Media Statistika

p-ISSN 1979 – 3693 e-ISSN 2477 – 0647
MEDIA STATISTIKA 14(2) 2021: 146-157
http://ejournal.undip.ac.id/index.php/media_statistika

GSTARI-ARCH MODEL AND APPLICATION ON POSITIVE CONFIRMED DATA FOR COVID-19 IN WEST JAVA

Mutik Alawiyah[1], Dianne Amor Kusuma[2], Budi Nurani Ruchjana[2]
[1]Master of Mathematics Study Program, Faculty of Mathematics and Natural Sciences Universitas Padjadjaran
[2]Department of Mathematics, Faculty of Mathematics and Natural Sciences Universitas Padjadjaran

e-mail: budi.nurani@unpad.ac.id

DOI: 10.14710/medstat.14.2.146-157

Article Info:
Received: 7 February 2021
Accepted: 7 June 2021
Available Online: 11 Januari 2022

Keywords:
ARCH, GSTARI, Positive Covid-19 Confirmed Cases, RMSE
.

Abstract: Time series model that is commonly used is the Box-Jenkins based time series model. Time series data phenomena based on Box-Jenkins can be combined with spatial data, it is called the space time model One model based on Box-Jenkins model with heterogeneous location characteristics is the Generalized Space Time Autoregressive Integrated (GSTARI) model for a model that assumes data is not stationary or has a trend. This paper discusses the development of the GSTARI model with the assumption that the error variance is not constant which is applied to positive data confirmed by Covid-19 in West Java Province, especially in 4 regencies/cities that have cases in the high category from 6 March 2020 until 31 December 2020. Four regencies/cities are Depok City, Bekasi City, Bekasi Regency, and Karawang Regency. Parameter estimation method for the assumption of non-constant error variance can use Autoregressive Conditional Heteroscedasticity (ARCH) method. GSTARI-ARCH modeling procedure followed three Box-Jenkins stages, namely the identification process, parameter estimation and checking diagnostic. Application of the GSTARI-ARCH Model to Covid-19 positive confirmed data in 4 regencies/cities has a minimum value of RMSE in Bekasi City. The plot of forecast results for the four regencies/cities has a similar pattern to the actual data only applicable for a short time for 1-2 days.

1. INTRODUCTION

A series of observations that are time indexed at the same interval is called time series data. The most widely used stationary and univariate time series models are the Autoregressive Moving Average (ARMA) Model. Meanwhile, Vector Autoregressive Moving Average (VARMA) Model is used for stationary and multivariate data. One of the characteristics of the VARMA Model is that it is quite high and shows the interactions between the observed variables. The time series used is a Box-Jenkins based time series. In the Box-Jenkins based time series model, three procedures are developed to model the time series, namely the indication process, parameters and diagnostic tests(Box & Jenkins, 1976).

Time series data can be combined with spatial data that can be combined with one of which is the Space Time Autoregressive (STAR) Model. STAR model assumes that each location has homogeneous data and has stationary data. For non-stationary data, STAR Model is further developed into the Space Time Autoregressive Integrated Moving Average (STARIMA) Model(Pfeifer & Deutsch, 1980).

The location of observing various phenomena that have symbols that are not homogeneous (heterogeneous) so that the Generalized Space Time Autoregressive (GSTAR) Model is developed (Ruchjana, 2002). GSTAR model assumes that each location is heterogeneous. In addition, GSTAR Model also assumes stationary data, but in some cases the phenomenon is confirmed to be positive for Covid-19 in West Java, the data is not stationary. Then to overcome GSTAR Model which has data that the Generalized Space Time Autoregressive Integrated Moving Average (GSTARIMA) model is not stationary was developed(Min et al., 2010).

The GSTAR model assumes that the variables are constant. As a result, in estimating the GSTAR Model, it was initially limited to the Ordinary Least Square (OLS) method. However, the variance error conditions sometimes change over time or the variance is not constant so that the study of the estimation method in the GSTAR Model can still be developed (Borovkova et al., 2008). In overcoming the problem of variance that is not constant, the GSTAR Model develops into the GSTARI-Generalized Least Square (GLS) Model.(Apriliawan et al., 2013; Suhartono et al., 2016), GSTAR-Seemingly Unrelated Regression (SUR) Model(Iriany et al., 2015; Habibie& Perdana, 2020)and GSTAR-Autoregressive Conditional Heteroscedasticity (ARCH) Model(Nainggolan et al., 2010; Bonar et al., 2017).

The space time phenomena that can be observed is a confirmed positive case of Covid-19 because it can be observed based on space and time simultanously. The Covid-19 phenomenon has previously been researched by Sifriyani by using the Susceptible Infected Recovered (SIR) model to estimate the Covid-19 Reproduction Rate (Sifriyani & Rosadi, 2020). Covid-19 has a high transmission rate and can be transmitted through close contact with a person infected with Covid-19 (Wang et al., 2020). The spread of Covid-19 from person to person happened quickly, both between locations and times (Fanelli & Piazza, 2020; Roosa et al., 2020; Shereen et al., 2020).

Positive confirmed cases in West Java often show a pattern of data that is assumed to be non-stationary due to fluctuations in mean and variance, so that the data is assumed to have a mean and variance that is not constant. If a positive positive result is confirmed not stationary in mean then we can use the differencing process. Furthermore, if it is not stationary with respect to variance, then ARCH is used. Data on 27 regencies/cities in West Java have a similar data pattern which has a non-constant variance. In the study, 4 regencies/cities in the nearest environment were selected and had a fairly high data pattern, but with the assumption that each location had different parameters or the location was heterogeneous. These locations are Bekasi Regency, Bekasi City, Depok City and Karawang Regency. The four locations were used as case data for the application of the GSTAR-ARCH model to data confirmed positive for Covid-19 in West Java.

2. LITERATURE REVIEW

2.1. GSTARI Model

GSTARI model is a development of GSTAR Model by Ruchjana (2002) based on STAR Model and STARI Model (Pfeifer & Deutsch, 1980).In various applications, STARI Model can be applied to the water discharge of the Citarum river (Alawiyah et al, 2020). Meanwhile, GSTARI Model with OLS estimation can be applied to data confirmed positive for Covid-19 in the area of Greater Bandung (Alawiyah et al, 2021). GSTAR model assumes that each location is heterogeneous (Ruchjana, 2002). The general form of GSTARI Model $(p; d; \lambda_1; \lambda_2; \lambda_3, \ldots, \lambda_k)$ is as follows:

$$\boldsymbol{y}(t) = \sum_{k=l}^{p} \sum_{l=0}^{\lambda_k} \left[\boldsymbol{\Phi}_{kl} \mathbf{W}^{(l)} \mathbf{y}(t-k) \right] + \mathbf{e}(t) \tag{1}$$

where $\mathbf{y}(t) = \mathbf{z}(t) - \mathbf{z}(t-1), \mathbf{y}(t-1) = \mathbf{z}(t-1) - \mathbf{z}(t-2)$ and assumption $\mathbf{e}(t) \sim iid\ N(\mathbf{0}, \sigma^2 \mathbf{I})$.

2.2. GSTARI-ARCH Model

STAR model, GSTAR model, and GSTARI model assume a constant error variance over time (homoscedastic), but usually, data phenomena that have error variance are not constant. The STAR model, GSTAR model, and GSTARI model can be developed using the GSTARI-ARCH Model with the assumption that the variance is not constant(heteroscedastic) (Nainggolan et al., 2010). GSTARI $(p; d; \lambda_1; \lambda_2; \lambda_3, \ldots, \lambda_k)$-ARCH (p) Model can be written as follows (Bonar et al., 2017):

$$\boldsymbol{y}(t) = \sum_{k=l}^{p} \sum_{l=0}^{\lambda_k} \left[\boldsymbol{\Phi}_{kl} \mathbf{W}^{(l)} \mathbf{y}(t-k) \right] + \mathbf{e}(t) \tag{2}$$

where $\mathbf{e}(t) = \mathbf{D}_t \boldsymbol{\eta}_t$, $\mathbf{y}(t) = \mathbf{z}(t) - \mathbf{z}(t-1), \mathbf{y}(t-1) = \mathbf{z}(t-1) - \mathbf{z}(t-2)$and assumption $\mathbf{e}(t) \sim iid\ N(\mathbf{0}, \boldsymbol{\Sigma}_t)$. $\boldsymbol{\Sigma}_t$ is covariance matrix of error vectors $\mathbf{e}(t)$ with conditional Z_{t-1}.

Estimation method of GSTARI-ARCH model consists of two stages, namely the Maximum Likelihood (MLE) method which is used to estimate variance equation parameters and Generalized Least Square (GLS) method which is used to estimate mean parameter equation(Nainggolan *et al.*, 2010).

2.3. Weight Matrix using Inverse Distance

The weight matrix based on the calculation of the actual distance between locations is the weighted distance matrix. Locations that are close to the weight reverse the distance of the greater weight value. The calculation of the inverse weight for the location is as follows:

$$w_{ij} = \frac{1}{d_{ij}} \tag{3}$$

where w_{ij} is the inverse weight matrix element of the distance of location i and j and is the distance of location i and location j. If there are 4 observation locations, then inverse distance weight matrix d_{ij} is:

$$\mathbf{W} = [w_{ij}] = \begin{bmatrix} 0 & w_{12} & w_{13} & w_{14} \\ w_{21} & 0 & w_{23} & w_{24} \\ w_{31} & w_{32} & 0 & w_{34} \\ w_{41} & w_{42} & w_{43} & 0 \end{bmatrix} \tag{4}$$

2.4. Estimation Procedure of GSTARI-ARCH Model

GSTARI-ARCH model estimation method consists of two stages, namely the variance equation parameter estimated by Maximum Likelihood (MLE) method and GSTARI mean equation parameter estimated by Generalized Least Square (GLS) method (Nainggolan *et al.*, 2010). The following is the method of estimating the variance equation and the mean equation:

a. MLE method of GSTARI-ARCH Model

For example, given a vector of observation data $\mathbf{z}(0), \mathbf{z}(1), \ldots, \mathbf{z}(T)$ which consists of $(T-1)$ observations with N locations, and F_t is the set of information available at t time. The ARCH regression model at location$-i$ with data $t = 1,2,\ldots,T$ is stated as follows:

$$\mathbf{y}_i(t) = \mathbf{X}_i^T(t)\boldsymbol{\beta}_i + \mathbf{e}_i(t) = \phi_{0i}\mathbf{z}(t-1) + \phi_{1i}\mathbf{V}_i(t-1) + \mathbf{e}_i(t) \quad (5)$$

$$\mathbf{e}(t) = \sqrt{h_i(t)}\eta_i(t) \quad (6)$$

$$h_i(t) = \alpha_{0i} + \alpha_{1i}\mathbf{e}_i^2(t-1) \quad (7)$$

The ARCH-Regression Model assumes that η_t is a variable that controls the independent properties of F_{t-1} and $\eta_t \sim iid\, N(0,1)$ so that the population density of $\mathbf{y}_t$conditional F_{t-1} is:

$$f(\mathbf{y}_t|F_{t-1}) = \frac{1}{\sqrt{2\pi h_t}}\exp\left(\frac{-(\mathrm{y}_t - \mathbf{X}_t^T\boldsymbol{\beta})^2}{2h_t}\right) \quad (8)$$

Futhermore, we used MLE method procedure in equation (8) until the covariance matrix $\boldsymbol{\Sigma}$ is obtained. The average regression parameters were estimated using the GLS method.

b. GLS Method of GSTARI-ARCH Model

The GLS method is a disaster method that can solve problems that arise between errors in various equations (locations) (Zellner, 1962). The following is a linear equation model:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e} \quad (9)$$

with assumption $\mathbf{e}|\mathbf{X} \sim (0, \sigma^2\mathbf{U})$ and $\mathbf{U} = \begin{bmatrix} u_1 & \cdots & \cdots \\ \vdots & \ddots & \vdots \\ \cdots & \cdots & u_T \end{bmatrix}$. Furthermore, the transformation of the linear model in equation (9) is carried out to produce:

$$\widehat{\boldsymbol{\beta}} = (\mathbf{X}^T\mathbf{L}^T\mathbf{L}\mathbf{X})^{-1}\mathbf{X}^T\mathbf{L}^T\mathbf{L}\mathbf{y} = (\mathbf{X}^T\mathbf{U}^{-1}\mathbf{X})^{-1}\mathbf{X}^T\mathbf{U}^{-1}\mathbf{y} \quad (10)$$

where $\mathbf{L}$ is the transformation matrix.

2.5. Root Mean Square Error (RMSE)

The criteria for selecting the best model can be seen from the value of Root Mean Square Error (RMSE). The RMSE formula for each can be defined as follows (Wei, 2006):

$$RMSE = \sqrt{\frac{\sum_{i=1}^{T} e_i^2}{T}} \quad (11)$$

where,

$e_i = Z_i - \hat{Z}_i$

T : the number of sample

Z_i : actual data in the outsample data
$\hat{Z}_i$: forecast data

2.6 Covid-19 Phenomenon in West Java

One of the infectious diseases caused by Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-Cov-2) is Covid-19. The Coronavirus is transmitted from human to human which occurs in the surrounding environment which is closely related to patients or carriers of Covid-19. All provinces in Indonesia have been affected by Covid-19, including in West Java. The curves of the Covid-19 cases in West Java occur in fluctuations so that the data is not stationary. In addition, seen from the variance of error, the daily positive confirm case data for Covid-19 has a tendency to the variance that is not constant. Then, the daily positive confirmed case data for Covid-19 in each location was observed using queen contiguity which can observe locations for the entire region for the case study, we choose 4 regencies city from 27 districts with criteria as a red zone which has an average case study highest confirmed positive and queen contiguity. The four regencies/cities are Depok City, Bekasi City, Bekasi Regency, and Karawang Regency which can be illustrated in Figure 1 below:

Source : https://id.wikipedia.org/wiki/Daftar_kabupaten_dan_kota_di_Jawa_Barat

Figure 1. Map Confirmed Positive Covid-19 in West Java

3. RESEARCH METHOD

In this research, the development of a Space Time Model in the form of the GSTARI-ARCH Model is carried out with the assumption that the error variance is not constant. The case study used daily data on the number of people who were confirmed positive for Covid-19 in 4 regencies/cities in West Java from March 6, 2020 to December 31, 2020. Of the 27 regencies/cities in West Java, the four regencies/cities are in the red zone and based on queen contiguity, the four regencies/cities are contiguous. The data was obtained from the website https://pikobar.jabarprov.go.id/data. The research steps are as following:

1. Determine for descriptive statistics and make plot data.
2. Determine the distance inverse weight matrix.
3. Stationary data goes through a differentiation process.
4. Modeling errors with the ARCH approach.
5. The parameter estimation of the GSTARI Model uses the ARCH method.
6. Checking diagnostic on the GSTARI-ARCH Model.
7. Forecasting the observation data of the GSTARI-ARCH model for 7 days.

4. RESULT AND DISUCSSION

4.1. Descriptive Statistics of Data Confirmed Positive Covid-19

Data were confirmed positive in West Java in a study consisting of 4 regencies/cities, namely Depok City, Bekasi City, Bekasi Regency and Karawang Regency. The following are descriptive statistics for positive confirmed data for Covid-19, which are presented in Table 1.

Table 1 Descriptive Statistics of Data Confirmed Positive Covid-19

Regencies/Cities	Mean	Min	Max	Var	St. Dev
Depok City	41	0	388	4,409	388
Bekasi City	47	0	259	5,871	678
Bekasi Regency	28	0	277	1,969	277
Karawang Regency	19	0	135	1,268	218

The descriptive statistics in Table 1 explain that the variance of each district / city has different variances around 1,268 to 5,871. Furthermore, you can see the time series plot used to see the phenomenon of positive data for Covid-19.

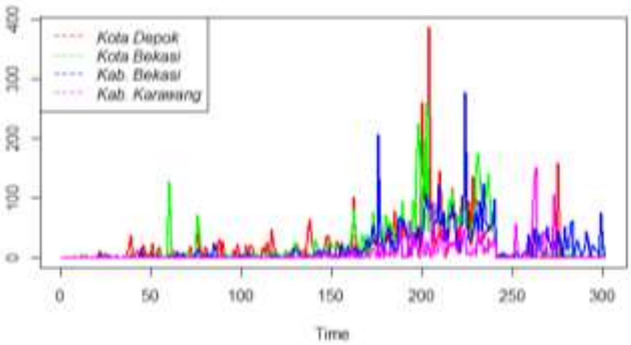


Figure 2 Time Series Data Confirmed Positive Covid-19

Time series plot in Figure 2 shows that on days 0 to 150 there is no fluctuation, but on days 150 to 301 the data fluctuates significantly. Data fluctuation in four cities and districts in West Java has a similar pattern. Furthermore, it is seen from the display between variables to see the positive relationship data confirmed by Covid-19 in each regencies/city. The values between variables for each regencies/city are presented in Table 2.

Table 2 Correlation Data between Variables Confirmed Positive Covid-19 at 4 Locations

	Depok City	Bekasi City	Bekasi Regency	Karawang Regency
Depok City	1	0.520	0.494	0.234
Bekasi City	0.520	1	0.592	0.256
Bekasi Regency	0.494	0.592	1	0.311
Karawang Regency	0.234	0.256	0.311	1

The correlation between variables based on Table 2 shows that each regency/city has a fairly high appearance. This means that there is an attachment between 4 regencies/cities in positive cases of Covid-19. The four regencies/cities in West Java are assumed to be in 1 spatial lag and the weight is determined based on the distance between locations. The following is the inverse distance weight matrix according to equation (4):

$$\mathbf{W}_{ij} = \begin{bmatrix} 0 & 0.4755 & 0.2921 & 0.2324 \\ 0.1088 & 0 & 0.7798 & 0.1114 \\ 0.0651 & 0.7596 & 0 & 0.1753 \\ 0.1543 & 0.3233 & 0.5223 & 0 \end{bmatrix} \quad (13)$$

4.2. Identification of GSTARI-ARCH Model

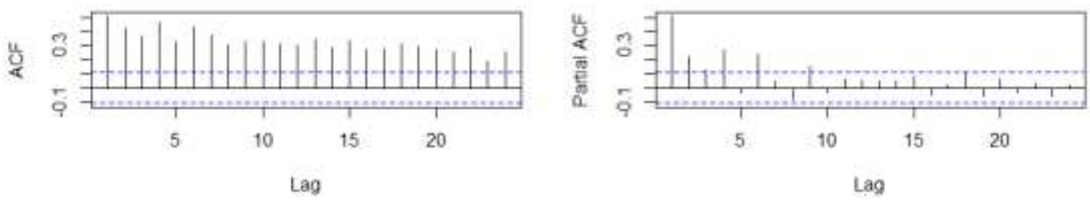
Stationary data is performed to identify whether the data is stationary or not. The data stationary test process can be done with the Augmented Dickey Fuller (ADF) test. The ADF test results for data confirmed positive for Covid-19 are presented in Table 3 below using Rstudio.

Table 3. Stationary Data Confirmed Positive for Covid-19

Regencies/Cities	Before *Differencing*		After*Differencing*	
	p-value	Conclusion	p-value	Conclusion
Depok City	0.0889	Not Stationary	0.01	Stationary
Bekasi City	0.0768	Not Stationary	0.01	Stationary
Bekasi Regency	0.2604	Not Stationary	0.01	Stationary
Karawang Regency	0.2142	Not Stationary	0.01	Stationary

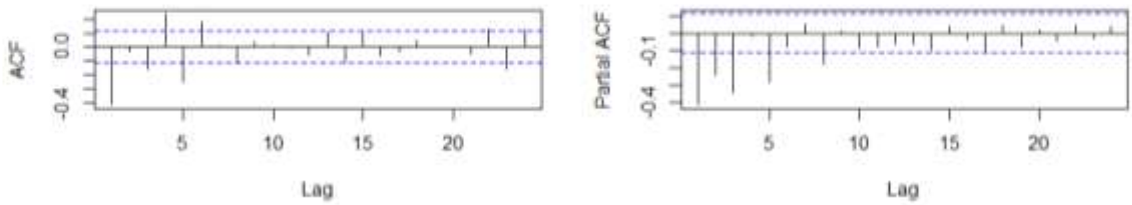
Based on Table 3, the data is confirmed to be positive for Covid-19, the four regencies/cities have a p-value of more than $\alpha = 5\% = 0.05$. If the p-value is more than 0.05 then H_0 is rejected, which means the data is not stationary. In connection with non-stationary data in four regencies/cities, it is necessary to do the differencing process once. Results obtained from completely different processes yielded a p-value of less than 0.05. This means that the four regencies/cities have stationary (Wei, 2019).

Identification process can be use ACF or PACF plots at each location. As an example, the ACF and PACF plots of Depok City are shown for the data before and after differentiation 1 time which is depicted in Figures 3a and 3b.



ACF Plot PACF Plot

Figure 3.a Plot of ACF and PACF Data for Covid-19 in Depok City Before Differentiation



ACF Plot PACF Plot

Figure 3.b Depok City ACF and PACF Covid-19 Plot Data After Differentiation

Based on the ACF and PACF plots for each city from Figure 3a and Figure 3b, it can be determined the order of the univariate time series model to be used in the space time model. Using the same method, it is also suitable for the selection of the ACF and PACF orders for the three districts / cities. In general, it appears that the ACF plots of the four cities are cut off at lag 1, which indicates the formation of the Moving Average (1) model. However, by using the invertible Moving Average (MA) it can be expressed as Autoregressive (AR). Furthermore, the AR model is carried out by a parsimony process so that the AR model is selected and the differencing process is carried out once so that the AR (1)model turns into an ARI (1,1) model. The ARI (1,1) model is combined with the location element and weight matrix to form the GSTARI (1,1,1) model. Assuming that the variance is not constant, the ARCH estimator order 1 is used so that the GSTARI (1,1,1) model changes to the GSTARI(1,1,1) –ARCH(1) model.

4.3. GSTARI-ARCH Model on Data Confirmed Positive for Covid-19

Initially, MLE method was used to estimate the model error conditional variance and the GLS method to estimate the model mean GSTARI-ARCH parameter. Before the parameter estimation process is carried out, the inverse weight matrix is first calculated in equation (13).

Furthermore, the parameter estimation of the GSTARI-ARCH Model is carried out. The results of parameter estimation with the GSTARI-ARCH Model are presented in Table 4 below using R:

Table 4. Parameter Estimator Value of GSTARI (1,1,1) –ARCH(1) Model

Parameter	Coefficient
$\hat{\phi}_{10}^{1}$	-0.2281
$\hat{\phi}_{10}^{2}$	-0.2801
$\hat{\phi}_{10}^{3}$	-0.2820
$\hat{\phi}_{10}^{4}$	-0.5285
$\hat{\phi}_{11}^{1}$	0.8560
$\hat{\phi}_{11}^{2}$	0.2280
$\hat{\phi}_{11}^{3}$	-0.0001
$\hat{\phi}_{11}^{4}$	-0.1890

The equation for GSTARI -ARCH model for each districts/cities can be obtained as follows:

1. GSTARI-ARCHModel for Depok City

$$\hat{Z}_1 = 0.7719Z_1(t-1) + 0.2281Z_1(t-2) + 0.9660Z_2(t-1) - 0.9660Z_2(t-2)$$
$$+0.5934Z_3(t-1) - 0.5934Z_3(t-2) + 0.4721Z_4(t-1) - 0.4721Z_4(t-2)$$

2. GSTARI-ARCHModel for Bekasi City

$$\hat{Z}_2 = 0.0219Z_1(t-1) - 0.0219Z_1(t-2) + 0.7199Z_2(t-1) + 0.2801Z_2(t-2)$$
$$+0.1571Z_3(t-1) - 0.1571Z_3(t-2) + 0.0230Z_4(t-1) - 0.0230Z_4(t-2)$$

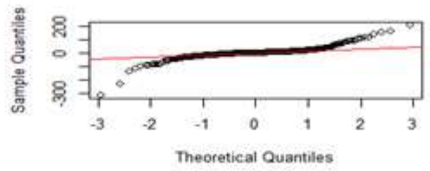
3. GSTARI-ARCH Model for Bekasi Regency

$$\hat{Z}_3 = 0.0061Z_1(t-1) - 0.0061Z_1(t-2) + 0.0716Z_2(t-1) - 0.0716Z_2(t-2) + 0.7180Z_3(t-1) + 0.2820Z_3(t-2) + 0.0165Z_4(t-1) - 0.0165Z_4(t-2)$$

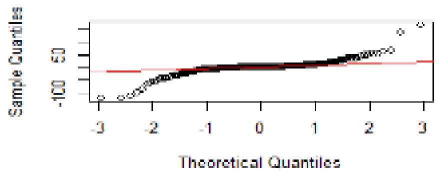
4. GSTARI-ARCHModel for Karawang Regency

$$\hat{Z}_4 = -0.0301Z_1(t-1) + 0.0301Z_1(t-2) - 0.0631Z_2(t-1) + 0.0631Z_2(t-2) - 0.1020Z_3(t-1) + 0.1020Z_3(t-2) + 0.4715Z_4(t-1) - 0.5285Z_4(t-2)$$

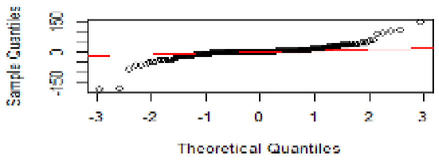
Checking diagnostic of the model, then the assumptions of the model assumptions have been met or not. The results from the Q-Q plot of each city show that the forecast service error results approach the normal distribution. The results of the error test using the Ljung-Box test have a p-value of more than α=5%=0.05. This means that the four regencies/cities have an uncorrelated error. The results of the Q-Q plot can be described in the four regencies/cities as illustrated in Figure 4 below.



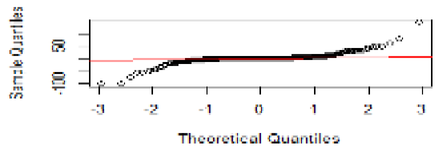
Depok City



Bekasi City



Bekasi Regency

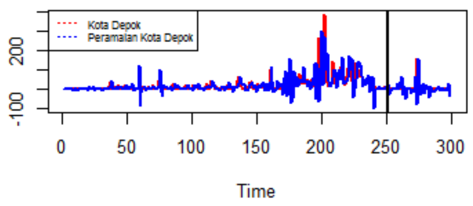


Karawang Regency

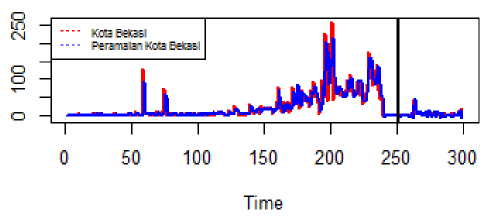
Figure 4. Q-Q Plot of GSTARI-ARCH Model

4.4. Forecasting of GSTARI(1,1,1)-ARCH(1) Model

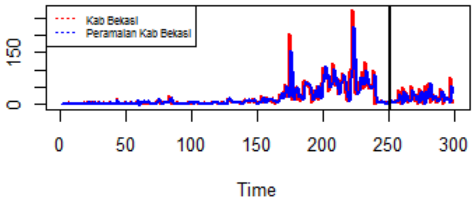
The GSTARI-ARCH equation model that has been obtained is then calculated the forecast value in each regency/city. The plot of forecasting results and actual data for each related regency/city is in Figure 5 below



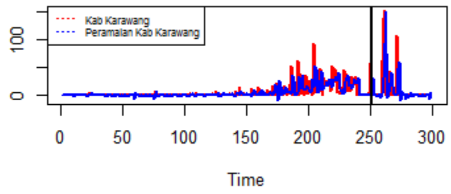
Depok City



Bekasi City



Bekasi Regency



Karawang Regency

Figure 5. Forecasting and Actual Plots of the GSTARI-ARCH Model

Figure 5 shows that the forecasted plots are similar to the actual data only for a short period of 1-2 days. Therefore, the next prediction to calculate the GSTARI-ARCH error model is carried out on the 7-day out sample data. This is supported by the plot of the forecast in a long time it shifts further away from the actual data. This causes the RMSE value of 301 observational data if taken from outside the sample 51 results in a very large error. Therefore, for the calculation of RMSE, an out sample data of 7 days or weeks was selected. The RMSE value data confirmed positive for Covid-19 is presented in Table 5 below:

Table 5. RMSE of Confirmed Positive Covid-19 Data

Regencies/Cities	Model GSTARI-ARCH
Depok City	16.22617
Bekasi City	1.24356
Bekasi Regency	4.47807
Karawang Regency	25.64952

RMSE value obtained in Table 5 shows that Bekasi City has a minimum RMSE value. Meanwhile, Bekasi Regency also still has a relatively small RMSE compared to Depok City and Karawang Regency. Even so, forecasting using the GSTARI-ARCH model for confirmed cases of Covid-19 in 4 districts can only be done for a short term, namely one to two days. This is because the data used is only data confirmed positive for Covid-19, while many external factors can affect forecasting positive Covid-19 data in West Java.Furthermore, the GSTARI-ARCH model can be developed by adding external factors such as the population density of a regencies/cities, the number of referral hospitals, the number of personal protective equipment, the level of population mobility and so on as exogenous variables, so the model is called GSTARIX-ARCH.

5. CONCLUSION

The GSTARI-ARCH model is a development of the space-time model for data that is not stationary or there is a trend with non-constant error variants. The GSTARI-ARCH modeling procedure follows 3 stages of Box-Jenkins, namely the identification model, parameter estimation and checking diagnostic. The results of the calculation of the total RMSE GSTARI-ARCH model for 4 regencies/cities simultaneously give a large enough value, but it can show that the four regencies/cities indicate that Bekasi City has a minimum RMSE. In total, the GSTARI-ARCH model for 4 regencies/cities shows that the prediction of the number of confirmed cases of Covid-19 can follow the actual data pattern, but only for the short term for 1-2 days. Further, these results may be recommended to the government and Satgas Covid-19 in West Java.

6. ACKNOWLEDGMENT

The author thanks the Rector of Universitas Padjadjaran who has provided fgrant for research dissemination for lecturer and student research through the Academic Leadership Grant with contract number 1959/UN6.3.1/PT.00/2021. The author also thanks to the reviewers for valuable comment to improve this paper.

REFERENCES

Alawiyah, M., Kusuma, D. A. and Ruchjana, B. N. (2020)Model Space Time Autoregressive Integrated (Stari(2,1,1)) Untuk Data Debit Air Sungai Citarum Di Provinsi Jawa Barat, *Barekeng: Jurnal Ilmu Matematika dan Terapan*, 14(1), pp. 147–158. doi: 10.30598/barekengvol14iss1pp147-158.

Alawiyah, M., Kusuma, D. A. and 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, p. 012035. doi: 10.1088/1742-6596/1722/1/012035.

Apriliawan, D., Tarno and Yasin, H. (2013)Pemodelan Laju Inflasi Di Provinsi Jawa Tengah Menggunakan Regresi Data Panel, *Jurnal Gaussian*, 2(18), p. 301–321.

Bonar, H., Nurani, B. N. and Darmawan, G. (2017) Development of Generalized Space Time Autoregressive Integrated with ARCH Error (GSTARI–ARCH) Model Based on Consumer Price Index Phenomenon at Several Cities in North Sumatera Province, *AIP Conference Proceedings*, 1827(1).

Borovkova, S., Lopuhaä, H. P. and Ruchjana, B. N. (2008) Consistency and Asymptotic Normality of Least Squares Estimators in Generalized STAR Models, *Statistica Neerlandica*, 62(4), p. 482–508.

Box, G. E. P. and Jenkins, G. M. (1976) Time Series Analysis: Forecasting and Control. 1st Edn'. Holden-Day, San Fransisco.

Fanelli, D. and Piazza, F. (2020) Analysis and Forecast of COVID-19 Spreading in China, Italy and France, *Chaos, Solitons and Fractals*, 134, p. 109761. doi: 10.1016/j.chaos.2020.109761.

Habibie, G. and Perdana, H. (2020) Penerapan GSTAR_SURpada Jumlah Penumpang, 09(2), p. 275–284.

Iriany, A. *et al.* (2015) GSTAR-SUR Model for Rainfall Forecasting in Tengger Region, East Java, *Int. Conf. Pure Appl. Res. Univ. Muhammadiyah Malang*, (1), p. 1–8.

Min, X., Hu, J. and Zhang, Z. (2010) Urban Traffic Network Modeling and Short-Term Traffic Flow Forecasting Based on GSTARIMA model, *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*, p. 1535–1540. doi: 10.1109/ITSC.2010.5625123.

Nainggolan, N. *et al.* (2010) GSTAR Models with ARCH Errors and the Simulation, (ICMNS), p. 1075–1084.

Pfeifer, P. E. and Deutsch, S. J. (1980) A Three-Stage Iterative Procedure for Space-Time Modeling, *Technometrics*, 22(1), p. 35–47.

Roosa, K. *et al.* (2020) Real-time Forecasts of the COVID-19 Epidemic in China from

February 5th to February 24th, 2020, *Infectious Disease Modelling*, 5, p. 256–263. doi: 10.1016/j.idm.2020.02.002.

Ruchjana, B. (2002) *A Generalized Space-Time Autoregressive Model and Its Application to Oil Production*.Ph.D. Dissertation, Unpublished, Bandung: ITB.

Shereen, M. A. *et al.* (2020) COVID-19 Infection: Origin, Transmission, and Characteristics of Human Coronaviruses, *Journal of Advanced Research*, 24, p. 91–98. doi: 10.1016/j.jare.2020.03.005.

Sifriyani, S. and Rosadi, D. (2020), Susceptible Infected Recovered (SIR) Model for Estimating Covid-19 Reproduction Number in East Kalimantan and Samarinda, Media Statistika, vol. 13, no. 2, p. 170-181.

Suhartono *et al.* (2016) GSTARX-GLS Model for Spatio-Temporal Data Forecasting, *Malaysian Journal of Mathematical Sciences*, 10(January), p. 91–103.

Wang, L. S. *et al.* (2020) A Review of the 2019 Novel Coronavirus (COVID-19) based on Current Evidence, *International Journal of Antimicrobial Agents*, 55(6), p. 105948. doi: 10.1016/j.ijantimicag.2020.105948.

Wei, W. W. S. (2006) Time Series Analysis: Univariate and Multivariate Methods (2nd Edition), in *Pearson: Addison Wesley*. doi: 10.1017/cbo9780511802447.020.

Wei, W. W. S. (2019) *Multivariate Time Series Analysis and Applications*. John Wiley & Son Ltd. Available at: http://repositorio.unan.edu.ni/2986/1/5624.pdf.

Zellner, A. (1962) An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias, *Journal of the American Statistical Association*, 57(298), p. 348–368. doi: 10.1080/01621459.1962.10480664.