model order, and parameter estimation and residual analysis. In the results and discussion section, modelling is first conducted using the GSTAR model, followed by the Transfer Function modelling with the input and output derived from the residuals of the GSTAR model. Residual diagnostics are then performed again using the GSTAR framework.

2.1 Data preprocessing

The data used depends on time and location. The dependent variables are monthly chili price data and rainfall as an exogenous variable. Each data set is centered by subtracting its mean. After that, both data sets must be stationary with respect to the mean and variance. If the data is not stationary, it must be differentiated or transformed. The differentiation process is carried out if the data is not stationary with respect to the mean. This process is done by subtracting the previous data. This process is recommended to be carried out only twice, in other words, the order of differentiation should be a maximum of second order. Meanwhile, stationarity with respect to variance is tested visually through data plots. If the data is not stationary with respect to variance, a Box-Cox transformation is performed. Once the data is stationary (weakly stationary), the process can continue with the GSTAR modeling stage.

2.2 Space-Time Autoregressive Model

The space–time model that accounts for past data at a given time in one location and across locations was first introduced by Cliff and Ord (1975) and is known as the STAR model. This model represents a space–time autoregressive (AR) process. A centred process {Y(t)} follows the STAR $(p;\ \lambda_1,\cdots,\lambda_p)$ model, it can be expressed as:

$$\boldsymbol{Y}(t) = \sum_{k=1}^{p}\sum_{\ell=0}^{\lambda_k} \phi_{k\ell}\boldsymbol{Y}(t-1) + \phi_{k\ell}\boldsymbol{W}^{(\ell)}\boldsymbol{Y}(t-k) + \boldsymbol{\varepsilon}(t) \tag{1}$$

Observation $\boldsymbol{Y}(t)$ is the vector of observations at N locations, with $\boldsymbol{Y}(t) = (Y_1(t), Y_2(t),\cdots, Y_N(t))'$. The notation $\phi_{k\ell}$ denotes the autoregressive parameters at temporal lag k and spatial lag ℓ. The matrix $\boldsymbol{W}^{(\ell)}$ represents the spatial weight matrix at spatial lag ℓ. When $\ell = 0$ the weight matrix reduces to the identity matrix, i.e., $\boldsymbol{W}^{(0)} = I$, so that the model simplifies to the standard autoregressive (AR) model in time series analysis. The AR(p) model indicates that the state of a time series at time t can be regressed on its own past values up to p previous periods [19]. Comprehensive discussions of the AR(p) model can be found in [20], [21], and [22].

If $\ell > 0$ the weight matrix is a matrix with zero elements on the main diagonal and row sums equal to one. The notation $\boldsymbol{\varepsilon}(t)$ represents the error vector at time t, which is assumed to be independently and identically distributed as usual, with zero mean and constant variance. The error vector is defined as $\boldsymbol{\varepsilon}(t) = (\varepsilon_1(t), \varepsilon_2(t),\cdots, \varepsilon_N(t))$. [23] later generalized the STAR model into the GSTAR model. This generalization is applied to the parameter assumptions across different locations. The GSTAR model is specifically designed for locations with heterogeneous geographical characteristics. Research on GSTAR has been widely conducted by Indonesian scholars due to Indonesia's diverse geographical conditions.

The GSTAR model of order$(p;\ \lambda_1,\cdots,\lambda_p)$ is adapted from Equation (1) by replacing the scalar parameter $\phi_{k\ell}$, with a diagonal matrix $\Phi_{k\ell} = diag(\phi_{k\ell}^{(1)},\dots,\phi_{k\ell}^{(N)}$. Based on its temporal and spatial orders, the simplest GSTAR model is of order (1;1). The GSTAR(1;1) model can be expressed as a Vector Autoregression VAR(1) form as follows [24]:

$$\boldsymbol{Y}(t) = \boldsymbol{\Phi}\boldsymbol{Y}(t-1) + \boldsymbol{\varepsilon}(t) \tag{2}$$

with the parameter matrix defined as $\boldsymbol{\Phi} = \boldsymbol{\Phi}_0 + \boldsymbol{\Phi}_1\boldsymbol{W}$. In the VAR(1) model, stationarity can be identified from the parameter matrix [22]. Consequently, the stationarity condition for the GSTAR (1;1) model can be determined by adapting the stationarity condition of VAR (1), namely that all eigenvalues of the parameter matrix must have absolute values less than one. Equation (2) is thus employed when examining the stationarity and invertibility of the GSTAR (1;1) model.

The notation of the parameter matrices for GSTAR (1;1), adapted from Equation (1), is given as $\Phi_0 = diag(\phi_{01},\dots,\phi_{0N})$, $\Phi_1 = diag(\phi_{11},\dots,\phi_{1N})$ and $W = (w_{ij})$. The GSTAR (1;1) model in matrix form can be expressed as:

$$\boldsymbol{Y}(t) = \boldsymbol{\Phi_0}\boldsymbol{Y}(t-1) + \boldsymbol{\Phi_1}\boldsymbol{W}\boldsymbol{Y}(t-1) + \boldsymbol{\varepsilon}(t) \tag{3}$$

Equation (3) is applied using the least squares method in the parameter estimation stage. The representation of the GSTAR (1;1) model for an individual observation, namely the observation at location i at time t, can be written as:

$$Y_i(t) = \phi_{0i} Y_i(t-1) + \phi_{1i} \sum_{j=1, j\neq i}^{N} w_{ij} Y_j(t-1) + \varepsilon_i(t) \quad (4)$$

with ϕ_{0i} and ϕ_{1i} denoting the temporal AR parameter and the spatial AR parameter, respectively, for location i. Equation (4) is employed in examining the statistical properties of the process $\{Y_i(t)\}$.

This space–time process belongs to the class of stochastic processes. The central assumption in space–time modelling is stationarity. Stochastic processes commonly recognize two types of stationarity: strong and weak. A process is considered strongly stationary if its distribution is invariant with time, meaning that all process moments remain constant over time. If only the first moment (mean) and second moments (variance/covariance) are invariant over time, then the process is said to be weakly stationary. Weak stationarity becomes equivalent to strong stationarity if the process is assumed to be identically and normally distributed. A detailed discussion of stationarity can be found in [25] and [22]. In this dissertation, the term stationarity refers specifically to weak stationarity.

The stationarity condition ensures that the GSTAR (1;1) model can be expressed as a linear combination of current and past error terms. In time series analysis, this form is known as the moving average (MA) model [26]. The equivalence of the GSTAR(1;1) model to the GSTMA representation is referred to as the invertibility property [27].

A distinctive feature of space–time models is the presence of both spatial and temporal dependence. The GSTAR (1;1) model represents spatial dependence through the spatial weight matrix. Studies on GSTAR generally assume that the first-order spatial weight matrix, W, it is fixed before modeling. There is, however, no standard technique for determining this matrix. Commonly, either uniform weighting based on the number of nearest neighbors or distance-based weighting schemes are used. Further discussions on the specification of nearest neighbors and distance-based weighting systems can be found in[28], [11], and [29].

Based on its definition, the spatial weight matrix $\boldsymbol{W}$ has the following characteristics [28].

1. For each location, the zeroth-order neighbourhood zone has only one member, namely the location itself; thus, the spatial weight matrix takes the form of the identity matrix.
$$\boldsymbol{W} = \boldsymbol{I_N}.$$
2. The diagonal elements of the spatial weight matrix are zero, w_{ii}, because the spatial interaction between a location indexed by i, it only occurs in the zeroth-order neighbourhood zone.
3. The sum of each row equals one $\sum_{j=1}^{N} w_{ij} = 1$. This property is used to preserve the mean value of $Y_i(t)$ in the GSTAR model.

Based on its construction, the spatial weight matrix has the following properties:

1. It may be symmetric, but it is generally not symmetric.
2. It is non-stochastic, since its determination is based on physical or geographical conditions of the observed locations.
3. Observational data do not determine it and do not vary over time.
4. It is subjective, as no standard method exists for its determination. Consequently, two researchers may adopt different weighting schemes for the same dataset.

These properties apply to a predetermined spatial weight matrix.

Parameter estimation is a central issue in modelling. The estimation method used is the least squares (LS) method, which estimates parameters by minimizing the sum of squared model errors. This is the best unbiased estimator when the model errors are mutually independent, have zero mean, and constant variance [30]. The method is particularly suitable for application in general linear models using RStudio software. Nevertheless, these assumptions are insufficient, and an additional assumption is required: the current error must be uncorrelated with the past process [22]. Furthermore, the normality assumption is often added to facilitate statistical inference, such as constructing confidence intervals and point estimation.

Note that for $t = 1,2,\ldots,T$, the GSTAR (1;1) model in Equation (2) can be expressed simultaneously as a linear model:

$$\boldsymbol{Y} = \boldsymbol{X\Phi} + \boldsymbol{\varepsilon}. \quad (5)$$

The parameter vector $\boldsymbol{\Phi}$ in Equation (5) can be estimated using the least squares method, which minimizes the sum of squared error vectors. The resulting estimator is given by:

$$\widehat{\boldsymbol{\Phi}} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{Y}. \quad (6)$$

The next step after parameter estimation is validation, also known as diagnostic testing. Diagnostic tests consist of two stages: the stationarity test of the model and the residual test. The stationarity test is conducted after obtaining parameter estimates, as the estimated parameters must satisfy stationarity conditions. [29] introduced a new approach to testing the stationarity of the GSTAR (1;1) model using the inverse of autocovariance matrix (IAcM). The stationarity condition in this approach is expressed through the IAcM, $\boldsymbol{M} = \mathbf{I}_N - \boldsymbol{\Phi}\boldsymbol{\Phi}'$, it must be a positive definite matrix. If the stationarity condition based on IAcM is satisfied, then the stationarity requirement described by [22] is also fulfilled. This stage is often referred to as model re-identification, since the IAcM provides analytical identification of the model [29].

If the GSTAR (1;1) model passes the stationarity test via IAcM, the next stage is to examine the model residuals. Residual diagnostics assess whether the residuals conform to the error assumptions. Several procedures are used for residual testing [29]:

1. Residual plot over time. If the residuals fluctuate steadily around zero within a specific range, they can be considered to originate from a random process.
2. STPACF plot. If the plotted values fall within the correlation bounds, the residuals are considered uncorrelated over time.
3. Histogram or normal Q-Q plot. If the histogram exhibits a unimodal and symmetric distribution, the residuals can be regarded as normally distributed. Similarly, if the Q-Q plot displays a linear pattern, the residuals are assumed to follow a normal distribution.
4. Scatter plot. A scatter plot of residuals can be used to examine residual correlations, either with respect to time or across locations.
5. Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE). The best model has the smallest RMSE, indicating lower residual variability.

2.3 ARIMA-Transfer Function (ARIMA-TF) Model

The transfer function model explains that the prediction of future values of a time series (output series, y_t) does not solely depend on its historical data but also considers one or more other time series (input series, x_t) related to the output series [31]. The construction of the transfer function model is based on the Autocorrelation Function (ACF) and the Cross-Correlation Function (CCF). The general form of the transfer function model for a single input (x_t) and a single output (y_t) is

$$y_t = v_0 x_t + v_1 x_{t-1} + v_2 x_{t-2} + \cdots + v_n x_{t-n} + n_t$$
$$y_t = (v_0 + v_1 B + v_2 B^2 + \cdots + v_n B^n) x_t + n_t \quad (7)$$

where y_t is a stationary output series, x_t is a stationary input series, n_t is the error term (noise series) that follows a certain time series model, such as ARIMA, and $v(B)$ represents the coefficients of the transfer function model or the impulse response weights.

The relationship between the output and input series can be expressed as:

$$y_t = v(B)x_t + n_t \quad (8)$$

where $v(B)$ the transfer function shows how the input affects the output through the backshift operator.

$$v(B) = \frac{\omega_s(B)B^b}{\delta_r(B)}$$

Here, n_t is the noise component. If the noise is modelled as an ARMA process with an AR parameter $\phi(B)$ and MA parameter $\theta(B)$, then:

$$n_t = \frac{\theta(B)}{\phi(B)} e_t$$

After representing the relationship between the input and output series through impulse response weights and modelling the noise component using an ARMA process, both components can be combined to form the transfer function model:

$$y_t = \frac{\omega_s(B)B^b}{\delta_r(B)} x_{t-b} + \frac{\theta(B)}{\phi(B)} e_t \quad (9)$$

With:

b = the number of periods before the input series begins affects the output series.

$\omega_s(B)$=$(\omega_0 - \omega_1 B - \omega_2 B^2 - \cdots - \omega_s B^s)$ an operator of orders, representing the number of past observations of x_t that influence y_t.

$\delta_r(B)$= $(1 - \delta_1 B - \delta_2 B^2 - \cdots - \delta_r B^r)$ an operator of order$r$, representing the number of past observations of the output series itself that influence y_t.

3. RESULT AND ANALYSIS

In this paper, the transfer function employed is the GSTAR transfer function, whereby the standard transfer function model's univariate input–output structure is extended to the multivariate case. The transfer function is commonly applied within the ARIMA framework. In this study, it is further developed under the spatial-temporal GSTAR (1;1) model. If ARIMA-TF modeling is applied to each location, it will be ineffective and inefficient. Through GSTAR-TF, the model can be estimated simultaneously through a matrix without having to perform modeling one by one for each location.

3.1 GSTAR-Transfer Function (GSTAR-TF) Modelling

The modelling process begins with preparing the input and output series. The input series, rainfall $\{X_i(t)\}$, is modelled using the GSTAR (1;1) model as expressed in Equation (4). Before this step, stationarity tests must be conducted for each series, and non-stationary series should be transformed accordingly. The GSTAR (1;1) model parameters for the input series, estimated using the OLS method (Equation (6)), are then applied to the output series, chili prices $\{Y_i(t)\}$. The next step involves prewhitening the input and output series, obtained as

$$a_i(t) = X_i(t) - \phi_{0i}X_i(t-1) + \phi_{1i}\sum_{j=1,j\neq i}^{N} w_{ij}X_j(t-1)$$

and

$$b_i(t) = Y_i(t) - \phi_{0i}Y_i(t-1) + \phi_{1i}\sum_{j=1,j\neq i}^{N} w_{ij}Y_j(t-1).$$

In GSTAR (1;1) modelling, autocorrelation is important in determining the cross-correlation coefficients. Subsequently, cross-correlations between the prewhitened series are computed. These prewhitened series serve as the basis for building the transfer function model.

As stated in Equation (8), the transfer function model is adapted so that $\{a_i(t)\}$ and $\{b_i(t)\}$ follow a GSTAR (1;1) process. The cross-correlation function between $\{a_i(t)\}$ dan $\{b_i(t)\}$ is then used to determine the order of the transfer function model. The estimated $v(B)$ of the transfer function, namely, $\omega_s(B)$ dan $\delta_r(B)$ indicates how the input influences the output under the backshift operator. The resulting transfer function model is subsequently employed for residual modelling using the GSTAR (1;1) framework. The GSTAR–TF obtained from this process is then applied as the final model, linking the output data $\{Y_i(t)\}$ and the input data $\{X_i(t)\}$. The transfer function through GSTAR (1;1) is thus represented as:

$$Y_i(t) = \sum_{j=1}^{3} \frac{\omega_s^{(j)}(B)B^{bj}}{\delta_r{}^{(j)}(B)} X_j(t-b) + \mu_i(t) \tag{10}$$

with $\mu_i(t)$ denoting the noise component modelled using GSTAR (1;1).

3.2 Case Study

This study utilizes rainfall and chili price data from three locations in West Kalimantan: Sambas, Ketapang, and Kapuas Hulu Regencies. In the GSTAR–TF modelling, the dependent variable (output) is the price of chili (Y), while the independent variable (input) is rainfall data (X). The temporal index covers monthly observations from January 2020 to December 2024, and the spatial index consists of the three regencies in West Kalimantan. Both rainfall and chili price data exhibit dependencies across time and space, supporting spatial-temporal modelling (GSTAR). The research locations are presented in Figure 1. Since all locations are within the same province and the monthly time dependence intuitively influences chili prices and rainfall, the model is specified with one temporal and one spatial lag, GSTAR (1;1).

The research data, comprising rainfall and chili prices across the three locations, are illustrated in Figure 2. Based on the plot in Figure 2(a), all three regencies experienced extreme chili price spikes in early 2021 and 2022 and a significant increase in December each year. Meanwhile, rainfall data (Figure 2(b)) appear relatively stable across years, consistent with West Kalimantan's location on the equator.

The stages of GSTAR spatial-temporal modelling, as described in the previous subsection, begin with examining data stationarity. This step is crucial to enable reliable forecasting. If the data are stationary, parameter estimates are expected to be unbiased, and forecasting results are more accurate. When the data are non-stationary, transformations can be applied to achieve stationarity. Stationarity can be assessed both visually and through inferential tests. Visually, Figure 2 suggests that rainfall (X) and chili prices (Y) are already reasonably stationary. Inferentially, the Augmented Dickey-Fuller (ADF) test yields p-values smaller than the 0.05 significance level, indicating that both rainfall and chili price series satisfy the weak stationarity condition.

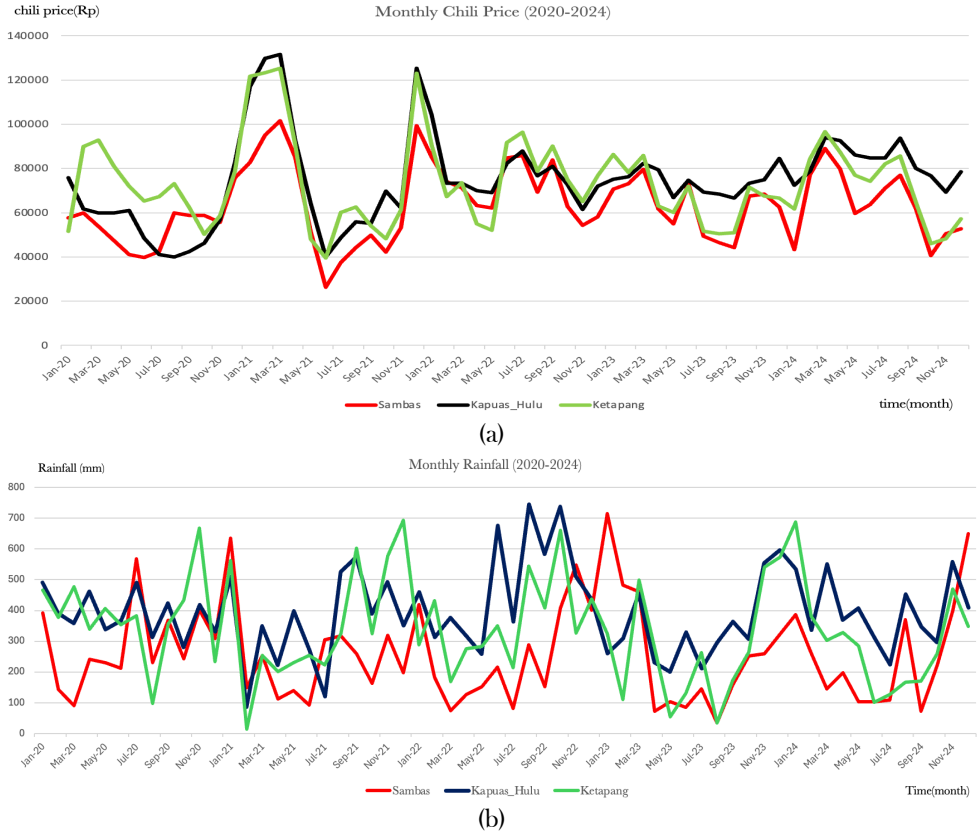


Figure 2. Plot of monthly chili prices (a) and rainfall data (b) for the period 2020-2024

After confirming data stationarity, the next step is determining the spatial weight matrix. Model order identification is not required here since the GSTAR (1;1) model has already been specified. A stochastic process $\{Y(t)\}$ that follows the GSTAR (1;1) model is expressed in Equation (3). In this equation, **W** denotes the spatial weight matrix, distinguishing this model from conventional time series models. Constructing the spatial weight matrix **W** is essential, as it captures spatial correlations in the GSTAR (1;1) model. Since only the first spatial lag is considered, the spatial weight matrix requires only one lag besides the identity matrix.

Distances between research locations were obtained using Google Maps, measured from each location's central government office (regency office). These distances, shown in Table 1, reflect substantial geographic heterogeneity. Using the inverse distance method, the spatial weight matrix **W** is obtained as:

$$\boldsymbol{W} = \begin{bmatrix} 0 & 0{,}46 & 0{,}54 \\ 0{,}47 & 0 & 0{,}53 \\ 0{,}52 & 0{,}48 & 0 \end{bmatrix}$$

This matrix satisfies the properties of a spatial weight matrix: each row sums to one and all elements are nonnegative.

The next stage is parameter estimation for the GSTAR (1;1) model using the Ordinary Least Squares (OLS) method. The first estimation uses rainfall data (X) as the input variable. Applying the linear form of the GSTAR (1;1) model yields the following estimates:

$$\Phi_0 = \begin{bmatrix} 0{,}51 & 0 & 0 \\ 0 & 0{,}42 & 0 \\ 0 & 0 & 0{,}21 \end{bmatrix} \text{ dan } \Phi_1 = \begin{bmatrix} -0.11 & 0 & 0 \\ 0 & -0{,}28 & 0 \\ 0 & 0 & 0{,}09 \end{bmatrix}$$

Table 1. Distance (Km) Between Research Locations

Regency	Sambas	Kapuas Hulu	Ketapang
Sambas	0	419,55	353,35
Kapuas Hulu	419,55	0	379,18
Ketapang	353,35	379,18	0

Table 2. Orders of the Transfer Function models and parameters for each location based on the lowest AIC

Location(i)	Location(j)	CCF Orders	Model
Sambas	Sambas	b=3, s=0, r=1	$b_1(t) = \frac{(\omega_0)a_1(t-3)}{(1-\delta_1 B)}$ + Noise model)
	Kapuas Hulu	b=3, s=1, r=0	$b_1(t) = (\omega_0 - \omega_1 B)a_2(t-3)$ + (noise model)
	Ketapang	b=3, s=1, r=0	$b_1(t) = (\omega_0 1\omega_1 B)a_3(t-3)$ + (noise model)
Kapuas Hulu	Sambas	b=10, s=0, r=1	$b_2(t) = (\frac{(\omega_0)}{1-\delta_1 B}a_1(t-10)$ + (noise model)
	Kapuas Hulu	b=4, s=0, r=1	$b_2(t) = \frac{(\omega_0)}{1-\delta_1 B}a_2(t-4)$ + (noise model)
	Ketapang	b=3, s=0, r=1	$b_2(t) = \frac{(\omega_0)}{(1-\delta_1 B)}a_3(t-3)$ + (noise model)
Ketapang	Sambas	b=2, s=0, r=1	$b_3(t) = \frac{(\omega_0)}{(1-\delta_1 B)}\ a_1(t-3)$ + (noise model)
	Kapuas Hulu	b=3, s=1, r=0	$b_3(t) = (\omega_0 + \omega_1 B)a_2(t-3)$ + (noise model)
	Ketapang	b=2, s=2, r=0	$b_3(t) = (\omega_0 + \omega_1 B + \omega_2 B^2)a_3(t-2)$ + (noise model)

These parameter estimates are significant under the Inverse of Autocorrelation Matrix (IAcM) test, as the eigenvalues of the parameter matrix do not exceed modulus 1. Thus, the parameter estimates are stable and yield a stationary process. The estimated parameters and spatial weight matrix can be expressed as the GSTAR (1;1) model for rainfall (X) at time ttt in Sambas, Kapuas Hulu, and Ketapang, respectively:

$$X_1(t) = 0{,}51X_1(t-1) - 0.05X_2(t-1) - 0.06X_3(t-1) + a_1(t)$$
$$X_2(t) = 0{,}42X_2(t-1) - 0.13X_1(t-1) - 0.15X_3(t-1) + a_2(t)$$
$$X_3(t) = 0{,}21X_3(t-1) + 0.05X_1(t-1) + 0.04X_2(t-1) + a_3(t).$$

These results indicate strong temporal autocorrelation at each location and cross-effects of varying signs (some positive, some negative) across regions. For example, rainfall in Sambas $\{X_1(t)\}$ and Kapuas Hulu $\{X_2(t)\}$. It is primarily influenced by past rainfall at its own locations. At the same time, the effects from neighboring regions are inverse: rainfall increases when the surrounding areas experienced lower rainfall in the previous period. In contrast, rainfall in Ketapang. In contrast, rainfall in Ketapang $\{X_3(t)\}$ shows consistent positive influence from its past values and neighboring regions. Since the model is stable (stationary), it is suitable for short-term forecasting.

The estimated parameters are then applied (transferred) to the output variable, chili prices (Y), yielding the GSTAR (1;1) model for the three locations:

$$Y_1(t) = 0{,}51Y_1(t-1) - 0.05Y_2(t-1) - 0.06Y_3(t-1) + b_1(t)$$
$$Y_2(t) = 0{,}42Y_2(t-1) - 0.13Y_1(t-1) - 0.15Y_3(t-1) + b_2(t)$$
$$Y_3(t) = 0{,}21Y_3(t-1) + 0.05X_1(t-1) + 0.04X_2(t-1) + b_3(t).$$

The residuals from both systems serve as the basis for prewhitening, which is then used as input for the cross-correlation function (CCF) to determine the order of the transfer function. The CCF results for $(a_i(t), b_j(t))$ across locations are used to identify the order of the transfer function models for each location. At Sambas, the CCF against its own series shows a significant peak at lag 3, giving order $b = 3$. No other lags are significant, so $s = 0$. The order r is chosen as either 0 or 1, depending on the smallest Akaike Information Criterion (AIC). A similar procedure is applied for Kapuas Hulu and Ketapang relative to Sambas, as summarized in Table 2.

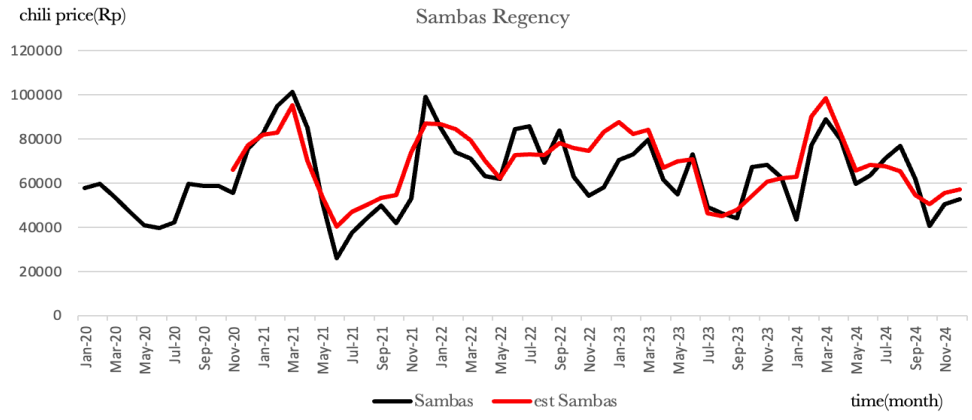
At Kapuas Hulu, the CCF of input and output with its own series shows significant correlation at lag 4 with order $b = 4, s = 0, r = 0$ or $r = 1$, In contrast, no correlation is observed with Sambas. Meanwhile, Ketapang's CCF with its own series suggests transfer function $b = 2$, s = 2, r =0 or $r = 1$. Using a similar approach, the transfer function orders for other locations are summarized in Table 2. Identifying these orders is a prerequisite for constructing each location's GSTAR–Transfer Function models.

Table 3. Estimated Parameters of the Best Transfer Function Models Based on the Lowest AIC

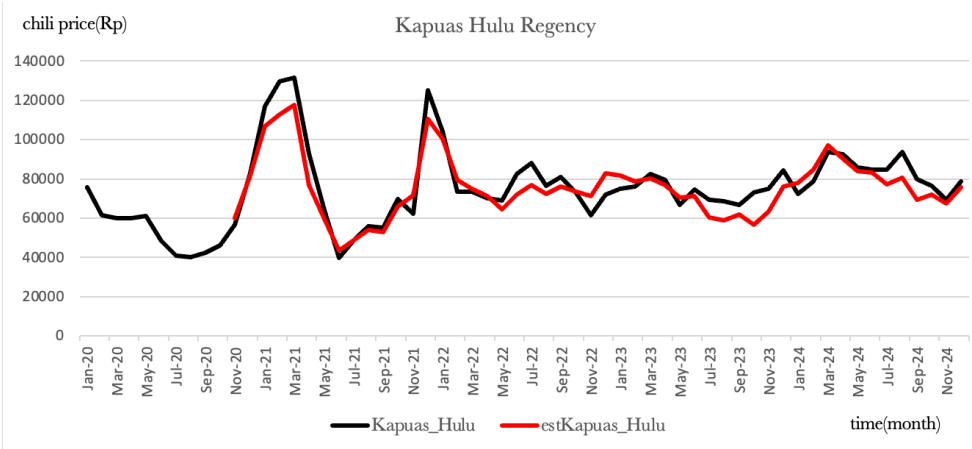
Location	Transfer Function Model (GSTAR-based)
Sambas	$Y_1(t) = (36.22)X_1(t-3) + (6.88)X_1(t-4)) + (35.51X_2(t-3) + (23.27X_2(t-4) + (31.77X_3(t-3) + (16.40X_3(t-4) + \mu_1(t)$
Kapuas Hulu	$Y_2(t) = 44.32X_2(t-4) - 2.22X_2(t-5) - 63.49X_1(t-10) - 9.52X_1(t-11) + 42.20X_3(t-3) + 10.13\,X_3(t-4) + \mu_2(t)$
Ketapang	$Y_3(t) = 19.64X_3(t-2) + 47.48X_3(t-3) + 17.77X_3(t-4) + 43.36X_1(t-3) + 16.91X_1(t-4) + 48.35X_2(t-3) + 32.99X_2(t-4) + \mu_3(t)$

Table 4. Noise Series Models of GSTAR (1;1)

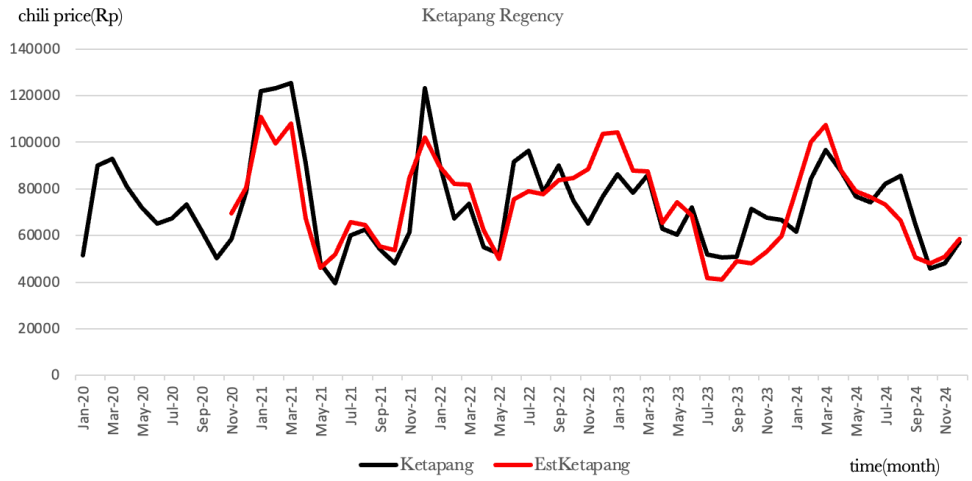
Location	Noise series model
Sambas	$\mu_1(t) = 0{,}63\mu_1(t-1) - 0.07\mu_2(t-1) - 0.08\mu_3(t-1)$
Kapuas Hulu	$\mu_2(t) = -0{,}08\mu_1(t-1) + 0.83\mu_2(t-1) - 0.09\mu_3(t-1)$
Ketapang	$\mu_3(t) = -0{,}23\mu_1(t-1) - 0.21\mu_2(t-1) + 0.80\mu_3(t-1)$



(a)



(b)

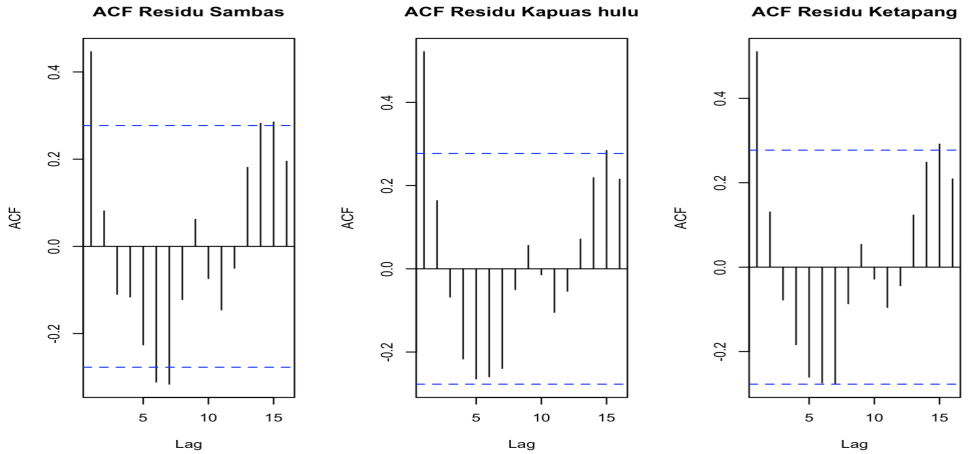


(c)

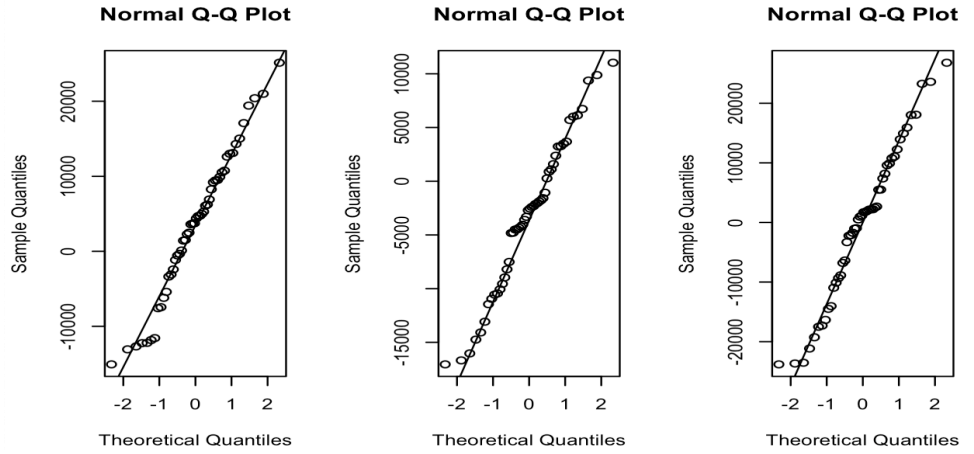
Figure 3. Plots of estimated data (red) versus actual data (black) for each location

Parameter estimation of the transfer function models for the three locations $i = 1,2,3$ based on Equation (10) the results are presented in Table 3, where the choice of r is determined by the lowest AIC value. The corresponding noise series models are shown in Table 4. Based on the models obtained for the three locations, rainfall influences the price of chili in addition to the effect of the chili price itself. The final GSTAR Transfer Function model was derived from the combination of Tables 3 and 4.

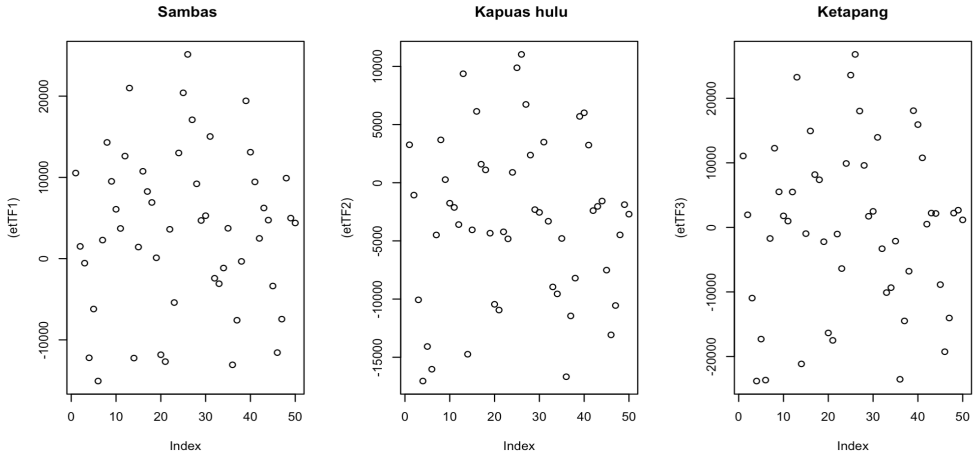
The plot of the actual and estimated data is presented in Figure 3. As shown, the calculated values closely follow the actual data pattern. The final step is the diagnostic test of the residuals from the GSTAR-Transfer Function model. Based on the diagnostic results (Figure 4), the residuals satisfy the classical assumptions. Although some correlation appears at the initial lag in Sambas, the ACF plot shows that it remains within the significance bounds. The normality test using the QQ plot also indicates that the residuals approximate a normal distribution. In contrast, the white noise test through the scatter plot confirms that the residuals are randomly distributed. Thus, overall, the diagnostic tests confirm that the model is suitable for forecasting (extrapolation) purposes. Furthermore, the Mean Absolute Percentage Error (MAPE) values for the GSTAR-Transfer Function model in Sambas, Kapuas Hulu, and Ketapang are 15%, 8%, and 14%, respectively, indicating that the model provides reliable performance.



(a) Heteroskedasticity test through ACF plot



(b) Normality test through qq-plot



(c) White noise test through Scatter plot

Figure 4. Residual diagnostic test results. All residuals meet the classical assumptions.

3.3 Interpretation and Analysis

The GSTAR–Transfer Function model demonstrated good predictive accuracy, with MAPE values of 15% for Sambas, 8% for Kapuas Hulu, and 14% for Ketapang. The model for Kapuas Hulu demonstrates superior accuracy, likely due to its inland location, where rainfall substantially influences the formation of chili prices. Kapuas Hulu exhibits a prolonged rainfall effect due to its unique geographical conditions, including hilly terrain, extensive forest areas, and large lakes that sustain soil moisture for longer periods. Heavy rainfall in this region often leads to flooding and poor drainage, intensifying the persistence of climatic shocks. Additionally, limited transport infrastructure creates distribution bottlenecks, as road networks are highly susceptible to damage and delays during extreme rainfall events. These combined geographic and logistical constraints explain why rainfall shocks in Kapuas Hulu longer-lasting impacts on chili price dynamics have compared to other regions. In Kapuas Hulu, rainfall effects are more prolonged, with significant negative impacts from Sambas rainfall 10–11 months earlier and from local and Ketapang rainfall 3–5 months earlier. These differences may be attributed to contrasting climatic patterns between inland Kapuas Hulu and the coastal regions of Sambas and Ketapang.

From a policy standpoint, addressing the prolonged impact of rainfall shocks in Kapuas Hulu requires targeted interventions that strengthen both infrastructure and market resilience. Investments in durable road networks and climate-resilient transportation systems are crucial for mitigating distribution bottlenecks during extreme rainfall events. In parallel, the development of more efficient commodity distribution strategies, such as regional storage facilities and coordinated supply chain management, would help stabilize market availability and mitigate price volatility. These measures are expected to enhance the adaptive capacity of local agricultural systems, ensuring more stable chili price dynamics in the face of increasing climate variability.

Overall, the results indicate that rainfall, as an input variable, significantly affects chili prices across the three regions, both temporally and spatially. Despite the absence of chili plantations in Sambas, Kapuas Hulu, and Ketapang, rainfall still exerts influence through its effect on transportation and distribution routes. In Sambas and Ketapang, rainfall in neighboring areas in the preceding 2–4 months positively correlates with chili prices, reflecting reliance on road and river transport. Prices in all three regions also depend on their own past values but are negatively influenced by lagged prices in neighboring locations.

4. CONCLUSION

Based on the study, rainfall significantly influenced chili pepper prices across the three research locations, both temporally and spatially. The main contribution of this study is to demonstrate GSTAR–TF as an alternative forecasting tool that explicitly accounts for rainfall shocks in agricultural commodity markets. The GSTAR–Transfer Function model demonstrated good predictive accuracy, with MAPE values of 15% for Sambas, 8% for Kapuas Hulu, and 14% for Ketapang. This reflects differences in seasonal patterns between inland (Kapuas Hulu) and coastal regions (Sambas and Ketapang). Furthermore, past chili prices also influenced current prices, with varying patterns across locations. The findings highlight rainfall as a crucial external factor when analysing chili price dynamics in West Kalimantan. For future research, the GSTAR model could be extended to higher orders to cover a broader spatial scope and include multiple exogenous variables. Additionally, parameter estimation could be refined using methods beyond OLS and maximum likelihood to improve robustness and accuracy.

ACKNOWLEDGEMENTS

We would like to thank the Department of Trade and the BMKG, Province of West Kalimantan, for providing the data used in this research. We also sincerely thank FMIPA Untan for funding this study through the DIPA research grant 2025.

5. REFERENCES

[1] D. Haryono, I. Sabiela, M. Riantini, V. P. Tambunan, and Firdasari, "Price fluctuations of curly red Chilli and their effects on farmers' welfare in Pringsewu regency, Lampung: The influence of government policy and household income," *Edelweiss Appl. Sci. Technol.*, vol. 9, no. 4, pp. 1399–1412, 2025, doi: 10.55214/25768484.v9i4.6303.

[2] S. Rofatunnisa, R. A. C. Putra, and S. Ghania, "Analysis Of Food Commodity Price Fluctuations On Inflation In West Kalimantan 2020-2023," *Agrimansion*, vol. 25, no. 3, pp. 640–652, 2024, [Online]. Available: https://agrimansion.unram.ac.id/index.php/Agri/article/view/1730.

[3] Pemprov Kalimantan Barat, "Profil Geografis." https://kalbarprov.go.id/profil/geografis (accessed Sep. 09, 2025).

[4] Yuliansyah, M. H. Kara, R. A. Masse, and Sumar'in, "Cross-border trade in the indonesia-malaysia border area at the plbn aruk of sambas regency," *Southeast Asia J. Grad. Islam. Bus. Econ.*, vol. 4, no. 1, pp. 7–16, 2025, [Online]. Available: https://journal.iaisambas.ac.id/index.php/SAJGIBE/article/view/3716/2681.

[5] Badan Penghubung Provinsi Kalimantan Barat, "Profil Daerah." https://penghubung.kalbarprov.go.id/profil-kalbar/?utm_source=chatgpt.com (accessed Sep. 09, 2025).

[6] N. R. Simanjuntak, U. Umar, D. Gunarto, and S. B. Soeryamassoeka, "Flood Hazard Level Analysis of Kapuas Hulu Sub-watershed Using Geographic Information System," *J. Tek. Sipil*, vol. 24, no. 4, pp. 1437–1448, 2025, doi: 10.26418/jts.v24i4.86222.

[7] E. Pusporani, M. A. D. P. Yuniar, S. A. N. Fajrina, V. A. Alexandra, and M. F. F. Mardianto, "Generalized Space Time Autoregressive (GSTAR) Modeling in Predicting the Price of Bird's Eye Chili in East Java, West Java, and Central Java," *CAUCHY J. Mat. Murni dan Apl.*, vol. 9, no. 2, pp. 187–198, 2024, doi: 10.18860/ca.v9i2.25730.

[8] U. Mukhaiyar, D. Widyanti, and S. Vantika, "The time series regression analysis in evaluating the economic impact of COVID-19 cases in Indonesia," *Model Assist. Stat. Appl.*, vol. 16, no. 3, pp. 197–210, 2021, doi: 10.3233/MAS-210533.

[9] N. Nurhayati, U. S. Pasaribu, and O. Neswan, "Application of generalized space-time autoregressive model on GDP data in West European countries," *J. Probab. Stat.*, vol. 2012, 2012, doi: 10.1155/2012/867056.

[10] Yundari, U. S. Pasaribu, and U. Mukhaiyar, "Error assumptions on generalized STAR model," *J. Math. Fundam. Sci.*, vol. 49, no. 2, pp. 136–155, 2017, doi: 10.5614/j.math.fund.sci.2017.49.2.4.

[11] Yundari, U. S. Pasaribu, U. Mukhaiyar, and M. N. Heriawan, "Spatial Weight Determination of GSTAR(1;1) Model by Using Kernel Function," *J. Phys. Conf. Ser.*, vol. 1028, no. 1, 2018, doi: 10.1088/1742-6596/1028/1/012223.

[12] U. S. Pasaribu, U. Mukhaiyar, M. N. Heriawan, and Y. Yundari, "Generalized Space-Time Autoregressive Modeling of the Vertical Distribution of Copper and Gold Grades with a Porphyry-Deposit Case Study," *Int. J. Adv. Sci. Eng. Inf. Technol.*, vol. 12, no. 5, pp. 2030–2038, 2022, doi: 10.18517/ijaseit.12.5.14835.

[13] N. M. Huda, N. Imro'ah, N. F. Arini, D. S. Utami, and T. Umairah, "Looking at GDP from a Statistical Perspective: Spatio-Temporal GSTAR(1;1) Model," *JTAM (Jurnal Teor. dan Apl. Mat.*, vol. 7, no. 4, p. 976, 2023, doi: 10.31764/jtam.v7i4.16236.

[14] C. Trinitha, N. Universitas, and I. Indonesia, "Enhancing Rice Price Forecasts with Generalized Space-Time Autoregressive (GSTAR) Models and Spatial Weighting Variations," pp. 306–317, 2025.

[15] Y. Hasanah, M. Herlina, and H. Zaikarina, "Flood Prediction using Transfer Function Model of Rainfall and Water Discharge Approach in Katulampa Dam," *Procedia Environ. Sci.*, vol. 17, pp. 317–326, 2013, doi: 10.1016/j.proenv.2013.02.044.

[16] T. Purwa, U. Nafngiyana, and S. Suhartono, "Comparison of arima, transfer function and var models for forecasting cpi, stock Prices, and indonesian exchange rate: accuracy vs explainability," *Media Stat.*, vol. 13, no. 1, pp. 1–12, 2020.

[17] U. Mukhaiyar, N. M. Huda, K. N. Sari, and U. S. Pasaribu, "Analysis of Generalized Space Time Autoregressive with Exogenous Variable (GSTARX) Model with Outlier Factor," *J. Phys. Conf. Ser.*, vol. 1496, no. 1, 2020, doi: 10.1088/1742-6596/1496/1/012004.

[18] Suhartono, S. R. Wahyuningrum, Setiawan, and M. S. Akbar, "GSTARX-GLS Model for Spatio-Temporal Data Forecasting," *Malaysian J. Math. Sci.*, vol. 10, no. January, pp. 91–103, 2016.

[19] U. Mukhaiyar, "The goodness of generalized STAR in spatial dependency observations modeling," 2015, p. 020008, [Online]. Available: https://pubs.aip.org/aip/acp/article-abstract/1692/1/020008/883852/The-goodness-of-generalized-STAR-in-spatial.

[20] G. E. P. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung, *Time Series Analysis: Forecasting and Control*, 5th Editio. New Jersey(US): John Wiley & Sons, Inc, 2016.

[21] J. D. Cryer and K.-S. Chan, *Time Series Analysis: With Applications to R*, 2nd ed. New York(US): Springer, 2008.

[22] W. W. S. Wei, *Time Series Analysis: Univariate and Multivariate Methods*, 2nd ed. United States (US): Pearson Education, 2006.

[23] S. Borovkova, H. P. Lopuhaä, and B. N. Ruchjana, "Consistency and asymptotic normality of least squares estimators in generalized STAR models," *Stat. Neerl.*, vol. 62, no. 4, pp. 482–508, 2008, doi: 10.1111/j.1467-9574.2008.00391.x.
[24] I. Melnyk and A. Banerjee, "Estimating structured vector autoregressive models," *33rd Int. Conf. Mach. Learn. ICML 2016*, vol. 2, pp. 1297–1330, 2016.
[25] T. Anderson, *The Statistical Analysis of Time Series*. New York: Willey, 1994.
[26] H. Lütkepohl, *New introduction to multiple time series analysis*. 2005.
[27] Y. Yundari and S. W. Rizki, "Invertibility of Generalized Space-Time Autoregressive Model with Random Weight," *CAUCHY J. Mat. Murni dan Apl.*, vol. 6, no. 4, pp. 246–259, 2021, doi: 10.18860/ca.v6i4.11254.
[28] C. Lam and P. C. L. Souza, "Estimation and Selection of Spatial Weight Matrix in a Spatial Lag Model," *J. Bus. Econ. Stat.*, vol. 38, no. 3, pp. 693–710, 2020, doi: 10.1080/07350015.2019.1569526.
[29] U. Mukhaiyar and U. Pasaribu, "A new procedure for generalized STAR modeling using IAcM approach," *ITB J Sci*, vol. 44, no. 2, pp. 179–192, 2012, doi: https://doi.org/ 10.5614/itbj.sci.2012.44.2.
[30] F. I. N. Fadlilah, U. Mukhaiyar, and F. Fahmi, "The generalized STAR(1,1) modeling with time correlated errors to red-chili weekly prices of some traditional markets in Bandung, West Java," *AIP Conf. Proc.*, vol. 1692, 2015, doi: 10.1063/1.4936442.
[31] Suci Pujiani Prahesti, Itasia Dina Sulvianti, and Yenni Angraini, "Peramalan Harga Batu Bara Acuan Menggunakan Metode Autoregressive Integrated Moving Average Dan Fungsi Transfer," *Xplore J. Stat.*, vol. 12, no. 1, pp. 1–11, 2023, doi: 10.29244/xplore.v12i1.1100.