Malaysian Journal of Mathematical Sciences 10(S) March : 91-103 (2016)
Special Issue: The 10th IMT-GT International Conference on Mathematics, Statistics and its Applications 2014 (ICMSA 2014)

MALAYSIAN JOURNAL OF MATHEMATICAL SCIENCES

Journal homepage: http://einspem.upm.edu.my/journal

GSTARX-GLS Model for Spatio-Temporal Data Forecasting

Suhartono*, Sri Rizqi Wahyuningrum, Setiawan and Muhammad Sjahid Akbar

Department of Statistics, Institut Teknologi Sepuluh Nopember, Kampus ITS Sukolilo Surabaya 60111, Indonesia

E-mail: suhartono@statistika.its.ac.id

*Corresponding author

ABSTRACT

Up to now, there have not been found a research about Generalized Space Time Autoregressive (GSTAR) that involve predictor. In fact, forecasting model in many cases involved predictor(s) both in univariate and multivariate cases such as ARIMAX and VARIMAX models. Moreover, most research about GSTAR models used Ordinary Least Squares (OLS) methods to estimate the parameters model. In many cases, the residuals of GSTAR model have correlation between locations and imply OLS method yields inefficient estimators. Otherwise, Generalized Least Squares (GLS) method that usually be used in Seemingly Unrelated Regression (SUR) model is an appropriate method for estimating parameters of multivariate models when the residuals between equations are correlated. The aim of this study is to propose GSTARX model with GLS method for estimating the parameters, known as GSTARX-GLS model. This research focuses on non metric predictor known as intervention variable. Theoretical study was carried out to develop new model building procedure for GSTARX-GLS model and the results were validated by simulation study. Then, the proposed model was applied for inflation forecasting at several cities in Indonesia. The results showed that GSTARX-GLS model yielded more efficient estimators than the GSTARX-OLS model. It was proved by the smaller standard error of GSTARX-GLS estimator. Additionally, GSTARX-GLS and GSTARX-OLS models gave more accurate inflation prediction in four cities in Indonesia than VARIMAX model.

Keywords: GSTARX, GLS, Predictor, Intervention, Spatio-temporal, Inflation.

1. Introduction

Due to computational advances and many open problems in forecasting could not be answered by univariate forecasting methods, it causes a lot of researches have been done on multivariate forecasting methods in recent years (De Gooijer and Hyndman (2006)). One of multivariate forecasting methods that frequently used in practical problem is VARIMA (Vector Autoregressive Moving Average). In daily activities, we often deal multivariate time series data that have relationship not only in time (with previous observations), but also in space (with observations at other location), known as spatio-temporal data (Ruchjana (2002)). Pfeifer and Deutsch (1980a, 1980b) are researchers who firstly introduce the space-time model, i.e. Space-Time Autoregressive or STAR model.

STAR model has disadvantage on the parameters flexibility that describes the relationship between space and time at spatio-temporal data. This limitation has been corrected by Ruchjana (2002) through a model known as Space-Time Model Generalized Autoregressive or GSTAR. Some researches about GSTAR have been done in many fields of application such as air pollution forecasting (Wutsqa and Suhartono (2010)) and tourism prediction (Wutsqa *et al*. (2010)).

In practice, forecasting activity both in univariate and multivariate cases frequently involves predictors. In time series analysis literatures, forecasting model which consists of predictors, usually notified by X, called ARIMAX (for univariate case) and VARIMAX (for multivariate case). Specifically, if predictors are metric then the ARIMAX is known as Transfer Function model (Box *et al*. (1994)), and for non metric predictors known as Intervention Analysis (Ismail *et al*. (2009); Lee *et al*. (2010)) or Calendar Variation models (Liu (2006)).

Literature survey showed that until now there is no research about space-time model that involves predictor variables. Moreover, most of GSTAR researches employed Ordinary Least Square (OLS) method to estimate the parameters model. OLS method assumes that residual of the model satisfies white noise and normally distributed condition. It means that residual in certain location has no correlation with residual in other locations. Unfortunately, the residual of GSTAR model in many cases tends to have correlation between locations, and it implies OLS yields inefficient estimators. Otherwise, Generalized Least Squares (GLS) is a parameter estimation method which could overcome the problem of correlation between residuals in different equations (locations). This method is usually applied to the Seemingly Unrelated Regression (SUR) model. The objective of this research is to develop GSTARX (GSTAR with a predictor) model for spatio-temporal data forecasting by implementing GLS method hereinafter written by GSTARX-GLS. This research focuses on non metric predictor variables. As acase study, the GSTARX-GLS model is applied for forecasting inflation in four major cities in Indonesia, i.e. Surabaya, Malang, Jember and Kediri, whereas the increase in fuel prices and Eid holiday as non metric predictors.

2. Methods

In this section, we describe the statistical method that is used for statistical estimations.

2.1 GSTAR Model

GSTAR is a generalization of the STAR models. Let $\{Z(t): t = 0, \pm 1, \pm 2, \cdots\}$ is a multivariate time series of *N* locations, then GSTAR with time order *p* and spatial order $\lambda_1, \lambda_1, \cdots, \lambda_p$, i.e. GSTAR$(p; \lambda_1, \lambda_1, \cdots, \lambda_p)$, in matrix notation can be written as follows (see Wutsqa *et al*., 2010):

$$\mathbf{Z}(t) = \sum_{s=1}^{p} \left(\mathbf{\Phi}_{s0} + \sum_{k=1}^{\lambda_s} \mathbf{\Phi}_{sk} \mathbf{W}^{(k)} \right) \mathbf{Z}(t-s) + \mathbf{e}(t) \quad (1)$$

where $\mathbf{\Phi}_{s0} = \text{diag}(\phi_{10}^{s}, \cdots, \phi_{N0}^{s})$, $\mathbf{\Phi}_{sk} = \text{diag}(\phi_{1k}^{s}, \cdots, \phi_{Nk}^{s})$, $\mathbf{e}(t)$ is residual model that satisfies identically, independent, distributed with mean $\mathbf{0}$ and covariance $\mathbf{\Sigma}$. For instance, GSTAR model with time and spatial order one for three locations $(p = 1,\ \lambda_p = 1,\ N = 3)$, is as follows:

$$\mathbf{Z}(t) = \mathbf{\Phi}_{10} \mathbf{Z}(t-1) + \mathbf{\Phi}_{11} \mathbf{W}^{(1)} \mathbf{Z}(t-1) + \mathbf{e}(t) \quad (2)$$

where:

$$\mathbf{Z}(t) = \begin{pmatrix} Z_1(t) \\ Z_2(t) \\ Z_3(t) \end{pmatrix},\ \mathbf{Z}(t-1) = \begin{pmatrix} Z_1(t-1) \\ Z_2(t-1) \\ Z_3(t-1) \end{pmatrix},\ \mathbf{e}(t) = \begin{pmatrix} e_1(t) \\ e_2(t) \\ e_3(t) \end{pmatrix},\ \mathbf{W}^{(1)} = \begin{pmatrix} 0 & w_{12} & w_{13} \\ w_{21} & 0 & w_{23} \\ w_{31} & w_{32} & 0 \end{pmatrix},$$

$$\mathbf{\Phi}_{10} = \text{diag}(\phi_{10}, \phi_{20}, \phi_{30}),\ \mathbf{\Phi}_{11} = \text{diag}(\phi_{11}, \phi_{21}, \phi_{31}).$$

Thus, equation (2) can be written in matrix form as follows:

$$\begin{pmatrix} Z_1(t) \\ Z_2(t) \\ Z_3(t) \end{pmatrix} = \begin{pmatrix} \phi_{10} & 0 & 0 \\ 0 & \phi_{20} & 0 \\ 0 & 0 & \phi_{30} \end{pmatrix} \begin{pmatrix} Z_1(t-1) \\ Z_2(t-1) \\ Z_3(t-1) \end{pmatrix} + \begin{pmatrix} \phi_{11} & 0 & 0 \\ 0 & \phi_{21} & 0 \\ 0 & 0 & \phi_{31} \end{pmatrix} \begin{pmatrix} 0 & w_{12} & w_{13} \\ w_{21} & 0 & w_{23} \\ w_{31} & w_{32} & 0 \end{pmatrix} \begin{pmatrix} Z_1(t-1) \\ Z_2(t-1) \\ Z_3(t-1) \end{pmatrix} + \begin{pmatrix} e_1(t) \\ e_2(t) \\ e_3(t) \end{pmatrix}. \quad (3)$$

There are several matrices of spatial weights or **W** that usually used in GSTAR model, i.e. uniform weight, weight based on inverse of distance between locations, weight based on normalization of cross correlation inference, and weight based on normalization of partial cross correlation inference (Suhartono and Subanar, 2006); Wutsqa *et al*., 2010).

2.2 Parameter Estimation

As in linear regression model, the estimator of GSTAR model could be obtained from OLS method by minimizing the sum of squares error, i.e. minimizing $\mathbf{e}'\mathbf{e} = (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})$. Thus, the OLS estimators $\boldsymbol{\beta}$ are as follows:

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y} \,. \tag{4}$$

In many cases, the residuals of GSTAR are correlated between locations and imply the OLS estimators become inefficient.

Otherwise, Generalized Least Squares (GLS) is an estimation method which could overcome the problem of correlation between residuals in different equations (locations). This method is usually applied to the Seemingly Unrelated Regression (SUR) model. SUR model consists of several equations and the relationships between variables are not in two-way relation, and there are correlations between equations that imply the residuals also have correlation between equations (Zellner (1962)). SUR models with M the dependent variables could be written as follows:

$$\mathbf{Z}_i = \mathbf{X}_i\boldsymbol{\beta}_i + \mathbf{e}_i \,, \quad i = 1, 2, \cdots, M \tag{5}$$

where $\mathbf{Z}_i$ is vector $T\times1$ of the sequences observation of the dependent variables, $\mathbf{X}_i$ is an observation matrix $T\times k$ of the independent variables, $\boldsymbol{\beta}_i$ is the parameter vector $k\times1$, and $\mathbf{e}_i$ is the residual vector $T\times1$. Equation (5) can also be written as follows:

$$\begin{pmatrix} \underset{(T\times1)}{\mathbf{Z}_1} \\ \underset{(T\times1)}{\mathbf{Z}_2} \\ \vdots \\ \underset{(T\times1)}{\mathbf{Z}_M} \end{pmatrix} = \begin{pmatrix} \underset{(T\times k)}{\mathbf{X}_1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \underset{(T\times k)}{\mathbf{X}_2} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \underset{(T\times k)}{\mathbf{X}_M} \end{pmatrix} \begin{pmatrix} \underset{(k\times1)}{\boldsymbol{\beta}_1} \\ \underset{(k\times1)}{\boldsymbol{\beta}_2} \\ \vdots \\ \underset{(k\times1)}{\boldsymbol{\beta}_M} \end{pmatrix} + \begin{pmatrix} \underset{(T\times1)}{\mathbf{e}_1} \\ \underset{(T\times1)}{\mathbf{e}_2} \\ \vdots \\ \underset{(T\times1)}{\mathbf{e}_M} \end{pmatrix}. \tag{6}$$

This equation is a SUR model which assuming $E[\mathbf{e} | \mathbf{X}_1, \mathbf{X}_2, \mathbf{L}, \mathbf{X}_M] = \mathbf{0}$ and $E[\mathbf{ee}' | \mathbf{X}_1, \mathbf{X}_2, \cdots, \mathbf{X}_M] = \mathbf{\Omega}$, where $\mathbf{\Omega}$ is the variance-covariance matrix, i.e. $\mathbf{\Omega} = \mathbf{\Sigma} \otimes \mathbf{I}$ (see Greene, 2002).

3. Research Design

Three studies are conducted in this research, i.e. the theoretical study on the development of GSTARX model building procedure using GLS estimation, simulation studies to validate the proposed modeling procedure, and applied study on forecasting inflation in four cities in Indonesia. Theoretical studies focus on determining the appropriate statistics that can be used to identify the order of GSTARX model and data structure on the GLS estimation. Simulation studies are designed for generating six GSTARX models and three scenarios of the effect of intervention as illustrated at Figure 1 and 2.

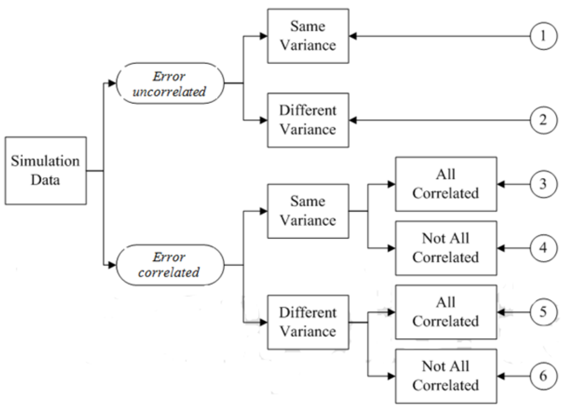

Figure 1: Design of Simulation Studies

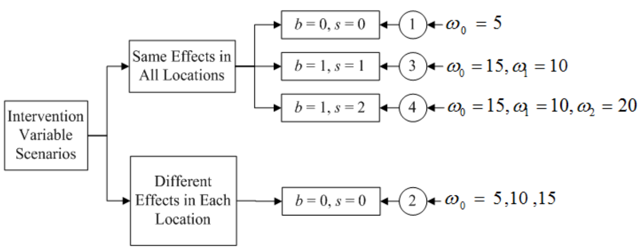

Figure 2: Three Scenarios of the Effect of InterventionVariable

In applied study, monthly inflation in four major cities at East Java Province, i.e. Surabaya, Malang, Jember and Kediri in the period 2000-2013 are used as case study. The data obtained from the Indonesian Central Bureau of Statistics. Both of an increase in the price of fuel and the presence of Eid during this period are used as predictor variables, i.e. intervention and dummy variable, respectively. Time series plot of monthly inflation in that four cities could be shown at Figure 3.

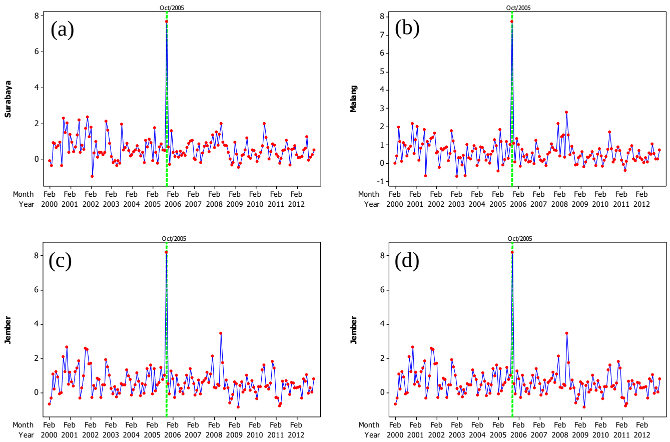

Figure 3: Time Series Plot of Inflation Data in (a) Surabaya, (b) Malang, (c) Jember, and (d) Kediri

4. Results

This section presents the results from theoretical study particularly about how to estimate the model parameter by using GLS method, simulation study about the advantage of GLS estimator, and applied studies about the forecast accuracy of GSTARX-GLS model for inflation forecasting, respectively.

4.1 Parameter Estimation $\hat{\boldsymbol{\beta}}$ of GSTARX-GLS Model

If $\mathbf{Y}_{i,t} = \mathbf{Z}_i(t)$ and a series $\{\mathbf{Z}(t): t = 0, \pm 1, \pm 2, \cdots, T\}$ is a multivariate time series of N locations, then GSTARX model with first order autoregressive, $p = 1$, spatial order 1, and the intervention order $b = r = s = 0$, i.e. GSTARX(1_1), can be written as follows:

$$\mathbf{Z}(t) = (\boldsymbol{\Phi}_0 + \boldsymbol{\Phi}_1 \mathbf{W}^1)\mathbf{Z}(t-1) + \boldsymbol{\beta}_{i,Int}\mathbf{P}_t + \mathbf{e}(t)\,. \tag{7}$$

The estimated parameters of GSTARX-GLS model obtained from:

$$\hat{\boldsymbol{\beta}} = \left(\mathbf{X}'\boldsymbol{\Omega}^{-1}\mathbf{X}\right)^{-1}\mathbf{X}'\boldsymbol{\Omega}^{-1}\mathbf{Y} \tag{8}$$

where $\boldsymbol{\Omega} = \boldsymbol{\Sigma} \otimes \mathbf{I}_T$, $\boldsymbol{\Sigma}$ = variance covariance matrix of size (N×N), i.e.

$$\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1N} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{N1} & \sigma_{N1} & \cdots & \sigma_{NN} \end{pmatrix}$$

and $\mathbf{I}_T$ = identity matrix of size (T×T). By using matrix representation, as an equation (7) then for each i location could be written asfollows:

$$\mathrm{Y}_{i,t} = \mathrm{Z}_i(t) = \begin{pmatrix} Z_i(1) \\ Z_i(2) \\ \vdots \\ Z_i(T) \end{pmatrix} \text{ and } \mathbf{X}_{i,t} = \left(\mathbf{Z}_i(t-1) \quad \mathbf{V}_i(t-1) \quad \mathbf{P}_i^{(T)}(t)\right), \ \boldsymbol{\beta}_i = \begin{pmatrix} \phi_{i0} \\ \phi_{i1} \\ \beta_{i,INT} \end{pmatrix},$$

where $i = 1,2,3,\cdots,N$.

4.2 Results of Simulation Study

In general, the simulated GSTARX(1_1) model with intervention order $b = r = s = 0$ could be written as follows:

$$\mathbf{Z}(t) = \left(\boldsymbol{\Phi}_0 + \boldsymbol{\Phi}_1\mathbf{W}^1\right)\mathbf{Z}(t-1) + \boldsymbol{\beta}_{i,Int}\mathbf{P}_t + \mathbf{e}(t)$$

and the matrix representation is

$$\begin{pmatrix} Z_1(t) \\ Z_2(t) \\ Z_3(t) \end{pmatrix} = \left[\begin{pmatrix} \phi_{10} & 0 & 0 \\ 0 & \phi_{20} & 0 \\ 0 & 0 & \phi_{30} \end{pmatrix} + \begin{pmatrix} \phi_{11} & 0 & 0 \\ 0 & \phi_{21} & 0 \\ 0 & 0 & \phi_{31} \end{pmatrix}\begin{pmatrix} 0 & w_{12} & w_{13} \\ w_{21} & 0 & w_{23} \\ w_{31} & w_{32} & 0 \end{pmatrix}\right]\begin{pmatrix} Z_1(t-1) \\ Z_2(t-1) \\ Z_3(t-1) \end{pmatrix} +$$

$$\begin{pmatrix} \beta_{11} & 0 & 0 \\ 0 & \beta_{21} & 0 \\ 0 & 0 & \beta_{31} \end{pmatrix}\begin{pmatrix} P_1^{(T)}(t) \\ P_2^{(T)}(t) \\ P_3^{(T)}(t) \end{pmatrix} + \begin{pmatrix} e_1(t) \\ e_2(t) \\ e_3(t) \end{pmatrix}$$

where $P_1^{(T)} = P_2^{(T)} = P_3^{(T)} = P^{(T)}$. The spatial weight that be used in the sixth simulation study is uniform weight for both OLS and GLS model.

The results of the first and second simulations (i.e. cases that no correlation between residuals at different locations) show that both GSTARX-OLS and GSTARX-GLS yield the same estimated parameters. It means that GSTARX-GLS estimator is GSTARX-OLS estimator for cases that residuals between locations have no correlation. Moreover, the GSTARX-GLS estimator for the third, fourth, fifth and sixth simulations yield lower standard error than GSTARX-OLS estimator. The comparison between GLS and OLS estimator of GSTARX model, particularly for scenario 3, could be seen in Figure 4 (for the spatio-temporal parameters).

Based on Figure 4, it can be seen that both estimators (OLS and GLS) are unbiased estimators. It is shown by the actual parameters (dashed blue line) are all inside of the plot (distribution of OLS and GLS estimators). Furthermore, the results also show that most of GLS estimators (red polygon) have lower standard error and indicate more efficient than OLS estimators (black polygon). It is illustrated by the distribution for GSTARX-GLS estimators (red polygon) tend to be narrower (or lower standard error) smaller than the distribution of GSTARX-OLS estimators (black polygon). Additionally, the results of the fourth, fifth, and sixth simulation studies also show the same conclusion with the third simulation study, i.e. GLS estimators have lower standard error than OLS estimators.

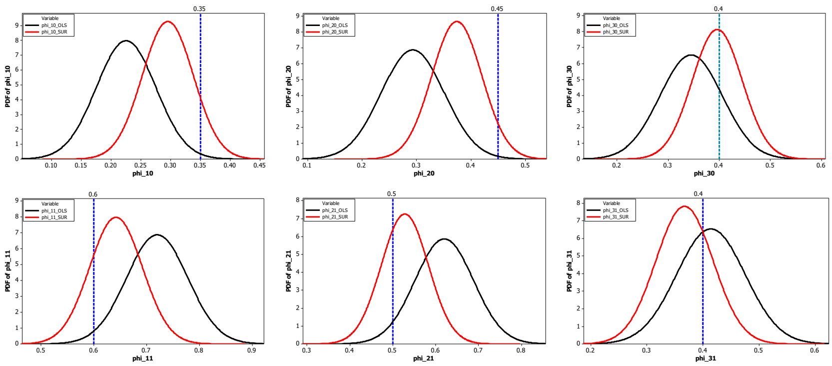

Figure 4: Distribution Plot of Spatio-Temporal Parameter in GSTARX Model on the 3rd simulation with Scenario3

4.3 Results of Applied Study

Figure 3 illustrate that the inflation data for 2000 to 2005 period in Surabaya, Malang, Jember and Kediri tend to be stable. However, in March and October 2005, there was high increase of inflation for all four locations due to an increase in fuel prices. The correlation between inflation in those four cities could be seen at Table1. It shows that all locations have high correlation with other locations.

TABLE 1: Correlation between Locations for Inflation Data

Location	Surabaya	Malang	Jember
Malang	0.855		
Jember	0.878	0.856	
Kediri	0.884	0.848	0.863

4.4 Determination of order b_l, s_l, r_l from Intervention Variable X

The order b_l, s_l, r_l of intervention variable X are determined by using response function plot as shown at Figure 5. This figure shows that the effect of interventions on inflation could be seen at the time T or (t=T). It means that the interventions at all locations have the order value of $b = s = r = 0$.

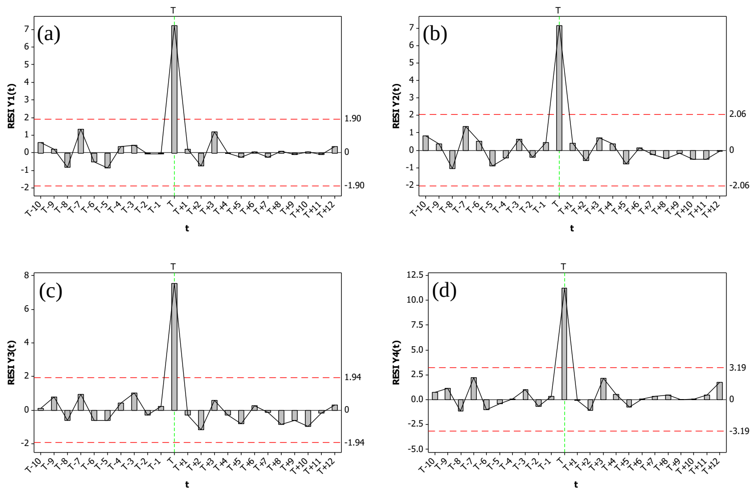

Figure5: Response Function Plot for Determining the Order of Intervention in GSTARX Model for Inflation at 4 Cities

4.5 Identification of the Autoregressive Order of GSTARX Model

As in VARIMAX model, identification of the autoregressive order is done by using plotMPCCF (Matrix Partial Cross Correlation Function) of stationary data and minimum value of AIC. Figure 6 and Table 2 illustrate the MPCCF and AIC values, respectively.

Variable/ Lag	1	2	3	4	5	6	7	8	9	10
y1	.+..	....	....	..+.	....	....	....	....	....	....
y2	....	....	....	....	...-	....	....	....	....	....
y3	....	....	....	..-.	....	....	....	....	....	....
y4	....	....	....	....	.-..	....	....	-...	....	....

Figure 6: MPCCF Plot of Inflation at FourCities

Based on the smallest AIC in Table 2, it shows that the best multivariate model involves the 1^{st} and 4^{th} order of autoregressive. Thus, the time order of the GSTARX model is GSTARX($[1,4]_1$). This study uses four types of spatial weight, i.e. uniform, based on inverse of distance, normalization of cross-correlation, and normalization of partial cross-correlation inference.

TABLE 2: The AIC of Several Tentative GSTARX Models

Lag	MA(0)	MA(1)
AR(0)	-5.656	-5.540
AR(1)	-6.051	-5.964
AR(2)	-5.983	-5.882
AR(3)	-6.211	-6.022
AR(4)	-6.325	-6.109
AR(5)	-6.218	-6.017

The best GSTARX-GLS model for inflation at four cities is using spatial weight based on normalization of partial cross-correlation inference, i.e.

$$\begin{pmatrix} Z_1(t) \\ Z_2(t) \\ Z_3(t) \\ Z_4(t) \end{pmatrix} = \left(\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0.118 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0.182 & 0 & 0 & 0 \\ 0 & 0.341 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.182 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0.322 & 0 & 0.345 & 0.333 \\ 0.333 & 0.333 & 0 & 0.333 \\ 0.333 & 0.333 & 0.333 & 0 \end{pmatrix} \right) \begin{pmatrix} Z_1(t-1) \\ Z_2(t-1) \\ Z_3(t-1) \\ Z_4(t-1) \end{pmatrix} +$$

$$\left(\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -0.149 & 0 \\ 0 & 0 & 0 & 0.118 \end{pmatrix} + \begin{pmatrix} 0.173 & 0 & 0 & 0 \\ 0 & 0.113 & 0 & 0 \\ 0 & 0 & 0.348 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0.333 & 0.333 & 0.333 \\ 0.333 & 0 & 0.333 & 0.333 \\ 0.333 & 0.333 & 0 & 0.333 \\ 0.333 & 0.333 & 0.333 & 0 \end{pmatrix} \right) \begin{pmatrix} Z_1(t-4) \\ Z_2(t-4) \\ Z_3(t-4) \\ Z_4(t-4) \end{pmatrix} +$$

$$\begin{pmatrix} 7.414 & 0 & 0 & 0 \\ 0 & 7.563 & 0 & 0 \\ 0 & 0 & 7.946 & 0 \\ 0 & 0 & 0 & 11.106 \end{pmatrix} \begin{pmatrix} P_1^{(T)}(t) \\ P_2^{(T)}(t) \\ P_3^{(T)}(t) \\ P_4^{(T)}(t) \end{pmatrix} + \begin{pmatrix} 0.692 & 0 & 0 & 0 \\ 0 & 0.428 & 0 & 0 \\ 0 & 0 & 0.806 & 0 \\ 0 & 0 & 0 & 0.964 \end{pmatrix} \begin{pmatrix} D_1(t) \\ D_2(t) \\ D_3(t) \\ D_4(t) \end{pmatrix} + \begin{pmatrix} e_1(t) \\ e_2(t) \\ e_3(t) \\ e_4(t) \end{pmatrix}.$$

Finally, the forecast accuracy between GSTAR-GLS, GSTAR-OLS and VARIMAX models are compared by implementing RMSE criteria at out sample data, i.e. January-December 2013. The result is shown in Table3.

Table 3 shows that the smallest RMSE for forecasting inflation in each city is yielded by different model. The best model for forecasting inflation in Surabaya obtained by GSTARX-GLS model with spatial weight based on normalized cross-correlation, whereas in Malang by GSTARX-GLS model with spatial weight based on normalized partial cross-correlation inference. Furthermore, the best model for forecasting inflation in Jember and Kediri is GSTARX-OLS models with spatial weight based on normalized partial cross-correlation inference and inverse distance, respectively.

TABLE 3: The Results of Forecast Accuracy Comparison between GSTAR-GLS, GSTAR-OLS and VARIMAX Models

Model	Spatial Weight	RMSE				Total RMSE
		Surabaya	Malang	Jember	Kediri	
VARIMAX		0.900	1.033	0.947	0.981	0.967
GSTARX-OLS	Uniform	0.808	0.946	0.834	0.956	0.889
	Inverse of distance	0.822	0.927	0.811	0.740*	0.828
	Normalized cross-correlation	0.831	1.112	0.847	0.742	0.894
	Normalized partial cross-correlation inference	0.713	0.923	0.800*	0.751	0.801*
GSTARX-GLS	Uniform	0.837	0.972	0.973	0.836	0.907
	Inverse of distance	0.828	1.196	0.876	0.813	0.941
	Normalized cross-correlation	0.666*	1.092	0.881	0.778	0.868
	Normalized partial cross-correlation inference	0.710	0.911*	0.886	0.782	0.826*

*Smallest RMSE

5. Conclusion

Based on the results of theoretical study it could be concluded that the model building of GSTARX-GLS has been proposed starting by identification step to determine the order of spatio-temporal and order of the effect (influence) of predictor variables. Moreover, the results of simulation study shows that estimators of GSTARX-GLS are more efficient than GSTAR-OLS, particularly shown by lower standard error of the estimators in case that residual between locations are correlated. Additionally, the empirical or applied study shows that GSTARX models yield more accurate forecast than VARIMAX model for forecasting inflation in four cities in Indonesia. It is shown by lower RMSE both of GSTARX-GLS and GSTARX-OLS than VARIMAX. Specifically, GSTARX-OLS yields more accurate forecast for inflation in Jember and Kediri, whereas GSTARX-GLS give more accurate forecast for inflation in Surabaya and Malang.

Further research is needed particularly for developing GSTARX model by involving both step function intervention and metric predictors as Transfer Function model. Moreover, other comparison study in other fields of forecasting is required to validate the proposed model.

References

Box, G.E.P., Jenkins, G.M. and Reinsel, G.C. (1994). *Time Series Analysis Forecasting and Control*. 3rded. Englewood Cliffs: Prentice Hall.

De Gooijer, J.G. and Hyndman, R.J. (2006). 25 years of time series forecasting. *International Journal of Forecasting*. **22**: 443-473.

Greene, W.H. (2002). *Econometric Analysis*. 5thed. New Jersey: Prentice Hall.

Ismail, Z., Suhartono, Yahaya, A. and Efendi, R. (2009). Intervention Model for Analyzing the Impact of Terrorism to Tourism Industry. *Journal of Mathematics and Statistics*. **5**: 322-329.

Lee, M.H., Suhartono and Sanugi, B. (2010). Multi input intervention model for evaluating the impact of the asian crisis and terrorist attacks on tourist arrivals. *Matematika*. **26**(1): 83-106.

Liu, L.M. (2006). *Time Series Analysis and Forecasting.* Illinois: Scientific Computing Associates.

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

Pfeifer, P.E. and Deutsch, S.J. (1980b). Identification and Interpretation of First Order Space-Time ARMA Models. *Technometrics*. **22**(1): 397-408.

Ruchjana, B.N. (2002). A Generalized Space-Time Autoregressive Model and Its Application to Oil Production, Unpublished Ph.D. Thesis, InstitutTeknologi Bandung.

Suhartono and Subanar. (2006).The Optimal Determination Of Space Weight in GSTAR Model by Using Cross-Correlation Inference. *Journal of Quantitative Methods*. **2**(2): 45-53.

Wutsqa, D.U. and Suhartono. (2010). Seasonal Multivariate Time Series Forecasting on Tourism Data by Using VAR-GSTAR Model. *Jurnal Ilmu Dasar*. **11**(1): 101-109.

Wutsqa, D.U., Suhartono and Ulama, B.S.S. (2010). Generalized Space-Time Autoregressive Modeling. *Proceedings of the 6th IMT-GT Conference on Mathematics, Statistics and its Application* (ICMSA 2010). Universiti Tunku Abdul Rahman, Malaysia.

Wei, W.W.S. (2006). *Time Series Analysis: Univarite and Multivariate*. 2nd ed. USA: Pearson Education Inc..

Zellner, A. (1962). An Efficient Methods of Estimation SUR and Test for Aggregation Bias. *Journal of American Statistical Association*. **57**: 348-368.