

Research Article

Application of Generalized Space-Time Autoregressive Model on GDP Data in West European Countries

Nunung Nurhayati,¹ Udjianna S. Pasaribu,² and Oki Neswan²

¹ Faculty of Science and Engineering, Jenderal Soedirman University, Purwokerto 53122, Indonesia

² Faculty of Mathematics and Natural Sciences, Bandung Institute of Technology, Bandung 40132, Indonesia

Correspondence should be addressed to Nunung Nurhayati, nurhayati.nunung@gmail.com

Received 28 September 2011; Revised 1 February 2012; Accepted 14 February 2012

Academic Editor: Shein-chung Chow

Copyright © 2012 Nunung Nurhayati et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper provides an application of generalized space-time autoregressive (GSTAR) model on GDP data in West European countries. Preliminary model is identified by space-time ACF and space-time PACF of the sample, and model parameters are estimated using the least square method. The forecast performance is evaluated using the mean of squared forecast errors (MSFEs) based on the last ten actual data. It is found that the preliminary model is GSTAR(2;1,1). As a comparison, the estimation and the forecast performance are also applied to the GSTAR(1;1) model which has fewer parameter. The results showed that the ASFE of GSTAR(2;1,1) is smaller than that of the order (1;1). However, the t -test value shows that the performance is significantly indifferent. Thus, due to the parsimony principle, the GSTAR(1;1) model might be considered as a forecasting model.

1. Introduction

Space-time data are frequently found in many areas of research, for example, monthly tea production from some plants, yearly housing price at capital cities, and yearly per capita GDP (gross domestic product) of several countries in some region. The generalized space-time autoregressive model of order $(p; \lambda_1, \dots, \lambda_p)$, shortened by GSTAR($p; \lambda_1, \dots, \lambda_p$), is one of space-time models characterized by autoregressive terms lagged in the p th order in time and the order of $(p; \lambda_1, \dots, \lambda_p)$ in space [1].

The term of generalization is associated with the model parameters. When a parameter matrix is diagonal, the GSTAR model is the same as space time autoregressive (STAR) model given by Martin and Oeppen [2] and Pfeifer and Deutsch [3]. The notion of generalization has also been used by Terzi [4] who also generalizes STAR models but in a different context. He

generalized the STAR(1;1) by adding the contemporaneous spatial correlation but still preserved the scalar parameters.

When $p = 1$ and $\lambda_1 = 1$, GSTAR(1;1) is called the first order of GSTAR model. The model has interpretation that the current observation in a certain location only depends on the immediate past observations recorded at the location of interest and at its nearest neighbourhood [5]. The order (1;1) is the simplest natural assumption if one wants to forecast future observations in a certain location. The STAR(1;1) model is another simple space time model which also has the same interpretation as GSTAR(1;1) model. However, contrary to the generalization model, its parameters of each spatial order are assumed to be the same for all location though; there is no *a priori* justification for this assumption [1]. Parameters of GSTAR model can be estimated by the method of least square. This method has been used to model the monthly oil production [5] and to model the monthly tea production [1]. However, when the model was applied to their data, none of the papers included a description about how to assess the model based on forecasting performance which is an important step when the modelling purpose is to build a forecasting model. In this paper, we attempt to put in the idea to optimize the goodness of fit in model selection.

This paper is presented as follows. In Section 2, the GSTAR model and the parameters least squares estimation is reviewed with the example given for GSTAR(1;1) and GSTAR(2;1,1). To illustrate the estimator properties for finite sample size, simulation study is discussed in Section 3. In the last section, the model is applied to the per capita GDP ratio data in West European countries for the period 1956–1996. The one step ahead forecasting is performed for each model for the period 1997–2006. As a comparison performance measure, it is used the empirical mean of squared forecast error (MSFE) where forecast error is defined as the difference between the actual value and the forecast value.

2. The Model

Let $\mathbf{Z}(t) = (Z_1(t), \dots, Z_N(t))'$ be an N -dimensional vector process with zero mean with N as is a fixed positive integer. GSTAR($p; \lambda_1, \dots, \lambda_p$) process is a space-time process $\mathbf{Z}(t)$ which satisfies

$$\mathbf{Z}(t) = \sum_{k=1}^p \sum_{\ell=0}^{\lambda_k} \Phi_{k\ell} \mathbf{W}^{(\ell)} \mathbf{Z}(t-k) + \mathbf{e}(t), \quad (2.1)$$

where p is the autoregressive order, λ_k is the spatial order of the k th autoregressive term, $\mathbf{W}^{(\ell)} = (w_{ij}^{(\ell)})$ is an $N \times N$ matrix of spatial weight for the spatial order ℓ which has a zero diagonal, sum of each row is equal to one, and matrix $\mathbf{W}^{(0)}$ is defined as the identity matrix \mathbf{I} . An $N \times N$ matrix $\Phi_{k\ell}$ is a diagonal parameter matrix of temporal lag k and spatial lag ℓ with the diagonal element $(\phi_{k\ell}^{(1)}, \dots, \phi_{k\ell}^{(N)})$. Finally, $\mathbf{e}(t)$ is an error vector at time t which is assumed to be independent normal with zero mean and constant variance.

Model parameters $\phi_{k\ell}^{(1)}, \dots, \phi_{k\ell}^{(N)}$ for $k = 1, 2, \dots$ and $\ell = 1, 2, \dots$ can be estimated by the least squares (LS) estimation. The procedure estimation and the asymptotic properties of LS estimators have been discussed extensively in Borovkova et al. [1]. The following examples are an illustration on how to find the LS estimator for GSTAR(1;1) and GSTAR(2;1,1) models, respectively.

Example 2.1 (LS estimation for GSTAR(1;1) model). From (2.1), GSTAR(1;1) model can be expressed as

$$\mathbf{Z}(t) = \Phi_{10}\mathbf{Z}(t-1) + \Phi_{11}\mathbf{W}^{(1)}\mathbf{Z}(t-1) + \mathbf{e}(t). \quad (2.2)$$

Suppose observation $\mathbf{Z}(t) = (Z_1(t), \dots, Z_N(t))'$ is given for $t = 1, 2, \dots, T$. By rearranging the component of $\mathbf{Z}(t)$ for $t = 1, 2, \dots, T$, by location then by time, model (2.2) can be expressed as a linear model

$$\mathbf{Z} = \mathbf{X}\Phi + \mathbf{e}, \quad (2.3)$$

where $\mathbf{Z} = (Z_1(1), \dots, Z_1(T), \dots, Z_N(1), \dots, Z_N(T))$, $\mathbf{X} = \text{diag}(\mathbf{X}_1, \dots, \mathbf{X}_N)$ with

$$\mathbf{X}_i = \begin{pmatrix} Z_i(0) & \sum_{j \neq i} w_{ij} Z_j(0) \\ Z_i(1) & \sum_{j \neq i} w_{ij} Z_j(1) \\ \vdots & \vdots \\ Z_i(T-1) & \sum_{j \neq i} w_{ij} Z_j(T-1) \end{pmatrix}, \quad (2.4)$$

$\Phi = (\phi_{10}^{(1)}, \phi_{11}^{(1)}, \dots, \phi_{10}^{(N)}, \phi_{11}^{(N)})'$, and $\mathbf{e} = (e_1(1), \dots, e_1(T), \dots, e_N(1), \dots, e_N(T))'$.

Then, the LS estimator for parameter matrix Φ is a solution of

$$(\mathbf{X}'\mathbf{X})\Phi = (\mathbf{X}'\mathbf{Z}). \quad (2.5)$$

Example 2.2 (LS estimation for GSTAR(2;1,1) model). For order $p = 2$, $\lambda_1 = 1$ and $\lambda_2 = 1$, model (2.1) is called GSTAR(2;1,1) which can be written as

$$\mathbf{Z}(t) = \Phi_{10}\mathbf{Z}(t-1) + \hat{\Phi}_{11}\mathbf{W}^{(1)}\mathbf{Z}(t-1) + \Phi_{20}\mathbf{Z}(t-2) + \Phi_{21}\mathbf{W}^{(2)}\mathbf{Z}(t-2). \quad (2.6)$$

By rearranging the component of $\mathbf{Z}(t)$ for $t = 2, 3, \dots, T$, by location then by time, model (2.6) can also be expressed as a linear model

$$\mathbf{Z} = \mathbf{X}\Phi + \mathbf{e}, \quad (2.7)$$

where \mathbf{Z} and \mathbf{e} are defined as in Example 2.1, $\mathbf{X} = \text{diag}(\mathbf{X}_1, \dots, \mathbf{X}_N)$ with

$$\mathbf{X}_i = \begin{pmatrix} Z_i(1) & \sum_{j \neq i} w_{ij}^{(1)} Z_j(1) & Z_i(0) & \sum_{j \neq i} w_{ij}^{(2)} Z_j(0) \\ Z_i(2) & \sum_{j \neq i} w_{ij}^{(1)} Z_j(2) & Z_i(1) & \sum_{j \neq i} w_{ij}^{(2)} Z_j(1) \\ \vdots & \vdots & \vdots & \vdots \\ Z_i(T-1) & \sum_{j \neq i} w_{ij}^{(1)} Z_j(T-1) & Z_i(T-2) & \sum_{j \neq i} w_{ij}^{(2)} Z_j(T-2) \end{pmatrix} \quad (2.8)$$

and $\Phi = (\phi_{10}^{(1)}, \phi_{11}^{(1)}, \phi_{20}^{(1)}, \phi_{21}^{(1)}, \dots, \phi_{10}^{(N)}, \phi_{11}^{(N)}, \phi_{20}^{(N)}, \phi_{21}^{(N)})'$.

Then, the LS estimator for parameter matrix Φ is a solution of

$$(\mathbf{X}'\mathbf{X})\Phi = (\mathbf{X}'\mathbf{Z}). \quad (2.9)$$

3. Simulation Study

Under the stationary assumption, the LS estimator for the GSTAR parameters is a consistent estimator [1]. To get insight into the LS properties for finite sample we performed a Monte Carlo simulation with 1000 replications. Artificial data were generated from GSTAR(1;1) model where the error $\mathbf{e}(t)$ is normally distributed with mean $\mathbf{0}$ and covariance matrix \mathbf{I}_4 . Spatial weight matrix and model parameters that used in the simulation, respectively, were

$$\mathbf{W}^{(1)} = \begin{pmatrix} 0 & 0.5 & 0.5 & 0 \\ 0.5 & 0 & 0 & 0.5 \\ 0.5 & 0 & 0 & 0.5 \\ 0 & 0.5 & 0.5 & 0 \end{pmatrix}, \quad (3.1)$$

$$\begin{aligned} \phi_{10}^{(1)} &= 0.2, & \phi_{10}^{(2)} &= 0.5, & \phi_{10}^{(3)} &= 0.3, & \phi_{10}^{(4)} &= 0.2, \\ \phi_{11}^{(1)} &= 0.4, & \phi_{11}^{(2)} &= 0.3, & \phi_{11}^{(3)} &= 0.5, & \phi_{11}^{(4)} &= 0.7 \end{aligned}$$

such that the parameter matrix for the generated model was

$$\Phi_{10} = \text{diag}(0.2, 0.5, 0.3, 0.2), \quad \Phi_{11} = \text{diag}(0.4, 0.3, 0.5, 0.7) \quad (3.2)$$

The LS estimator vector $\hat{\phi} = (\hat{\phi}_{10}^{(1)}, \hat{\phi}_{10}^{(2)}, \hat{\phi}_{10}^{(3)}, \hat{\phi}_{10}^{(4)}, \hat{\phi}_{11}^{(1)}, \hat{\phi}_{11}^{(2)}, \hat{\phi}_{11}^{(3)}, \hat{\phi}_{11}^{(4)})$ was calculated using R functions in Algorithm 1. Simulation was carried out for sample size $T = 40, 50, 100, 500, 1000, 10000$, and for each T , simulations were repeated 1000 times to obtain the estimates average

$$\bar{\hat{\phi}}_T = \frac{1}{1000} \sum_{r=1}^{1000} \hat{\phi}_T^{(r)} \quad (3.3)$$

and empirical mean squared error (MSE)

$$\text{MSE}_T = \frac{1}{1000} \sum_{r=1}^{1000} \left\| \hat{\phi}_T^{(r)} - \phi \right\|^2, \quad (3.4)$$

where $\left\| \hat{\phi}_T^{(r)} - \phi \right\|^2 = \sum_{i=1}^4 (\hat{\phi}_{10}^{(i)} - \phi_{10}^{(i)})^2 + (\hat{\phi}_{11}^{(i)} - \phi_{11}^{(i)})^2$.

The result is presented in Table 1. It can be seen that the parameters estimates (in average) approaches the true parameters as T increases while the empirical MSE is getting smaller and approaching 0 as T increasing. It means that behavior of the LS estimator in the simulation exhibits the consistent property. In general, we can notice that the LS estimation could give fair estimates even for moderate sample size such as $T = 40$ and $T = 50$.

```

# =====
# [FUNCTION]: OLS estimation for GSTAR(p;L1,..,Lp) models
# =====
# 3 dimension zeros matrix
# -----
zeros <-function(m,n,p){
W<-rep(0, m*n*p)
dim(W)<-c(m, n, p)
W}
# "vec" operator
vec<-function(X){
a<-dim(X)
Y<- t(X[1, ])
for (i in 2:a[1]){
Y<- cbind(Y,t(X[i, ]))}
t(Y)}
# -----
# Inverse of matrix
# -----
inv<-function(X){
if(dim(X)[1]!= dim(X)[2]) stop("THE MATRIX MUST BE SYMMETRIC!!!")
else{
if (det(X)==0) stop("THE MATRIX IS SINGULAR!!!")
else { n<-dim(X)[1]
solve(X)}}}
# -----
# Construction of vector Zi, for each i=1,..,N
# -----
# Construction of vector Zi, for each i=1,..,N
# -----
# suppose x = c(p,L1,.., Lp) represent the model order
Zi<-function(Zt, x){
N<-dim(Zt)[1] #number of sites
T<-dim(Zt)[2]-1 #number of time periods
p<-x [1]
Zi<-matrix(0, T-p+1, N)
for (i in 1: N)
Zi[,i]<-Zt[i,(p+1):(T+1)]
Zi}
# -----
# Construction of matrix Xi, for each i= 1,..,N
# -----
Xi<-function(Zt,x)
{N<-dim(Zt)[1] #number of sites
T<-dim(Zt)[2]-1 #number of time periods
p<-x[1]
La<-x[2: length(x)]
r<-lmd+1 # where lmd = the greatest order for weight matrices
WZ<-zeros(N, T, r)
for (k in 1: r)
WZ[, k]<-W [, k]% * %Zt[, 1: T]
Xi<-zeros((T-p+1),sum(La+1), N)
if (p==1)
{ for (i in 1:N){

```

Algorithm 1: Continued.

```

      TR<-WZ[i,p:T,1:(La[1]+1)]
      Xi[,i]<-TR}} if (p>=2){
    for (i in 1:N){
      TR<-WZ[i,p:T,1:(La[1]+1)]
      for (s in 2:p)
        TR<-cbind(TR,WZ[i,(p-s+1):(T-s+1),1:(La[s]+1)])
    Xi[,i]<-TR}}
Xi} # _____
# OLS parameter of GSTAR model
# _____
gstar<-function(Zt,x){
  p<-x[1]
  La<-x[2:length(x)]
  r<-lmd+1 # where lmd = the greatest order for weight matrices
  N<-dim(Zt)[1] #number of sites
  T<-dim(Zt)[2]-1 #number of time periods
  Xi<-Xi(Zt,x)
  Zi<-Zi(Zt,x)
  coef.OLS<-matrix(0,sum(La+1),N)
  col.name<-array(0,N)
  for (i in 1:N){
    coef.OLS[,i]<-inv(t(Xi[,i])%*%Xi[,i])%*%t(Xi[,i])%*%Zi[,i]
    col.name[i]<-paste("site",i)}
  colnames(coef.OLS)<-col.name
  round(coef.OLS,4)}
# _____
# Residuals of GSTAR model
# _____
# (1). To find the LS estimates only, for example GSTAR(2;1,1), use
# the command:
# > gstar(Zt,c(2,1,1))
# where Zt is data matrix.
# (2). To find the estimates, prediction values, and residuals
# vector respectively, call the function by the following
# commands:
# > as.2<-res(Zt,c(2,1,1))
# > as.2$coef
# > as.2$pred
# > as.2$res
# _____
res<-function(Zt,x){
  coef<-gstar(Zt,x)
  Xi<-Xi(Zt,x)
  p<-x[1]
  La<-x[2:length(x)]
  N<-dim(Zt)[1] #number of sites
  T<-dim(Zt)[2]-1 #number of time periods
  Z.OLS<-matrix(0,T-p+1,N)
  res.OLS<-matrix(0,N,T-p+1)
  if (p==1){
    for (i in 1:N){
      if (La[1]!=0)Z.OLS[i]<-Xi[,i]%*%coef[i]
      else Z.OLS[i]<-Xi[,i]*coef[i]
      res.OLS[i,<-t(Z.OLS[i]-Zt[i,(p+1):(T+1)])]}

```

Algorithm 1: Continued.

```

if (p!=1){
for (i in 1:N){
Z.OLS[,i]<-Xi[,i]% * %coef[,i]
res.OLS[i,]<-t(Z.OLS[,i]-Zt[i,(p+1):(T+1)])}
az<-new.env()
az$Xi<-Xi # matrix Xi
az$coef<-coef
az$pred<-t(Z.OLS)
az$res<-res.OLS
ax<-as.list(az)
# _____

```

Algorithm 1: R codes for least square estimation of GSTAR model.

Table 1: The LS estimated values (in average) for the data generated from GSTAR(1;1) model with 1000 replications for various sample sizes T compared to the theoretical parameter.

	$T = 40$	$T = 50$	$T = 100$	$T = 500$	$T = 1000$	$T = 10000$
$\phi_{10}^{(1)} = 0.2$	0.1799	0.1822	0.1891	0.1996	0.2004	0.2000
$\phi_{10}^{(2)} = 0.5$	0.4639	0.4684	0.4819	0.4971	0.4987	0.5002
$\phi_{10}^{(3)} = 0.3$	0.2786	0.2815	0.2942	0.2966	0.2983	0.2998
$\phi_{10}^{(4)} = 0.2$	0.1802	0.1828	0.1929	0.1983	0.1986	0.1998
$\phi_{11}^{(1)} = 0.4$	0.4000	0.4025	0.4028	0.3999	0.4002	0.4005
$\phi_{11}^{(2)} = 0.3$	0.2961	0.2944	0.2970	0.005	0.2998	0.2999
$\phi_{11}^{(3)} = 0.5$	0.4880	0.4906	0.4931	0.4988	0.5001	0.5002
$\phi_{11}^{(4)} = 0.7$	0.6989	0.6980	0.6961	0.7008	0.7002	0.7002
MSE	0.0279	0.0219	0.0105	0.0002	0.0001	0.0001

4. Application of GSTAR Model to the Ratio of Per Capita GDP Data

In this section, we apply the GSTAR model to the ratio of per capita GDP data in 16 West European countries. The data was kindly given by Maddison [6] from Faculty of Economics, University of Groningen, the Netherland, who passed away on April 24 2010. The per capita GDP of a country is the country GDP value divided by its population, and the per capita GDP of total West Europe is the sum of each West European country GDP divided by the total population in West Europe. The ratio of the per capita GDP of a country is the country per capita GDP value divided by the per capita GDP of total West Europe, multiplied by 100. Hence, the unit of the per capita GDP ratio is the percentage. For the data analysis in the following subsection, we will use the ratio of the per capita GDP data and for simplicity it will be called the GDP ratio data.

4.1. Dataset and Preliminary Model Building

The dataset is the GDP ratio data for periods 1955–2006. It consists of 52 observations of 16 dimensional vectors. For the purpose of forecasting the data was grouped into the training data

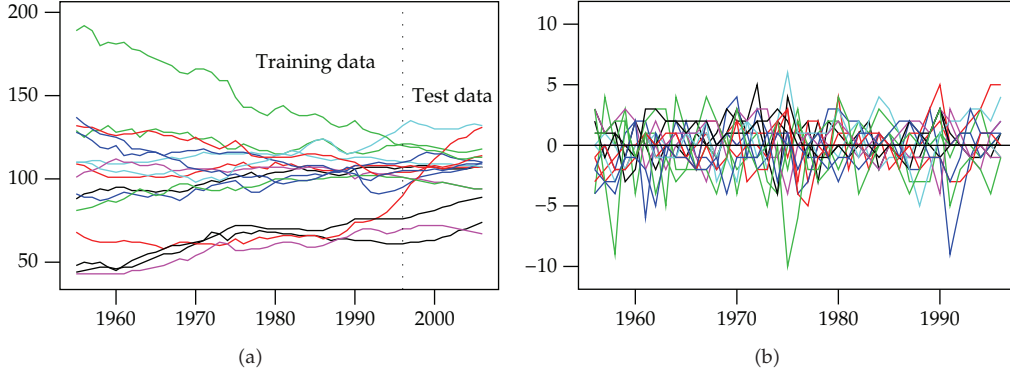


Figure 1: (a) The ratio of per capita GDP in 16 West European countries for the period 1955–2006 and (b) plot of the centralized data of the difference data.

set and test data set. The training data is the first 42 observations that will be used for model building and the test data is the last ten data that will be used in forecasting performance comparison.

Clearly, the 42 observations in the training data, depicted in Figure 1(a), are not stationary though they tend to converge to the value between 50% and 150%. Therefore, to achieve the zero mean stationary data, the first difference transformation and data centralization must be applied.

Suppose $Y_i(t)$ represents the per capita GDP ratio for country i , $i = 1, 2, \dots, 16$ at time t , $t = -1, 0, 1, \dots, 50$. The first difference transformation of $Y_i(t)$ denoted by $D_i(t)$ is

$$D_i(t) = Y_i(t) - Y_i(t - 1) \quad (4.1)$$

and the centralization of the differenced data is

$$Z_i(t) = D_i(t) - \bar{D}_i, \quad (4.2)$$

where \bar{D}_i is the average of $Z_i(t)$ at location i , for $i = 1, 2, \dots, 16$. Plot of the 16 series of this transformation is displayed in Figure 1(b) and their behaviour has seemed to represent stationary series.

As a preliminary model building, we set some notations used in model (2.1). The length of time period is $T = 40$, and the number of sites is $N = 16$. Time period $t = 0$ will correspond to 1956, $t = 1$ to 1957, $t = 2$ to 1958, and so on. Since the time dimension is one, the time lag can be ordered naturally by the sequence of $k = 1, 2, \dots$. On the other hand, the spatial order ℓ may be defined in a different ways because in a two-dimensional space there is no specific order just as in one-dimensional space. For the GDP data, there are 16 countries. The first order and the second order neighbours of the countries are given in Table 2. These are defined based on the geographical location of the countries. The first order neighbours of a country are those which have a common border with the country or within a close distance along a sea route. A second order neighbour of a country is the union of all first order neighbourhood countries of its first order neighbours, excluding itself.

Table 2: Geographical neighbourhood of order 1 and order 2.

No.	Country	Countries in the 1st neighborhood	Countries in the 2nd neighborhood
1	Austria	6,9,15	2,3,5,7,10,14
2	Belgium	5,6,10,16	1,3,8,9,13,14,15
3	Denmark	6,10,11,14	1,2,4,5,15,16
4	Finland	11,14	3,6
5	France	2,6,9,13,15,16	1,3,7,8,10,12,14
6	Germany	1,2,3,5,10,14,15	4,9,11,13,16
7	Greece	9	1,5,15
8	Ireland	16	2,5,10
9	Italy	1,5,7,15	2,6,13,16
10	Netherland	2,3,6,16	1,5,8,11,14,15
11	Norway	3,4,14	6,10
12	Portugal	13	5
13	Spain	5,12	2,6,9,15,16
14	Sweden	3,4,6,11	1,2,5,10,15
15	Switzerland	1,5,6,9	2,3,7,10,13,14,16
16	United Kingdom	2,5,8,10	3,6,9,13,15

Suppose $n_i^{(\ell)}$, $\ell = 1, 2$ represents the number of countries which are the ℓ th neighbours of a country i . The spatial weight of order ℓ between countries i and j can be defined as

$$w_{ij}^{(\ell)} = \begin{cases} \frac{1}{n_i^{(\ell)}}, & \text{if } j \text{ is the } \ell\text{th neighbours of } i, \\ 0, & \text{if } j \text{ is not the } \ell\text{th neighbours of } i. \end{cases} \quad (4.3)$$

For example, from Table 2 it can be seen that Austria has 3 the first order neighbours, Germany, Netherland and Switzerland, and has 6 second order neighbours, Belgium, Denmark, France, Greece, Netherland, and Sweden. Then, the first order of spatial weight between Austria and each nearby country is

$$w_{16}^{(1)} = w_{19}^{(1)} = w_{1,15}^{(1)} = \frac{1}{3} \quad (4.4)$$

and the second order of spatial weight is $w_{12}^{(2)} = w_{13}^{(2)} = w_{15}^{(2)} = w_{17}^{(2)} = w_{1,10}^{(2)} = w_{1,14}^{(2)} = 1/6$.

4.2. GSTAR Model Building for the GDP Data

After transforming the data and constructing spatial weight matrix, the next step is identification of the model order. In STAR model-building, [3] used the sample space time autocorrelation function (STACF) and space time partial autocorrelation function (STPACF) as the primary tools for model identification. The order $(p; \lambda_1, \dots, \lambda_p)$ model can be characterized by the tail-off behaviour of the autocorrelations and the cut-off behaviour of the partial after p time lag and λ_p spatial lag. Since STAR model is a special case of GSTAR model, these autocorrelation functions will be adopted to identify the order of GSTAR model.

Figure 2 presents sample STACF and STPACF for the differenced data $\mathbf{Z}(t)$ up to time lag 10 and spatial lags 0, 1, and 2. The pattern is not clearly suggested an exact order.

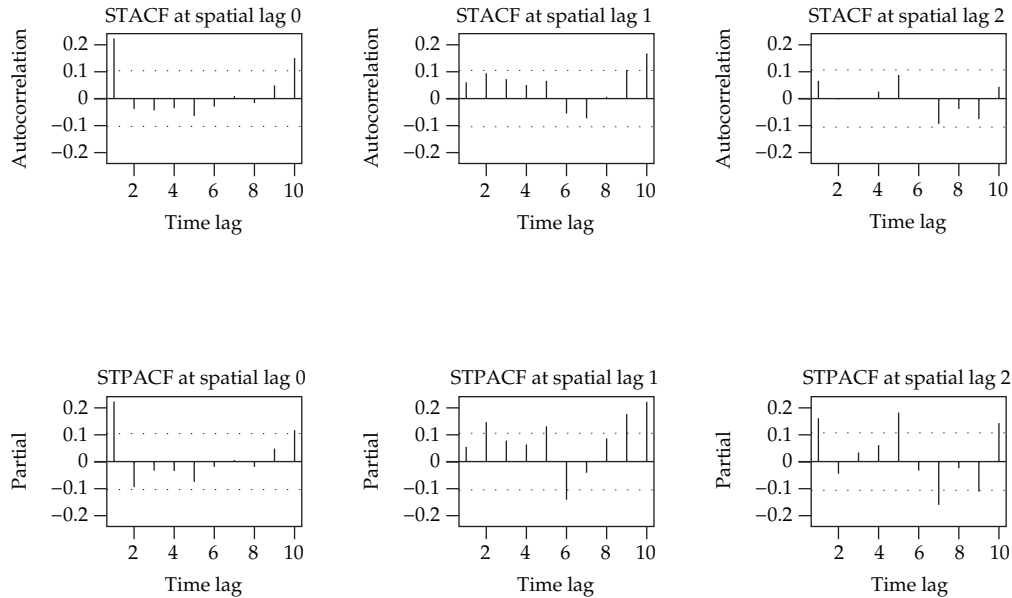


Figure 2: Sample space-time ACF and PACF of the differenced data.

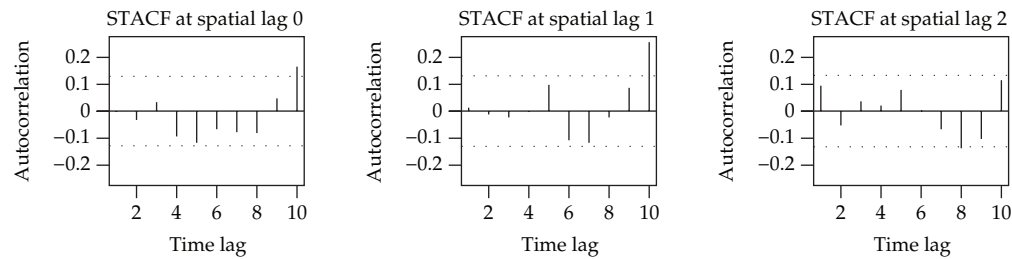


Figure 3: Sample space-time autocorrelation of the residuals from the GSTAR(2;1,1) model.

However, since the sample STPACF cut off after time lag 2 and spatial lag 1, order (2;1,1) can be considered as the space-time order candidate. In addition, the space-time partials also cut off at time lag 5 and spatial 2, but applying this GSTAR model to the data will result too many estimated parameters because there will be at least 160 parameters which have to be estimated.

The 64 parameters and the error variance in this model were estimated using the least square method and the result is presented in Table 3. The empirical MSE for GSTAR(2;1,1) model is 2.735 counted based on 16×39 or 624 values. The residuals histogram and normal probability plot on Figures 4(a) and 4(b) show that the GSTAR(2;1,1) residuals are approximately normal distributed with zero mean and constant variance. Meanwhile, fitted value versus residuals plot in Figure 4(c) exhibits that the residuals do not show a significant pattern. From Figure 3 we can observe that the STACF of the residuals is significantly almost zero except for time lag 5 and 10, and spatial lag 2. The exception points at time lags 5 and 10, suggesting that the seasonal difference of order 5 might be useful for further model analysis. But the seasonal model analysis is not discussed here because it is out of the research scope.

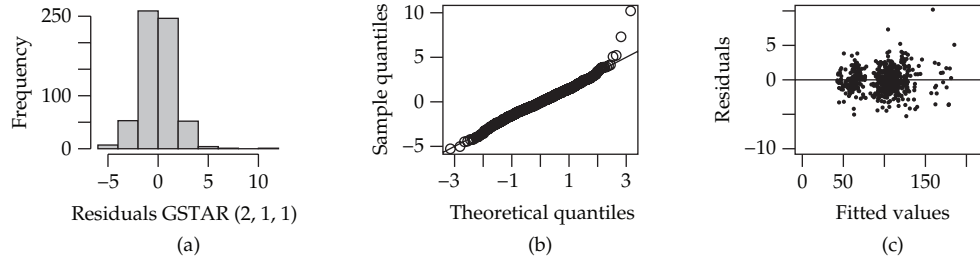


Figure 4: Histogram and normal probability plot of the GSTAR(2,1,1) residuals (a) and (b) and plot of residuals versus fitted value (c).

Table 3: Least square estimates for GSTAR(2;1,1) parameter.

Site i	$\hat{\phi}_{10}^{(i)}$	$\hat{\phi}_{11}^{(i)}$	$\hat{\phi}_{20}^{(i)}$	$\hat{\phi}_{21}^{(i)}$
1	-0.083	0.058	0.045	-0.046
2	-0.092	0.028	0.404	0.052
3	-0.218	0.730	-0.113	0.638
4	0.566	-0.534	-0.262	0.253
5	0.152	0.563	-0.030	0.255
6	0.176	0.211	-0.052	0.046
7	0.106	-0.048	-0.018	0.261
8	0.332	-0.126	-0.114	0.540
9	0.271	-0.424	-0.217	0.461
10	-0.026	-0.250	0.097	-0.332
11	0.549	-0.278	-0.166	0.239
12	0.442	0.198	-0.254	-0.204
13	0.376	0.242	-0.054	0.037
14	0.390	-0.120	-0.457	0.304
15	0.042	0.564	-0.250	-0.949
16	0.293	0.226	0.119	-0.318

$\hat{\sigma}_{1S}^2 = 2.969$; MSE = 2.735.

For forecasting purpose, the estimated parameters in Table 3 can be used to predict the h -step-ahead forecast value at time forecast origin T defined by

$$\hat{Z}_T(h) = \hat{\Phi}_{10} \hat{Z}_T(h-1) + \hat{\Phi}_{11} \mathbf{W}_T^{(1)} \hat{Z}_T(h-1) + \hat{\Phi}_{20} \hat{Z}_T(h-2) + \hat{\Phi}_{21} \mathbf{W}_T^{(2)} \hat{Z}_T(h-2). \quad (4.5)$$

5. GSTAR(1;1) Modeling

GSTAR(1;1) model is the simplest model of GSTAR($p; \lambda_1, \dots, \lambda_p$) model class defined in (2.1) because it is only characterized by the autoregressive terms lagged in time and spatial of order one. From (2.1) we can write GSTAR(1;1) as

$$Z_i(t) = \phi_{0i} Z_i(t-1) + \phi_{1i} \sum_{j=1}^N w_{ij} Z_j(t-1) + e_i(t), \quad (5.1)$$

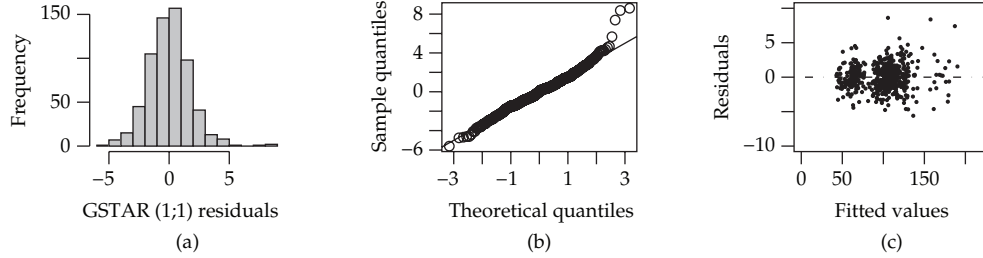


Figure 5: GSTAR(1;1) residuals histogram (a), normal probability (b), and residuals versus fitted values (c).

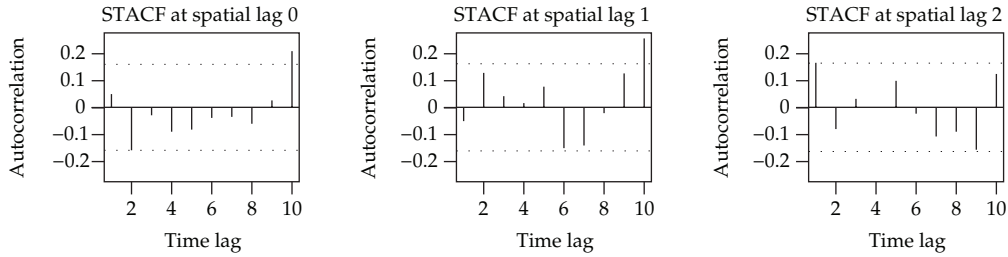


Figure 6: Space-time autocorrelation of the GSTAR(1;1) residuals.

or in matrix notation

$$\mathbf{Z}(t) = \mathbf{\Phi}_{10}\mathbf{Z}(t-1) + \mathbf{\Phi}_{11}\mathbf{W}^{(1)}\mathbf{Z}(t-1) + \mathbf{e}(t), \quad (5.2)$$

where $\mathbf{Z}(t) = (Z_1(t), \dots, Z_N(t))$, $\mathbf{\Phi}_{10} = \text{diag}(\phi_{10}^{(1)}, \dots, \phi_{10}^{(N)})$, and $\mathbf{\Phi}_{11} = \text{diag}(\phi_{11}^{(1)}, \dots, \phi_{11}^{(N)})$. To simplify the notation, $\mathbf{\Phi}_{10}$ and $\mathbf{\Phi}_{11}$ are frequently written, respectively, by $\mathbf{\Phi}_0$ and $\mathbf{\Phi}_1$.

The model has an interpretation that the current observation in a certain location only depends on the immediate past observations recorded at the location of interest and at the nearest locations. The GSTAR(1;1) model has $2N$ parameters since it is assumed that the parameter for each location is allowed to be unequal. For the per capita GDP data the least square estimator for GSTAR(1;1) model is presented in Table 4. The residuals distribution of the model, presented in Figure 5, give the conclusion that the residuals are approximately normally distributed with zero mean and the variance is nearly constant. In general, the STACF plot in Figure 6 shows that the residuals are significantly uncorrelated. The insignificant value was only found at the last time lag.

For the GDP data case, GSTAR(2;1,1) model has 64 parameters while GSTAR(1;1) has 32 parameters. Hence, it is not a surprise if the empirical MSE of GSTAR(2;1,1) is less than that of GSTAR(1;1). However, though the number of parameters of GSTAR(1,1) is half of the other one, the empirical MSE of GSTAR(1,1) is only decreasing 0.257 compared to the GSTAR(2;1,1).

The distribution of the MSE difference for each country is presented in a bubble plot (Figure 7). Center of the bubble is the value of MSE difference and the radius is its absolute value. The bubble placed under the zero axis indicates that the GSTAR(2,1,1) MSE for the associated country is smaller than that of the GSTAR(1;1) MSE. From the figure, we can see that the decrease of MSE values is mostly contributed by five countries, Belgium, Denmark,

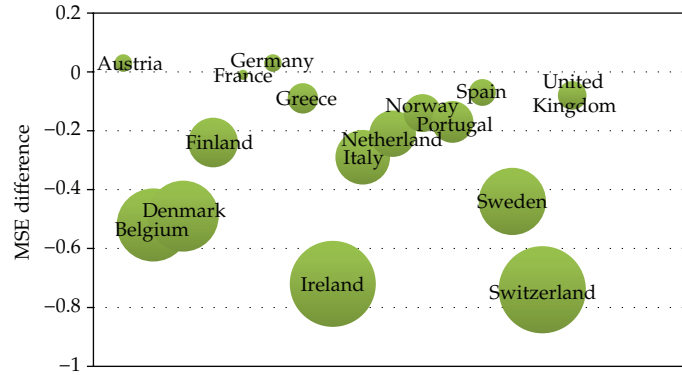


Figure 7: Bubble plot for the difference value of empirical MSE between GSTAR(2;1,1) and GSTAR(1;1). The bubble placed under the zero axis indicates that GSTAR(2,1,1) has a smaller value of MSE difference than that of the GSTAR(1;1) model.

Table 4: Least square estimates for GSTAR(1;1) parameters.

Site i	$\hat{\phi}_{10}^{(i)}$	$\hat{\phi}_{11}^{(i)}$
1	-0.073	0.086
2	-0.090	0.250
3	-0.201	0.871
4	0.422	-0.452
5	0.192	0.582
6	0.187	0.221
7	0.116	0.011
8	0.355	0.077
9	0.183	-0.261
10	0.012	-0.123
11	0.465	-0.175
12	0.358	0.054
13	0.317	0.287
14	0.285	-0.162
15	0.006	0.119
16	0.340	0.203

$\hat{\sigma}_{LS}^2 = 3.1556$; MSE = 2.992.

Ireland, Sweden, and Switzerland. Though the MSE value seems different, the paired t -test for the result gives the P values < 0.678 . It means that the MSE values for both models are significantly indifferent.

6. Comparison of Forecast Performance

For the purpose of forecasting model, it would be useful if we also consider their forecast performance. Therefore, in this section we will examine the one-step-ahead forecasting performance for each model candidate using the last ten actual data points of the per capita

GDP ratio data set. Result of this section is expected to become a supplementary reference in finding the most parsimony space-time model for the case of per capita GDP ratio.

In (4.5) we have defined the h -step-ahead forecasting for GSTAR(2;1,1) model. For GSTAR(1;1) model, it can be estimated by

$$\widehat{\mathbf{Z}}_T(h) = \Phi_{10}\widehat{\mathbf{Z}}_T(h-1) + \Phi_{11}\mathbf{W}^{(1)}\widehat{\mathbf{Z}}_T(h-1), \quad (6.1)$$

$$\widehat{\mathbf{Z}}_T(h) = 0.1674\widehat{\mathbf{Z}}_T(h-1) + 0.0453\mathbf{W}^{(1)}\widehat{\mathbf{Z}}_T(h-1), \quad (6.2)$$

where $\Phi_{10} = \text{diag}(\phi_{10}^{(1)}, \dots, \phi_{10}^{(N)})$, $\Phi_{11} = \text{diag}(\phi_{11}^{(1)}, \dots, \phi_{11}^{(N)})$, and for $h \leq 0$, $\widehat{\mathbf{Z}}_T(h) = \mathbf{Z}(T-h)$.

The one-step-ahead forecast $\widehat{\mathbf{Z}}_{T+j-1}(1)$ of each data points can be calculated using (4.5), (6.1), and (6.2) for $h = 1$. Note that calculation of the forecast value for each time $j = 1, 2, \dots, 10$, is performed without updating the parameter estimates and average data. The one-step-ahead forecast for GSTAR(2;1,1) model is

$$\begin{aligned} \widehat{\mathbf{Z}}_{T+j-1}(1) &= \Phi_{10}\mathbf{Z}(T+j-1) + \Phi_{11}\mathbf{W}^{(1)}\mathbf{Z}(T+j-1) + \Phi_{20}\mathbf{Z}(T+j-2) \\ &\quad + \Phi_{12}\mathbf{W}^{(1)}\mathbf{Z}(T+j-1), \end{aligned} \quad (6.3)$$

where $j = 1, 2, \dots, 10$. Meanwhile the one-step-ahead forecast for GSTAR(1;1) model is

$$\widehat{\mathbf{Z}}_{T+j-1}(1) = \Phi_{10}\mathbf{Z}(T+j-1) + \Phi_{11}\mathbf{W}^{(1)}\mathbf{Z}(T+j-1). \quad (6.4)$$

To compare the forecast performance, we use mean of square forecast error (MSFE) which is defined by

$$\text{MSFE} = \frac{1}{N} \sum_{i=1}^N \bar{E}_i, \quad (6.5)$$

where

$$\bar{E}_i = \frac{1}{10} \sum_{j=1}^{10} \left[Y_i(T+j) - \widehat{Y}_{i,T+j-1}(1) \right]^2, \quad (6.6)$$

$$\widehat{Y}_{i,T+j-1}(1) = \left(\widehat{\mathbf{Z}}_{i,T+j-1}(1) + D_i \right) + Y_i(T+j-1),$$

and $\widehat{\mathbf{Z}}_{i,T+j-1}(1)$ is the element of $\widehat{\mathbf{Z}}_{i,T+j-1}(1)$.

To measure the performance closeness between two models, we also calculate the MSFE difference between model M_1 and model M_2 for country i which is defined by

$$C_i = \bar{E}_i(M_1) - \bar{E}_i(M_2). \quad (6.7)$$

The performance of M_1 and M_2 model is said to be close if the value of C_i is near to zero. The negative value of C_i indicates that for location i , the M_1 model is better than M_2

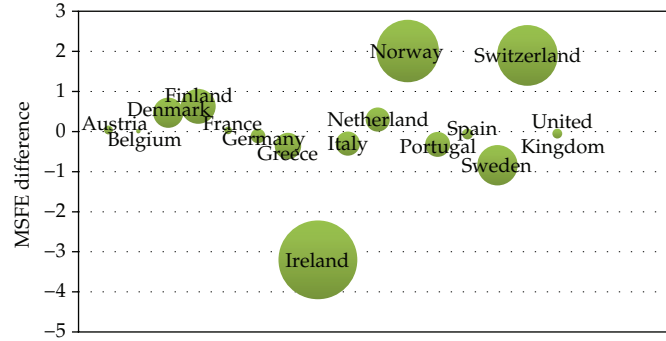


Figure 8: Bubble plot for the difference of mean of square forecast error (MSFE) between $GSTAR(2;1,1)$ and $GSTAR(1;1)$. The bubble under the zero axis indicates that the $GSTAR(2,1,1)$ has a smaller value of MSFE difference than that of the $GSTAR(1;1)$ model.

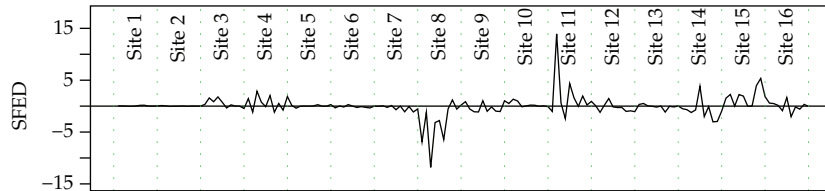


Figure 9: Square forecast error difference (SFED) for $GSTAR(2;1,1)$ and $GSTAR(1;1)$.

model. Since the residual is approximately normally distributed, the paired t -test can also be applied to test the null hypothesis that model M_1 and model M_2 have the same forecast performance.

Suppose M_1 and M_2 in (6.7) are represented by $GSTAR(2;1,1)$ and $GSTAR(1;1)$ model, respectively. The MSFE difference values are displayed in a bubble plot (Figure 8). Center of the bubble is the value of MSFE difference and the radius is its absolute value. The bubble placed under the zero axis indicates that the MSFE value for model M_1 is smaller than that of the model M_2 . From the figure, it can be seen that most of the countries have the MSFE value which is scattered around the zero axis. The high value of the P value < 0.9858 for the paired t -test gives the evidence that the forecast performance of each model is not significantly different.

The behavior of both differences for each time is displayed in Figure 9. From the figure, it can be observed that most points in Ireland (site 8) have a negative value. It means that in Ireland $GSTAR(2;1,1)$ model has a better performance than $GSTAR(1;1)$. Conversely, the most points in Norway (site 11) and Switzerland (site 15) are positive when the difference is counted for the pair of $GSTAR(2;1,1)$ and $GSTAR(1;1)$. On the other hand, Figure 7 also shows that in Norway there is a point with high jump difference value. It means that the worst performance of $GSTAR(2;1,1)$ occurred at this point. In addition, the ratio of the negative value and the positive value is almost the same. It gives the interpretation that forecast performance of $GSTAR(2;1,1)$ and $GSTAR(1;1)$ model is almost similar.

7. Conclusions

GSTAR modeling has been built to the ratio of per capita GDP data in West European countries. The model of order (2;1,1) has been identified as the candidate model. However, when the forecast performance is compared with GSTAR(1;1), it is found that the performance is significantly indifferent. Due to parsimony principle, we recommend that the GSTAR(1;1) might be considered as a forecasting model.

Acknowledgments

The research of the first author is supported by the Scholarship of Sandwich-Like Program no. 1995.1/D4.4/2008 of the Ministry of National Education of Republic of Indonesia c/q Directorate General of Higher Education (DIKTI). The authors are very grateful to Professor Gopalan Nair from the School of Mathematics and Statistics, the University of Western Australia, for his supervision during the first author's visit in 2008. Helpful comments from earlier manuscript's reviewer are also thankfully acknowledged.

References

- [1] S. Borovkova, H. P. Lopuhaä, and B. N. Ruchjana, "Consistency of asymptotic normality of least squares estimators in generalized STAR models," *Statistica Neerlandica*, vol. 62, no. 4, pp. 482–508, 2008.
- [2] R. L. Martin and J. E. Oeppen, "The identification of regional forecasting," *Transactions of the Institute of British Geographers*, vol. 66, pp. 119–128, 1975.
- [3] P. E. Pfeifer and S. J. Deutsch, "A three-stage iterative procedure for space-time modeling," *Technometrics*, vol. 22, pp. 397–408, 1980.
- [4] S. Terzi, "Maximum likelihood estimation of a GSTAR(1;1) model," *Statistical Methods and Applications*, vol. 3, pp. 377–393, 1995.
- [5] B. Nurani, *Suatu model generalisasi space-time autoregresi dan penerapannya pada produksi minyak bumi*, Dr. dissertation, Institut Teknologi Bandung, Indonesia, 2002.
- [6] A. Maddison, *Statistics on World Population, GDP and Per Capita GDP, 1-2006 AD*, 2008, <http://www.ggdc.net/maddison/>.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

