Simulation of Generalized Space-Time Autoregressive with Exogenous Variables Model with X Variable of Type Metric

Reza Mubarak[1], Suhartono[1]

Abstract* – *One of the models time series which also involves spatial aspects (spatio-temporal) is Generalized Space Time Autoregressive (GSTAR). Until now, GSTAR modelling don't involve metric-type, which is called GSTARX. Parameter estimation for spatio temporal modeling is still limited by using Ordinary Least Square (OLS) which is less efficient because the residuals are correlated. Generalized Least Square (GLS) is one of the alternative methods for parameter estimation residuals are correlated. In this study would like to looking at the efficiency of GLS estimation method is compared with OLS to correlated data in GSTARX model. Simulation results show that the estimation GLS method is more efficient than using OLS if residual correlated.

Index Terms* – *GSTRAX, OLS, GLS, Simulation.

INTRODUCTION

Forecasting has become an important part of human life in various aspects. One of the models time series which also involves spatial aspects (spatio-temporal) is Generalized Space Time Autoregressive (GSTAR). However, in its application to the development GSTAR involving an exogenous variable is still much to do, especially with the exogenous variables of type metrics.

Spatio temporal modeling parameter estimation is still limited by using Ordinary Least Square or OLS [1] and Maximum Likelihood [2]. Terzi argues parameter estimation by OLS on GSTAR less efficient due to residual correlated. Generalized Least Square (GLS) is one of the alternative methods for parameter estimation of residual cross-correlated.

The purpose of this research is looking at the efficiency of GLS estimation method is compared with OLS to correlated data in GSTARX model.

LITERATURE REVIEW

A. GSTAR

STAR model assumes autoregressive parameter is the same for each location, so that the STAR model can only be used on the same location or homogeneous. Whereas, on the assumption that there is GSTAR states allowed different parameters for each location, so GSTAR used at research sites that are heterogeneous [3].

GSTAR model of order autoregressive (time) and spatial order $\lambda_1, \lambda_2, \dots, \lambda_p$, GSTAR $(p; \lambda_1, \lambda_2, \dots, \lambda_p)$ in matrix notation can be written as follows :

$$\boldsymbol{Y}(t) = \sum_{k=1}^{p} \left[\boldsymbol{\Phi}_{k0} \boldsymbol{Y}(t-k) + \sum_{l=1}^{\lambda_p} \boldsymbol{\Phi}_{k1} \boldsymbol{W}^{(l)} \boldsymbol{Y}(t-k) \right] + \boldsymbol{\varepsilon} + \boldsymbol{\varepsilon}(t) \quad (2)$$

Where,

$\boldsymbol{\Phi}_{k0}$ = $diag\left(\phi_{k0}^{(1)}, \dots, \phi_{k0}^{(N)}\right)$ is time parameter matrix

$\boldsymbol{\Phi}_{kl}$ = $diag\left(\phi_{kl}^{(1)}, \dots, \phi_{kl}^{(N)}\right)$ is spatial parameter matrix

ε(t) = noise vector size (N x 1) is an independent, identical, multivariate normal distribution with mean zero and variance-covariance matrix $\sigma^2 \boldsymbol{I}_N$.

Weighting values are chosen so that, to qualify $w_{ii}^{(k)} = 0$ and $\sum_{i \neq j} w_{ij}^{(k)} = 1$.

MATERIAL AND METHOD

Steps in the modeling GSTARX-OLS and GSTARX-GLS on this simulation data is as follows:

1) Generating residual data multivariate normal distribution with mean zero and variance-covariance matrix $E(\varepsilon\varepsilon')$ for the three locations and *n* = 300.
2) Coefficient parameters used in the model GSTARX ([1]$_1$) in accordance with the terms of stationary parameters GSTAR, namely eigenvalues parameter is less than one, then be written $|\lambda \boldsymbol{I} - \Phi|$, $|\Phi| = |\lambda \boldsymbol{I}|$, with $\lambda_i < 1$. The parameters used can be seen in the following matrix equation:

$$\Phi_1 = \begin{bmatrix} 0{,}35 & 0{,}30 & 0{,}30 \\ 0{,}25 & 0{,}45 & 0{,}25 \\ 0{,}20 & 0{,}20 & 0{,}40 \end{bmatrix}$$

3) Data generated residual formed into a VAR(1) model which these data will be exogenous variables Xt.
4) Perform again step 1) and forms into VAR(1) model and coefficient parameter is :

$$\Phi_1^* = \begin{bmatrix} 0{,}15 & 0{,}22 & 0{,}22 \\ 0{,}12 & 0{,}20 & 0{,}12 \\ 0{,}18 & 0{,}18 & 0{,}10 \end{bmatrix}$$

and variance covariance matrix as follows :

a. Residual not correlated between all location, same variance (Simulation 1)
b. Residual not correlated between all location, different variance (Simulation 2)
c. All residual correlated between all locations with the same variance (Simulation 3)
d. All residual correlated between all locations with the different variance (Simulation 4)
e. Residual correlated between some locations with the same variance (Simulation 5)

[1]Reza Mubarak, Suhartono are with Department of Statistics, Faculty of Mathematics and Science, Institut Teknologi Sepuluh Nopember, Surabaya. Email: reza.mubarak@msn.com; suhartono@statistika.its.ac.id

   f. Residual correlated between some locations with the different variance (Simulation 6)

   This step of formed the variable y_t.

5) Getting the data series $Y_{i,t}$ three locations with the effects of calendar variations as well as a predictor of type metric with two scenarios. Effect all the same locations on the predictors of type metrics in the following scenarios:
   a. Order is b = 1, r = 0, s = 1
   b. Order is b = 1, r = 0, s = 2
6) GSTARX estimating model parameters with OLS and GLS method.

Comparing the results of model estimation GSTARX-OLS and GSTARX-GLS and calculate the efficiency of the method GLS.

RESULT

Table 1. Provides the results of the sixth simulation.

Simulation	Φ^*_{ij}	Efficiency 1st	Efficiency 2nd	Simulation	Φ^*_{ij}	Efficiency 1st	Efficiency 2nd
1	Φ^*_{10}	0.04	0.12	4	Φ^*_{10}	31.99	31.96
	Φ^*_{20}	0.04	0.33		Φ^*_{20}	20.72	21.18
	Φ^*_{30}	0.05	0.42		Φ^*_{30}	21.35	17.92
	Φ^*_{11}	0.02	0.07		Φ^*_{11}	28.02	28.16
	Φ^*_{21}	0.02	0.16		Φ^*_{21}	13.99	14.29
	Φ^*_{31}	0.03	0.21		Φ^*_{31}	17.60	13.90
2	Φ^*_{10}	0.37	0.50	5	Φ^*_{10}	4.09	1.08
	Φ^*_{20}	0.10	0.72		Φ^*_{20}	3.74	8.37
	Φ^*_{30}	0.45	0.27		Φ^*_{30}	8.63	9.86
	Φ^*_{11}	0.22	0.29		Φ^*_{11}	2.59	0.62
	Φ^*_{21}	0.05	0.40		Φ^*_{21}	2.07	5.38
	Φ^*_{31}	0.31	0.18		Φ^*_{31}	6.41	7.10
3	Φ^*_{10}	9.66	14.52	6	Φ^*_{10}	10.22	12.41
	Φ^*_{20}	4.45	8.50		Φ^*_{20}	23.28	37.63
	Φ^*_{30}	9.09	9.08		Φ^*_{30}	70.28	73.51
	Φ^*_{11}	7.33	11.80		Φ^*_{11}	4.70	6.62
	Φ^*_{21}	2.93	5.62		Φ^*_{21}	6.12	26.89
	Φ^*_{31}	6.19	6.36		Φ^*_{31}	67.01	70.81

Table 1 that the value of efficiency has to be all positive, which means that estimates using GLS better than the OLS.

CONCLUSION

The results of simulation 1 and 2 show that if the locations is correlated so parameter estimation GLS is not more efficient than OLS. However, the different results obtained in simulation 3 to 6 provide GLS more efficient if between locations correlated.

REFERENCES

[1] N. Ruchjana, “Pemodelan Kurva Produksi Minyak Bumi Menggunakan Model Generalisasi STAR”. Forum Statistika dan Komputasi IPB. Bogor, 2002.

[2] S. Terzi, “Maximum Likelihood Estimation A Generalized STAR (p; 1_p) Model,” pp.154–178. Roma: Universitas La Sapienza. 1995.

[3] S. A. Borovkova, H. P. Lopuhaä, and Ruchjana, B.N, “Least Square Estimation of Generalized Space Time Autoregressive (GSTAR) Model and Its Properties,” The 5th International Conference on Research and Education in Mathematics, 2008.