Application of GSTAR(1,1) model for layer peat soil predicted based on resistivity log data

R Jonathan¹*, Yundari¹, Nurhasanah², O Y E Nada²

¹ Mathematics Department, Universitas Tanjungpura, Pontianak, Indonesia
² Physics Department, Universitas Tanjungpura, Pontianak, Indonesia

*Email: ryanjonathan8819@student.untan.ac.id

Abstract. In this study, GSTAR modeling was carried out with the inverse of distance weight matrix obtained from Geoelectrical Resistivity data at several peatland locations around the Universitas Tanjungpura, Pontianak. This data can identify the subsurface layer of the soil through the electric current that binds into the soil. However, due to the limitation of the tool to measure the resistivity value, it can only measure 1/5 of the depth of the observation length. To overcome this problem, predictions are made at the next depth using the GSTAR model. The study began by measuring the resistivity value of the land using the geoelectric method and mapping it. Through this GSTAR modeling, predictions are made for the unobserved subsurface to determine the type of soil layer. Knowing the type of deeper soil layer can help contractors build plant concrete stakes to keep buildings safe on peatland. The results of the GSTAR(1.1) model are not accurate enough to estimate the resistivity value data. This is possible because the correlation between rock ages is not the same, so further analysis is required.

1. Introduction

Soils play an important role in infrastructure and construction projects such as buildings, roads, dams, and other structural developments. Different types of soil are used in construction projects, such as soft soils (peat soils). This type of soil exists all over the world and covers almost 8% of the earth's surface [1]. It mainly occurs in the tropical countries of the world [2], [3]. West Kalimantan is the 4th largest peatland in Indonesia after Papua, Central Kalimantan, and Riau, which is 1.8 million hectare [4]. However, the use of peat soil has been reported to be problematic due to its extreme softness, unconsolidation, low shear strength, stiffness, and high water content [5], [6]. This leads to serious problems such as long-term and over-billing during or after construction and ultimately leads to cost overruns in time and costs in construction projects [7]. Besides, construct a building on peat soils becomes harder, especially in planting concrete piles for a building. The planted concrete stakes must reach a layer of soil/rock which is hard enough for the buildings to be sturdy and durable.

The tools for analyzing the subsurface are through well logs and seismic. The well logs data that can identify sediment or rock are resistivity logs, spontaneous potential (SP) logs, Gamma Ray logs, Neutron logs, density logs, and sonic logs [8]. These well logs data can help geologists to determine layers sediment or rock without drilling.
Several studies have used well log data such as analyze rock resistivity using the concept of anisotropy [9], Gamma Ray logs using the GSTAR model [10], identification of hard soil depth using resistance type of geoelectrical method [11], [12], and so on. For stochastic modeling using resistivity data, this has not been done so in this study the data is modeled using the GSTAR(1,1) space-time model. The selection of the GSTAR(1,1) model is based on both time and spatial data which tend to be homogeneous, so that the model identification stage is not carried out. The parameter index used in this model is rock depth. This is in accordance with the superposition principle of stratigraphic analysis which says that the lower the rock layer, the older the rock/layer. With the assumption that the GSTAR model is stationary, it is hoped that this data can be used to estimate subsurface rock types through resistivity data.

2. Method

The following represents the linear regression of the GSTAR(1;1) model. The notation of the parameter matrix for the GSTAR(1;1) model is \( \Phi = diag(\phi_1, \ldots, \phi_N) \) and \( W = (w_{ij}) \), therefore the GSTAR(1;1) model in matrix form can be written as follows

\[
Y(t) = \Phi_0 Y(t-1) + \Phi_1 W Y(t-1) + \varepsilon(t)
\]

The form of the model in Equation (1) makes it easier for the parameter estimation stage using the least squares method. The part that characterizes the GSTAR model is its spatial weight matrix, namely \( W \). In this study, the inverse distance is used as the spatial weight. The formula is

\[
w_{ij} = \frac{1}{\sum_{j=1}^{N} d_{ij}} d_{ij}^{-1} \quad j \neq i
\]

\( w_{ij} = 0 \) for \( j = i \). In equation (2), \( d_{ij} \) is the distance from location \( i \) to location \( j \) [13].

The stages of modeling GSTAR(1,1) can be seen in Figure 2. Software that used for analyze the data was RStudio.
3. Result and Discussion

3.1. Data Description

The data used in this study are primary data from the resistivity data of the Auditorium Universitas Tanjungpura, Pontianak, West Kalimantan in Figure 3. The data were obtained using the electrical resistivity method with Schlumberger configuration. The Schlumberger setup is often used for data acquisition because it provides good results for data acquisition with deep penetration [14].

The summary of statistics for resistivity data at three locations are shown in Table 1. From the table, we find that location 3 has the highest mean resistivity, while location 2 has the lowest mean resistivity.
The range values is quite wide. Compared to other locations, the standard deviation at location 3 is also quite large. It turns out that the data distribution at location 3 is very widespread.

**Figure 3.** Study location in Auditorium Universitas Tanjungpura Pontianak.

| Statistic         | Location 1 | Location 2 | Location 3   |
|-------------------|------------|------------|--------------|
| Average           | 28.51      | 21.88      | 8610.40      |
| Standar Deviation | 50.54      | 31.40      | 11863.98     |
| Min               | 2.50       | 0.57       | 3.00         |
| Max               | 198.38     | 215.22     | 55038.87     |

3.2. **GSTAR Modeling on Resistivity Data**

The first step, the data is centered so that each observation reduced by its mean. The second step, we check stationary of data. The data are stationary if the mean and variance are constant[15]. The data is "stationary on mean" if the data is stable and fluctuates around the mean. The data is "varying stationary" when the data varies from time to time, but the mean does not have to be constant. Stationarity can be verified by observing the time series data graph and the ADF test. The data graph can be seen in Figure 4 and assumption of elevation angle of three location is same. Figure 4 shows that the variation of the plot is not around the mean value and the fluctuation is not constant. It can be concluded that the resistivity data at three locations are not stationary therefore, the data need to be differencing.

Otherwise, the stationarity of the data can be identified using the ADF test. The ADF test was carried out with the tseries package in RStudio software. The ADF test results of the resistivity data after differencing are shown in Table 2. Table 2 shows that each variable has a p-value <0.05, so it can be decided that the data is stationary.
Figure 4. Plot of the resistivity data at (a) Location 1, (b) Location 2, and (c) Location 3.

Table 2. ADF test

| No | Variable     | p-value | Decision   |
|----|--------------|---------|------------|
| 1  | Location 1   | 0.010   | Stationary |
| 2  | Location 2   | 0.010   | Stationary |
| 3  | Location 3   | 0.010   | Stationary |

The next step is determination of weight matrix. The weight matrix of the GSTAR model used in this research is the inverse of distance weight matrix. The element of an inverse distance weight matrix is designed based on the real distance (see Figure 3).

The weight matrix is obtained

\[
W = \begin{bmatrix}
0 & 0.59091 & 0.40909 \\
0.70968 & 0 & 0.29032 \\
0.62857 & 0.37143 & 0
\end{bmatrix}
\]

Based on the weight matrix \( W \), we use the GSTAR(1,1) model. The order of the GSTAR model can be identified based on the space-time autocorrelation function (STACF) plot and the space-time partial autocorrelation function (STPACF) plot for the locations. But, in this research we use the first of time and spatial order, because we assumption time and spatial lag is homogenous.

The model we obtained is written by Equation(1) where

\[
\Phi_0 = \begin{bmatrix}
3.67 \times 10^{-1} & 0 & 0 \\
0 & 3.06 \times 10^{-1} & 0 \\
0 & 0 & 6.29 \times 10^{-1}
\end{bmatrix} \quad \text{and} \quad \Phi_1 = \begin{bmatrix}
-7.27 \times 10^{-5} & 0 & 0 \\
0 & 1.25 \times 10^{-4} & 0 \\
0 & 0 & -975 \times 10^{-3}
\end{bmatrix}
\]
Figure 5. Plot the results of the original (black) and estimated (red) data using GSTAR(1,1) for each location (top) and Interpretation of rock layers through resistivity (bottom) (a) location 2 (b) location 1 (c) location 3.

If this resistivity data is interpreted in rock layers, it is obtained as in Figure 5. It can be seen from Figure 5 that each location produces a different layer. This is possible because the layers used for each location are not the same relative time. As a result, the simultaneous modeling using GSTAR(1,1) is not good in estimation results.

4. Conclusion

The estimation results obtained show that the GSTAR(1.1) model for this resistivity data is not yet accurate. This can be seen from the large MSE values at locations 2 and 3. So, that predictions below the soil layer (which have not been observed) cannot be carried out. Further research is needed for the development of this model. The steps that should be done before doing the GSTAR modeling are looking for the same time correlation for each location. The results of the discussion on the GSTAR (1,1) model with inverse distance weights indicate that at location 1, at a depth of 17 meters, quite hard rock layers have been found. At location 2, hard rock layers were found at a depth of 4 meters, while at location 3, at 17.5 meters deep hard rocks were found.
References
[1] Dehghanbanadaki A, Sotoudeh M A, Golpazir I, Keshtkarbanaeemoghadam A and Ilbeigi M, 2019 Bull. Eng. Geol. Environ. 78 pp. 1345–1358
[2] Abdel-Slam AE 2018 HBRC J. 14 (3) pp. 294–299
[3] Leng L, Ahmed O and Jalloh M 2019 Geosci. Front, 10(2) pp. 373–380
[4] Katadata TP, “Luas Gambut Indonesia Terbesar Kedua di Dunia” 2019. https://katadata.co.id/timpublikasikatadata/infografik/5e9a519433cb1/luas-gambut-indonesia-terbesar-kedua-di-dunia.
[5] Ghareh S, Kazemian S and Shahin M 2020 Geomech. Eng, 21(4) pp. 337–348
[6] Deboucha S, Hashim R and Alwi A 2008 Electron, J. Geotech. Eng. 13 pp. 1–9
[7] Elufowoju F 2019 Niger J. Technol. 38 (2) pp. 289–297
[8] Kendall C, “The Tools of Subsurface Analysis,” Fall 2005
[9] Prameswari F W, Bahri A S and Parnadi W 2012 J. Sains dan Seni ITS 1(1), p. B-15-B-20
[10] Yundari, Pasaribu U S, Mukhaiyar U and Heriawan M N, “Spatial Weight Determination of GSTAR(1;1) Model by Using Kernel Function,” in Journal of Physics: Conference Series, 2018, vol. 1028, no. 1, doi: 10.1088/1742-6596/1028/1/012223.
[11] Muliadi M, Zulfian Z and Muhardi M 2019 Positron 9(2), p. 86
[12] Masudi, Nurhasanah and Muhardi 2021 J. Ilmu dan Inov. Fis. 05(01) pp. 59–64
[13] Yundari, Pasaribu US and Mukhaiyar U 2017 J. Math. Fundam. Sci., 49(2) pp. 136–155
[14] Dhahir M and Nsaif N 2016 Diyala J. Eng. Sci. 9(2) pp. 50–61
[15] Box G, Jenkins G and Reinsel G 1994 Time Series Analysis Forecasting and Control 3rd ed. (New Jersey: Prentice-hall Internasional).