

## FOREST FIRE ANALYSIS FROM PERSPECTIVE OF SPATIAL-TEMPORAL USING GSTAR ( $p; \lambda_1, \lambda_2, \dots, \lambda_p$ ) MODEL

Hesty Pratiwi<sup>1</sup>, Nurfitri Imro'ah<sup>2\*</sup>, Nur'ainul Miftahul Huda<sup>3</sup>

<sup>1,2</sup>Department of Statistics, Faculty of Mathematics and Natural Sciences, Universitas Tanjungpura

<sup>3</sup>Department of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Tanjungpura  
Jln. Prof. Dr. H. Hadari Nawawi, Pontianak, 78124, Indonesia

Corresponding author's e-mail: \* [nurfitriimroah@math.untan.ac.id](mailto:nurfitriimroah@math.untan.ac.id)

### ABSTRACT

#### Article History:

Received: 18<sup>th</sup> November 2024

Revised: 3<sup>rd</sup> February 2025

Accepted: 6<sup>th</sup> March 2025

Published: 1<sup>st</sup> April 2025

#### Keywords:

GSTAR;

Hotspot;

Queen Contiguity.

West Kalimantan is particularly susceptible to the devastating effects of forest fires, among the natural disasters that have a significant impact. One of the indicators that can be used to identify forest fires is the presence of hotspots. The term "hotspot" refers to data that has both spatial and temporal characteristics. Using the Generalized Space-Time Autoregressive (GSTAR) model combined with the Queen Contiguity weight matrix, this research aims to model and forecast the confidence level of hotspots in Kubu Raya Regency and its surrounding areas. We chose the GSTAR model because of its ability to model spatial interactions between locations and temporal change patterns over time. According to NASA FIRMS, the data used in this study were confidence level hotspot data, covering the period from January 2014 to August 2024. To define locations for modeling, the study area was divided into grids measuring  $1 \times 1$  degrees. The maximum confidence level value in each grid was used to represent the highest potential fire risk. The research process consists of the following stages: data preparation, stationarity testing, calculation of the Queen Contiguity spatial weight matrix, identification of model orders based on STACF and STPACF plots, and estimation of model parameters to predict hotspot confidence levels. The GSTAR (3;1) model was selected as the best model because it satisfies the white-noise assumption and has a MAPE value of 14.78%. Based on the MAPE, the GSTAR (3;1) model can provide reasonably accurate predictions for the confidence level of fire points over the following three periods. The prediction results indicate a decline in the fire point confidence level across all locations during the following three periods. The findings of this study can support the optimization of resource allocation in the prevention of forest fires.



This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution-ShareAlike 4.0 International License](https://creativecommons.org/licenses/by-sa/4.0/).

#### How to cite this article:

H. Pratiwi, N. Imro'ah and N. M. Huda., "FOREST FIRE ANALYSIS FROM PERSPECTIVE OF SPATIAL-TEMPORAL USING GSTAR ( $p; \lambda_1, \lambda_2, \dots, \lambda_p$ ) MODEL," *BAREKENG: J. Math. & App.*, vol. 19, iss. 2, pp. 1379-1392, June, 2025.

Copyright © 2025 Author(s)

Journal homepage: <https://ojs3.unpatti.ac.id/index.php/barekeng/>

Journal e-mail: [barekeng.math@yahoo.com](mailto:barekeng.math@yahoo.com); [barekeng\\_journal@mail.unpatti.ac.id](mailto:barekeng_journal@mail.unpatti.ac.id)

Research Article · Open Access

## 1. INTRODUCTION

Forest fires are one of the environmental disasters that frequently occur in Indonesia, particularly in West Kalimantan. The situation in question is driven by a confluence of natural forces, including extended periods of drought and the El Niño weather phenomenon, as well as human activity, such as clearing land through burning strategies. Especially in peatlands, the risk of forest fires is increased when the climate is dry and the temperatures are high. Forest fires have severe repercussions, including the destruction of the environment, the ill effects of smoke on human health (such as acute respiratory illness), and the interruption of transportation and the economy [1]. Additionally, the haze that is produced frequently travels to neighbouring countries, threatening diplomatic ties at the time.

One of the regions in West Kalimantan most prone to forest and land fires is the Kubu Raya Regency [2]. The case occurred by the vast peatlands prone to fire during the dry season for various reasons. According to information obtained from SiPongi, the forest and land fire monitoring system of the Ministry of Environment and Forestry, the area of forest fires in West Kalimantan experienced a significant increase in 2023, reaching more than 111,000 hectares from the previous year. Because Kubu Raya Regency contains many hotspots, it is among the regions with the highest potential for sparking fires. The fact that these hotspots have been identified through satellite technology is a significant signal for determining the likelihood of forest fires.

A comprehensive analysis is required to comprehend the patterns of forest fires that occur in regions such as Kubu Raya. Because these patterns are influenced by relationships between locations (spatial) and changes in time (temporal), they are considered quite complicated. For instance, hotspots in a particular area are influenced by the conditions and the presence of hotspots in other locations, especially if the locations are directly adjacent and have environmental conditions encouraging fire spread, such as dry peatlands [3]. Additionally, temporal patterns demonstrate that hotspots tend to arise at specific times, such as during the dry season; consequently, the extent to which they expand might vary daily or monthly. A model capable of simultaneously combining spatial and temporal aspects is required to represent this dynamic accurately. Several research have been conducted to investigate forest fires and the GSTAR model. One of these studies, conducted by Imro'ah et al. [4], utilized binary logistic regression to examine the probability of forest fires occurring in West Kalimantan. According to the findings of this study, the danger of forest fires was significantly impacted by factors such as the distance to rivers and highways, the types of land, and the climate variables. This study suggests that the fire extinguishing system should be strengthened and that public education should be increased to reduce the risk of forest fires. Alkaff and Yulianto [5] used Seasonal ARIMA with monthly data from NASA's Terra satellite to make a prediction about the frequency of hotspots in Kalimantan for the period spanning from 2001 to 2018. Although SARIMA (1,0,1)(1,0,1)<sup>12</sup> was the best model, it was not very accurate in predicting significant spikes. For example, in September 2018, 36 hotspots were only projected to be 2. The assumption of stationarity, which helps to eliminate big swings in the data, is connected to this constraint.

Modeling the connectedness between regions based on directly intersecting geographic boundaries is more robust by utilizing the Queen Contiguity weight matrix in the GSTAR model. This model helps to strengthen spatial analysis. This matrix captures the spatial influence between hotspots more flexibly since it considers the relationship between nearby regions in all three directions: horizontally, vertically, and diagonally [6]. Additionally, [7] utilized the GSTAR model to make predictions regarding the number of cases of COVID-19 in Java Island by using the Queen Contiguity weight matrix. Forty-one districts and cities in Banten, DKI Jakarta, and West Java were the data sources. The GSTAR (1;1) model showed a high level of accuracy in forecasting the progression of cases over the next five days, which aided in formulating policies to deal with the pandemic.

This study's application of a 1 × 1 degree grid for the selected areas distinguishes this study from several other studies conducted in the past. This grid is utilized to ensure that the setup procedure and the places that are described correspond to a consistent size standard. Specifically for regions containing a significant number of hotspots that change over time, dividing the area into grid units enables a more extensive spatial-temporal analysis in this study. The GSTAR model in this investigation utilizes the Queen Contiguity weight matrix. This matrix connects the relationship between sites geographically based on direct neighbors in a horizontal, vertical, and diagonal manner. In comparison to other research that made use of equal weights or inverse distances, this offers a further moment in the process of capturing spatial influences.

By applying the GSTAR method with the Queen Contiguity weight matrix, this study aims to model fire spots in the Kubu Raya geographic region. When predicting when and where fire spots will appear, this method is anticipated to provide more accurate responses. As part of the efforts to prevent and manage forest fires, this model will help reduce the number of forest fires in the region. Additionally, it will be utilized to optimize the distribution of resources within the context of fire prevention and management. When local governments have better predictions, they can take more effective and expedient action if forest fires occur. The statement above holds particularly true in regions such as Kubu Raya, susceptible to danger.

The findings of this research are anticipated to make significant contributions to developing strategies for mitigating the effects of forest and land fire disasters. Furthermore, the developed model can also be utilized in other regions at a high risk of fire, strengthening the policy on the prevention and mitigation of forest fires in Indonesia as a whole.

## 2. RESEARCH METHODS

### 2.1. Generalized Space-Time Autoregressive Model (GSTAR)

The Generalized Space-Time Autoregressive (GSTAR) model captures geographical and temporal relationship patterns in time series data involving many geographic locations. A generalization of the autoregressive (AR) and space-time autoregressive (STAR) models, the generalized spatial-temporal autoregressive (GSTAR) model considers interactions between locations in space as well as changes over time. The following is the general equation that defines the GSTAR ( $p; \lambda_1, \lambda_2, \dots, \lambda_p$ ) model [8]:

$$\mathbf{Y}_t = \left( \sum_{k=1}^p \sum_{\ell=0}^{\lambda_k} \Phi_{k\ell} \mathbf{W}^{(\ell)} \mathbf{Y}_{t-k} \right) + \mathbf{e}_t \quad (1)$$

where

$$\mathbf{Y}_t = \begin{bmatrix} Y_t^{(1)} \\ Y_t^{(2)} \\ \vdots \\ Y_t^{(N)} \end{bmatrix}; \Phi_{kp} = \begin{bmatrix} \phi_{kp}^{(1)} & 0 & \dots & 0 \\ 0 & \phi_{kp}^{(2)} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \phi_{kp}^{(N)} \end{bmatrix}; \Phi_{k\lambda_k} = \begin{bmatrix} \phi_{k\lambda_k}^{(1)} & 0 & \dots & 0 \\ 0 & \phi_{k\lambda_k}^{(2)} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \phi_{k\lambda_k}^{(N)} \end{bmatrix}; \mathbf{e}_t = \begin{bmatrix} e_t^{(1)} \\ e_t^{(2)} \\ \vdots \\ e_t^{(N)} \end{bmatrix}$$

and  $\mathbf{W}^{(\ell)}$  is spatial weight matrix at lag  $\ell$

Nevertheless, suppose the matrix  $\Phi$  is transformed into a scalar or vector, which is no longer a matrix. In that case, it signifies that every location is regarded as having the same autoregressive coefficient. This implies that the diversity of interactions between places in space is no longer explicitly taken into consideration. In this instance, the model is changed into a STAR model, which stands for space-time autoregressive [9]. As a result, the equation for the STAR ( $p; \lambda_1, \lambda_2, \dots, \lambda_p$ ) model can be stated as follows [10]:

$$\mathbf{Y}_t = \left( \sum_{k=1}^p \sum_{\ell=0}^{\lambda_k} \phi_{k\ell} \mathbf{W}^{(\ell)} \mathbf{Y}_{t-k} \right) + \mathbf{e}_t \quad (2)$$

The STAR formulation incorporates spatial effects using a spatial weight matrix to describe the interactions between different locations. On the other hand, if the spatial component of the STAR model is disregarded, then the model can be simplified into a Vector Autoregressive (VAR) model. The VAR model is a multivariate statistical model that solely considers the temporal link between variables at different locations. Other than that, it does not consider the impact of spatial factors. VAR ( $p$ ) model can be defined as follows [11]:

$$\mathbf{Y}_t = \Phi_1 \mathbf{Y}_{t-1} + \Phi_2 \mathbf{Y}_{t-2} + \dots + \Phi_k \mathbf{Y}_{t-k} + \mathbf{e}_t \quad (3)$$

where  $Y_t$  is the vector of observations at period  $t$ ,  $e_t$ , residual at period  $t$ , and  $\Phi_k$ , and the parameter matrix of the VAR model for each  $k = 1, 2, 3, \dots, p$ .

The VAR model is a model that evaluates the interaction between numerous variables in a system, where each variable is influenced by its previous values as well as the past values of other variables. Such a model is known as a VAR model. If this model is reduced to a single variable, the interaction between the variables is removed, and the model transforms into an autoregressive (AR) model. In the AR model, the variable solely depends on its historical values, and no other variables influence it. An equation that can be used to describe the AR ( $p$ ) model in its entirety is as follows [12]:

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_k Y_{t-k} + e_t \quad (4)$$

where  $Y_t$  is the observation at period  $t$ ,  $e_t$  is the residual at period  $t$ , and  $\phi_k$  is the AR parameter for  $k = 1, 2, 3, \dots, p$ .

## 2.2. Queen Contiguity Weight Matrix

The spatial weight matrix is an essential component in the development of the GSTAR model, as it aids in the description of the relationship that exists between sites in the spatial dimension [13]. Through geographical closeness, such as geographic distance, this matrix exerts control over the influence of one site on another. In the case of forest fires, for instance, the matrix elements show the degree to which there is interaction between several different areas. The weight conditions that must be satisfied are as follows:  $w_{ij} = 0$  where  $i = 1, 2, \dots, N$ . The following form ( $\mathbf{W}$ ) can generally represent the  $N \times N$  weight matrix [14].

$$\mathbf{W} = \begin{bmatrix} 0 & w_{12}^{(\ell)} & \dots & w_{1N}^{(\ell)} \\ w_{21}^{(\ell)} & 0 & \dots & w_{2N}^{(\ell)} \\ \vdots & \vdots & \ddots & \vdots \\ w_{N1}^{(\ell)} & w_{N2}^{(\ell)} & \dots & 0 \end{bmatrix}$$

where  $w_{ij}$  is the weight between regions  $j$  and  $i$ .

The Queen Contiguity weight matrix is available among the several types of weight matrices. The Queen Contiguity spatial weight matrix is utilized in the GSTAR model to describe the relationship between sites based on their geographical proximity. When two sites share a boundary, whether horizontally, vertically, or diagonally, they are considered neighbors [15]. This is analogous to the movement of the "Queen" piece in the game of chess. In this matrix, the neighboring locations are given a weight of 1, while the non-neighboring locations are provided with 0. Because it considers all the neighbors, even those who share corners, the Queen Contiguity technique is more adaptable than other ways, like the Rook Contiguity method. The utilization of this matrix is highly appropriate for conducting research that calls for a thorough description of spatial interactions. To ensure that the value of  $\sum w_{ij} = 1$  is achieved, it is necessary to first normalize this matrix by employing the following formula:

$$w_{ij} = \frac{c_{ij}}{c_i} \quad (5)$$

where  $w_{ij}$  the weight of  $i$ -th location respect to  $j$ -th,  $c_{ij}$  is element of  $i$ -th row,  $j$ -th column, and  $c_i$  is the total elements of  $i$ -th row. **Figure 1** illustrates how spatial weights are calculated using Queen Contiguity weights.



**Figure 1. Queen Contiguity Weight Illustration**

According to **Figure 1**, it is possible to observe that location 8 has first neighbors, also known as direct neighbors, at locations 3, 5, 7, 10, 11, and 12. Location 8 has direct neighbors that intersect it from all directions, including horizontally, vertically, and diagonally. Additionally, location 8 has second neighbors at locations 2, 6, 9, 13, and 14, which are one step further away but continue to be under spatial influence according to the Queen Contiguity concept. These second neighbors are located at locations 2, 6, 9, 13, and 14. As a result, location 8 interacts with both its immediate neighbors and its neighbors further away, demonstrating a more intricate pattern of interrelationships in the context of spatial analysis.

**2.3. Parameter Estimation**

Ordinary least square (OLS) is used for parameter estimation within the GSTAR model. The following is an example of the OLS method in its generic form [16].

$$Y = X\beta + e \tag{6}$$

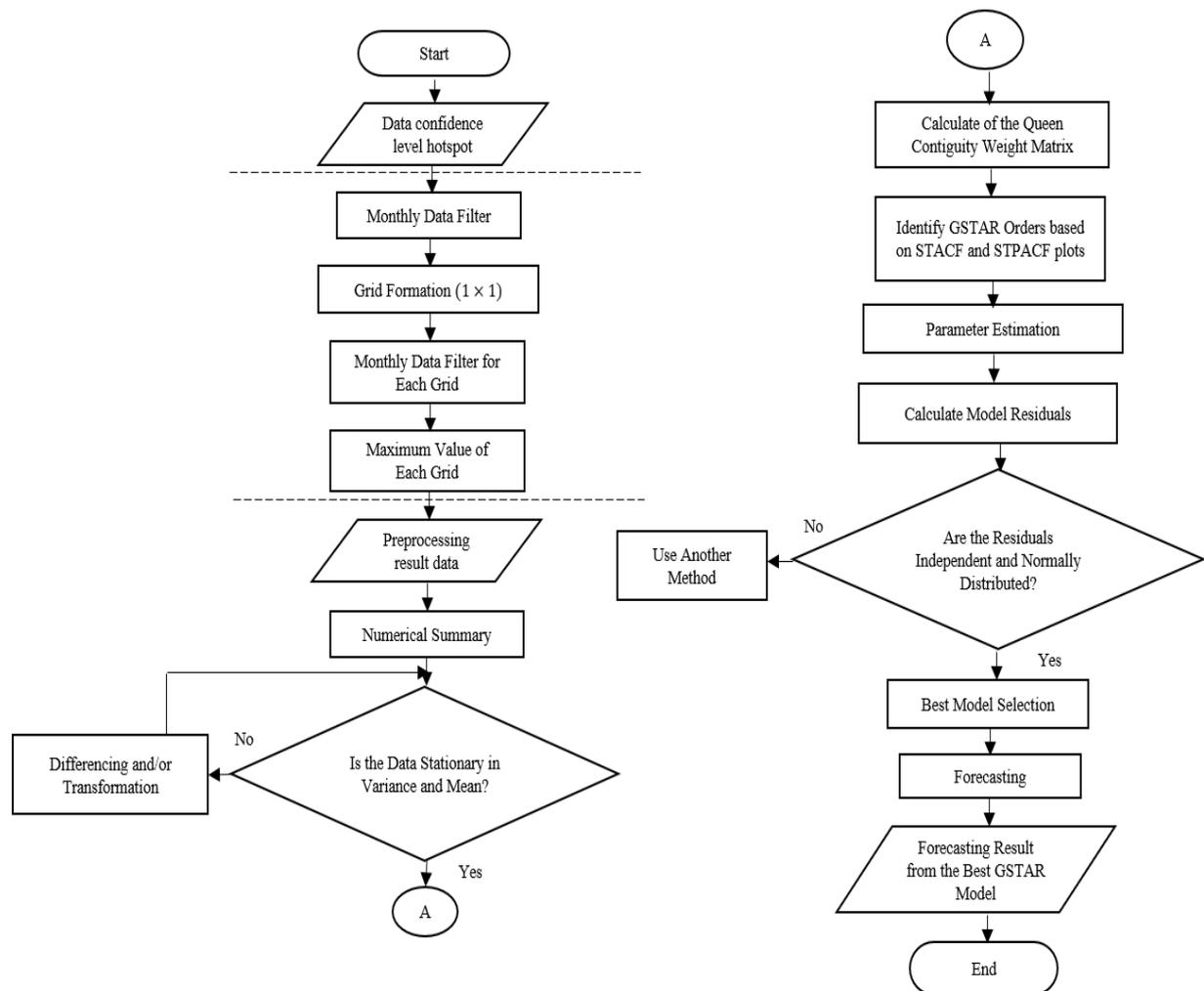
**Equation 6** can also be written as follows for GSTAR (1;1) Model:

$$\begin{bmatrix} Y_t^{(1)} \\ Y_t^{(2)} \\ \vdots \\ Y_t^{(N)} \end{bmatrix} = \begin{bmatrix} Y_{(t-1)}^{(1)} & V_{(t-1)}^{(1)} & 0 & 0 & \dots & \dots & 0 & 0 \\ 0 & 0 & Y_{(t-1)}^{(2)} & V_{(t-1)}^{(2)} & \dots & \dots & 0 & 0 \\ 0 & 0 & 0 & 0 & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & \dots & Y_{(t-1)}^{(i)} & V_{(t-1)}^{(i)} \end{bmatrix} + \begin{bmatrix} \phi_{10}^{(1)} \\ \phi_{11}^{(1)} \\ \vdots \\ \phi_{10}^{(N)} \\ \phi_{11}^{(N)} \end{bmatrix} + \begin{bmatrix} e_t^{(1)} \\ e_t^{(2)} \\ \vdots \\ e_t^{(N)} \end{bmatrix}$$

Where  $V_{(t-1)}^{(i)} = \sum_{j=1}^N w_{ij} Y_{(t-1)}^{(j)}$  for  $i \neq j$ . Therefore, **Equation 7** can be utilized for parameter estimation.

$$\beta = [X^T X]^{-1} [X^T Y] \tag{7}$$

In general, the steps involved in the research are as follows: (1) Order identification, (2) Estimation of the parameters, (3) Diagnostic test of the residual, and (4) forecasting the future. The following flowchart in **Figure 2** uses the Queen Contiguity weight matrix to provide a more in-depth illustration of the phases of forming a GSTAR model from hotspot data.

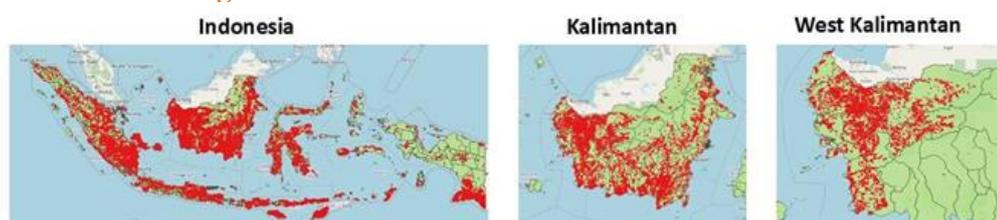


**Figure 2.** Flowchart of GSTAR Model

### 3. RESULTS AND DISCUSSION

#### 3.1 Data Preprocessing

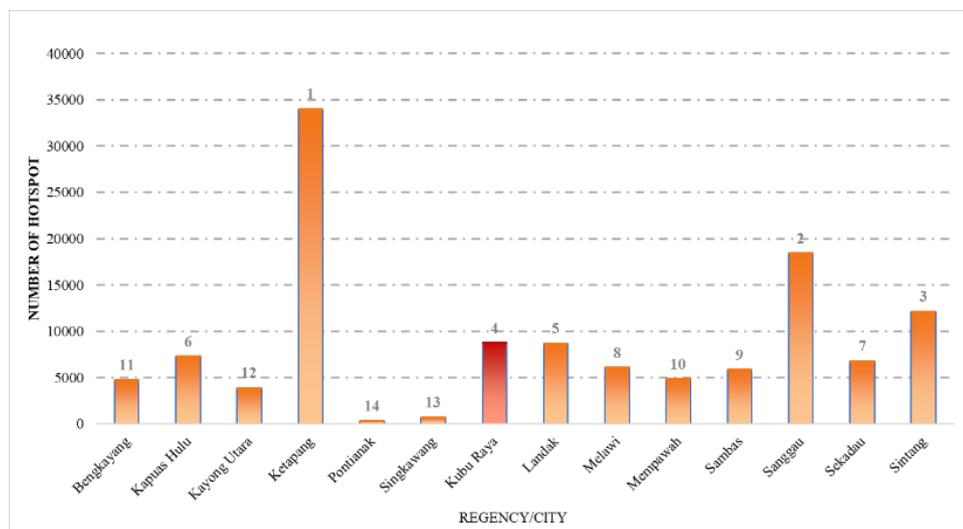
We carried out data pretreatment activities before forming the GSTAR model to verify that the data was ready to be analyzed and that the assumptions underpinning the model were agreed upon. For this research, the data utilized were daily confidence-level data from hotspots collected from the NASA FIRMS website (Fire Information for Resource Management System). The data covered the period from January 2014 to August 2024. This information consists of several primary characteristics, including the date, the longitude, the latitude, and confidence level. As a result of the fact that the data that was obtained encompassed the entirety of Indonesia, the initial step that was conducted was to filter the data according to geographic area. After preprocessing, the data were aggregated into 128 monthly observations, consisting of maximum confidence-level hotspot data ranging from 0% to 100%. Specifically, the West Kalimantan region is the focus of this research. The sorting of the data was therefore carried out based on the coordinates of latitude and longitude that correspond to the administrative boundaries of West Kalimantan. An illustration of this process is shown in **Figure 3**.



**Figure 3.** Filtering the Data of Hotspot

The research location was decided to be in West Kalimantan because this region contains a large amount of, makes peatlands and is highly exposed to fires [17]. The composition of peatlands, characterized by a high concentration of dry organic matter, makes them combustible. Therefore, peatlands can potentially cause catastrophic fires in agricultural and forest settings. The Kubu Raya Regency is widely regarded as having one of the most extensive peatlands in West Kalimantan compared to the other regencies and cities in the region. The composition of peatlands, characterized by a high concentration of dry organic matter, makes them combustible. Therefore, peatlands can potentially cause catastrophic fires in both forests and on land. The Kubu Raya Regency is widely regarded as having one of the most extensive peatlands in West Kalimantan compared to the other regencies and cities in the region [18]. Figure 4 illustrates the number of hotspots in each district and town in West Kalimantan during 2014 – 2023.

According to the data presented in Figure 4, Kubu Raya has been identified as the fourth-highest hotspot during the past ten years. Furthermore, this study concentrates on the Kubu Raya region because it is directly adjacent to Pontianak, the capital city of West Kalimantan, and peatlands predominate in the Kubu Raya region. Pontianak is the hub of the capital city, and as a result, it has a high population and significant levels of economic and governmental activity. The condition can raise the likelihood of forest fires spreading their effects to the surrounding area, including the release of haze.



**Figure 4.** Number of Hotspots in Regency/City in West Kalimantan Period 2014-2023

The following are data preprocessing steps.

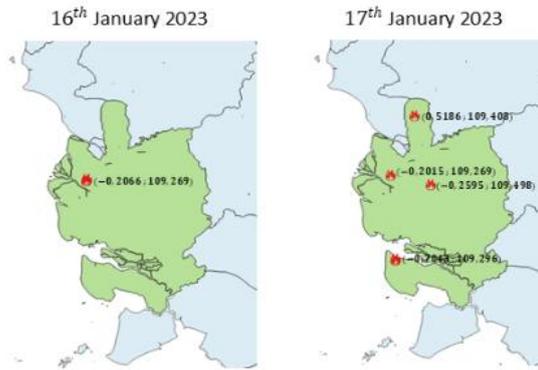
a. Filtering Monthly Data

The data that was utilized is annual data, comprised of twelve months. Immediately following filtering the hotspot data for the West Kalimantan region, it is essential to filter it into monthly data for each year.

b. Formating the Grid Location

The next step involves the creation of a location grid of the West Kalimantan region that is  $1 \times 1$  degrees in size. We used a grid size  $1 \times 1$  degrees to simplify the calculating process. Any reduction in the size of the grid will increase the number of locations that are generated, which will ultimately result in the calculation being complicated. Following the formation of the grid, the grid unit that provides coverage of the Kubu Raya Regency region has been chosen as the research area. Removing hotspots and grid units located outside of the Kubu Raya area allows the analysis to concentrate solely on pertinent locations.

A grid is created because the number of hotspots in a region will be different today compared to the number that existed the day before. As a result, it is essential to create a grid to record these variations. In the illustration in Figure 5, this is indeed made.



**Figure 5. Visualization Comparison of Hotspot Maps**

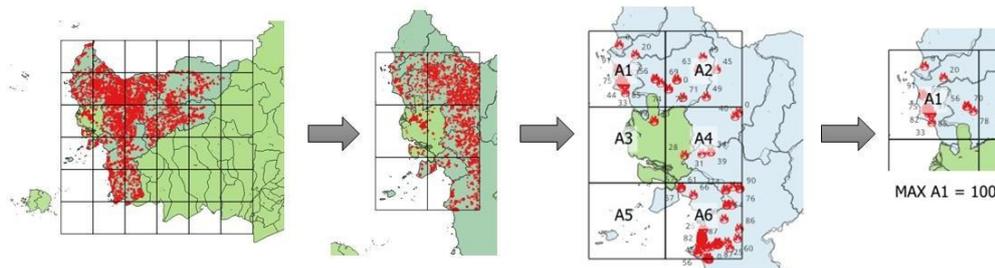
**Figure 5** shows data samples from hotspots collected on two different days. As of January 17, 2023, four hotspots were spread across four other locations, whereas on January 16, 2023, there was only one hotspot. When the latitude and longitude of each point are considered, these locations can be seen. The GSTAR model necessitates a predetermined number of locations throughout the period [19]. Because of this, the study area was partitioned into grids of equal sizes to standardize the location.

c. Filtering Monthly Data in Each Grid

After the location grid is formed, the next step is to filter the monthly data based on the grid units that have been formed. Each grid unit will, therefore, consist of fire point data gathered over one year. Following the formation of each grid, each such grid is regarded as a location that we utilized for GSTAR analysis.

d. Aggregating of Hotspot Values in Each Grid

Aggregating the data contained within each grid is the last step in the grid formation process. A representative value for each grid is determined by taking the highest possible value of the fire confidence level at each time point (for example, each day). This value is taken from each grid. Suppose there are multiple fires in a grid on the same day. In that case, the purpose of this step is to ensure that the highest value, which indicates the most significant fire potential, is selected as the representative of the highest value. If a grid does not contain any fires, we reset that grid's value to zero (0). **Figure 6** provides a comprehensive analysis of the steps involved in data preprocessing.



**Figure 6. Illustration of Preprocessing Data Steps**

**Figure 6** reveals that six grids have been formed, which indicates that 6 locations will be utilized for GSTAR modeling. The maximum value is extracted from each grid and used to represent specific locations for the modeling process. The districts and sub-districts involved in each grid are listed in **Table 1**.

**Table 1. Summary of Location in Each Grid**

Grid Location	Regency	Subdistrict
A1	Bengkayang	Bengkayang, Capkala, Ledo, Monterado, Betung River, Bawang Valley, Lumar, Samalantan, Raya River, and Raya Islands River.
	Landak	Sompak, Banyuke Hulu, Mandor, Mempawah Hulu, and Menyuke.

Grid Location	Regency	Subdistrict
	Singkawang	East Singkawang, South Singkawang, North Singkawang, Central Singkawang, and West Singkawang.
	Mempawah	Anjongan, Mempawah Hilir, East Mempawah, Sadaniang, Segedong, Kunyit River, Toho, Pinyuh River.
	Kuburaya	Kuala Mandor B
A2	Landak	Jagoi Babang, Suti Semarang, Teriak, Tujuh Belas, Siding, Seluas, Sanggau Ledo
	Sanggau	Tayan Hulu, Bonti, Noyan, Kembayan, Balai, Parindu, Sekayam, Beduai, and Entikong.
	Bengkayang	Air Besar, Jelimpo, Kuala Behe, Meranti, Menyuke, Ngabang, Sengah Temila, and Sebangki.
A3	Kubu Raya	Kubu, Teluk Pakedai, Sungai Kakap, Rasau Jaya, Ambawang River, Terentang, Batu Ampar, and Kuala Mandor B.
	Pontianak	West Pontianak, Pontianak City, East Pontianak, South Pontianak, North Pontianak, and Southeast Pontianak
A4	Sanggau	Toba, Tayan Hilir, Meliau
	Kubu Raya	Sungai Raya, Sungai Ambawang, Terentang, Kubu, Batu Ampar
	Kayong Utara	Seponti
	Ketapang	Simpang Hulu, Simpang Dua,
A5	Kayong Utara	Pulau Maya
A6	Ketapang	Muara Pawan, Benua Kayong, Sungai Melayu Rayak, Pemahan, Matan Hilir Utara, Simpang Dua, Landai, Laur River
	Kayong Utara	Sukadana, Teluk Batang,

### 3.2 Numerical Summary

After the data has been collected, the following phase is a descriptive analysis to obtain an overall picture of the data distribution [20]. The average, standard deviation, minimum, and maximum values of the number of hotspots on each grid are some of the fundamental statistics included in this descriptive study. The conclusions drawn from the descriptive analysis of the hotspot data on each grid can be seen in **Table 2**.

**Table 2. Numerical Summary of Data Hotspot**

Grid	A1	A2	A3	A4	A5	A6
Maximum	100	100	100	100	82	96
Mean	54	45	42	47	13	42
Standard Deviation	20	30	27	24	25	32

**Table 2** shows descriptive statistics for the data maximum of hotspots in each cell grid gathered in six different grid cells (A1 to A6), each containing 128 observations. Every single grid cell has a value of zero as its minimum. Grid cells A5 and A6 have maximum values slightly lower than the average, at 82 and 96, respectively. The maximum value for most grid cells is 100. January 2023 was the month that the value of 82 (red text in **Table 2**) in the A5 grid unit occurred. This hotspot is the only one found in the A5 grid cell region. When compared to the other grid cells, this hotspot has the maximum value, which is the lowest. Because of the state of the A5 region, which is the boundary between the Kubu Raya area and Karimata Island, the humidity level is higher. As a result, the potential for fire is lower than in other grid cells, which tend to be dry and combustible. Grid cell A1 has the highest average of 54 observations, while grid cell A5 has the lowest average of 13 (red text in **Table 2**). The average observation value varies from grid to grid across the board. This is in addition to the fact that the standard deviation, which indicates the degree of variance in the data, differs between each grid cell. Compared to grid cell A4, which has the lowest standard deviation of 25, grid cell A6 has the highest standard deviation of 32, indicating greater variety in the data. For the most part, **Table 2** summarizes the properties and distribution of the data on each grid cell, which helps conduct additional analysis.

#### 3.2.1 Data Stationary Test

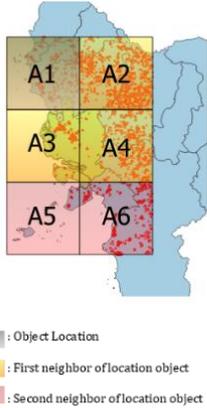
While developing the GSTAR model, one of the processes involves ensuring that the data is steady in terms of both the mean and the variance [21]. In addition, an ADF test, also known as an Augmented Dickey-Fuller test, is carried out to determine whether the data is stationary in the mean [22]. The ADF test was

conducted using the RStudio software, resulting indicated for each cell grid in a p-value of  $0.01 \leq \alpha$  (0.01). This p-value indicates that the data is stationary in the mean, allowing it to be utilized for the subsequent step.

### 3.3 Queen Contiguity Weight Matrix

The Queen Contiguity rule will be utilized for the next stage to calculate the weight matrix. Using the Queen Contiguity rule on the generated location grid, the neighborhood is depicted in the following table, also known as **Table 3**.

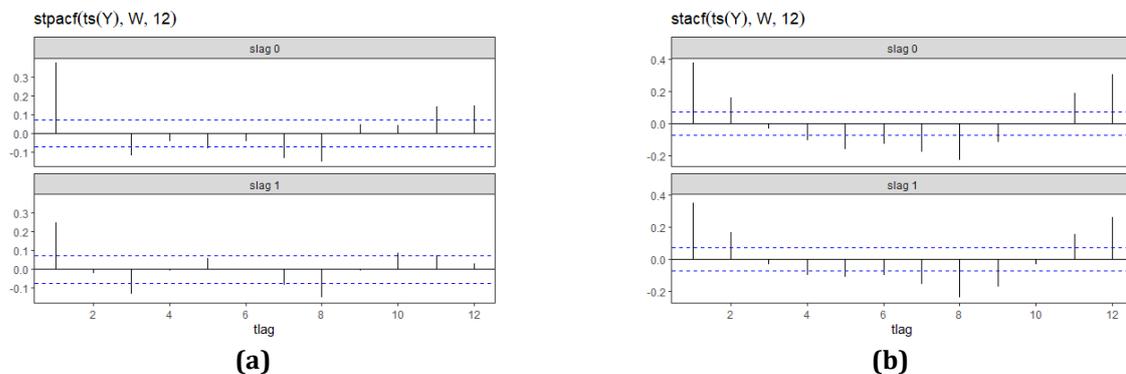
**Table 3. Queen Contiguity Weight Matrix**

Illustration	Weight	Before	After
		Normalization	Normalization
 <p>Legend:                      ■ Object Location                      ■ First neighbor of location object                      ■ Second neighbor of location object</p>	$W^{(0)}$	$I$ , for identity matrix size $N \times N$	
	$W^{(1)}$	$\begin{bmatrix} 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 \end{bmatrix}$	$\begin{bmatrix} 0 & 0.33 & 0.33 & 0.33 & 0 & 0 \\ 0.33 & 0 & 0.33 & 0.33 & 0 & 0 \\ 0.2 & 0.2 & 0 & 0.2 & 0.2 & 0.2 \\ 0.2 & 0.2 & 0.2 & 0 & 0.2 & 0.2 \\ 0 & 0 & 0.33 & 0.33 & 0 & 0.33 \\ 0 & 0 & 0.33 & 0.33 & 0.33 & 0 \end{bmatrix}$
	$W^{(2)}$	<b>UNDEFINED</b> Grid cells A3 and A4 have no second-order neighbors in this spatial grid, meaning grid cells A3 and A4 have no second-order spatial locations in any direction. Therefore, it is impossible to create precisely the queen continuity weight matrix.	

According to **Table 3** (first column), it is possible to observe that grid cell A1 is surrounded by grid cells A2, A3, and A4 as its first neighbors. This means that the  $W_1$  matrix of each grid cell has a value of 1 for the first neighbor and 0 for the other neighbors (see column 3). In the created grid, grid cells A3 and A4 do not have any neighbors in the second spatial lag. The condition indicates that in the second spatial order, no locations are exactly adjacent to grid cells A3 and A4 from any possible orientations. Because of this, it is impossible to create the queen contiguity weight matrix in its entirety. Every site must have nearby neighbors, including diagonal neighbors, for every order of spatial lag to implement the queen contiguity technique. Because the same columns that render the parameter estimation matrix incapable of producing an inverse are present, we cannot continue estimating parameters if the imperfect queen contiguity weight matrix is generated. Therefore  $W^{(\ell)}$  undefined for  $\ell \geq 2$ . The method for determining the weight of the queen contiguity matrix can be seen in **Figure 1**.

#### 3.3.1 Order Identification and Establishment of the GSTAR Model

The STPACF plot of the data can identify the GSTAR model order [23]. **Figure 7** shows STACF and STPACF plots of the data acquired. These plots are based on the weight matrix established in **Table 3**, which was obtained using the RStudio software.



**Figure 7. Plot (a) STACF, (b) STPACF**

Figure 7 shows that the STPACF first slag plot has a cutting-off lag, specifically at lag first, second, and third. GSTAR (1;1), GSTAR (2;1), and GSTAR (3;1) are the possible GSTAR models that can be produced during the process. Following the formation of the GSTAR model, it is possible to proceed to the next step.

### 3.4 Parameter Estimation and Diagnostics Residual Test

Table 4 shows the results of the independence and normality tests of each model's residuals, along with the results of the GSTAR model estimation produced by utilizing Queen Contiguity weights with parameter estimation using the least squares (LS) method. Additionally, the results of the independence test (Ind.) and residual normality (Norm.) can be seen in Table 4. It is important to note that the white noise assumption means that the residuals are normally distributed and independent of each other. The parameter estimation results indicate varied values, with some parameters having positive values and others having negative values, as described in Table 4. The parameter is not utilized in the modeling process, as indicated by the red text in Table 4, which shows that the parameter is insignificant. As a result, the model does not contain parameters that are not significant. The residual diagnostic results have determined that the GSTAR (3;1) model has a white noise assumption with an error percentage of 14.78%. This indicates that this model offers the most favorable outcomes regarding normalcy and independence in each grid.

Table 4. Parameter Estimation and Diagnostics Residual Test

Model	Parameter	Location						MAPE (%)
		(1)	(2)	(3)	(4)	(5)	(6)	
(1;1)	$\phi_{10}$	.578	.372	.457	.447	.312	.198	13.51
	$\phi_{11}$	.408	.465	.452	.568	.191	.815	
	Ind.	Yes	Yes	Yes	Yes	No	Yes	
	Norm.	Yes	Yes	Yes	Yes	No	Yes	
(2;1)	$\phi_{10}$	.377	.336	.403	.314	.359	.14	13.88
	$\phi_{20}$	.325	<b>-0.17</b>	.211	.338	<b>-0.065</b>	-.13	
	$\phi_{11}$	.363	.177	.36	.475	.187	.656	
	$\phi_{21}$	-.078	.355	<b>-0.074</b>	-.136	<b>.014</b>	.409	
	Ind.	Yes	Yes	Yes	Yes	No	Yes	
	Norm.	Yes	Yes	Yes	Yes	No	Yes	
(3;1)	$\phi_{10}$	.317	.370	.405	.283	.319	.136	14.78
	$\phi_{20}$	.219	<b>.009</b>	.196	.309	<b>-0.009</b>	-.125	
	$\phi_{30}$	.191	-.040	<b>.031</b>	<b>-0.058</b>	<b>-0.021</b>	<b>-0.035</b>	
	$\phi_{11}$	.370	.146	.320	.499	.254	.658	
	$\phi_{21}$	<b>-0.040</b>	.454	-.233	<b>-0.115</b>	.154	.413	
	$\phi_{31}$	<b>-0.062</b>	<b>.013</b>	.198	<b>-0.032</b>	-.222	<b>.031</b>	
	Ind.	Yes	Yes	Yes	Yes	Yes	Yes	
	Norm.	Yes	Yes	Yes	Yes	Yes	Yes	

The GSTAR (3;1) model was the most effective approach. Consequently, we used the following formula to predict each grid cell.

$$Y_t = \Phi_{10}Y_{t-1} + \Phi_{11}W^{(1)}Y_{t-1} + \Phi_{20}Y_{t-2} + \Phi_{21}W^{(1)}Y_{t-2} + \Phi_{30}Y_{t-3} + \Phi_{31}W^{(1)}Y_{t-3} + e_t$$

Here is a formula that can be used to estimate hotspots in grid cell A1 until A6.

$$\hat{Y}_t^{(1)} = 0.317Y_{t-1}^{(1)} + 0.123(Y_{t-1}^{(2)} + Y_{t-1}^{(3)} + Y_{t-1}^{(4)}) + 0.219Y_{t-2}^{(1)} + 0.191Y_{t-3}^{(1)}$$

$$\hat{Y}_t^{(2)} = 0.370Y_{t-1}^{(2)} + 0.049(Y_{t-1}^{(1)} + Y_{t-1}^{(3)} + Y_{t-1}^{(4)}) + 0.151(Y_{t-2}^{(1)} + Y_{t-2}^{(3)} + Y_{t-2}^{(4)}) - 0.140Y_{t-3}^{(2)}$$

$$\hat{Y}_t^{(3)} = 0.405Y_{t-1}^{(3)} + 0.017Y_{t-1}^{(5)} + 0.064(Y_{t-1}^{(1)} + Y_{t-1}^{(2)} + Y_{t-1}^{(4)} + Y_{t-1}^{(6)}) + 0.196Y_{t-2}^{(3)} - 0.047(Y_{t-2}^{(1)} + Y_{t-2}^{(4)} + Y_{t-2}^{(5)} + Y_{t-2}^{(6)}) + 0.040(Y_{t-3}^{(1)} + Y_{t-3}^{(2)} + Y_{t-3}^{(4)} + Y_{t-3}^{(5)} + Y_{t-3}^{(6)})$$

$$\hat{Y}_t^{(4)} = 0.283Y_{t-1}^{(4)} + 0.1(Y_{t-1}^{(1)} + Y_{t-1}^{(2)} + Y_{t-1}^{(3)} + Y_{t-1}^{(5)} + Y_{t-1}^{(6)}) + 0.309Y_{t-2}^{(4)}$$

$$\hat{Y}_t^{(5)} = 0.319Y_{t-1}^{(5)} + 0.085(Y_{t-1}^{(3)} + Y_{t-1}^{(4)} + Y_{t-1}^{(6)}) + 0.051(Y_{t-2}^{(3)} + Y_{t-2}^{(4)} + Y_{t-2}^{(6)}) - 0.074(Y_{t-3}^{(3)} + Y_{t-3}^{(4)} + Y_{t-3}^{(6)})$$

$$\hat{Y}_t^{(6)} = 0.136Y_{t-1}^{(6)} + 0.219(Y_{t-1}^{(3)} + Y_{t-1}^{(4)} + Y_{t-1}^{(5)}) - 0.125Y_{t-2}^{(6)} + 0.138(Y_{t-2}^{(3)} + Y_{t-2}^{(4)} + Y_{t-2}^{(5)})$$

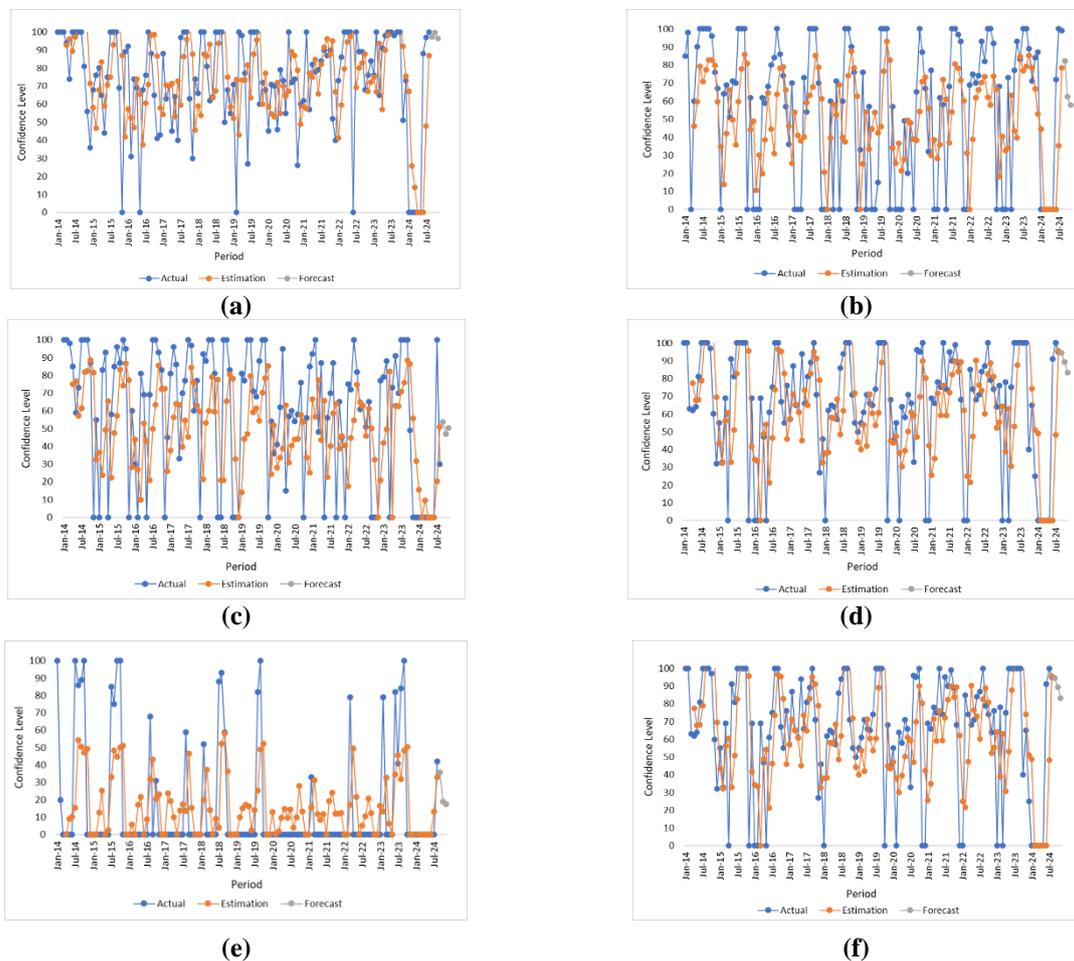
### 3.5 Forecast

The GSTAR (3;1) model is utilized to forecast the hotspot data for the subsequent three months in each grid cell.

**Table 5. Result Prediction of Confidence Level Maximum for Each Location**

Period	A1	A2	A3	A4	A5	A6
Sep-24	97%	84%	54%	94%	36%	65%
Okt-24	100%	64%	47%	89%	19%	60%
Nov-24	96%	58%	50%	83%	17%	59%

Figure 8 shows a plot of the forecast data versus the estimate versus the forecast from each grid cell.



**Figure 8. Comparison Plot of the Actual, Estimation, and Forecast Values for each Location (a) A1, (b) A2, (c) A3, (d) A4, (e) A5, (f) A6**

As can be observed from the significant match between the estimated line and the real data at most sites, Figure 8 shows that the estimating model can accurately reproduce the actual data. In general, the forecasts continue to successfully follow the primary trend at most locations, even though the results show some deviations from the actual data. Nevertheless, there are some discrepancies at specific spots, which indicates that the model can be further enhanced to improve the prediction accuracy at certain areas about certain sites.

## 4. CONCLUSIONS

The conclusion from the analysis that has been carried out is that the GSTAR (3;1) model is the most effective. This conclusion is based on the fact that the development of the GSTAR model ( $p; \lambda_1, \lambda_2, \dots, \lambda_p$ ) utilizing the confidence level hotspot data, the Queen Contiguity weight matrix, and a  $1 \times 1$  degree location grid results in the formation of the GSTAR (3;1) model. With a MAPE value of 14.78%, the GSTAR (3;1) model is the most accurate for predicting the fire point confidence level at the Kubu Raya location and its environs, which satisfies the white noise assumption. However, the result indicates that the model's accuracy level falls into the good category. A decline was observed in most places, as indicated by the confidence level hotspot findings for September, October, and November 2024. It was in September 2024 that Location A1 had the maximum degree of trust, 97%. In October 2024, the confidence level grew to 100%; in November 2024, it declined to 96%. The fact that this is the case demonstrates that location A1 is in dire need of preventative mitigation and stringent supervision, such as positioning firefighters in vulnerable areas. Location A1 includes all the sub-districts in Singkawang City and Mempawah Regency, as well as some sub-districts in Bengkayang Regency, Landak Regency, and Kuburaya Regency, specifically Kuala Mandor B District. This location encompasses all the sub-districts in the aforementioned districts. According to the findings of this research, the GSTAR model can accurately forecast the level of confidence associated with hotspots. However, it also helps optimize resource allocation in forest fire prevention. This study only uses one weight matrix and one variable. Future research might use other weight matrices and other elements that influence the incidence of hotspots, such as the amount of rainfall or the wind speed, to produce more accurate forecast findings.

## REFERENCES

- [1] N. P. Sambodo, M. Pradhan, R. Sparrow, and E. van Doorslaer, "WHEN THE SMOKE GETS IN YOUR LUNGS: SHORT-TERM EFFECTS OF INDONESIA'S 2015 FOREST FIRES ON HEALTH CARE USE," *Environmental Health*, vol. 23, no. 1, p. 44, May 2024, doi: 10.1186/s12940-024-01079-x.
- [2] L. N. Adelia, I. P. Ash Shidiq, and F. Zulkarnain, "IDENTIFICATION OF FOREST FIRE HAZARD AREA IN KUBU RAYA REGENCY, WEST KALIMANTAN PROVINCE," *IOP Conf Ser Earth Environ Sci*, vol. 1190, no. 1, p. 012037, Jun. 2023, doi: 10.1088/1755-1315/1190/1/012037.
- [3] T. F. Dicelebica, A. A. Akbar, and D. R. Jati, "IDENTIFIKASI DAN PENCEGAHAN DAERAH RAWAN BENCANA KEBAKARAN HUTAN DAN LAHAN GAMBUT DI KALIMANTAN BARAT," *Jurnal Ilmu Lingkungan*, vol. 20, no. 1, pp. 115–126, Jan. 2022, doi: 10.14710/jil.20.1.115-126.
- [4] N. Imro, D. Kusnandar, and N. N. Debatara, "WEST KALIMANTAN FOREST FIRE PROBABILITY MAPPING USING BINARY LOGISTIC REGRESSION," vol. 8, no. 2, pp. 204–214, 2024.
- [5] M. Alkaff and N. E. Yulianto, "PREDIKSI JUMLAH KEJADIAN TITIK API MELALUI PENDEKATAN DERET WAKTU MENGGUNAKAN MODEL SEASONAL ARIMA," *Jurnal ELTIKOM*, vol. 3, no. 2, pp. 54–63, 2019, doi: 10.31961/eltikom.v3i2.122.
- [6] 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," *Spat Stat*, vol. 54, p. 100734, Apr. 2023, doi: 10.1016/j.spasta.2023.100734.
- [7] N. M. Huda and N. Imro'ah, "COVID-19 CASE MODELING IN JAVA ISLAND USING A SPATIAL MODEL, GSTAR(1;1), WITH MODIFIED SPATIAL WEIGHTS: QUEEN CONTIGUITY WEIGHT MATRIX," 2024, p. 090009. doi: 10.1063/5.0201676.
- [8] Y. Widyaningsih and N. Ayu, "GENERALIZED SPATIO-TEMPORAL AUTOREGRESSIVE (GSTAR) MODELING ON CORONAVIRUS DISEASE (COVID-19) CASES IN DKI JAKARTA," 2024, p. 030003. doi: 10.1063/5.0212710.
- [9] N. Ilmi, A. Aswi, and M. K. Aidid, "GENERALIZED SPACE TIME AUTOREGRESSIVE INTEGRATED MOVING AVERAGE (GSTARIMA) DALAM PERAMALAN DATA CURAH HUJAN DI KOTA MAKASSAR," *Inferensi*, vol. 6, no. 1, p. 25, Mar. 2023, doi: 10.12962/j27213862.v6i1.14347.
- [10] U. Mukhaiyar, F. T. Nabilah, U. S. Pasaribu, and N. M. Huda, "THE SPACE-TIME AUTOREGRESSIVE MODELLING WITH TIME CORRELATED ERRORS FOR THE NUMBER OF VEHICLES IN PURBALEUNYI TOLL GATES," *J Phys Conf Ser*, vol. 2243, no. 1, p. 012068, Jun. 2022, doi: 10.1088/1742-6596/2243/1/012068.
- [11] D. R. Febrianti, M. A. Tiro, and S. Sudarmin, "METODE VECTOR AUTOREGRESSIVE (VAR) DALAM MENGANALISIS PENGARUH KURS MATA UANG TERHADAP EKSPOR DAN IMPOR DI INDONESIA," *VARIANSI: Journal of Statistics and Its application on Teaching and Research*, vol. 3, no. 1, p. 23, 2021, doi: 10.35580/variansiunm14645.
- [12] J. Fattah, L. Ezzine, Z. Aman, H. El Moussami, and A. Lachhab, "FORECASTING OF DEMAND USING ARIMA MODEL," *International Journal of Engineering Business Management*, vol. 10, Jan. 2018, doi: 10.1177/1847979018808673.
- [13] 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, p. 012223, Jun. 2018, doi: 10.1088/1742-6596/1028/1/012223.

- [14] Y. Yundari, U. S. Pasaribu, and U. Mukhaiyar, "ERROR ASSUMPTIONS ON GENERALIZED STAR MODEL," *Journal of Mathematical and Fundamental Sciences*, vol. 49, no. 2, p. 136, Oct. 2017, doi: 10.5614/j.math.fund.sci.2017.49.2.4.
- [15] A. Tsanawafa, D. A. Kusuma, and B. N. Ruchjana, "SPATIAL WEIGHT MATRIX COMPARISON OF SAR-X MODEL USING CASETTI APPROACH," *CAUCHY: Jurnal Matematika Murni dan Aplikasi*, vol. 9, no. 1, pp. 129–137, May 2024, doi: 10.18860/ca.v9i1.25579.
- [16] M. Sobari, A. Desiyanti, D. Yanti, P. Monika, A. S. Abdullah, and B. N. Ruchjana, "COMPARISON OF SPATIAL WEIGHT MATRICES IN SPATIAL AUTOREGRESSIVE MODEL: CASE STUDY OF INTANGIBLE CULTURAL HERITAGE IN INDONESIA," *JTAM (Jurnal Teori dan Aplikasi Matematika)*, vol. 7, no. 1, p. 244, Jan. 2023, doi: 10.31764/jtam.v7i1.10757.
- [17] A. Schmidt *et al.*, "FIRE FREQUENCY, INTENSITY, AND BURN SEVERITY IN KALIMANTAN'S THREATENED PEATLAND AREAS OVER TWO DECADES," *Frontiers in Forests and Global Change*, vol. 7, Feb. 2024, doi: 10.3389/ffgc.2024.1221797.
- [18] A. N. Pradana, A. Djuraidah, and A. M. Soleh, "LAND USE CHANGE MODELLING USING LOGISTIC REGRESSION, RANDOM FOREST AND ADDITIVE LOGISTIC REGRESSION IN KUBU RAYA REGENCY, WEST KALIMANTAN," *Forum Geografi*, vol. 37, no. 2, Dec. 2023, doi: 10.23917/forgeo.v37i2.23270.
- [19] M. S. Akbar *et al.*, "A GENERALIZED SPACE-TIME AUTOREGRESSIVE MOVING AVERAGE (GSTARMA) MODEL FOR FORECASTING AIR POLLUTANT IN SURABAYA," *J Phys Conf Ser*, vol. 1490, no. 1, p. 012022, Mar. 2020, doi: 10.1088/1742-6596/1490/1/012022.
- [20] T. C. Guetterman, "BASICS OF STATISTICS FOR PRIMARY CARE RESEARCH," *Fam Med Community Health*, vol. 7, no. 2, p. e000067, Mar. 2019, doi: 10.1136/fmch-2018-000067.
- [21] U. Mukhaiyar, "THE GOODNESS OF GENERALIZED STAR IN SPATIAL DEPENDENCY OBSERVATIONS MODELING," 2015, p. 020008. doi: 10.1063/1.4936436.
- [22] A. Roza, E. S. Violita, and S. Aktivani, "STUDY OF INFLATION USING STATIONARY TEST WITH AUGMENTED DICKEY FULLER & PHILLIPS-PERON UNIT ROOT TEST (CASE IN BUKITTINGGI CITY INFLATION FOR 2014-2019)," *EKSAKTA: BERKALA ILMIAH BIDANG MIPA*, vol. 23, no. 02, pp. 106–116, Jun. 2022, doi: 10.24036/eksakta/vol23-iss02/303.
- [23] A. B. Salsabila, B. N. Ruchjana, and A. S. Abdullah, "DEVELOPMENT OF THE GSTARIMA(1,1,1) MODEL ORDER FOR CLIMATE DATA FORECASTING," *International Journal of Data and Network Science*, vol. 8, no. 2, pp. 773–788, 2024, doi: 10.5267/j.ijdns.2024.1.001.