Media Statistika

p-ISSN 1979 – 3693 e-ISSN 2477 – 0647
MEDIA STATISTIKA 15(2) 2022: 186-197
http://ejournal.undip.ac.id/index.php/media_statistika

ESTIMATING AND FORECASTING COVID-19 CASES IN SULAWESI ISLAND USING GENERALIZED SPACE-TIME AUTOREGRESSIVE INTEGRATED MOVING AVERAGE MODEL

Sukarna[1], Nurul Fadilah Syahrul[1], Wahidah Sanusi[1], Aswi Aswi[2], Muhammad Abdy[1], Irwan[1]
[1]Mathematics Department, Faculty of Mathematics and Natural Science, Universitas Negeri Makassar, Indonesia
[2]Statistics Department, Faculty of Mathematics and Natural Science, Universitas Negeri Makassar, Indonesia

e-mail: sukarna@unm.ac.id

DOI: 10.14710/medstat.15.2.186-197

Article Info:
Received: 3 April 2022
Accepted: 31 March 2023
Available Online: 4 April 2023

Keywords:
estimating, forecasting, GSTARIMA, covid-19

Abstract: A range of spatio-temporal models has been used to model Covid-19 cases. However, there is only a small amount of literature on the analysis of estimating and forecasting Covid-19 cases using the Generalized Space-Time Autoregressive Integrated Moving Average (GSTARIMA) model. This model is a development of the GSTARMA model which has non-stationary data. This paper aims to estimate and forecast the daily number of Covid-19 cases in Sulawesi Island using GSTARIMA models. We compared two models namely GSTARI and GSTIMA considering the root mean square error (RMSE). Data on a daily number of Covid-19 cases (from April 10, 2020, to May 07, 2021) were used. The location weight used is the inverse distance weight based on the distance between airports in the capital cities of each province. The appropriate models obtained based on the data are the GSTARIMA (1;0;1;1) model and the GSTARIMA (1;1;1;0) model. The results showed that the forecast for the number of new Covid-19 cases is accurate and reliable only for the short term.

1. INTRODUCTION

The first case of *Corona Virus* disease-19 (Covid-19) was found in the city of Wuhan, Hubei Province, China at the end of December 2019 (Shereen, et al., 2020). Covid-19 is a type of coronavirus beta or abbreviated as β-*coronavirus* caused by *Severe Acute Respiratory Syndrome-Corona Virus* 2 (SARS-CoV-2) (Ahmar & Boj, 2020; Djalante et al., 2020; Sohrabi et al., 2020; Torrealba-Rodriguez, et al., 2020). Covid-19 spread to other countries and the World Health Organization (WHO) declared Covid-19 as a pandemic on March 11, 2020 (Djalante et al., 2020; WHO, 2020). In Indonesia, President Joko Widodo announced two confirmed people who were first infected with Covid-19 on March 2, 2020 (WHO, 2020). By May 7, 2021, there were 1,703,632 positive confirmed cases of Covid-19 in Indonesia, while the recovered and dead cases were 1,558,423 cases and 46,663 cases, respectively (Ministry of Health, 2021).

Sulawesi Island has the highest number of positive confirmed cases of Covid-19 in the Eastern part of Indonesia. This is based on the final update reported by (the Ministry of

Health, 2021) on 7 May 2021, with the following details: Sulawesi Island has a total of 111,134 positive confirmed cases of Covid-19, 15,672 cases in North Sulawesi, 12,506 cases in Central Sulawesi, 61,608 cases in South Sulawesi, 10,450 cases in Southeast Sulawesi, 5,422 cases in Gorontalo, and 5,476 cases in West Sulawesi.

Data on new cases of Covid-19 is presented in the form of time series data in various locations. The data presented in this form is called space-time or spatio-temporal data. As the Covid-19 cases are presented in space and time, it is suitable to be modeled using the spatio-temporal model (Abdullah et al., 2018). The spatio-temporal model was first introduced by Cliff and Ord in 1975 (Cliff & Ord, 1975) known as the Space-Time Autoregressive (STAR) model, and expanded by Pfeifer and Deutsch in 1980 (Pfeifer & Deutsch, 1980) by including the Moving Average in the model to be the Space-Time Autoregressive Moving Average (STARMA) model. These models have a weakness in autoregressive parameters as it is assumed to have the same characteristics in each location et al., 2018). This is what prompted scientists to propose ideas for improving the STAR model. The generalized Space-Time Autoregressive (GSTAR) model was proposed to improve the STAR model (Borovkova, et al., 2002) and expanded by Di Giacinto in 2006 (Di Giacinto, 2006) known as the Generalized Space-Time Autoregressive Moving Average (GSTARMA) model. The GSTARMA model can be expanded into the Generalized Space-Time Autoregressive Integrated Moving Average (GSTARIMA) model (Wei, 2019). This model allows the assumption of different autoregressive parameters in each location (Zewdie et al., 2018).

Recently, several studies used spatio-temporal models in modeling Covid-19 (Lee, et al., 2021; Wang et al., 2021). Multivariate spatio-temporal (MVST) was introduced (Lee et al., 2021). Autocorrelation spatial, hot spots and spatio-temporal scanning statistics were used (Wang et al., 2021). They used Spearman rank correlation analysis and multiple linear regression to explore the relationship between the factors that influence Covid-19 cases.

To our knowledge, there are only three studies that used the Space-Time Autoregressive model to model Covid-19 (Alawiyah, et al., 2021; Awwad, et al., 2021; Pasaribu, et al., 2021). The GSTAR model for daily Covid-19 cases with the inverse weight of the distance from the average length of travel between trains from one province to others on Java Island was used (Pasaribu et al., 2021). They modified the inverse distance weight using the ratio population in the 6 provinces on Java Island. The GSTARI model for daily Covid-19 cases in the city of Bandung Raya using standardized queen contiguity weights was also used (Alawiyah et al., 2021). STARIMA for daily Covid-19 cases in three areas, that is, Makkah, Jeddah, and Taif using standardized queen contiguity weight for the spatial were applied (Awwad et al., 2021).

Our study differs from the previous three studies. We focused on estimating and forecasting Covid-19 cases in Sulawesi Island using the GSTARIMA model which has not been done by other studies. We modified the location weight matrix by using inverse distance weighting. The distance used is the distance between airports in the capital city of each province. The reason behind that is that moving from one province to another has a significant possibility of transmitting the spread of Covid-19 and public transportation is one of the places that are vulnerable to the spread (Pasaribu et al., 2021). The main public transportation that allows all tourists to access from around the world is air transportation (Ricardianto, et al., 2017). Our research focuses on the provinces in Sulawesi Island as it has the highest total of confirmed positive Covid-19 in Eastern Indonesia based on data from the Ministry of Health of the Republic of Indonesia (Ministry of Health, 2021).

2. LITERATURE REVIEW

2.1. GSTARIMA Model

GSTARIMA (λ_k; p; d; q) is modelling of some observations $Z_i(t)$ at each location N in space (i = 1, 2, ..., N) for t periods. Time effects and spatial effects are defined as time series models and space weight matrices, respectively (Borovkova, et a2008).

The GSTARIMA model with spatial order 1, 2, …, λ_k, autoregressive order (p), differencing (d), and moving average (q), denoted by GSTARIMA (λ_k; p; d; q) is defined as follows (Wei, 2019):

$$\begin{aligned}\Delta Z_{i(N\times 1)}(t) = &\sum_{k=1}^{p}\sum_{l=0}^{\lambda_k} \phi_{kl(N\times N)} W_{(N\times N)}^{(l)}\, \Delta Z_{i(N\times 1)}(t-k) + a_{i_{(N\times 1)}}(t) \\ &-\sum_{k=1}^{q}\sum_{l=0}^{m_k} \Theta_{kl(N\times N)} W_{(N\times N)}^{(l)}\, a_{i_{(N\times 1)}}(t-k)\end{aligned} \tag{1}$$

where $\Delta Z_{i(N\times 1)}(t)$ is the observation vector at time t, t =1, 2, 3,…, K and the i^{th} location ($N \times 1$) after first differencing. While $p, q, \lambda_k, m_k, N, W_{(N\times N)}^{(l)}, W^0, \phi_{kl(N\times N)}, \Theta_{kl(N\times N)}$, and $a_{i_{(N\times 1)}}(t)$ are autoregressive (AR) order, the moving average (MA) order, spatial order of the k-th autoregressive term, spatial order of the k-th moving average term, space weight matrix ($N \times N$) at the spatial lag l, identity matrix, the autoregressive parameter at temporal lag k and spatial lag l, the moving average parameter at temporal lag k and spatial lag l, and the random normally distributed error vector at the time (t), respectively, with $\Delta Z_{i(N\times 1)}(t) = Z_{i(N\times 1)}(t) - Z_{i(N\times 1)}(t-1)$, so that $\Delta Z_{i(N\times 1)}(t-k) = Z_{i(N\times 1)}(t-k) - Z_{i(N\times 1)}(t-k-1)$, and E($a_i(t)$) = 0.

2.2. GSTARIMA(1;1;1;0) or GSTARI (1;1;1) Model

Based on equation (1) GSTARIMA(1;1;1;0) or GSTARI (1;1;1) model is defined as follows:

$$\begin{aligned}\Delta Z_{i(N\times 1)}(t) = &\ \phi_{10(N\times N)}\Delta Z_{i(N\times 1)}(t-1) \\ &+\phi_{11(N\times N)} W_{(N\times N)}^{(1)} \Delta Z_{i(N\times 1)}(t-1) + a_{i_{(N\times 1)}}(t)\end{aligned} \tag{2}$$

2.3. GSTARIMA(1;0;1;1) or GSTIMA(1;1;1) Model

Based on equation (1) GSTARIMA(1;0;1;1) or GSTIMA(1;1;1) model is defined as follows:

$$\begin{aligned}\Delta Z_{i(N\times 1)}(t) = &\ a_{i_{(N\times 1)}}(t) - (\Theta_{10(N\times N)} a_{i_{(N\times 1)}}(t-1) \\ &+\Theta_{11(N\times N)} W_{(N\times N)}^{(1)} a_{i_{(N\times 1)}}(t-1))\end{aligned} \tag{3}$$

2.4. Spatial Weight Matrix

Spatial weight matrix $\boldsymbol{W} = (w_{ij})$ is defined as follows (Prillantika, et al., 2018):

$$\boldsymbol{W} = (w_{ij}) = \begin{bmatrix} 0 & w_{12} & \cdots & w_{1j} \\ w_{21} & 0 & \cdots & w_{2j} \\ \vdots & \vdots & \ddots & \vdots \\ w_{i1} & w_{i2} & \cdots & w_{ij} \end{bmatrix} \tag{4}$$

2.4.1. Inverse Distance Weight Matrix

The inverse distance weight matrix between the i-th location and j-th location is defined as follows: (Maria, Budiman, Haviluddin, & Taruk, 2020; Prillantika et al., 2018):

$$w_{ij} = \begin{cases} \dfrac{\frac{1}{d_{ij}}}{\sum_{j=1}^{N} \frac{1}{d_{ij}}}, & i \neq j \\ 0, & i = j \end{cases} \tag{5}$$

and

$$d_{ij} = \sqrt{(u_i - u_j)^2 + (v_i - v_j)^2}, i \neq j \tag{6}$$

where d_{ij} is the distance from location i to j, (u_i, u_j) is the latitude coordinate, and (v_i, v_j) is the longitude coordinate, and d_{ij} is assumed to be the same as d_{ji}.

2.5. Root Mean Square Error (RMSE)

The best model was selected based on the accuracy of predicting out-sample data (Setiawan, Suhartono, & Prastuti, 2016) using the Root Mean Square Error (RMSE). The model with the smallest RMSE is preferred.

3. MATERIAL AND METHOD

The daily number of Covid-19 cases was gathered from the official Indonesian website https://covid19.go.id, covering daily periods from April 10, 2020, to May 7, 2021, in six provinces in Sulawesi Island. There are six spaces in this study and are defined as follows:

Z_1 = data on new cases of Covid-19 in North Sulawesi
Z_2 = data on new cases of Covid-19 in Central Sulawesi
Z_3 = data on new cases of Covid-19 in South Sulawesi
Z_4 = data on new cases of Covid-19 in Southeast Sulawesi
Z_5 = data on new cases of Covid-19 in Gorontalo
Z_6 = data on new cases of Covid-19 in West Sulawesi

Data on Covid-19 cases are taken through the R package httr, function GET in R Version 4.1.0 (R Core Team, 2021). The data is divided into two groups, namely data in sample: April 10, 2020 – February 17, 2021 (314 observations) and data “out sample”: February 18, 2021 – 07 May 2021 (78 observations). Augmented Dickey-Fuller (ADF) test is done through the tseries function adf test package. The plot of the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) can be identified through the package stats function acf and pacf. The weight of the inverse distance matrix is built through the stats function dist package using the "Euclidean" method. The parameters of the GSTARIMA model are estimated using the Ordinary Least Square (OLS) method through the package stats function lm. All functions of the package use R Version 4.1.0 (R Core Team, 2021). The map of the predicted results is visualized through Tableau Public 2020.4 (Tableau Public, 2020).

4. RESULTS AND DISCUSSION

4.1. Descriptive Analysis

Descriptive Statistics of in-sample Covid-19 data in Sulawesi Island can be seen in Table 1. Based on Table 1, the data used in every location are the same (314 data). South Sulawesi Province has the highest mean of new Covid-19 cases (168.60) followed by North

Sulawesi (46.58) and Southeast Sulawesi (31.39). Gorontalo Province has the lowest mean of new Covid-19 cases 14.77.

Table 1. Descriptive Statistics of in-sample COVID-19 Data on Sulawesi Island

Location	N	Mean	Median	Standard Deviation	Minimum	Maximum
Z_1	314	46.58	30.50	45.52	0	241
Z_2	314	29.89	5.00	52.34	0	275
Z_3	314	168.60	112.50	177.93	0	679
Z_4	314	31.39	23.50	32.54	0	213
Z_5	314	14.77	5.00	23.97	0	146
Z_6	314	15.90	4.00	33.89	0	241

4.2. GSTARIMA Model

We estimated the GSTARIMA model using the procedure of three-stage (Pfeifer & Deutsch, 1980). This procedure is an extension of the methodology of Box-Jenkins in space-time setup, namely, model identification, parameter estimation, and diagnostic checking. The steps taken at the identification stage are time series plots, ACF, and PACF plots. The parameter estimation stage involves checking whether the parameters are statistically significant. The diagnostic checking stage consists of checking whether the residuals are independent or not (white noise process).

4.2.1. Time Series Plot

A time series plot of the number of Covid-19 cases in six provinces on Sulawesi Island is displayed in Figure 1. Based on Figure 1, the province of South Sulawesi has the highest number of daily Covid-19 cases. The time series plots visually show that the data are not stationary.

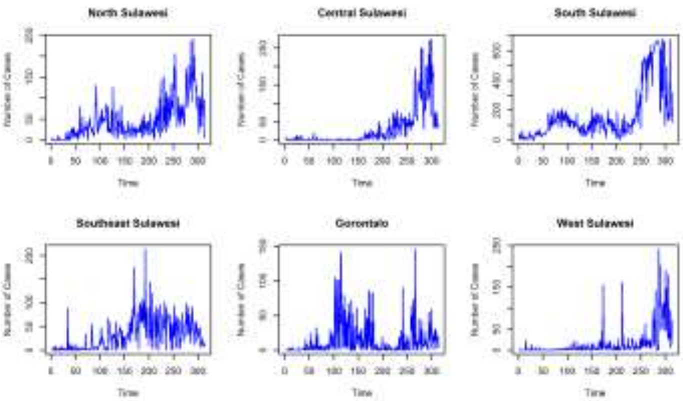

Figure 1. Time Series Plot of the Number Of Covid-19 Cases in Six Provinces on Sulawesi Island

4.2.2. Data Stationarity

Detecting stationarity in time series data is formally carried out with the Augmented Dickey-Fuller (ADF) test. Based on the ADF test, data on the number of Covid-19 new cases in five provinces in Sulawesi Island has a p-value ≥ 0.05 suggesting that the data is not stationary. However, only data on the number of Covid-19 from West Sulawesi province is stationary with a p-value < 0.05 (0.047). One possible method for making time-series data stationary is differencing. After applying the first differencing, the data for each location is stationary. Time series plots after the first differencing can be shown in Figure 2.

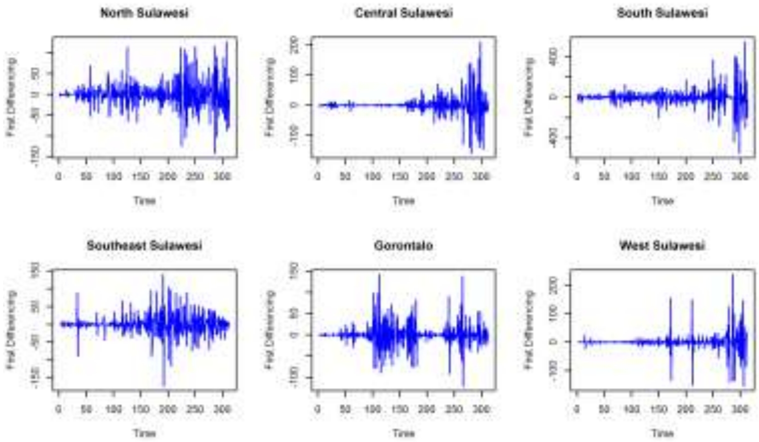

Figure 2. Time Series Plots After First Differencing

The temporal order of the GSTARIMA model is based on the ACF and PACF plots from the stationary data. The ACF and PACF plots from the stationary data are given in Figures 3 and 4, respectively.

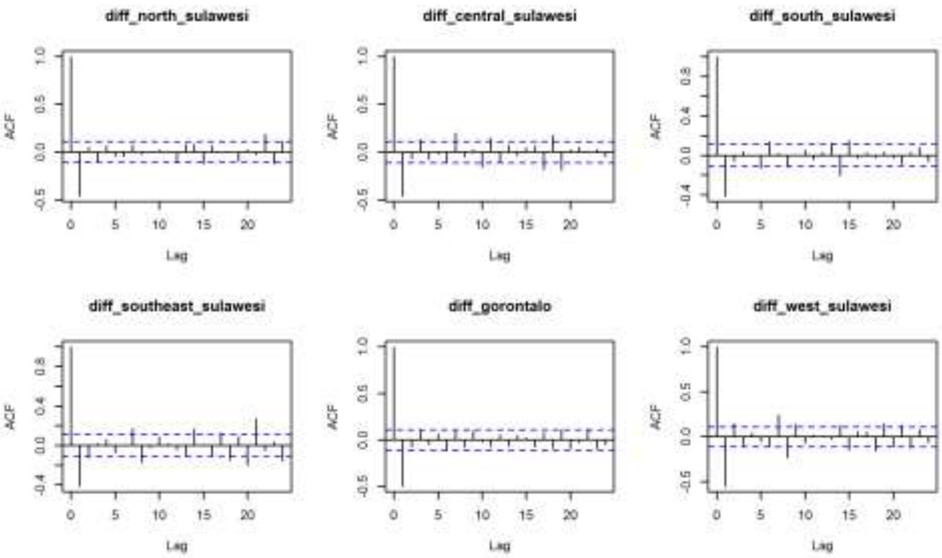

Figure 3. ACF Plots from Stationary Data

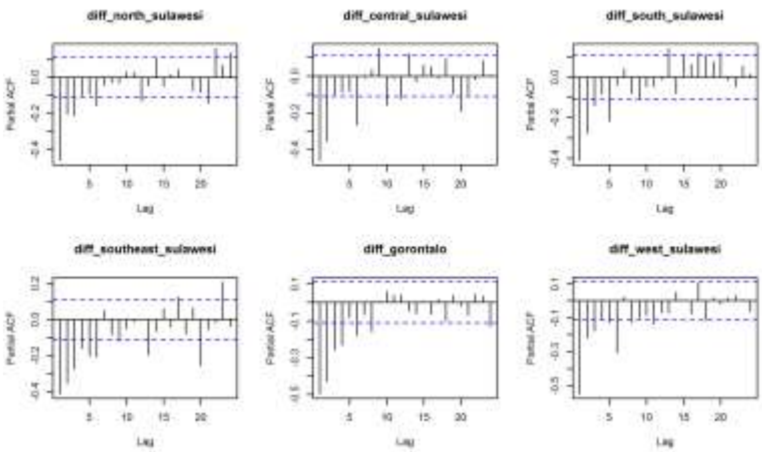

Figure 4. PACF Plots from Stationary Data

The order of autoregressive (AR) and moving average (MA) can be determined based on ACF and PACF plots. Based on Figures 3 and 4, the first autoregressive order AR(1) or the first moving average order MA(1) is preferred. While space order one is chosen as all six locations are within one area on Sulawesi Island (Astuti, Ruchjana, & Soemartini, 2017). As a result, the GSTARIMA (1;1;1;0) or GSTARIMA (1;0;1;1) models are appropriate models for the Covid-19 data.

4.2.3. Inverse Distance Weight

The inverse distance weight is obtained based on the Euclidean distance between locations using the latitude and longitude coordinates of each location. The location points used for each province are the airports of each capital city in each province, namely Sam Ratulangi airport for North Sulawesi province, Mutiara airport for Central Sulawesi province, Sultan Hasanuddin airport for South Sulawesi province, Haluoleo airport for the province Southeast Sulawesi, Djalaluddin airport for Gorontalo province, and Tampa Padang airport for West Sulawesi province. The latitude and longitude coordinate values of each province in degree units are available from https://maps.google.go.id. By using latitude and longitude coordinates, the inverse distance weight matrix can be seen as follows:

$$\boldsymbol{W} = \begin{bmatrix} 0 & 0.1723 & 0.1130 & 0.1565 & 0.4244 & 0.1338 \\ 0.1194 & 0 & 0.1600 & 0.1656 & 0.2009 & 0.3541 \\ 0.0952 & 0.1946 & 0 & 0.2675 & 0.1231 & 0.3196 \\ 0.1330 & 0.2030 & 0.2697 & 0 & 0.1729 & 0.2214 \\ 0.3377 & 0.2306 & 0.1162 & 0.1619 & 0 & 0.1535 \\ 0.0906 & 0.3458 & 0.2567 & 0.1764 & 0.1305 & 0 \end{bmatrix} \tag{7}$$

4.2.4. Parameter Estimation

The parameter estimations of the GSTARIMA (1;1;1;0) model, and GSTARIMA (1;0;1;1) model using the Ordinary Least Square (OLS) method given in Tables 5 and 6, respectively.

Table 5. Parameter Estimation for GSTARIMA (1;1;1;0) Model

Parameters	Estimations	*Stad. Errors*	*t-values*	*p-values*
ϕ_{10}^{1}	−0.4615	0.0713	−6.4690	1.26×10^{-10}*
ϕ_{10}^{2}	−0.4509	0.0741	−6.0810	1.44×10^{-9}*
ϕ_{10}^{3}	−0.4176	0.0265	−15.7790	$< 2 \times 10^{-16}$*
ϕ_{10}^{4}	−0.4175	0.0807	−5.1760	2.51×10^{-7}*
ϕ_{10}^{5}	−0.4953	0.0862	−5.7460	1.07×10^{-8}*
ϕ_{10}^{6}	−0.5516	0.0737	−7.4840	1.11×10^{-13}*
ϕ_{11}^{1}	0.0008	0.1282	0.0060	0.9950
ϕ_{11}^{2}	−0.1462	0.1144	−1.2790	0.2010
ϕ_{11}^{3}	−0.0035	0.1444	−0.0240	0.9810
ϕ_{11}^{4}	0.0198	0.0886	0.2240	0.8230
ϕ_{11}^{5}	−0.0537	0.1227	−0.4370	0.6620
ϕ_{11}^{6}	−0.0005	0.0868	−0.0060	0.9950

*indicate statistical significance at the 5%

Table 6. Parameter Estimation for GSTARIMA (1;0;1;1) Model

Parameters	Estimations	Stad. Errors	t-values	p-values
Θ_{10}^{1}	−0.6724	0.0792	−8.4880	$< 2 \times 10^{-16}$*
Θ_{10}^{2}	−0.7024	0.0837	−8.3920	$< 2 \times 10^{-16}$*
Θ_{10}^{3}	−0.6338	0.0287	−22.0890	$< 2 \times 10^{-16}$*

Table 6. Parameter Estimation for GSTARIMA (1;0;1;1) Model

Parameters	Estimations	Stad. Errors	t-values	p-values
Θ_{10}^{4}	−0.8266	0.0964	−8.5730	$< 2 \times 10^{-16}$*
Θ_{10}^{5}	−0.9379	0.1083	−8.6610	$< 2 \times 10^{-16}$*
Θ_{10}^{6}	−0.8424	0.0884	−9.5260	$< 2 \times 10^{-16}$*
Θ_{11}^{1}	0.0027	0.1497	0.0180	0.9850
Θ_{11}^{2}	−0.0535	0.1317	−0.4060	0.6850
Θ_{11}^{3}	−0.1115	0.1601	−0.6960	0.4860
Θ_{11}^{4}	0.0816	0.1005	0.8120	0.4170
Θ_{11}^{5}	−0.0613	0.1353	−0.4530	0.6510
Θ_{11}^{6}	0.0256	0.0968	0.2650	0.7910

*indicate statistical significance at the 5%

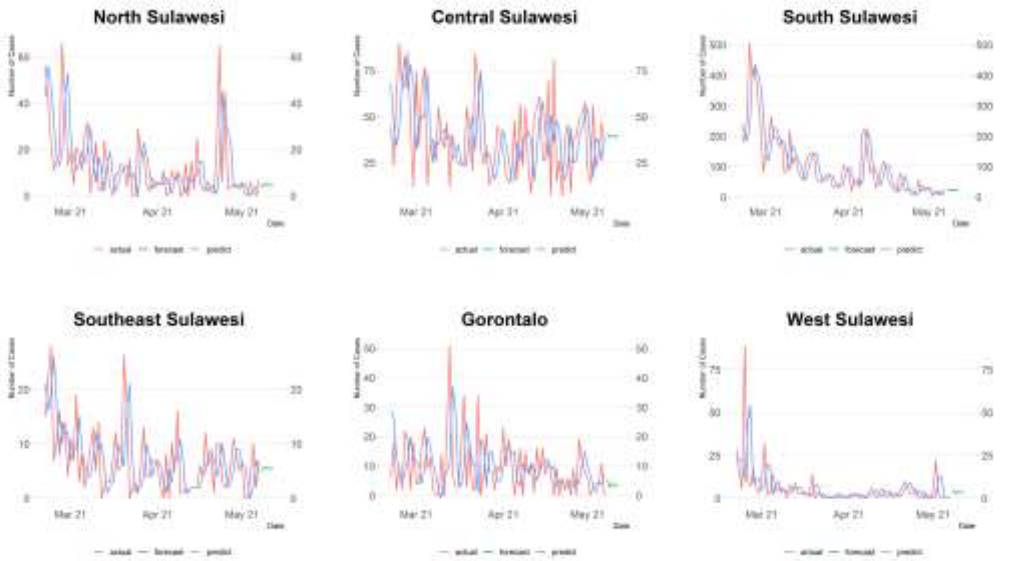

Figure 5. A plot of Actual Data, Prediction, and Forecast GSTARIMA (1;1;1;0) Model

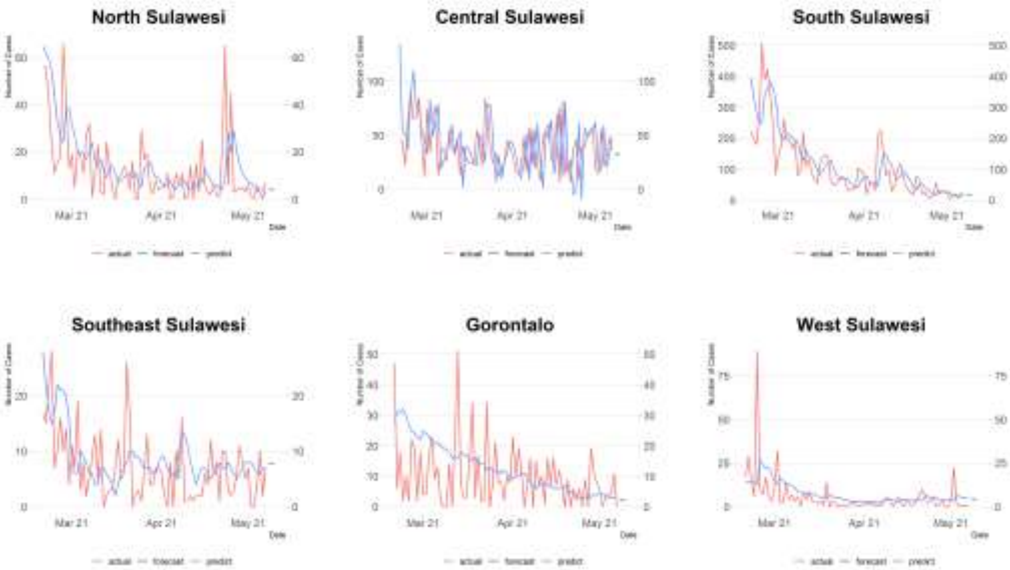

Figure 6. A plot of Actual Data, Prediction, and Forecast of GSTARIMA (1;0;1;1) Model

4.3. Comparison of Actual Data, Prediction Results and Forecasting of GSTARIMA (1;1;1;0) and GSTARIMA (1;0;1;1) Models

The plots of actual data, prediction, and forecast of the GSTARIMA (1;1;1;0) and GSTARIMA (1;1;0;1) models are given in Figures 5 and 6, respectively. Figure 5 depicts that the prediction results follow the pattern of the actual data. Similarly, in Figure 6, the predicted results are around the mean of the actual data.

4.4. The Best Model Selection

The accuracy measures of GSTARIMA (1;1;1;0) and GSTARIMA (1;0;1;1) models used the root mean square error (RMSE) and are given in Table 9.

Table 9. RMSE Value of GSTARIMA (1;1;1;0) and GSTARIMA (1;0;1;1) Models

Location	RMSE	
	GSTARIMA (1;1;1;0)	GSTARIMA (1;0;1;1)
North Sulawesi	14.8966	14.1819
Central Sulawesi	24.8389	31.5988
South Sulawesi	62.8879	66.8529
Southeast Sulawesi	6.4335	6.3063
Gorontalo	11.9153	11.8635
West Sulawesi	12.8801	10.8155

Based on Table 9, Central Sulawesi and South Sulawesi have the smallest RMSE values in the GSTARIMA (1;1;1;0) model. However, North Sulawesi, Southeast Sulawesi, and Gorontalo have relatively the same RMSE values for both models. The RMSE value of the GSTARIMA (1;0;1;1) model for West Sulawesi is smaller than the GSTARIMA (1;1;1;0) model. The map of prediction of the GSTARIMA (1;1;1;0) model is given in Figure 7.

A number of previous studies have explored the space-time autoregressive model with Covid-9 (Alawiyah et al., 2021; Awwad et al., 2021; Pasaribu et al., 2021). For instance, a GSTARI model using a standardized queen contiguity location weight matrix has been discussed (Alawiyah et al., 2021). They used one model namely the GSTARI model to compare the predicted results with the actual data on Covid-19 in four locations in Bandung Raya. One study used the GSTARI model and compared two weighted matrices (the inverse distance weight matrix and modified the inverse distance weight matrix) (Pasaribu et al., 2021). In their model, the spatial weight matrix is based on the distance between public transport (train) (Alawiyah et al., 2021). These studies still have not explored the GSTIMA model (GSTARIMA (1;0;1;1)).

Another study estimated the parameters of three different STARIMA models and selected the best STARIMA model (Awwad et al., 2021). Their research tried to estimate the parameters of the model but had not yet considered the prediction and forecasting stage. Unlike the GSTARIMA model, the STARIMA model assumes to have the same characteristics in each location. Similar to the first study, a standardized queen contiguity location weight matrix has also been used (Awwad et al., 2021). This is certainly different from our study. In our study, we compared two models: the GSTARIMA (1;1;1;0) and GSTARIMA (1;0;1;1) models as an extension of the STARIMA model which allows having different autoregressive parameters in each location. We used the distance between air transportation for the spatial weight matrix.

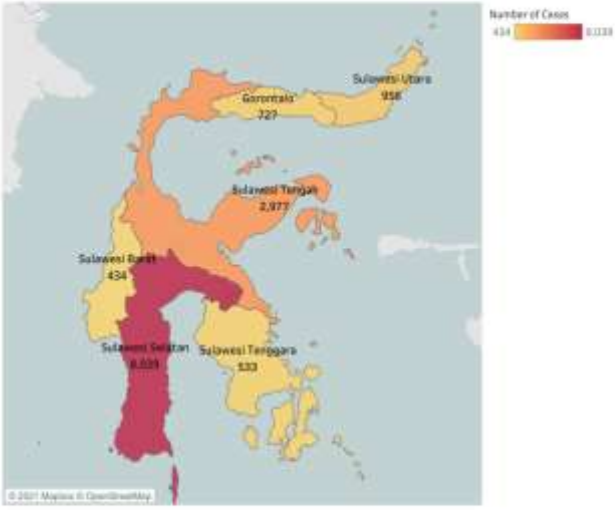

Figure 7. The Map of Prediction Results using GSTARIMA (1;1;1;0) Model from 20 February 2021 to 07 May 2021

Our results showed that the GSTARIMA (1;1;1;0) model is more suitable for two provinces, namely Central Sulawesi and South Sulawesi as the RMSE is smaller than GSTARIMA (1;0;1;1) model. However, for West Sulawesi, GSTARIMA (1;0;1;1) model is more suitable than the GSTARIMA (1;1;1;0) model. The three other provinces: North Sulawesi, Southeast Sulawesi, and Gorontalo have relatively the same RMSE values for both models. Furthermore, the prediction results of the GSTARIMA (1;1;1;0) model in each province have followed the actual time series plot of the data. Likewise, with the GSTARIMA (1;0;1;1) model, the prediction results obtained are around the average value of the actual data.

The map of prediction for the GSTARIMA (1;1;1;0) model for total cases (February 20, 2021 – May 7, 2021) is relatively closer to the actual data than the GSTARIMA (1;0;1;1) model. This may be because the GSTARIMA (1;1;1;0) model relies on data from the previous two days in several locations, while the GSTARIMA (1;0;1;1) model only depends on the residual value generated by the model.

The forecast results for the GSTARIMA (1;1;1;0) model could be used for the next 3 days, while the forecast results for the GSTARIMA (1;0;1;1) model could be used only for the next one day. This indicates that the forecast for the spread of Covid-19 is reliable only for the short term. This result is also in line with research stating that Covid-19 did not have good predictive results if it is predicted in the long term (Alawiyah et al., 2021)) The GSTARIMA model can certainly be applied in other locations and other cases.

5. CONCLUSION

In summary, our results showed that the GSTARIMA model can represent the relationship between time and location so that the estimates from the model can be used to predict and forecast the future. One of the data that can be applied to the GSTARIMA model is the daily data on new cases of Covid-19 in Sulawesi. The forecast for the spread of Covid-19 is reliable only for the short term. In the future, it is highly recommended to estimate the parameters of the GSTARIMA model using different estimation methods such as the Kalman-Filter method.

REFERENCES

Abdullah, A. S., Matoha, S., Lubis, D. A., Falah, A. N., Jaya, I. G. N. M., Hermawan, E., & Ruchjana, B. N. (2018). Implementation of Generalized Space Time Autoregressive (GSTAR)-Kriging Model for Predicting Rainfall Data at Unobserved Locations in West Java. *Applied Mathematics and Information Sciences*, *12*(3), 607–615.

Ahmar, A. S., & Boj, E. (2020). Will COVID-19 Confirmed Cases in The USA Reach 3 Million? A Forecasting Approach by using SutteARIMA Method. *Current Research in Behavioral Sciences*, *1*(September), 100002. Elsevier B.V. Retrieved from https://doi.org/10.1016/j.crbeha.2020.100002

Alawiyah, M., Kusuma, D. A., & Ruchjana, B. N. (2021). Application of Generalized Space Time Autoregressive Integrated (GSTARI) Model in The Phenomenon of Covid-19. *Journal of Physics: Conference Series*, *1722*(1).

Astuti, D., Ruchjana, B. N., & Soemartini. (2017). Generalized Space Time Autoregressive With Exogenous Variable Model and its Application. *Journal of Physics: Conference Series, 893*(1).

Awwad, F. A., Mohamoud, M. A., & Abonazel, M. R. (2021). Estimating COVID-19 Cases in Makkah Region of Saudi Arabia: Space-time ARIMA Modeling. *Plos ONE*, *16*(4 April), 1–16. Retrieved From Http://Dx.Doi.Org/10.1371/Journal.Pone.0250149

Borovkova, S., Lopuhaä, H. P., & Ruchjana, B. N. (2002). Generalized STAR Models With Experimental Weights. *Proceeding Of The 17th International Workshop On Statistical Modelling, Chania, Greece*, 139–147.

Borovkova, S., Lopuhaä, H. P., & Ruchjana, B. N. (2008). Consistency And Asymptotic Normality Of Least Squares Estimators In Generalized STAR Models. *Statistica Neerlandica*, *62*(4), 482–508.

Cliff, A. ., & Ord., J. (1975). Model Building and the Analysis of Spatial Pattern in Human Geography. *Journal of the Royal Statistical Society Series B*, *37*, 297–348.

Djalante, R., Lassa, J., Setiamarga, D., Sudjatma, A., Indrawan, M., Haryanto, B., Mahfud, C., et al. (2020). Review and Analysis of Current Responses to COVID-19 in Indonesia: Period of January to March 2020. *Progress in Disaster Science*, *6*, 100091.

Di Giacinto, V. (2006). A Generalized Space Time ARMA Model with an Application to Regional Unemployment Analysis in Italy. *International Regional Science Review*, *29*, 159–198.

Lee, D., Robertson, C., & Marques, D. (2021). Quantifying The Small-Area Spatio-Temporal Dynamics of The Covid-19 Pandemic in Scotland During A Period with Limited Testing Capacity. *Spatial Statistics*, (xxxx), 100508. Elsevier B.V. Retrieved from https://doi.org/10.1016/j.spasta.2021.100508

Maria, E., Budiman, E., Haviluddin, & Taruk, M. (2020). Measure Distance Locating Nearest Public Facilities using Haversine and Euclidean Methods. *Journal of Physics: Conference Series*, *1450*(1).

Ministry of Health. (2021). COVID 19 in Indonesia. Retrieved May 7, 2021, from https://covid19.go.id

Pasaribu, U. S., Mukhaiyar, U., Huda, N. M., Sari, K. N., & Indratno, S. W. (2021).

Modelling COVID-19 Growth Cases of Provinces in Java Island by Modified Spatial Weight Matrix GSTAR through Railroad Passenger's Mobility. *Heliyon*, *7*(2), e06025. Elsevier Ltd. Retrieved from https://doi.org/10.1016/j.heliyon.2021.e06025

Pfeifer, P. ., & Deutsch, S. . (1980). A Three Stage Iterative Procedure for Space Time Modelling. *Technometrics*, *22*(1), 35–47.

Prillantika, J. R., Apriliani, E., & Wahyuningsih, N. (2018). Comparison between GSTAR and GSTAR-Kalman Filter Models On Inflation Rate Forecasting in East Java. *Journal of Physics: Conference Series*, *974*(1), 0–10.

R Core Team. (2021). R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.r-project.org

Ricardianto, I. P., Djajaputra, D. G., & Martono, P. D. (2017). Air Transport and Tourism in Indonesia. *IOSR Journal of Applied Chemistry*, *10*(05), 01–19.

Setiawan, Suhartono, & Prastuti, M. (2016). S-GSTAR-SUR Model For Seasonal Spatio Temporal Data Forecasting. *Malaysian Journal of Mathematical Sciences*, *10*, 53–65. Universiti Putra Malaysia.

Shereen, M. A., Khan, S., Kazmi, A., Bashir, N., & Siddique, R. (2020). COVID-19 Infection: Origin, Transmission, and Characteristics of Human Coronaviruses. *Journal of Advanced Research*, *24*, 91–98. Cairo University. Retrieved from https://doi.org/10.1016/j.jare.2020.03.005

Sohrabi, C., Alsafi, Z., O'Neill, N., Khan, M., Kerwan, A., Al-Jabir, A., Iosifidis, C., et al. (2020). World Health Organization Declares Global Emergency: A Review Of The 2019 Novel Coronavirus (COVID-19). *International Journal of Surgery*, *76*(February), 71–76. Elsevier. Retrieved from https://doi.org/10.1016/j.ijsu.2020.02.034

Tableau Public. (2020). Tableau Public is a Free Platform to Publicly Share and Explore Data Visualizations Online. Tableau Software, LLC, a Salesforce Company. Retrieved from https://public.tableau.com/

Torrealba-Rodriguez, O., Conde-Gutiérrez, R. A., & Hernández-Javier, A. L. (2020). Modeling and Prediction of COVID-19 in Mexico Applying Mathematical and Computational Models. *Chaos, Solitons and Fractals*, *138*. Elsevier Ltd.

Wang, Q., Dong, W., Yang, K., Ren, Z., Huang, D., Zhang, P., & Wang, J. (2021). Temporal and Spatial Analysis Of COVID-19 Transmission In China And Its Influencing Factors. *International Journal of Infectious Diseases*, *105*, 675–685. International Society for Infectious Diseases. Retrieved from https://doi.org/10.1016/j.ijid.2021.03.014

Wei, W. W. S. (2019). *Multivariate Time Series Analysis and Applications*. United State of America: John Wiley & Sons, Inc.

WHO. (2020). Coronavirus Disease 2019 (COVID-19) World Health Situation Report - 1. *WHO Indonesia Situation Report*, *2019*(March), 1–6.

Zewdie, M. A., Wubit, G. G., & Ayele, A. W. (2018). G-STAR Model for Forecasting Space-Time Variation of Temperature in Northern Ethiopia. *Turkish Journal of Forecasting*, *02*(1), 9–19.