Proceedings of the Jangjeon Mathematical Society
22 (2019), No. 1. pp. 43 – 50

www.jangjeon.or.kr
http://dx.doi.org/10.17777/pjms2019.22.1.43

# MONTE CARLO SIMULATION OF ERROR ASSUMPTIONS IN GENERALIZED STAR(1;1) MODEL

DEBBY MASTERIANA AND UTRIWENI MUKHAIYAR

Abstract. This study is aimed to compare several definitions of error assumption in Generalized Space Time Autoregressive (GSTAR) (1;1) model. This definition is needed since the assumption of normal and identically distributed (*iid*) is difficult to be satisfied. Four assumptions are simulated in this article respectively spatial correlated error, time correlated error, spatial and time correlated error with martingale difference, and spatial and time correlated error with beta estimator. By using Monte Carlo simulation, a thousand replications in thirteen locations are done towards GSTAR (1;1) model with varied amount of random data. The result shows that convergence of generalized least square (GLS) estimation approaches the real parameter is faster gained by spatial and time correlated error with martingale difference assumption. This new definition of error assumption provides a more precise GSTAR (1;1) model.

2010 Mathematics Subject Classification. 62M10, 91B72, 91G60.

Keywords and phrases. GSTAR (1;1), Error Assumption, Monte Carlo Simulation.

## 1. Introduction

Many events which occur in everyday life are not only related to the event in previous time but also with the location or the surrounding area. Initially, time series analysis and spatial analysis studies were discussed separately. If multiple locations are fixed, then a time series analysis is used. Conversely, if multiple locations are not fixed and time is fixed, then a spatial analysis is used. However, along with the development of science and technology, space-time analysis was emerged.

Many researchers initiate the idea of space-time modelling. For example, Craigmile and Guttorp in 2011 were using wavelet-based space-time hierarchical Bayesian models for modelling the temperature trends in last five decades from central Sweden [2]. Meanwhile, Huser and Davison in 2014 were using a pairwise censored likelihood to obtain a consistent estimation of the extremes space-time data under mild mixing conditions [5]. In this article, we use Generalized Space Time Autoregressive (GSTAR) model. Previously, Terzi in 1995 purposed GSTAR as Space Time Autoregressive (STAR) model with spatial correlation which occurred at the same time and parameters in each location [9]. Contrast with her, Ruchjana in 2002 introduced GSTAR model as STAR model for heterogeneous locations with different parameter values in each location [8]. Since this definition is prevailed, then the definition of GSTAR in this article refers to Ruchjana's.

This work was supported by Institut Teknologi Bandung.

Generally, errors in GSTAR model are assumed to be normal, independent and identically distributed (*iid*). These assumptions allow researchers to process the data and estimate parameter easier. However, the assumptions of normal and *iid* error are difficult to be satisfied because of varied data. This fact brings out an idea of weakening the error assumption. In 2017, Yundari et al argued that spatial dependency between locations is possibly resulting dependency which does not happen at time and location only but also in the errors. Consequently, the assumptions of normal and *iid* errors are not always fit on data with time or location errors depedency [10].

Recently, some researchers developed weaken error assumptions in the GSTAR model. Attempts to define these error assumptions have more difficulty than normal and *iid* error assumptions, especially in the parameter estimation stage. For example, Borovkova et al in 2008 weakened the error assumption in GSTAR model became time correlated error by using martingale difference [1]. This was a process with a zero mean which produced zero conditional mean value when historical error is given. Moreover, Nurhayati in 2010 also developed an assumption of spatial correlation error in GSTAR (1;1) by using Newton-Raphson approach [7]. Inspired by these researches, this article combines both of assumption become spatial and time correlated error. Furthermore, Monte Carlo simulation will be done towards several assumptions including our proposed idea. The result of this simulation will provide a more precise model of GSTAR (1;1) in spatial and time series analysis for some restricted cases.

## 2. Generalized Space Time Autoregressive (GSTAR)

Given process of $Y_i(t) = (Y_1(t), Y_2(t), , Y_N(t))^t$ with $i = 1, 2, ..., N$ locations and time $t$. $Y_i(t)$ is following GSTAR model with time order $p$ and spatial lag $\lambda_1, ..., \lambda_p$, or written as GSTAR $(p; \lambda_1, ..., \lambda_p)$ if satisfies:

$$Y_i(t) = \sum_{k=1}^{p} \left[ \phi_{k0} Y(t-k) + \sum_{s=1}^{\lambda_k} \phi_{ks} W^{(s)} Y(t-k) \right] + \epsilon(t) \quad (1)$$

with $\lambda_k$ representing spatial order for the $k$-th time lag and $\phi_{ks}$ representing regression parameter of $k$-th time order and $s$-th spatial order. Meanwhile, weight matrix $W^{(s)}$ representing $s$-th order spatial weight with size $N \times N$ and total of each row is one. $\epsilon(t)$ is error vector in time $t$ with dimension $N$ [1]. In matrix notation, equation 1 can be written as:

$$Y(t) = \sum_{k=1}^{p} \sum_{s=1}^{\lambda_k} \Phi_{ks} W^{(s)} Y(t-k) + \epsilon(t) \quad (2)$$

with $Y(t) = \begin{bmatrix} Y_1(t) \\ Y_2(t) \\ \vdots \\ Y_N(t) \end{bmatrix}$, $\epsilon(t) = \begin{bmatrix} \epsilon_1(t) \\ \epsilon_2(t) \\ \vdots \\ \epsilon_N(t) \end{bmatrix}$, $W = \begin{bmatrix} 0 & w_{12} & \dots & w_{1N} \\ w_{21} & 0 & \dots & w_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ w_{N1} & w_{N2} & \dots & 0 \end{bmatrix}$ for every $s$-th spatial order which satisfies $\Phi_k = diag(\phi_{k1}, \phi_{k2}, ..., \phi_{kN})$ where $k = 1, 2, ..., p$ for every $s$-th spatial order and $\sum_{j=1}^{N} w_{ij} = 1$.

In this article, we are using GSTAR (1;1) model. Consider $Y_i(t)$ following GSTAR (1;1) model which given by:

$$Y(t) = (\Phi_0 + \Phi_1 W)Y(t-1) + \epsilon(t). \tag{3}$$

Since the process of $Y(t)$ is assumed to be centered, then $E[Y(t)] = 0$ for every $t$. Consequently, STAR (1;1) is a special case of GSTAR(1;1) model when $\Phi_0 = \phi_{10}\mathbb{I}$ and $\Phi_1 = \phi_{11}\mathbb{I}$. Furthermore, stationary requirement of GSTAR (1;1) model is given by this following theorem.

**Theorem 2.1.** *If $\phi_{10}^{(i)}$ and $\phi_{11}^{(i)}$ satisfy $\left|\phi_{10}^{(i)} + \phi_{11}^{(i)}\right| \leqslant 1$ and $\left|\phi_{10}^{(i)} - \phi_{11}^{(i)}\right| \leqslant 1$ for every $i = 1, 2, ..., N$ then GSTAR (1:1) model is stationary.*

*Proof.* Noted that GSTAR (1;1) model is a special form of STAR (1). Let $r_s$ satisfies the equation:

$$|r_s\mathbb{I} - (\Phi_{10}\mathbb{I} + \Phi_{11}W)| = 0. \tag{4}$$

We will prove that the solution of 4 lies inside the unit circle such that $|r_s| < 1$. Based on Gershgorin theorem [4], every eigenvalue of matrix A is located in one disc such that the following inequality below satisfied.

$$|r - a_{ii}| \leqslant \sum_{i \neq j} |a_{ij}|, i = 1, 2, ..., N. \tag{5}$$

Let $A = \Phi_0 + \Phi_1 W$, where each of $\Phi_0$ and $\Phi_1$ is diagonal matrix of GSTAR (1;1). Since $w_{ii} = 0$ and $\sum_{j=1}^{N} w_{ij} = 1$, then $a_{ii} = \phi_{10}^{(i)}$, $a_{ij} = \phi_{11}^{(i)}$ for every $i \neq j$, and $\sum_{i \neq j} \left|\phi_{11}^{(i)}\right|$ for $i = 1, 2, ..., N$. If $r_s$ is the solution of 4, then at least one of location $i$ satisfies:

$$\left|r - \phi_{10}^{(i)}\right| \leqslant \phi_{11}^{(i)}. \tag{6}$$

By quadrating both side of inequality 6, then we have:

$$\left(r - \phi_{10}^{(i)}\right)^2 \leqslant \left(\phi_{11}^{(i)}\right)^2$$
$$r^2 - 2r\phi_{10}^{(i)} + \left(\phi_{10}^{(i)}\right)^2 - \left(\phi_{11}^{(i)}\right)^2 \leqslant 0$$

with the root of equation $r_{1,2}$ is given by:

$$\begin{aligned} r_{1,2} &= \frac{2\phi_{10}^{(i)} \pm \sqrt{\left(2r\phi_{10}^{(i)}\right)^2 - 4\left(\left(r\phi_{10}^{(i)}\right)^2 - \left(r\phi_{11}^{(i)}\right)^2\right)}}{2} \\ &= \phi_{10}^{(i)} \pm \phi_{11}^{(i)}. \end{aligned}$$

Since $|r| \leqslant 1$, then $\left|\phi_{10}^{(i)} \pm \phi_{11}^{(i)}\right| \leqslant 1$ for every $i = 1, 2, ..., N$. This means $\left|\phi_{10}^{(i)} + \phi_{11}^{(i)}\right| \leqslant 1$ and $\left|\phi_{10}^{(i)} - \phi_{11}^{(i)}\right| \leqslant 1$. Hence, these inequalities give sufficient condition for stationarity of GSTAR (1;1) model. □

## 3. Monte Carlo Simulation of GSTAR (1;1)

The main principal of Monte Carlo simulation is using repeated random sampling to generate simulated data and obtain numerical results. Generally, the mathematics model which is used in Monte Carlo simulation originated from statistical analysis, such as a designed experiment or time series analysis. In this paper, we apply Monte Carlo simulation with a thousand replications towards several error assumptions in GSTAR (1;1) model at thirteen locations. Selection of locations was done because the result of this research will be used for further research which contains thirteen locations.

In whole assumptions, we try to compare the root mean square error (RMSE) which is obtained by calculating the difference between selected assumption in model and in-sample model. The in-sample model here is assumed to have spatial and time correlated error. The goal of this simulation is to find out how fast the convergence of parameters in GSTAR (1;1) obtained if the error has spatial and time correlation. Four error assumptions which used in this simulation are explained as follows.

3.1. **Spatial Correlated Error Assumption.** This kind of error assumption refers to error in such locations which is influenced by surrounding errors. The influence of surrounding errors is varied. According to Nurhayati (2010), if the error is correlated linearly, then GSTAR (1;1) model with spatial correlated error (SCE) assumption is defined as [7]:

$$a_i(t) = \rho \sum_{j \in J} w_{ij} \epsilon_j(t) + \eta_i(t). \tag{7}$$

In vector notation, equation (7) can be written as:

$$a(t) = \rho W \epsilon(t) + \eta(t) \text{ or } (\mathbb{I}_N - \rho W)\,\epsilon(t) = \eta(t). \tag{8}$$

Noted that spatial correlation of error is explained by $\rho$ which multiplied by weight matrix $W$ in equation (8). As implication of this assumption, $\eta(t)$ is error of whole model with normal distribution. Furthermore, Nurhayati also developed parameter estimation procedure which was named as Generalized Least Square (GLS). In this article, we are using the same procedure to obtain the estimated parameters of GSTAR (1;1) model. Monte Carlo simulation is done by fixing the value of correlation $\rho$, weight matrix $W$, and covariance matrix $\sum$. Moreover, we are also fixing the parameter $\Phi_0$ and $\Phi_1$ which satisfy the stationary condition in Therorem 2.1. By having fixed parameters value, we are able to calculate the correct result of in-sample model. Conversely, consider that we do not have any fixed parameters, then we will estimate the parameters by using GLS such that the result of spatial correlated error in model is gained. Lastly, RMSE of this assumption is calculated. Procedure to have RMSE will be repeated in next assumptions.

3.2. **Time Correlated Error Assumption.** This error assumption refers to error in such locations which is influenced by errors in previous time. The influence is varied and its definition is subjective as long as satisfy martingale difference condition [1]. Suppose the error is correlated linearly, then GSTAR (1;1) model with time correlated error (TCE) is defined as:

$$a(t) = \epsilon(t) \times \epsilon(t-1). \tag{9}$$

**Theorem 3.1.** *Suppose* $\epsilon(t) \sim N(0, \sum)$*. Vector error* $a(t) = \epsilon(t) \times \epsilon(t-1)$ *satisfies martiangle difference.*

*Proof.* According to Hamilton (1994), if vector series with size $(N \times 1)$ of $Y_t$ satisfies $E[Y_t] = 0$ and $E[Y_t|Y_t, Y_{t-1}, ...., Y_1] = 0$, then $Y_t$ forms a vector sequence with martingale difference [3]. Now, we will show that $a(t)$ satisfies martingale difference by proving that $a(t)$ satisfies Hamilton's. Generate $\epsilon(t) \sim (0, \sum)$ with $E[\epsilon(t-1)|F_{t-1}] = 0$. Let $a(t) = \epsilon(t) \times \epsilon(t-1)$, then

$$\begin{aligned} E[a(t)|F_{t-1}] &= E[\epsilon(t)\epsilon(t-1)|F_{t-1}] \\ &= E[\epsilon(t)|F_{t-1}]\dot{E}[\epsilon(t-1)|F_{t-1}] \\ &= E[\epsilon(t)|F_{t-1}]\dot{0} \\ &= 0. \end{aligned}$$

Since $E[a(t)|F_{t-1}] = 0$, then $a(t)$ satisfies martiangle difference. □

Since equation (9) satisfied martingale difference process, then according to Mukhaiyar (2007), the parameter estimation is executable by using least square method [6]. Noted that time correlated error is explained by $a(t)$ which defined as multiplication of error in time $t$ and previous time $t-1$. Monte Carlo simulation is done by fixing the value of weight matrix $W$ and covariance matrix $\sum$. By doing the same procedure with spatial correlated error assumption, RMSE of this model is obtained.

3.3. **Spatial and Time Correlated Error Assumption with Martingale Difference.** This part is our new idea to be proposed. Spatial and time error correlated assumption refers to spatial correlated error with error vector $\eta(t)$ in equation (8) satisfies martingale difference process. In another words, this error is influenced by surrounding errors in previous time. This error assumption in GSTAR (1;1) model is defined as:

$$a(t) = \rho W \epsilon(t) + (\epsilon(t) \times \epsilon(t-1)). \tag{10}$$

Monte carlo simulation in this error assumption is done by fixing the value of correlation $\rho$, weight matrix $W$ and covariance matrix $\sum$. By doing the same procedure with spatial correlated error assumption, RMSE of this model is obtained.

3.4. **Spatial and Time Correlated Error Assumption with $\beta$.** This idea was proposed by Nurhayati in 2010 [7]. This kind of error assumption has difficulty in parameter estimation since there is new parameter $\beta$. The parameter of $\beta$ here denotes the correlation index which describes relationship between error and previous error. In another words, the error of error in GSTAR (1;1) model is influenced by surrounding errors in previous time. This error assumption in GSTAR (1;1) model is defined as:

$$a(t) = \rho W \epsilon(t) + (\beta \times \eta(t-1)) + \epsilon(t). \tag{11}$$

Noted that time correlated error is explained by multiplication of $\beta$ and $\eta(t-1)$. Monte carlo simulation of this error assumption is done by fixing the value of correlation $\rho$ and $\beta$, weight matrix $W$ and covariance matrix $\sum$. By doing the same procedure in previous assumption, then RMSE of this model is obtained.

## 4. Result of Monte Carlo Simulation in GSTAR (1;1)

The Monte Carlo simulation was done under the condition of spatial and time correlated error with parameter $\beta$. Furthermore, this assumption is chosen to be compared with another assumptions. There are several amount of data time, respectively 40, 50, 60, 70, ..., 200. As a note, the parameter of $\Phi_0$ and $\Phi_1$ are chosen close to zero but still satisfy the stationary requirement. We expect this selection will fasten the convergence of parameter. Our hypothesis is GSTAR (1;1) model with spatial and time correlated error and parameter $\beta$ will give smaller error than another assumptions, because the data is generated with this condition. Neverthless, the result of Monte Carlo simulation shows it is wrong. Table 1 below gives the result of simulation of four error assumptions in GSTAR (1;1) model.

Table 1. Result of Monte Carlo Simulation in GSTAR (1;1) Model.

| | | RMSE in time T | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 40 | 50 | 60 | 70 | 80 | 90 | 100 | 110 | 120 | 130 | 140 | 150 | 160 | 170 | 180 | 190 | 200 |
| Assumption of Error | SCE | 1.52 | 1.50 | 1.66 | 1.58 | 1.31 | 1.45 | 1.33 | 1.30 | 1.15 | 1.17 | 1.10 | 0.88 | 1.00 | 0.93 | 1.55 | 1.54 | 1.83 |
| | TCE | 1.36 | 1.35 | 1.51 | 1.42 | 1.13 | 1.27 | 1.21 | 1.18 | 1.06 | 1.08 | 0.99 | 0.83 | 0.90 | 0.84 | 1.42 | 1.41 | 1.62 |
| | STCEMD | 1.51 | 1.50 | 1.64 | 1.60 | 1.36 | 1.30 | 1.34 | 1.28 | 1.60 | 1.58 | 1.48 | 1.51 | 1.38 | 1.45 | 1.51 | 1.64 | 1.66 |
| | STCEB | 1.65 | 1.75 | 2.01 | 1.88 | 1.87 | 1.79 | 1.87 | 1.62 | 1.57 | 1.62 | 1.87 | 1.94 | 1.78 | 1.70 | 1.87 | 1.97 | 2.07 |

Based on the table above, time correlated error (TCE) assumption looks like giving smaller error rather than others. Neverthless, the value of RMSE between TCE and spatial and time correlated with martingale difference (STCEMD) assumptions is not quite difference when the amount of time data reaches 200. The interesting fact is the biggest root mean square error (RMSE) is given by spatial and time correlated error with parameter $\beta$ (STCEB) itself. This facts shows that the convergence of generalized least square (GLS) estimation approaches the real parameter is not influenced by the kind of generated error. However the kind of error, the true definition of error can cover it well. Here, we agree to claim that spatial and time correlated with martingale difference (STCEMD) assumption is useable in GSTAR (1;1) model. The remain problem is how to formulate the suitable parameter estimation since we using GLS method in this observation.

Now, take a look for the stability of RMSE which gained of each error assumption in different amount of data. Noted that STCEMD gives the smallest variance rather than others. As seen in Figure 1, fluctuation of STCEMD is not too significant and produces this assumption has the best stability result of RMSE rather than other assumption. This result concludes that the convergence of generalized least square (GLS) estimation approaches the real parameter is faster gained by STCEMD.

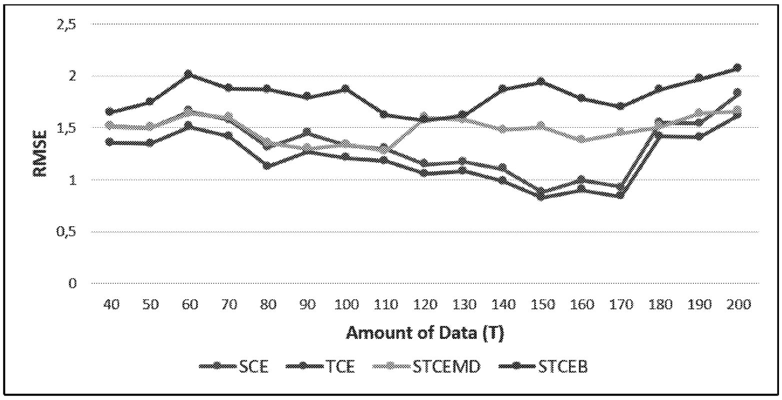


FIGURE 1. RMSE of Error Assumptions in GSTAR (1;1) Model.

## 5. CONCLUSION

Based on the result of Monte Carlo simulation towards several error assumptions in GSTAR (1;1) model, the convergence of generalized least square (GLS) estimation approaches the real parameter is faster gained by spatial and time correlated error with martingale difference (STCEMD) assumption. Furthermore, the root mean square error (RMSE) which obtained by this assumption is not quite different when amount of time data reaches 200. This new definition of error assumption provides a more precise GSTAR (1;1) model. Moreover, forecasting by using this error assumption inovates spatial and time series model.

## REFERENCES

[1] Borovkova, S., Lopuha, Hendrik P., and Ruchjana, B. N., *Consistency and asymptotic normality of least squares estimators in generalized STAR models*, Statistica Neerlandica. 62 (2008), 482-508.

[2] Craigmile, P. F. and Guttorp, P., *Space-time modelling of trends in temperature series*, Journal of Time Series Analysis. 32 (2011), 378-395.

[3] Hamilton, J. D., *Time series analysis*, Princeton University Press. (1994).

[4] Horn, A. R. and Johnson, C. R., *Matrix analysis second edition*, Cambridge University Press. (2013).

[5] Huser, R. and Davison, A.C., *Space-time modelling of extreme events*, Journal of the Royal Statistical Society B. 76 (2014), 439-461.

[6] Mukhaiyar, U. *Kekonsistenan lemah penaksir kuadrat terkecil model space time GSTAR(1;1) melalui proses beda martingale: studi kasus pada produksi bulanan perkebunan teh di wilayah jawa barat*, Thesis of Magister Program in Institut Teknologi Bandung. (2007).

[7] Nurhayati, N., *Analisis model generalized STAR(1;1) dengan galat berkorelasi spasial*, Dissertation of Doctoral Program in Institut Teknologi Bandung. (2010).

[8] Ruchjana, B. N., *Suatu model generalisasi space-time autoregresi dan penerapannya pada produksi minyak bumi*, Dissertation of Doctoral Program in Institut Teknologi Bandung. (2002).
[9] Terzi, S., *Maximum likelihood estimation of a generalized STAR($p$;$1_p$) model*, Journal of The Italian Statistical Society. 3 (1995), 377-393.
[10] Yundari., Pasaribu, U. S., and Mukhaiyar, U., *Error assumption on generalized STAR model*, Journal of Math Fund Science. 48 (2017), 136-155.

Mathematics Department, Institut Teknologi Bandung., Jl. Ganesha 10 Bandung, 40132. Indonesia.
*E-mail address*: debbymath.itb@gmail.com

Statistics Research Division, Institut Teknologi Bandung., Jl. Ganesha 10 Bandung, 40132. Indonesia.
*E-mail address*: utriweni@math.itb.ac.id