



The Generalized STAR Modeling with Heteroscedastic Effects

Utriweni Mukhaiyar^{1,*}, Syahri Ramadhani²

¹Statistics Research Division, Faculty of Mathematics and Natural Sciences, Institut Teknologi Bandung, Indonesia

²Undergraduate Programme in Mathematics, Faculty of Mathematics and Natural Sciences, Institut Teknologi Bandung, Indonesia

*Corresponding Author

Email: utriweni@math.itb.ac.id

ABSTRACT

Most of the Generalized Space Time Autoregressive (GSTAR) models assume the constant error variance. In fact, there are many space-time observations whose variability is changing over the times. In this study, a GSTAR model is built with an error variance that is not constant or has a heteroscedasticity effect, namely the combination of GSTAR–Autoregressive Conditional Heteroscedasticity (ARCH). The parameters of the GSTAR–ARCH model are estimated using the Generalized Least Square (GLS) method to obtain the efficient parameter estimation. As a case study, the GSTAR–ARCH model is applied to the daily mean wind speed data of New Orleans, Florida and Mississippi, in order to predict the occurrence of Hurricane Katrina that occurred in 2005. It is obtained that the heteroscedastic involvement in GSTAR modeling gives the better results in predictions, compared to the homoscedastics approach. Furthermore, as the order of model is higher, the GSTAR model performances is better, which is shown by the least Mean Squared Errors (MSE) and Mean Absolute Percentage Error (MAPE). The obtained results show that the GSTAR model (3;0,0,1)–ARCH(1) predicts the Hurricane Katrina better than the GSTAR(3;0,0,1) and GSTAR(1;1)–ARCH(1) models.

Keywords: GSTAR; ARCH; Conditional Variance; Generalized Least Squares; Heteroskedasticity

INTRODUCTION

A hurricane is a natural phenomenon in the form of wind gusts with a speed exceeding 119 km/hour. Hurricanes are a type of tropical cyclone that usually forms on warm sea surfaces around the equator. The wind speed at one location is influenced by the wind speed of the previous time at that location and also influenced by the average wind speed at other locations. This means that the wind speed can be modeled with space-time models such as STARMA(p,q), STAR(p), STMA(q), GSTAR(p,q) and STARMAG(p,q). The model used in this study is the GSTAR model which uses a certain weight matrix according to the location conditions.

The arrival of a storm which is usually predicted by weather satellites will be predicted using the GSTAR model. By using the daily average wind speed data one year before the storm, it is hoped that the arrival of the storm can be predicted earlier and reduce the number of victims, both property and life, due to the storm. However, a large increase in wind speed when a storm occur, causes the data to have heteroscedastic effect, so that the error generated by the GSTAR model has a non-constant variance. It makes the estimation of the initial model parameters being no longer efficient, thus a model that explains the variance of the error is not constant, namely the ARCH model is should be developed. The use of the ARCH is to model the variance of errors, such that it is expected to eliminate heteroscedastic elements and more efficient parameters can be produced.

The development of GSTAR model in Indonesia is very fast, both theoretically and in application. Theoretically, it includes the stationary properties of the process using the Inverse Autocovariance Matrix [1] as well as the kernel approach [2, 3] and minimum spanning tree approach [4], GSTAR with correlated errors [5, 6], GSTAR with outliers [7], GSTAR for discrete data [8], and invertibility of kernel GSTAR model [9]. The application of the GSTAR model has been carried out on economic data [10], tea plantations [11], palm oil production [12], red chili commodity prices [13], number of dengue fever cases [14], predictions of robbery cases in Medan, North Sumatra [15], the spread of Covid-19 cases in Java [16], and copper and gold grades vertical distribution [17]. On the other hand, the ARCH model, which accommodates the element of heteroscedasticity and exogenous variable which make the high volatility of process, is widely developed in economic problems. The impact of Covid-19 as the exogenous factor to the economic sector be explored by [18]. Sometimes the exogeneous factor cause a point of change happen and it should be detected [19]. However, its application has also been carried out to predict electric current [20], caterpillar pests in oil palm plantations [21], and rice prices [22].

The development of the GSTAR model with an error variance by considering the heteroscedasticity effect, has been investigated by [23] on the GSTAR(1,1) model with application to stock prices. The construction of GSTAR model with the ARCH effect and estimate the parameters using the maximum likelihood method approach be explored by [24]. In this study, the GSTAR-ARCH model was developed by estimating the parameters using the Generalized Least Square (GLS) approach.

METHODS

Generalized STAR Model

The GSTAR model is a generalization of the STAR model where the model parameters for each location that were initially considered homogeneous can be different. An observation at location i at time t are expressed as $Z_{i,t}$. If the observations between locations are related, then these observations can be modeled using the GSTAR model. The general form of GSTAR is,

$$Z_t = \sum_{j=1}^k \sum_{\ell=0}^{\lambda_k} \Phi_{j\ell} W^{(\ell)} Z_{t-j} + e_t$$

with \mathbf{Z}_t is a N -dimensional column vector $(Z_{1,t}, Z_{2,t}, \dots, Z_{N,t})'$, $\mathbf{W}^{(\ell)}$ is a N -dimensional weight matrix for spatial lag- ℓ^{th} , Φ_{jl} is N -dimensional matrix of autoregressive parameters for spatial lag ℓ and time lag j , and \mathbf{e}_t is a N -dimensional vector of errors respective to the observations inside vector \mathbf{Z}_t . For homoscedastic GSTAR, the \mathbf{e}_t is a white noise vector whose mean and variance are constant, and follows normal multivariate distribution, meanwhile for heteroscedastics case, the variance is not constant.

The GSTAR modeling stage follows the Box-Jenkins iteration [1], consist of model identification, parameter estimation and diagnostic checking. Before doing GSTAR modeling, the data of process must have stationary properties. If it is not stationary, a differentiation process should be performed on the data until the data is stationary. In estimating the parameters of the GSTAR model, it can be done using the Ordinary Least Square (OLS) method by constructing the GSTAR model into a linear form $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e}$, so that the OLS estimator obtained is $\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$ [1]

ARCH(1) Model

Consider a process $\{Y_t\}$ which follows AR(p) model such that can be written as:

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} + e_t$$

which e_t is uncorrelated errors but has inconstant variance or depend on time. Based on Engle (1982), the error e_t can be expressed as,

$$e_t = a_t \sigma_t \tag{1}$$

with a_t is random sample which independent and has identical standard normal distribution, and

$$\sigma_t^2 = \alpha_0 + \alpha_1 e_{t-1}^2 + \dots + \alpha_p e_{t-p}^2 \tag{2}$$

If the errors are known until $(t-1)$ the the conditional variance of e_t is stated as:

$$Var_{t-1}(e_t) = E_{t-1}[e_t^2] = E[e_t^2 | e_{t-1}^2, e_{t-2}^2, \dots] = \sigma_t^2$$

From Eq. (2), it can be said that the conditional variance of e_t depends on squares of the past errors and inconstant. Tthis condition is named as ARCH(p) model [25].

The simplest form of the ARCH(p) model and used in this study is the ARCH(1) model. In this model, the error variance at time t is affected by the square of the error of the previous one time lag. The ARCH(1) model is formulated as:

$$e_t = a_t \sigma_t \text{ and } \sigma_t^2 = \alpha_0 + \alpha_1 e_{t-1}^2$$

with α_0 and α_1 are non-negative parameters of ARCH(1) model.

The variance of ARCH(1) errors is,

$$\begin{aligned}\text{Var}(e_t) &= E[e_t^2] - (E[e_t])^2 = E[e_t^2] = E[a_t^2]E[\sigma_t^2] = E[\alpha_0 + \alpha_1 e_{t-1}^2] \\ &= E[\alpha_0] + E[\alpha_1 e_{t-1}^2] = \alpha_0 + \alpha_1 \text{Var}(e_t)\end{aligned}$$

Then it is obtained,

$$\text{Var}(e_t) = \frac{\alpha_0}{1-\alpha_1} \quad (3)$$

Since the variance is positive then, based on Eq. (3), $\alpha_0 > 0$ and $0 \leq \alpha_1 < 1$, be the stationary condition of ARCH(1) model.

GSTAR(1;1) – ARCH (1) Model

Consider $\mathbf{Z}_t = (Z_{1,t}, Z_{2,t}, \dots, Z_{N,t})'$ as a vector of observations in N location at time t , can be modeled as GSTAR(1;1)-ARCH(1), if it can be expressed as:

$$\mathbf{Z}_t = (\Phi_0 + \Phi_1 \mathbf{W})\mathbf{Z}_{t-1} + \mathbf{e}_t \quad (4)$$

where $\mathbf{e}_t \sim N(\mathbf{0}, \mathbf{\Omega}_t)$, is vector of errors which follows normal distribution with zero mean and inconstant variance over the time. The covariance matrix $\mathbf{\Omega}_t$ is defined as $\mathbf{\Omega}_t = \text{diag}(\mathbf{h}_{1,t}, \mathbf{h}_{2,t}, \dots, \mathbf{h}_{N,t})$ and $\mathbf{h}_{i,t}$ is a vector of errors variance on location i at time t , which can be modeled as ARCH(1), that is

$$\mathbf{h}_{i,t} = \alpha_{0i} + \alpha_{1i} e_{i,t-1}^2$$

with α_{ki} is parameter of model for location i and $k = 0, 1$.

The assumption used in this model is that the errors between locations are uncorrelated with each other so that the error variance at location i at time t is affected by the square of the errors on that location at time $(t - 1)$, but is not affected by the errors on the other locations. Meanwhile, the observation value of location i is influenced by the observations on that location and also the neighbor locations. The method used to estimate this model is the Generalized Least Square (GLS) method.

Suppose that the matrix $\mathbf{\Omega}$ has eigenvalues, $\lambda_1, \lambda_2, \dots, \lambda_T$. By Cholesky's decomposition, it can be written as $\mathbf{\Omega} = \mathbf{C}\mathbf{\Lambda}\mathbf{C}'$, with $\mathbf{\Lambda} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_T)$ is a diagonal matrix, and \mathbf{C} is an orthogonal matrix.

Consider the matrix $\mathbf{C} = \mathbf{P}^{-1}\mathbf{\Lambda}^{-\frac{1}{2}}$, with $\mathbf{\Lambda}^{\frac{1}{2}} = \text{diag}(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \dots, \sqrt{\lambda_T})$. Thus,

$$\mathbf{\Omega}^{-1} = \mathbf{P}^{-1}\mathbf{\Lambda}^{-1}(\mathbf{P}')^{-1} = \mathbf{P}^{-1}\mathbf{\Lambda}^{-\frac{1}{2}}\mathbf{\Lambda}'^{-\frac{1}{2}}(\mathbf{P}')^{-1} = \mathbf{C}\mathbf{C}'$$

Let the transformed linear model, $\mathbf{Y}^* = \mathbf{X}^*\boldsymbol{\beta} + \mathbf{e}^*$ be defined with $\mathbf{Y}^* = \mathbf{C}\mathbf{Y}$, $\mathbf{X}^* = \mathbf{C}\mathbf{X}$, and $\mathbf{e}^* = \mathbf{C}\mathbf{e}$. Then, unbiased estimator of GSTAR-ARCH model parameters are presented as:

$$\hat{\boldsymbol{\beta}}_{\text{GLS}} = (\mathbf{X}'\mathbf{\Omega}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{\Omega}^{-1}\mathbf{Y} \quad (5)$$

with $\Omega = \text{diag}(h_1(1), \dots, h_1(T), \dots, h_N(1), \dots, h_N(T))$ is NT -dimensional of diagonal [26]. Furthermore, the stages of GSTAR-ARCH modeling is illustrated in a flow chart as presented in Fig. 1.

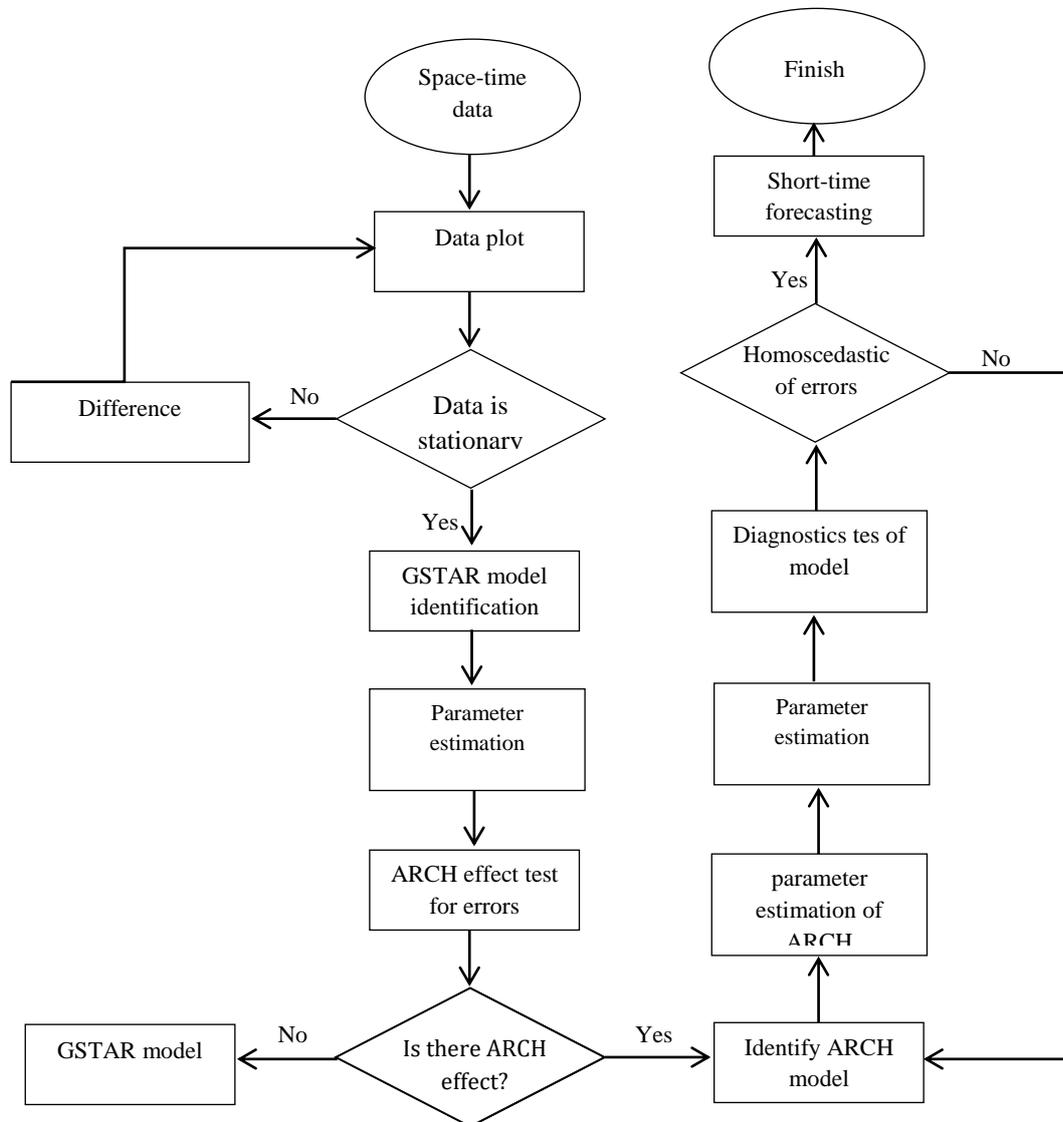


Figure 1. Flowchart of the GSTAR-ARCH modeling stage. Modeling is carried out to obtain a homoscedastic error. The equation for the mean is modeled by the GSTAR model while the variance is modeled by the ARCH model.

RESULTS AND DISCUSSION

As a case study, the data used are the average daily wind speed in three states of the United States ($N = 3$) from September 1, 2004 to August 24, 2005 ($T = 358$) obtained from the National Oceanic and Atmospheric Administration (NOAA) that belongs to the United States. Geographically, the United States is located around the equator, so some areas of the country are often hit by storms. Hurricane Katrina was one of the deadliest hurricanes that occurred in 2005. According to the United States Department of Oceanic and Atmospheric Research, the total loss caused by Hurricane Katrina was over a

thousand million dollars and more than 1,800 people died. The states of interest are New Orleans (Louisiana), Florida, and Mississippi, each of which can be seen in Fig. 2.

The modeling is carried out with the help of the R application. The data will be modeled with the space-time model and must meet the stationary properties first. The stationary data can be seen from the plot of the row of observations at each location as in Fig. 3(a). From the figure, it can be seen that there are several wind speed values that are higher than other observations. In addition, there is also a slight downward and rising pattern, which indicates a data pattern that is not stationary on average. Therefore, the data differentiation is done first. The series plot after one-time differentiation can be seen in Fig. 3(b). The stationary data is then centered so that it has a zero mean (centralized process). The process variability which occasionally increases, indicates that the variance is not constant. This will be accommodated in the modeling with heteroscedastic effect.

In GSTAR modeling, one of the important elements that characterizes the relationship between locations is the presence of a weight matrix. The weight matrix has entries w_{ij} , which represents the weight of location- j to location- i . This matrix has zeros entries in the main diagonal and the total weight in one row is equal to one. In this study, the weight matrix used is uniform and binary weights. The spatial lag used is limited to only one spatial lag. For simplicity, the uniform and binary weight matrix be used, respectively, are

$$W^{(1)} = \begin{pmatrix} 0 & 0.5 & 0.5 \\ 0.5 & 0 & 0.5 \\ 0.5 & 0.5 & 0 \end{pmatrix} \text{ and } W^{(1)} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

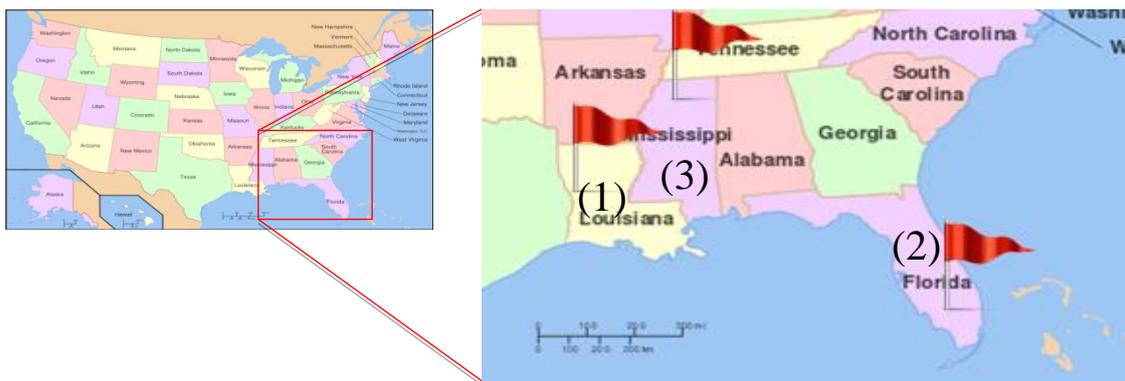


Figure 2. Map of the United States and the states where the observations are located (Source: <https://greatpersie.wordpress.com>). New Orleans (Louisiana), Florida, and Mississippi are referred to as Location 1, Location 2, and Location 3, respectively

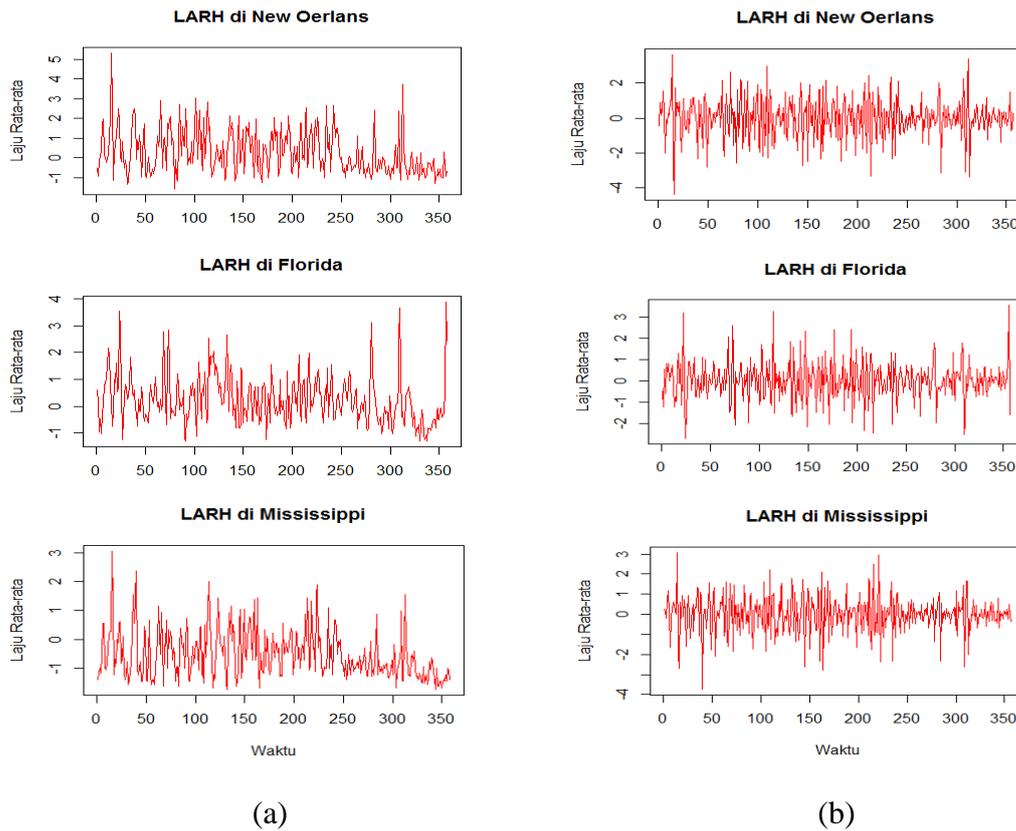


Figure 3. Plot series for each location, (a) before differentiation and (b) after differentiation by one time lag. After differentiation, the data become centered with a more stationary pattern in the mean

The first stage in the modeling is model identification with the help of Space-Time ACF and PACF plots, called STACF and STPACF. However, because the model has been determined at the beginning, namely the GSTAR(1;1) model, the model identification stage can be skipped. The STACF and STPACF plots obtained are used to see whether there is a relationship between time and location from the daily average wind rate data. The plots of STACF and STPACF can be seen in Fig. 4.

From Fig. 4 it can be seen that the data have time and spatial dependence, although the GSTAR(1;1) model is not very appropriate to model this data. From the STPACF plot, the best possible model for the data is GSTAR(3;0,0,1). Thus, the modeling be considered are GSTAR(1;1) and GSTAR(3;0,0,1) model. First, the obtained estimated parameters using Ordinary Least Squares (OLS) method for the GSTAR(1;1) model can be seen in Table 1.

Table 1. The OLS parameter estimation of GSTAR(1;1) model

Parameter	ϕ_{01}	ϕ_{11}	ϕ_{02}	ϕ_{12}	ϕ_{03}	ϕ_{13}
Estimation	-0.15	0.08	-0.08	-0.04	-0.28	0.20

The next step is to test the presence of the ARCH effect on the error of each location. The existence of the ARCH effect can be detected from the plot of the squared error of each location which can be seen in Fig. 5.

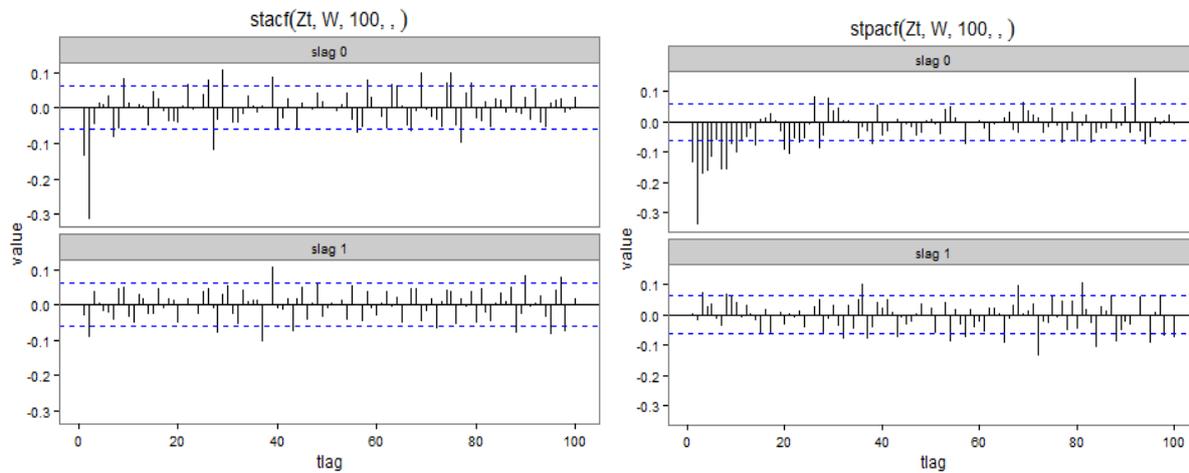


Figure 4. STACF (left) and STPACF (right) plot based on uniform weight matrix. The STACF has more patterned values than STPACF. Thus the autoregressive model is more appropriate

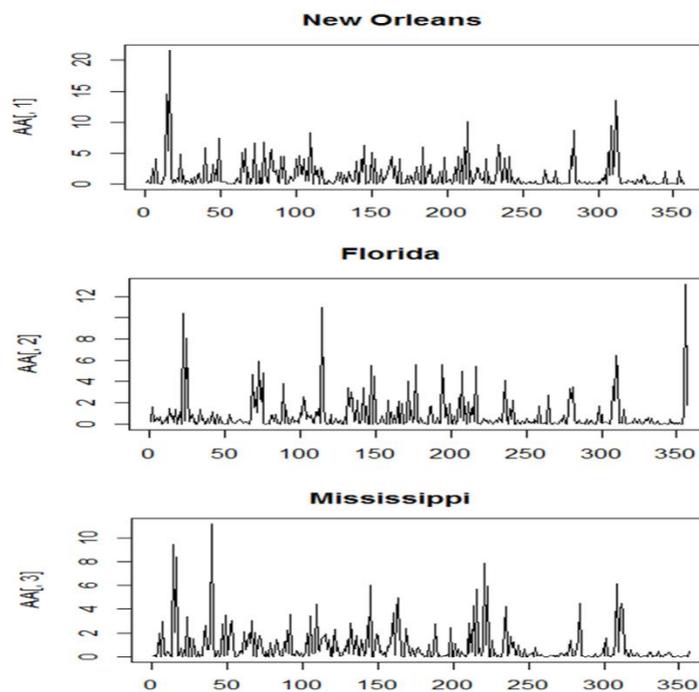


Figure 5. The square errors plot of the GSTAR(1;1) model. It can be seen that the error variance is not constant, indicating the presence of the ARCH

To confirm the existence of the ARCH effect, the ARCH-LM test will be used (Engle, 1982). The presence of ARCH effect on the error is indicated by the p-value which is smaller than $1\% \leq \alpha \leq 10\%$. The ARCH-LM test results for the first six time lags in Table 2, show that the p-value is smaller in almost all locations and time lags. So it can be concluded that there is an ARCH effect on the GSTAR(1;1) model, means that the variance of errors is not constant over time. A slight difference found at Location 2, Florida. The p-values obtained are less than 9.2% until the third time lag, indicates that the wind speed value in this area tends to be more constant in average and variance than the other two observation locations. At a time lag of more than three, the wind speed in Florida did not show any ARCH effect on the process. However, the presence of heteroscedasticity effects in two other locations, also in Florida until the first three-time lags, be the reason to consider the inconstant variance in this case.

Table 2. The p-values of heteroscedastic effect existence using ARCH-LM test

Time lag	New Orleans ($\times 10^{-4}$)	Florida	Mississippi ($\times 10^{-2}$)
1	0.0019	0.0304	0.0300
2	0.0085	0.0731	0.0980
3	0.0067	0.0917	0.2000
4	0.0224	0.1550	0.4300
5	0.0666	0.2170	0.8000
6	0.1490	0.2950	0.9300

Next, the inconstant errors variance of GSTAR(1;1) be modeled by ARCH(1). The obtained parameter estimation of model ARCH(1) model using Maximum Likelihood (ML) method, can be seen in Tabel 3.

Table 3. The ML parameter estimation of ARCH(1) model. The i -th parameter of location- j is presented by α_{ij} for $i = 0,1$ and $j = 1,2,3$.

Parameter	α_{01}	α_{11}	α_{02}	α_{12}	α_{03}	α_{13}
Estimation	1.01	0.25	0.66	0.18	0.60	0.30

Thus, the variances of errors for each location are obtained as::

$$\sigma_{1,t}^2 = 1.01 + 0.25e_{1,t-1}^2$$

$$\sigma_{2,t}^2 = 0.66 + 0.18e_{2,t-1}^2$$

$$\sigma_{3,t}^2 = 0.60 + 0.30e_{3,t-1}^2$$

Furthermore, the variances of errors of each location, for $t=1,2,\dots, T$ be the entries of NT dimensional of diagonal matrix, $\mathbf{\Omega}(t) = \text{diag}(\sigma_{1,t}^2, \sigma_{2,t}^2, \sigma_{3,t}^2)$. This matrix is used for parameter estimation of GSTAR(1;1) – ARCH (1) model by using the GLS mehod. The estimated parameters are presented in Table 4.

Table 4. The GLS parameter estimation of GSTAR(1;1) model with the errors variance follows ARCH(1) model

Parameter	ϕ_{01}	ϕ_{11}	ϕ_{02}	ϕ_{12}	ϕ_{03}	ϕ_{03}
Estimation	-0.11	0.08	-0.05	-0.05	-0.27	0.22

The estimated parameters of the model obtained using the GLS method in Table 4, are not much different from the estimated parameters obtained using the OLS method in Table 1. This is probably because the ARCH(1) model is not the right model to model the error variance of the GSTAR(1;1) model.

For comparison, the modeling with the same steps was carried out again using a binary weight matrix and a weight matrix based on wind direction. Determination of the best model is done by comparing the value of the Mean Squared Error (MSE) of each model. From Table 5 it can be concluded that the use of a uniform weight matrix in the GSTAR(1;1)-ARCH (1) modeling is slightly better than the binary weight matrix. The three locations involved in modeling give the composition of the uniform and binary weight matrix are quite similar.

Table 5. The MSE values of GSTAR(1;1) – ARCH(1) model with two types of thw weight matrix. The results using uniform weight matrix is slightly better than binary, although the difference is not significant

Weight Matrix	MSE
Uniform	0.975
Binary	0.978

The comparison between the original and estimated data using the GSTAR(1;1)-ARCH(1) model can be seen in Fig. 6. From this figure, it can be seen that there is a big difference between the original data and estimated results. This is because the STPACF plot in Fig. 4 indicates that the GSTAR(1;1) model is not appropriate for the data.

From the STPACF plot, the possible next space-time model is the GSTAR (3;0,0,1) with a fixed error variance be modeled by the ARCH(1) model. This model explains that the condition of location i at t is influenced by its own condition at- $(t - 1)$, $(t - 2)$, $(t - 3)$ and the conditions of other locations which are its closest neighbors at $(t - 3)$. The modeling is carried out following the similar steps in modeling GSTAR(1;1) – ARCH (1) until the parameters are obtained before and after the ARCH element is calculated. The estimation parameter results can be seen in Table 6. The estimation parameters obtained from the GSTAR(3;0,0,1) modeling are not much different from the GSTAR estimation parameters GSTAR (3;0,0,1)-ARCH(1) so that the estimation results obtained will also not far different.

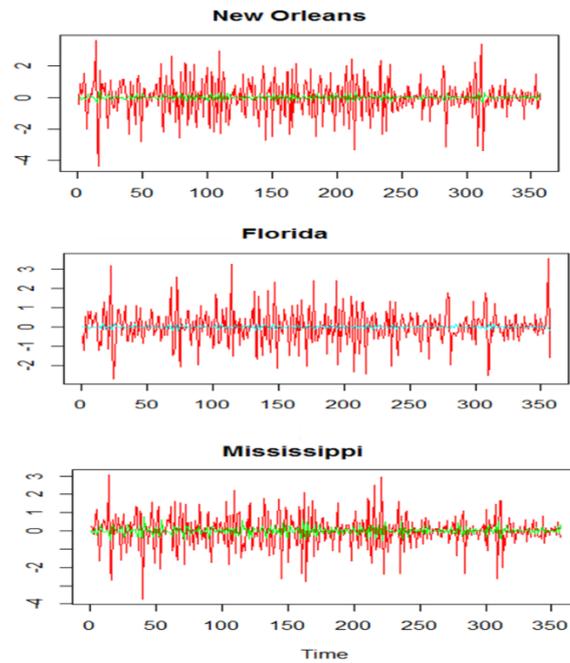


Figure 6. Plot of original series (red) and estimation (blue, green) using GSTAR(1;1) – ARCH(1) model. The large errors indicate that the GSTAR(1;1) model is not suitable.

Table 6. The comparison of GSTAR(3;0,0,1) parameter estimation before and after the errors variance be modeled by ARCH(1). Generally, there is no difference of both models.

Parameter	ϕ_{10}^1	ϕ_{20}^1	ϕ_{30}^1	ϕ_{31}^1	ϕ_{10}^2	ϕ_{20}^2	ϕ_{30}^2	ϕ_{31}^2	ϕ_{10}^3	ϕ_{20}^3	ϕ_{30}^3	ϕ_{31}^3
Before	-0.24	-0.40	-0.15	-0.07	-0.17	-0.32	-0.27	-0.02	-0.31	-0.37	-0.13	-0.07
After	-0.21	-0.36	-0.13	-0.09	-0.13	-0.33	-0.24	-0.03	-0.30	-0.34	-0.10	-0.03

The comparison of the original and estimated data using the GSTAR(3;0,0,1)-ARCH(1) model can be seen in Fig. 7. By comparing the plots in Fig. 6 and Fig. 7, it can be seen that the estimation results generated by this model are better than the GSTAR(1;1)-ARCH(1) model. The GSTAR(3;0,0,1)-ARCH(1) model can capture the pattern of process variability. However, this model is not the best model for the data. From Fig. 7 it can be seen that the estimation results cannot reach the very high either low value of the original data. This may be due to the inaccurate selection of the ARCH(1) model as a model that explains the error variance for each location.

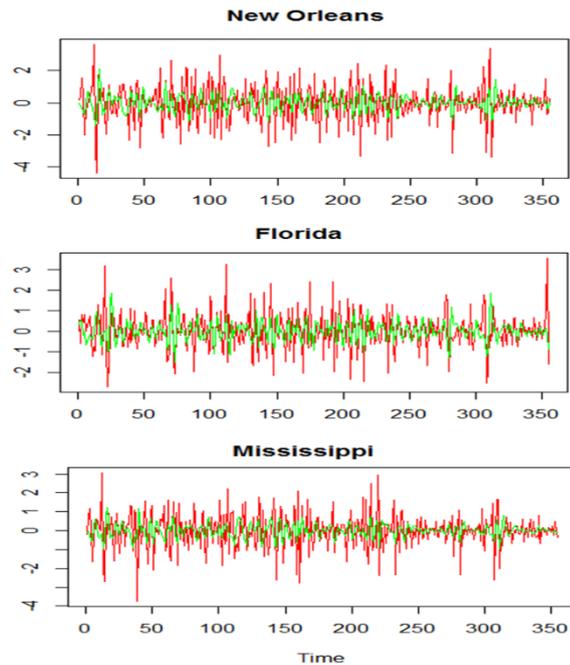


Figure 7. Plot of the original series (red) and estimation (green) GSTAR(3;0,0,1) – ARCH(1) model

The MSE value for the GSTAR(3;0,0,1)–ARCH(1) model is 0.875. This value is smaller than the MSE value of the GSTAR(1;1)–ARCH (1) model, so the model that will be used for short-term prediction is the GSTAR(3;0,0,1)–ARCH(1) model. To make sure the selection of the GSTAR(3;0,0,1)–ARCH(1) model, a comparison was made with the GSTAR(3;0,0,1) model without the effect of heteroscedasticity. Table 7 shows the comparison of MSE, MAD and MAPE values for the two models. Although those values were not significant different, the model used for short-term prediction is the GSTAR(3;0,0,1)–ARCH(1) model, since it is slightly better.

Table 7. The comparison of GSTAR(3;0,1,1) and GSTAR(3;0,0,1)–ARCH(1) model

	GSTAR(3;0,0,1)	GSTAR(3;0,0,1)–ARCH(1)
MSE	0.88	0.86
MAD	0.70	0.70
MAPE	7.48	6.86

From Table 8, it can be concluded that the GSTAR (3;0,0,1)–ARCH (1) model is not very good for estimating the changes in the daily average wind speed, which are too large. In a relatively short period of time, August 25 – 29, 2005, there was an increase in the daily average wind speed. If this speed increase is assumed to be the beginning of Hurricane Katrina, then Hurricane Katrina is expected to hit New Orleans and Mississippi on September 1, 2005, which is three days later than the actual time of hurricane landfall in New Orleans and Mississippi.

Table 8. The comparison of real data and its estimation using GSTAR(3;0,0,1)–ARCH(1) model

Year 2005	New Orleans		Florida		Mississippi	
	Est.	Real	Est.	Real	Est.	Real
25 Aug	6.55	6.26	15.09	15.43	2.67	4.25
26 Aug	7.05	7.61	15.73	12.08	3.15	5.59
27 Aug	7.06	10.96	17.04	9.17	3.54	6.49
28 Aug	7.04	25.05	17.53	5.59	3.82	13.87
29 Aug	7.024	47.65	16.90	2.68	3.43	28.19

CONCLUSIONS

The daily average wind speed data in New Orleans, Florida, and Mississippi of the United States are not only influenced by the wind speed at the previous days in the same location, but the influenced of the wind speed in neighbor states can not be ignored. Through the previous modeling, it was found that the GSTAR(3;0,0,1)–ARCH(1) model is a better model than the GSTAR(3;0,0,1) model. It means that, the wind speed in three-previous days of the closest neighbors will influence today’s wind speed in the reference location. This model can capture the pattern of the wind speed volatilities compare to other observed models. However, the GSTAR(3;0,0,1)–ARCH(1) model is still not good enough to predict wind speeds that are extrme high nor low. This can be caused because the ARCH(1) model is not the right model to model the model error variance. Further analysis of the error variance model such as GARCH(p,q) is also needed so that the error variance can be modeled better. As for the application of the GSTAR model with the ARCH effect, it can also be developed in weather cases in Indonesia, as well as in other fields of science.

ACKNOWLEDGMENTS

This research is supported by P3MI 2020 and P2MI 2021 research grant of ITB.

REFERENCES

- [1] U. Mukhaiyar, and U.S. Pasaribu, A new procedure for generalized STAR modeling using IAcM approach. *ITB J. Sci* 44(2) : 179-192, 2012.
<https://journals.itb.ac.id/index.php/jmfs/article/view/106>
- [2] Yundari, U.S. Pasaribu, U. Mukhaiyar, M.N. Heriawan, Spatial Weight Determination of GSTAR(1;1) Model by Using Kernel Function, *Journal of Physics: Conference Series* 1028 (012223): 1-8, 2018.
<https://iopscience.iop.org/article/10.1088/1742-6596/1028/1/012223>
- [3] Yundari, N.M. Huda, U. Mukhaiyar, U.S. Pasaribu, and K.N. Sari, Stationary Process in GSTAR(1;1) through Kernel Function Approach: AIP Conference Proceedings, 2268, 020010, 2020. <https://aip.scitation.org/doi/10.1063/5.0016808>

- [4] U. Mukhaiyar, B. I. Bilad, and U. S. Pasaribu, The Generalized STAR Modelling with Minimum Spanning Tree Approach of Weight Matrix for COVID-19 Case in Java Island, *Journal of Physics: Conference Series*, **2084** 012003, 2021.
- [5] D. Masteriana, and U. Mukhaiyar, Monte Carlo Simulation of Error Assumptions in Generalized STAR(1;1) Model. *Proceedings on the Jangjeon Mathematical Society* 22(1) : 43-50, 2019. DOI: 10.17777/pjms2019.22.1.43
- [6] Yundari, U.S. Pasaribu, U. Mukhaiyar, Error Assumptions on Generalized STAR Model, *Journal of Mathematical and Fundamental Sciences* 49(2): 136-155, 2017. <https://journals.itb.ac.id/index.php/jmfs/article/view/3285>
- [7] U. Mukhaiyar, N.M. Huda, K.N. Sari, and U.S. Pasaribu, Analysis of Generalized Space Time Autoregressive with Exogenous Variable (GSTARX) Model with Outlier Factor, *Journal of Physics: Conference Series*, 1496(1), 012004, 2020. DOI: 10.1088/1742-6596/1496/1/012004
- [8] N.M. Huda, U. Mukhaiyar, and U.S. Pasaribu, The approximation of GSTAR model for discrete cases through INAR model: *Journal of Physics: Conference Series*, 1722(1), 012100, 2021. <https://iopscience.iop.org/article/10.1088/1742-6596/1722/1/012100/pdf>
- [9] Yundari and S.W. Rizki, Invertibility of Generalized Space-Time Autoregressive Model with Random Weight, *CAUCHY*, 6(4):246-259, 2021. DOI: 10.18860/ca.v6i4.11254
- [10] N. Nurhayati, U.S. Pasaribu, and O. Neswan, Application of generalized STAR model on GDP data in West European Countries. *J. of Probability and Statistics* 2012 : 1-16, 2012. <https://www.hindawi.com/journals/jps/2012/867056/>
- [11] U. Mukhaiyar, The Goodness of generalized STAR in spatial dependency observations modeling, *AIP Conf. Proc.* 1692 (020012), 2015. <https://aip.scitation.org/doi/abs/10.1063/1.4936436>
- [12] R.F. Nugraha, S. Setyowati, U. Mukhaiyar, Prediction of Oil Palm Production using The Weighted Average of Fuzzy Sets Concept Approach. *AIP Conf. Proc.* 1692 (020008), 2015. <https://aip.scitation.org/doi/abs/10.1063/1.4936440>
- [13] N.F.I. Fadlilah, U. Mukhaiyar, F. Fahmi, The Generalized STAR(1;1) Modeling with Time Correlated Errors to Red-Chili Weekly Prices of some Traditional Markets in Bandung, West Java. *AIP Conf. Proc.* 1692(020014), 2015. <https://iopscience.iop.org/article/10.1088/1742-6596/1722/1/012100>
- [14] U. Mukhaiyar, N.M. Huda, K.N. Sari, and U.S. Pasaribu, Modeling Dengue Fever Cases by Using GSTAR(1;1) Model with Outlier Factor: *Journal of Physics: Conference Series*, 1366(1), 012122, 2019. <https://iopscience.iop.org/article/10.1088/1742-6596/1366/1/012122>
- [15] D. Masteriana, M.I. Riani, and U. Mukhaiyar, Generalized STAR (1;1) Model with Outlier - Case Study of Begal in Medan, North Sumatera: *Journal of Physics: Conference Series*, 12456(1), 012046, 2019. <https://iopscience.iop.org/article/10.1088/1742-6596/1245/1/012046/pdf>
- [16] U.S. Pasaribu, U. Mukhaiyar, N.M. Huda, K.N. Sari, and S.W. Indratno, 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, 2021. <https://www.sciencedirect.com/science/article/pii/S2405844021001304>

- [17] Pasaribu U S, Mukhaiyar U, Heriawan M N and Yundari 2021 Generalized Space-Time Autoregressive Modeling of the Vertical Distribution of Copper and Gold Grades with a Porphyry-Deposit Case Study *IJASEIT* **11**(6) (to be published)
- [18] U. Mukhaiyar, D. Widyanti, and S. Vantika, The Time Series Regression Analysis in Evaluating the Economic Impact of Covid-19 Cases in Indonesia. *Journal of Model Assisted Statistics and Applications*, 16(3), 2021a.
<https://ieeexplore.ieee.org/document/7005725>
- [19] S.S. Sholihat, S.W. Indratno, and U. Mukhaiyar, The Role of Parameters in Bayesian Online Change Point Detection: Detecting Early Warning of Mount Merapi Eruptions, *Heliyon*, e07482, 2021. <https://doi.org/10.1016/j.heliyon.2021.e07482>
- [20] U.S. Pasaribu, U. Mukhaiyar, S. Setiyowati, An ARCH Model the Electric Power of Extra High Voltage (EHV) Transmission Substation Forecasting in Cawang, Jakarta, Indonesia: *The Proceedings of IEEE INAGENTSYS*, 1589, 484, 2014.
- [21] S. Setiyowati, R.F. Nugraha, U. Mukhaiyar, Non-stationary time series modeling on caterpillars pest of palm oil for early warning system, *AIP Conference Proceedings* 1692, 020011, 2015, doi: 10.1063/1.4936439
- [22] S. Setiyowati, U.S. Pasaribu, and U. Mukhaiyar, Non-Stationary Model for Rice Prices in Bandung, Indonesia: *The Proceedings of 2012 IEEE Conference on Control, System & Industrial Informatics, (ICA2013)*, 2013.
- [23] S. Borovkova, Svetlana and R. Lopuhaa, Spatial GARCH: A Spatial Approach to Multivariate Volatility Modeling (November 9, 2012). Available at SSRN: <https://ssrn.com/abstract=2176781> or <http://dx.doi.org/10.2139/ssrn.2176781>.
- [24] N. Nainggolan, and J. Titaley, Development of generalized space time autoregressive (GSTAR) model, *AIP Conference Proceedings* **1827**, 020034, 2017. <https://doi.org/10.1063/1.4979450>
- [25] R.F. Engle, Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation, *Econometrica*, Vol 50 (4), pp. 987-1007, 1982. <https://www.jstor.org/stable/1912773>
- [26] S. Ramadhani, Prakiraan Kedatangan Badai Katrina dengan Model Ruang Waktu GSTAR($p; \lambda_1, \lambda_2, \dots, \lambda_p$)-ARCH(1), Laporan Tugas Akhir, Institut Teknologi Bandung, 2016.