GSTAR-SUR Modeling With Calendar Variations And Intervention To Forecast Outflow Of Currencies In Java Indonesia

M S Akbar1,2, Setiawan1, Suhartono1, B N Ruchjana3, and M A A Riyadi4
1Department of Statistics, Institut Teknologi Sepuluh Nopember (ITS)
2PhD Student, Department of Statistics, Institut Teknologi Sepuluh Nopember (ITS)
3Master Student, Department of Statistics, Institut Teknologi Sepuluh Nopember (ITS)
Arief Rahman Hakim Street, Surabaya 60111, Indonesia
4Department of Mathematics, Faculty of Mathematics and Natural Sciences
Universitas Padjadjaran (UNPAD), Bandung, Indonesia. Raya Bandung Street-Sumedang Km. 21 Jatinangor Sumedang 45363, Indonesia

E-mail: 1) m_svahid_a@statistika.its.ac.id; 2)setiawan@statistika.its.ac.id; 3)suhartono@statistika.its.ac.id; 4) budinr@unpad.ac.id; 5) m.alfan.alfian.riyadi@gmail.com

Abstract. Ordinary Least Squares (OLS) is general method to estimates Generalized Space Time Autoregressive (GSTAR) parameters. But in some cases, the residuals of GSTAR are correlated between location. If OLS is applied to this case, then the estimators are inefficient. Generalized Least Squares (GLS) is a method used in Seemingly Unrelated Regression (SUR) model. This method estimated parameters of some models with residuals between equations are correlated. Simulation study shows that GSTAR with GLS method for estimating parameters (GSTAR-SUR) is more efficient than GSTAR-OLS method. The purpose of this research is to apply GSTAR-SUR with calendar variation and intervention as exogenous variable (GSTARX-SUR) for forecast outflow of currency in Java, Indonesia. As a result, GSTARX-SUR provides better performance than GSTARX-OLS.

1. Introduction
There are many applications of time series analysis with location effects. For example tea productions, daily oil productions, air pollution measurements per hour at several cities [1] [2] [3]. These examples constitute to spatial time series, which is indexed data by time and location. In spatial applications, the locations can be correlated each other. The Space Time Autoregressive (STAR) method is one of the time series methods involving location.

STAR model was first developed by Cliff and Ord [4]. The STAR model applies to a homogeneous location, because all locations have the same time for series time and space time values. Ruchjana developed STAR model for parameter values of each location can be different, both time series and space time parameter values. The development of this model was known as Generalized Space Time Autoregression (GSTAR). Research about GSTAR has been applied in various fields, such as Application of Generalized Space-Time Autoregressive Model on GDP Data in West European Countries [5], Forecasting oil production data at volcanic layer Jatibarang, West Java Indonesia [6]. In forecasting research, there are many forecasting cases involving exogenous variables matric and non-
metric. Forecasting models that involve exogenous variables are usually coupled with the suffix - X, such that ARIMAX or VARIMAX. The GSTAR method also has been studied method GSTAR-X, that is GSTARX-SUR Model for Spatio-Temporal Data Forecasting [7]. This research involves intervention as exogenous variable in cases of inflation in four cities of East Java.

This article will discuss forecasting with the GSTAR method that involves the effects of calendar variations and interventions simultaneously on exogenous variables. The application of this method is to forecast rupiah outflow in five provinces of Java island, there are Jakarta, West Java, Central Java, Yogyakarta, and East Java. As an exogenous variables, the case of rupiah outflow on Java island influenced by Eid Fitr (calendar variation) and government policy in January 2007 and January 2010. So that forecasting of rupiah outflow data consists of two exogenous variables, there are calendar variation and intervention. The aim of this research is to apply GSTAR-SUR with calendar variation and intervention as exogenous variable (GSTARX-SUR) for forecast outflow of currency in five provinces of Java, Indonesia.

2. Method
In this section, we describe the statistical methods used to estimate GSTAR model parameters and the concept of calendar variations and interventions.

2.1. Generalized Space Time Autoregression (GSTAR)
The Space Time Autoregressive (STAR) model was first developed by Cliff and Ord [4]. Ruchjana developed STAR model into GSTAR. GSTAR is more flexible spatial weights when compared with STAR [2]. GSTAR model with λp spatial order and p-time order (GSTAR (λ, p) is [7]:

\[ Z(t) = \sum_{s=1}^{p} \Phi_x^s + \sum_{k=1}^{m} \Phi_w^k W(t-k)Z(t-s) + a(t), \]

(1)

with \( \Phi_x = \text{diag} (\phi_{10}^{(1)}, \ldots, \phi_{10}^{(N)}) \), \( \Phi_w = \text{diag} (\phi_{10}^{(1)}, \ldots, \phi_{10}^{(N)}) \), weight \( W(t) \) error model that are identically independent and normal distribution with zero mean and covariance \( \Sigma \). If spatial order 1, time order 1 and three locations \( (\lambda_p = 1, p = 1, N = 3) \) then equation (1) becomes [6]:

\[ Z(t) = \Phi_1 Z(t-1) + \Phi_1 W(t)Z(t-1) + a(t). \]

(2)

Equation (2) can be written in matrix form:

\[
\begin{pmatrix}
Z_1(t) \\
Z_2(t) \\
Z_3(t)
\end{pmatrix} = \begin{pmatrix}
\phi_{10}^{(1)} & 0 & 0 \\
0 & \phi_{20}^{(2)} & 0 \\
0 & 0 & \phi_{30}^{(3)}
\end{pmatrix}
\begin{pmatrix}
Z_1(t-1) \\
Z_2(t-1) \\
Z_3(t-1)
\end{pmatrix}
+ \begin{pmatrix}
\phi_{11}^{(1)} & 0 & 0 \\
0 & \phi_{21}^{(2)} & 0 \\
0 & 0 & \phi_{31}^{(3)}
\end{pmatrix}
\begin{pmatrix}
w_{12} & w_{13} \\
w_{21} & w_{23} \\
w_{31} & w_{32}
\end{pmatrix}
\begin{pmatrix}
Z_1(t-1) \\
Z_2(t-1) \\
Z_3(t-1)
\end{pmatrix}
+ \begin{pmatrix}
a_1(t) \\
a_2(t) \\
a_3(t)
\end{pmatrix}.
\]

(3)

Equation (3) can be written like linear model:

\[ Y = X\beta + a. \]

(4)

So \( \beta \) can be obtained by OLS method given by:

\[ \hat{\beta}_{\text{GSTAR,OLS}} = (X'X)^{-1}X'Y. \]

(5)

If each locations are correlated, then estimations GSTAR parameters can be done using GLS method. Because its produce efficient estimators. This will be proven in simulation result. Estimated GSTAR parameters using GLS is [8]:

\[ \hat{\beta}_{\text{GSTAR-SUR}} = (X'\Omega X)^{-1}X'\Omega^{-1}Y. \]

(6)

Impact of weight on GSTAR is significant. Therefore weight that is used are uniform weight and normalization cross correlation.
2.2. Calendar variation and intervention

Calendar effect in time series was first researched by Liu (1980) and Cleveland and Devlin (1980). Liu discussed effects of calendar variations on ARIMA model identification and applied it to Taiwan highway traffic volumes [9]. Cleveland and Devlin identifies the effects of calendar variations with two methods, namely spectrum analyzes and time domain graphical displays. This method applies to international airline passengers series [10]. In 2010 Lee and Suhartono forecasted the sale of Muslim clothing from Muslim garment companies in Indonesia. This study used Eid Fitr as effect of calendar variation [11]. This is indicated by the high level of Bank Indonesia expenditure during Eid Fitr as shown in Figure 2. Date of Eid Fitr from 2003 to 2011 as shown in Table 1.

| Year | Date          | Year    | Date          |
|------|---------------|---------|---------------|
| 2003 | 25-26 November| 2008    | 01-02 October |
| 2004 | 14-15 November| 2009    | 21-22 September|
| 2005 | 03-04 November| 2010    | 10-11 September|
| 2006 | 23-24 October | 2011    | 30-31 August  |
| 2007 | 12-13 October |         |               |

Calendar variation model can be written as linear regression equation:

$$y_t = \beta_0 + \beta_1 V_{t1} + \beta_2 V_{t2} + \cdots + \beta_p V_{tp} + \epsilon_t,$$

with $V_{tp}$ as dummy variable for $p$ calendar variation effects. Identification of calendar variation effects can be seen from time series plot data. $\epsilon_t$ is error model that are identical, independent and normal distribution with zero mean and covariance $\sigma^2$.

The intervention model is a model used for external events beyond estimates as well as internal events that are expected to affect predicted variables. The general form of the intervention model is [12]:

$$Z_t = \frac{\omega_t(B)B^0}{\delta_t(B)} I_t + \eta_t,$$

with $\omega_t(B) = (\omega_0 - \omega_1 B - \omega_2 B^2 - \cdots - \omega_p B^p)$, $\delta_t(B) = (1 - \delta_1 B - \delta_2 B^2 - \cdots - \delta_p B^p)$. $I_t = \text{intervention variables}$, $\eta_t = \text{error model}$ that are identical, independent and normal distribution with mean 0 and covariance $\sigma^2$.$\eta_t$. In general, there are two types of intervention variables, which are step function and pulse function. The step function is an intervention that occurs over a long period of time. The notation of step function is:

$$I_t = I_s = \begin{cases} 0, & t < T, \\ 1, & t \geq T. \end{cases}$$

Pulse function is an intervention that affects the time of intervention and does not continue at a later time. Pulse function notation is given by:

$$I_t = I_p = \begin{cases} 0, & t \neq T, \\ 1, & t = T. \end{cases}$$

3. Result

There are two results in this research, which are simulation study and real case. Simulation studies are used to show that parameter estimation using GLS method is more efficient than OLS. Application of GSTAR-X SUR model for the accuracy of forecast of outflow data of rupiah in each province in Java of island.
3.1. Simulation
The simulation study was designed by generating two types of error, i.e., same variance, no correlation (1) and same variance, correlation (2) for three locations. Each variance covariance matrix and simulation parameter coefficients are:

\[
\Sigma_{(1)} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \Sigma_{(2)} = \begin{pmatrix} 1 & 0.5 & 0.45 \\ 0.5 & 1 & 0.55 \\ 0.45 & 0.55 & 1 \end{pmatrix} \quad \phi = \begin{pmatrix} 0.7 & 0 & 0 \\ 0 & 0.6 & 0 \\ 0 & 0 & 0.8 \end{pmatrix}
\]

The simulation process was iterated 1000 times. So that obtained 1000 parameters for each location. Based on Figure 1, if between each location are not correlated, the estimated value of GSTAR-OLS (Blue line) and GSTAR-SUR (Red line) parameters are same (as Figure 1 (a)). If between location are correlated then GSTAR-SUR more efficient. It is shown that standard deviation of GSTAR-SUR is smaller than GSTAR-OLS (as Figure 1 (b)).

3.2. Case study
Case study uses Rupiah outflow data in Java with five regions since January 2003 until December 2012, which are Jakarta, West Java, Central Java, Yogyakarta, and East Java. Data are divided into in-sample and out-sample. January 2003 until December 2011 as in-sample data and January 2012 until December 2012 as out-sample data. Figure 2 shows two kind of events, which are calendar variation effect on Eid Fitr and intervention in January 2007 and January 2010. Therefore forecasting Rupiah outflow are divided into two stages. Stage one is modeling calendar variation effect on Eid Fitr and intervention as predictors with regression GLS. R-square for five models are shown at Table 1. Residuals from stage one are input for stage two. Stage two builds GSTAR model for forecasting outflow data in 2012. Table 3 shows that there are high correlation between each locations. Identification of GSTAR model use cross correlation function (MCCF) matrix as Figure 3. It shows that autoregressive order is GSTAR(1,[1,12]). The first order of autoregressive is chosen with the consideration that the outflow data is affected by the conditions of the previous month. In this study, we use two kinds of weight, which are uniform and normalization cross correlation. Estimated GSTAR model parameters use two methods, which are OLS (GSTAR-OLS) and GLS (GSTAR-SUR). Criteria of model evaluation use Mean Absolute Percentage Error (MAPE).

| Table 2. R-Square Model for Five Location |
|------------------------------------------|
| Location | Jakarta | West Java | Central Java | Yogyakarta | East Java |
|----------|---------|-----------|--------------|------------|-----------|
| Jakarta  | 0.533   | 0.699     | 0.742        | 0.610      | 0.680     |

| Table 3. Correlation Between Location Outflow Rupiah |
|------------------------------------------------------|
| Location   | Jakarta | West Java | Central Java | Yogyakarta |
|------------|---------|-----------|--------------|------------|
| West Java  | 0.786   |           |              |            |
| Central Java | 0.747  | 0.890     |              |            |
| Yogyakarta | 0.736   | 0.758     | 0.806        |            |
| East Java  | 0.815   | 0.931     | 0.883        | 0.778      |
Figure 1. Histogram Parameters $\phi_{10}$, GSTAR(1,1) with 1000 iterations,
(a) Type Variance Covariance Matrix 1 and (b) Type Variance Covariance Matrix 2
Table 4 shows in general that the best model is GSTARX-SUR (1, [1,12]) with normalization cross correlation weight. This is indicated by the three smallest MAPE value.
4. Conclusion

Based on the result of research, it is found that simulation of GSTARX-SUR estimator is more efficient than GSTAR-OLS. This is indicated by the standard error value of GSTARX-SUR estimators are smaller than GSTAR-OLS in cases residuals correlated between locations. In the case study of rupiah outflow forecast in five provinces in Java with effect of calendar variation (Eid Fitr) and intervention (government policy January 2007 and January 2010), GSTAR-X SUR with normalization weights are more accurate in West Java, Central Java and Yogyakarta provinces. GSTAR-X OLS with normalization weight produces more accurate in East Java province. GSTAR-X OLS with uniform weight produces more accuracy in the province of Jakarta. This is indicated by a lower MAPE value. Further research is development of GSTAR-X model by considering exogenous variable that are metric. In addition, other comparative studies in other areas of forecasting need to validate the proposed model.

5. Acknowledgments

I would like to thanks to The Directorate General of Higher Education, The Ministry of Education and Culture of the Republic of Indonesia for BPPS grant and Intitut Teknologi Sepuluh Nopember for financial support.
References

[1] Mukhaiyar U and Pasaribu U S 2012 A New Procedure for Generalized STAR Modeling using IAcM Approach, *ITB Journal of Science*, vol. 44 A, no. 2, pp. 179-192.

[2] Ruchjana B N 2002 *A Generalized Space-Time Autoregressive Model and Its Application to Oil Production*, Ph.D. Thesis, ITB, Bandung.

[3] Ruchjana B N, Abdullah A S, Toharudin T, and Jaya I G N M 2013 Clustering Spatial on the GSTAR Model for Replacement New Oil Well, in *Padjadjaran International Physics Symposium 2013 (PIPS-2013)*, Bandung, pp. 205-210.

[4] Cliff A D and Ord J K 1975 Model Building and the Analysis of Spatial Pattern in Human Geography, *Journal of the Royal Statistical Society, Series B (Methodological)*, vol. 37, no. 3, pp. 297-348.

[5] Nurhayati N, Pasaribu U S and Neswan O 2012 Application of Generalized Space-Time Autoregressive Model on GDP Data in West European Countries, *Journal of Probability and Statistics*, vol. 2012, pp. 1-16.

[6] Ruchjana B N, Borovkova S A, and Lopuhaa H P 2012, Least Squares Estimation of Generalized Space Time Autoregressive (GSTAR) Model and Its Properties, *The 5th International Conference on Research and Education in Mathematics*, pp. American Institute of Physics pp. 61-64.

[7] Suhartono, Wahyuningrum S R, Setiawan, and Akbar M S 2016 GSTARX-GLS Model for Spatio-Temporal Data forecasting, *Malaysian Journal Of Mathematical Sciences*, vol. 10(S), pp. 91-103.

[8] Setiawan, Suhartono, and Prastuti M. 2016 S-GSTAR-SUR Model for Seasonal Spatio Temporal Data Forecasting, *Malaysian Journal of Mathematical Sciences*, vol. 10(S), pp. 53-65.

[9] Liu L. M 1980 Analysis of Time Series with Calender Effects,” *Management Science*, Vol.26, No.1.

[10] Cleveland W S and Devlin S J 1980 Calendar Effects in Monthly Time Series: Detection by Spectrum Analysis and Graphical Methods, *Journal of the American Statistical Association*, vol. 75, no. 371, pp. 487-496.

[11] Lee M H, Suhartono, and Hamzah N A 2010 Calender Variation Model Based on ARIMAX for Forecasting Sales Data with Ramadhan Effect, *Proceedings of The Regional Conference on Statistical Sciences 2010 (RCSS’10)*, Malaysia Institute of Statistics, Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA (UiTM), Malaysia, pp. hal349-361.

[12] Wei W W 2006 *Time Series Analysis Univariate and Multivariate* (Second Edition). USA: Pearson Education, Inc.,