

Received: 12 November 2018

Revised: 27 November 2018

Accepted: 3 November 2018

Published: 30 December 2018

## **Pemodelan *Generalized Space Time Autoregressive (GSTAR)* dan Penerapannya pada Penderita TB Paru (BTA+) di DKI Jakarta**

Ana Nur Islamiyah<sup>1, a)</sup>, Widyanti Rahayu<sup>2, b)</sup>, Eti Dwi Wiraningsih<sup>2, c)</sup>

<sup>1</sup>*Program Studi Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Jakarta*

<sup>2</sup>*Program Studi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Jakarta*

Email: <sup>a)</sup>ananurislamiyah40@gmail.com, <sup>b)</sup>widyanti.rahayu@gmail.com, <sup>c)</sup>etidwi@gmail.com

### **Abstract**

Generalized Space Time Autoregressive (GSTAR)( $p; \lambda_k$ ) model is one of space time model that used for modeling time series with location relevance ( $p$ ) is the autoregressive order and  $\lambda_k$  for the first order of spatial. Location relevance can be represented with a weighted. The number of patient TB Paru (BTA+) in DKI Jakarta is one of time series data with location relevance so can be used for GSTAR model. In this research, the number of patient TB Paru (BTA+) in DKI Jakarta were modeled with GSTAR model using inverse distance weighted and normalization of cross correlation weighted. Furthermore, a better model GSTAR is chosen by the smallest RMSE. The result of this research for the number of patient TB Paru (BTA+) is GSTAR (1;1) model. GSTAR (1;1) model by using normalization of cross correlation weighting is better than that model using inverse distance weighting. Because GSTAR (1;1) model by using normalization of cross correlation weighting has root mean square error (RMSE) 73,57728 less than GSTAR (1;1) model by using inverse distance weighting.

**Keywords:** GSTAR, inverse distance, normalization of cross correlation, RMSE, pulmonary tuberculosis.

### **Abstrak**

Model *Generalized Space Time Autoregressive* (GSTAR)( $p; \lambda_k$ ) merupakan salah satu model deret waktu yang digunakan untuk peramalan data yang mengandung unsur lokasi dan waktu dengan orde autoregressive ( $p$ ) dan orde spasial ( $\lambda_k$ ). Keterkaitan antar lokasi dinyatakan dalam suatu nilai berdasarkan pembobot lokasi tertentu. Data banyaknya penderita TB Paru (BTA+) di DKI Jakarta merupakan salah satu data deret waktu yang memiliki keterkaitan antar lokasi sehingga dapat dimodelkan dengan model GSTAR. Pada penelitian ini model GSTAR menggunakan pembobot lokasi invers jarak dan pembobot lokasi normalisasi korelasi silang diterapkan pada data banyaknya penderita TB Paru (BTA+) di DKI Jakarta. Setelah itu, memilih model GSTAR yang lebih baik menggunakan nilai RMSE terkecil. Hasil dari penelitian ini dengan menerapkan data banyaknya penderita TB Paru (BTA+) di DKI Jakarta diperoleh model GSTAR(1;1) dengan pembobot lokasi normalisasi korelasi silang memiliki \textit{root mean square error}(RMSE) 73,57728 yang lebih kecil dari model GSTAR(1;1) dengan pembobot lokasi invers jarak. Sehingga, model dengan pembobot lokasi normalisasi korelasi silang lebih baik dibandingkan dengan pembobot lokasi invers jarak.

**Kata-kata kunci:** GSTAR, invers jarak, normalisasi korelasi silang, RMSE, TB Paru.

## PENDAHULUAN

Seiring dengan fenomena yang terjadi dalam kehidupan sehari-hari dan semakin berkembangnya kajian-kajian mengenai analisis deret waktu, muncul pemikiran adanya dugaan bahwa ada beberapa data yang dicatat dari suatu kejadian yang tidak hanya dipengaruhi dengan kejadian pada waktu-waktu sebelumnya, tetapi juga dipengaruhi oleh kondisi lokasi disekitarnya (pengaruh spasial). Menurut (Borovkova *et al.*, 2008, h. 61) Model dengan keterkaitan deret waktu dan lokasi disebut model (*space time*).

Model *space time* atau model yang memiliki keterkaitan waktu dan lokasi pertama kali diperkenalkan oleh Pfeifer dan Deutsch (1980) yang dikenal dengan Model *Space Time Autoregressive* (STAR). Model STAR memiliki kelemahan yaitu adanya asumsi parameter *autoregressive* yang bernilai sama pada semua lokasi, sehingga model STAR dikembangkan oleh Borovkova *et al.* (2002) dengan mengusulkan model yang juga dapat menggabungkan keterkaitan waktu dan lokasi yaitu model *Generalized Space Time Autoregressive*. Berbeda dengan model STAR, model GSTAR lebih fleksibel karena asumsi parameter *autoregressive* pada model ini berbeda setiap lokasi, sehingga sesuai diterapkan pada lokasi yang memiliki karakteristik heterogen Wutsqa *et al.* (2010).

Pengaruh waktu pada model GSTAR ditunjukkan dengan parameter *autoregressive* dan untuk pengaruh spasial ditunjukkan dengan pembobot lokasi yang menyatakan besarnya keterkaitan antar lokasi. (Suhartono dan Subanar, 2006, h. 46) menyatakan terdapat empat pembobot lokasi dalam model GSTAR yaitu pembobot lokasi seragam, biner, invers jarak, dan normalisasi korelasi silang. Keterkaitan lokasi kurang sesuai apabila dilihat dari kedekatan lokasi sehingga diperlukan pertimbangan jarak sebenarnya. Oleh karena itu, pembobot lokasi yang sesuai menggunakan invers jarak. Pembobot lokasi normalisasi korelasi silang dapat digunakan karena mempertimbangkan korelasi pada data yang memiliki keterkaitan lokasi dan waktu.

Dalam bidang kesehatan model GSTAR dapat digunakan untuk memprediksi banyaknya penderita penyakit tertentu pada periode selanjutnya sehingga dapat dijadikan pertimbangan dalam melakukan tindakan antisipasi dan menentukan kebijakan. Salah satu jenis penyakit yang dapat dimodelkan dengan model *space time* yaitu penyakit TB Paru (BTA+). Tuberkulosis (TB) merupakan penyakit menular yang disebabkan oleh bakteri *Mycobacterium Tuberculosis* (Kemenkes RI, 2016).

DKI Jakarta merupakan salah satu daerah yang banyak mengalami kejadian TB Paru (BTA+). Dinas Kesehatan Provinsi DKI Jakarta (2016), merujuk pada hasil survei terakhir tahun 2016 tentang prevalensi penyakit TB Paru didapatkan angka 55.503 penderita. Jakarta Timur, Barat dan Selatan merupakan wilayah dengan jumlah TB Paru (BTA+) terbesar di Provinsi DKI Jakarta, yaitu rata-rata sebanyak 2000 penderita. Kota Jakarta Timur, Barat, Selatan, Utara, Pusat dan Kabupaten Kepulauan Seribu merupakan kota dengan luas wilayah yang besar dengan kepadatan penduduk yang tinggi serta mobilitas penduduknya cukup tinggi dan lokasi antar wilayah berdekatan sehingga tidak menutup kemungkinan peluang terjadinya penyebaran penyakit TB Paru (BTA+) dari waktu ke waktu serta lokasi saling berkaitan.

Pada Penelitian ini model *space time* diterapkan di bidang kesehatan, dengan menggunakan model GSTAR dan penerapannya pada banyaknya penderita TB Paru (BTA+) di DKI Jakarta. Model yang diharapkan yaitu model yang menggambarkan keterkaitan waktu dan lokasi dengan membandingkan pembobot lokasi invers jarak dan pembobot lokasi normalisasi korelasi silang, berdasarkan pembobot lokasi akan dilihat pembobot lokasi mana yang menghasilkan model GSTAR terbaik dengan melihat nilai kesalahan peramalan terkecil.

**METODE**

**Model Generalized Space Time Autoregressive (GSTAR)**

Model *Generalized Space Time Autoregressive* (GSTAR)( $p; \lambda_k$ ) merupakan salah satu perluasan dari model *Space Time Autoregressive* (STAR), model ini cenderung lebih fleksibel dalam menentukan orde *autoregressive* ( $p$ ). Keterkaitan spasial atau lokasi ( $\lambda_k$ ) dinyatakan oleh matriks pembobot lokasi.

Model GSTAR merupakan pemodelan dari sejumlah pengamatan  $Z_i(t)$  yang terdapat pada tiap  $N$  lokasi dalam suatu ruang ( $i= 1, 2, \dots, N$ ) terhadap  $t$  periode waktu. Efek waktu dirumuskan sebagai model deret waktu dan efek spasial dirumuskan sebagai matriks pembobot lokasi. Model STAR pertama kali diperkenalkan oleh Pfeifer dan Deutsch (1980) namun memiliki kelemahan yaitu adanya asumsi parameter *autoregressive* yang bernilai sama pada semua lokasi, sehingga model STAR lebih sesuai untuk lokasi dengan karakteristik sama (homogen) (Borovkova *et al.* 2008).

Kelemahan model STAR ini kemudian diperbaiki oleh Borovkova *et al.* (2002) melalui model GSTAR. Berbeda dengan model STAR, model GSTAR lebih fleksibel karena asumsi parameter *autoregressive* pada model ini berbeda setiap lokasi, sehingga sesuai diterapkan pada lokasi yang memiliki karakteristik heterogen (Wutsqa *et al.* 2010).

Model GSTAR dengan orde *autoregressive* ( $p$ ) dan orde spasial  $\lambda_1, \lambda_2, \dots, \lambda_k$  dinotasikan dengan GSTAR ( $p; \lambda_k$ ) didefinisikan dengan menggunakan persamaan sebagai berikut:

$$Z_{i(N \times 1)}(t) = \sum_{k=1}^p \sum_{l=0}^{\lambda_k} \Phi_{kl(N \times N)} W_{(N \times N)}^{(l)} Z_{(N \times 1)}(t - k) + e_{(N \times 1)}(t) \tag{1}$$

dengan,  $Z_i(t)$  merupakan vektor pengamatan pada waktu ke- $t$  lokasi ke- $i$  yang berukuran ( $N \times 1$ ).  $P$  merupakan orde *autoregressive* (AR).  $\lambda_k$  adalah orde spasial ke- $k$ .  $N$  merupakan banyaknya lokasi.  $W^l$  merupakan matriks pembobot lokasi ukuran ( $N \times N$ ) pada lag spasial  $l$  dengan  $l = 0, 1, 2, \dots, N$ . Matriks  $W^0$  adalah matriks identitas I. Matriks  $\Phi_{kl(N \times N)}$  merupakan diagonal matriks parameter dari waktu lag ke- $k$  lokasi ke- $l$  dengan elemen diagonal ( $\Phi_{kl}^1, \dots, \Phi_{kl}^N$ ).  $e_{(N \times 1)}(t)$  adalah vektor residual pada waktu ke- $t$ , dalam hal ini  $e(t) \overset{iid}{\sim} N(0, \sigma^2)$ .

**Pembobot Lokasi Invers Jarak**

Pembobotan dengan menggunakan metode invers jarak dilakukan berdasarkan jarak sebenarnya antar lokasi. Pembobot invers jarak memberikan nilai bobot yang besar untuk jarak yang lebih dekat dan memberikan pembobot yang kecil untuk jarak yang jauh (Anggraeni *et al.*, 2013, h. 73) Jarak yang digunakan pada pembobot ini mempertimbangkan koordinat lintang dan bujur. Derajat lintang dan bujur yang selanjutnya akan dikonversikan ke kilometer. Pembobot lokasi invers jarak dinyatakan sebagai berikut

$$w_{ij} = \frac{w_{ij}^*}{\sum_{j=1}^n w_{ij}^*} \tag{2}$$

dengan,

$$w_{ij}^* = \begin{cases} \frac{1}{d_{ij}}; & i \neq j \\ 0; & i = j \end{cases} \quad (i, j = 1, 2, \dots, n)$$

$$d_{ij} = \sqrt{(u_i - u_j)^2 + (v_i - v_j)^2}, i \neq j$$

dengan,  $d_{ij}$  merupakan jarak dari lokasi  $i$  ke  $j$ ,  $(u_i, u_j)$  adalah koordinat dari garis lintang, dan  $(v_i, v_j)$  adalah koordinat dari garis bujur.

**Pembobot Lokasi Normalisasi Korelasi Silang**

Suhartono dan Subanar (2006) mengenalkan pembobot normalisasi korelasi silang untuk model GSTAR. Pembobot normalisasi korelasi silang ini tidak mengisyaratkan aturan tertentu, seperti bergantung

pada jarak antar lokasi. Secara umum korelasi silang antar lokasi ke- $i$  dan ke- $j$  pada lag waktu ke- $k$ ,  $corr[Z_i(t), Z_j(t - k)]$  didefinisikan sebagai berikut

$$\rho_{ij}(k) = \frac{\gamma_{ij}(k)}{\sigma_i \sigma_j}, k = 0, \pm 1, \pm 2, \dots$$

dengan  $\gamma_{ij}(k)$  merupakan kovarian silang antar pengamatan pada lokasi ke- $i$  dan lokasi ke- $j$  pada lag waktu ke- $k$ ,  $\sigma_i$  dan  $\sigma_j$  merupakan standar deviasi pada lokasi ke- $i$  dan lokasi ke- $j$ . Penduga korelasi silang pada data sampel adalah

$$r_{ij}(k) = \frac{\sum_{t=k+1}^n [Z_i(t) - \bar{Z}_i] [Z_j(t - k) - \bar{Z}_j]}{\sqrt{(\sum_{t=1}^n [Z_i(t) - \bar{Z}_i]^2) (\sum_{t=1}^n [Z_j(t) - \bar{Z}_j]^2)}}$$

Penentuan pembobot lokasi dapat diselesaikan dengan menormalisasikan korelasi silang antar lokasi pada lag yang sesuai. Proses ini menghasilkan pembobot lokasi sebagai berikut

$$w_{ij} = \frac{r_{ij}(1)}{\sum_{k \neq i} |r_{ik}(1)|} \tag{3}$$

dengan,  $i \neq j$ , dan  $\sum_{j \neq i} |w_{ij}| = 1$ .

Pembobot lokasi dengan normalisasi dari besaran-besaran korelasi silang antar lokasi pada waktu yang bersesuaian memungkinkan adanya kemungkinan hubungan antar lokasi. Bobot ini juga memberikan fleksibilitas pada besar dan tanda hubungan antar lokasi yang berlainan yaitu positif dan negatif.

### Estimasi Parameter Model GSTAR

Menurut (Borovkova *et al.*, 2002, h. 62) model *Generalized Space Time Autoregressive* (GSTAR) dapat dinyatakan sebagai suatu model linier dan estimasi dari parameter autoregresif model GSTAR dapat dilakukan dengan menggunakan metode kuadrat terkecil yaitu dengan meminimumkan jumlah kuadrat sisaan.

Model GSTAR merupakan pemodelan dari sejumlah pengamatan  $Z_i(t)$  yang terdapat pada tiap  $N$  lokasi dalam suatu ruang ( $i = 1, 2, \dots, N$ ) terhadap  $t$  periode waktu..

$$Z_{i(N \times 1)}(t) = \sum_{k=1}^p \sum_{l=0}^{\lambda_k} \Phi_{kl(N \times N)} W_{(N \times N)}^{(l)} Z_{(N \times 1)}(t - k) + e_{(N \times 1)}(t)$$

Jika diberikan pengamatan  $Z_i(t)$ ,  $t = 0, 1, \dots, t$ , untuk lokasi  $i = 1, 2, \dots, N$  dengan,  $V_i(t) = \sum_{j=1}^N w_{ij} Z_j(t)$ .

Untuk  $i \neq j$ ,

Maka model untuk lokasi ke- $i$  dalam model linear dapat dinyatakan dengan

$$Z_i = Z_i^* \Phi + e_i \tag{4}$$

Dalam hal ini,

$$Z_i = \begin{bmatrix} Z_i(1) \\ Z_i(2) \\ \vdots \\ Z_i(t) \end{bmatrix}, Z_i^* = \begin{bmatrix} Z_i(0) & V_i(0) \\ Z_i(1) & V_i(0) \\ \vdots & \vdots \\ Z_N(t-k) & Z_N(t-k) \end{bmatrix}, \Phi = \begin{bmatrix} \Phi_{kl}^i \\ \Phi_{kl}^i \\ \vdots \\ \Phi_{kl}^N \end{bmatrix}, e_i = \begin{bmatrix} e_i(1) \\ e_i(2) \\ \vdots \\ e_N(t) \end{bmatrix}.$$

Estimasi parameter model dengan metode kuadrat terkecil digunakan dengan meminimumkan jumlah kuadrat sisaannya, yaitu  $\hat{\Phi} = (Z^{*'} Z^*)^{-1} (Z^{*'} Z)$ .

### Uji Asumsi Residual

Uji kelayakan model diperlukan sebagai langkah selanjutnya setelah mendapatkan parameter dan model yang signifikan, dengan cara memeriksa asumsi residual (*white noise*). Model GSTAR dikatakan memenuhi asumsi jika residual yang terbentuk menyebar dengan sebaran normal ganda artinya model tidak berkorelasi. dapat dilakukan dengan menggunakan uji *Ljung Box Pearce* (Wei, 2006) sebagai berikut:

1. Hipotesis  
 $H_0 : \rho_1 = \rho_2 = \dots = \rho_k = 0$  (Sisaan tidak *white noise*).  
 $H_1 : \text{Terdapat } \rho_k \neq 0, k = 1, 2, \dots, K$  (Sisaan *white noise*).
2. Statistik Uji

$$LB = n(n + 2) \sum_{k=1}^K \frac{\widehat{\rho}_k^2}{n - k}$$

dengan  $n$  adalah banyaknya pengamatan,  $k$  adalah banyaknya lag, dan  $\widehat{\rho}_k$  adalah autokorelasi duga pada lag ke- $k$ .

3. Kriteria Pengujian  
 Taraf signifikan sebesar  $\alpha = 0,05$ , tolak  $H_0$  jika  $LB > x_{1-\alpha, k}^2$  tabel.

### Uji Ketepatan Model

Tujuan dari model peramalan adalah memodelkan untuk meramalkan nilai yang akan datang dengan *error* sekecil mungkin, salah satu alternatif untuk pemilihan model berdasarkan nilai *error* dapat dilakukan dengan RMSE atau (*Root Mean Square Error*) untuk setiap model. RMSE dirumuskan sebagai

$$RMSE = \sqrt{MSE} = \sqrt{\frac{1}{M} \sum_{t=1}^M (Z(t) - \hat{Z}(t))^2}$$

keterangan:

$M$  = banyaknya data ramalan yang dilakukan.

$Z(t)$  = data yang sebenarnya pada waktu ke-  $t$ .

$\hat{Z}(t)$  = data hasil ramalan pada waktu ke-  $t$ .

Nilai RMSE berkisar antara 0 sampai  $\infty$ . Semakin kecil nilai RMSE maka model yang digunakan semakin bagus (Wei, 2006).

Sedangkan menurut Pendawa *et al.* (2015) MAPE atau *Mean Absolute Percentage Error* merupakan ukuran kesalahan nilai dugaan model yang dinyatakan dalam bentuk rata-rata persentase absolut residual. Persamaan MAPE dirumuskan sebagai berikut

$$MAPE = \frac{100\%}{M} \sum_{t=1}^M \left| \frac{Z(t) - \hat{Z}(t)}{Z(t)} \right|$$

keterangan:

$M$  = banyaknya data ramalan yang dilakukan.

$Z(t)$  = data yang sebenarnya pada waktu ke-  $t$ .

$\hat{Z}(t)$  = data hasil ramalan pada waktu ke-  $t$ .

Kemampuan peramalan sangat baik jika memiliki nilai MAPE kurang dari 10% dan mempunyai kemampuan peramalan yang baik jika nilai MAPE kurang dari 20%.

### Sumber Data

Data yang digunakan pada penelitian ini adalah data banyaknya penderita TB Paru (BTA+) di lima Kotamadya dan satu Kabupaten DKI Jakarta, berupa data triwulan dari bulan Januari-Maret 2007 sampai Juli-September 2017. Data diperoleh dari P2PML Kementerian Kesehatan RI. Terdapat lima Kotamadya dan satu Kabupaten yang berada di DKI Jakarta yaitu,

$Z_1(t)$  : Banyaknya penderita TB Paru (BTA+) di Kota Jakarta Pusat.

$Z_2(t)$  : Banyaknya penderita TB Paru (BTA+) di Kota Jakarta Barat.

$Z_3(t)$  : Banyaknya penderita TB Paru (BTA+) di Kota Jakarta Timur.

$Z_4(t)$  : Banyaknya penderita TB Paru (BTA+) di Kota Jakarta Utara.

$Z_5(t)$  : Banyaknya penderita TB Paru (BTA+) di Kota Jakarta Selatan.

$Z_6(t)$  : Banyaknya penderita TB Paru (BTA+) di Kabupaten Kepulauan Seribu.

### Tahapan Penelitian

Penelitian ini merupakan penelitian terapan yang menerapkan model GSTAR pada banyaknya penderita TB Paru (BTA+) di DKI Jakarta dengan menggunakan pembobot lokasi invers jarak dan pembobot lokasi normalisasi korelasi silang.

Berikut beberapa tahapan atau langkah-langkah yang dilakukan untuk mencapai tujuan penelitian.

1. Penginputan data penderita TB Paru (BTA+) di DKI Jakarta.
2. Analisis data deret waktu yang digunakan adalah analisis deret waktu univariat, hal ini dilakukan untuk mendapatkan data yang mengandung AR.
3. Untuk proses stasioner diperlukan *mean* dan *varians* yang konstan, juga kovarian  $\gamma_k = Cov(Z(t), Z(t+k)) = E[(Z(t) - \mu)(Z(t+k) - \mu)]$  dengan kovarian antara dua data deret waktu hanya bergantung pada jarak waktu antara dua periode waktu tersebut (Wei, 2006:10).
4. Identifikasi orde *autoregressive* (p) dalam model GSTAR diperoleh melalui deret waktu univariat dengan melihat korelogram ACF dan PACF pada masing-masing lokasi, selain itu dapat ditentukan dengan menentukan nilai *AIC*. Untuk memilih model terbaik dapat dilihat dari nilai *AIC* terkecil.
5. Menghitung pembobot lokasi pada model GSTAR dengan menggunakan pembobot lokasi invers jarak dan pembobot lokasi normalisasi korelasi silang.
6. Menduga parameter model GSTAR untuk data penderita TB Paru (BTA+) sesuai orde *autoregressive* dan orde spasial 1 untuk masing-masing pembobot lokasi.
7. Menguji kelayakan model model GSTAR, model dikatakan layak jika memenuhi asumsi *varians* konstan (*white noise*).
8. Melakukan validasi model GSTAR untuk data penderita TB Paru (BTA+) pada masing-masing pembobot lokasi menggunakan RMSE. Model dengan RMSE terkecil menyatakan model terbaik.
9. Salah satu tujuan dalam menganalisis data deret waktu untuk meramalkan nilai pengamatan dimasa mendatang.

## HASIL DAN PEMBAHASAN

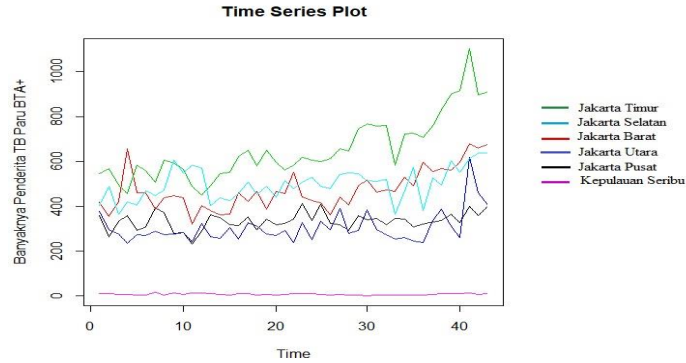
### Statistika Deskriptif Data Banyaknya Penderita TB Paru (BTA+) di DKI Jakarta

Statistika deskriptif secara umum digunakan untuk mempermudah pemahaman dan penjelasan tentang data banyaknya penderita TB Paru (BTA+) di DKI Jakarta. Statistika deskriptif dirangkum pada Tabel 1.

**TABEL 1. STATISTIKA DESKRIPTIF DATA BANYAKNYA PENDERITA TB PARU (BTA+)**

Lokasi	JakPus	JakBar	JakTim	JakUt	JakSel	K. Seribu
<b>Mean</b>	333,05	472,70	657,21	304,02	498,74	7,49
<b>Std. Deviasi</b>	38,06	91,34	138,76	70,73	71,60	3,55
<b>Minimum</b>	232	320	450	233	363	2
<b>Maximum</b>	411	679	1102	621	637	17

Berdasarkan Tabel 1 diketahui bahwa data banyaknya penderita TB Paru (BTA+) dari bulan Januari 2007 sampai September 2017 memiliki nilai rata-rata terbesar di Kota Jakarta Timur dan terkecil di Kabupaten Kepulauan Seribu, hal ini juga dipengaruhi dengan luas wilayah dan jumlah penduduk yang cukup besar di Jakarta Timur dibandingkan dengan kota lain yaitu sebanyak 2,8 juta penduduk dengan luas wilayah yang cukup besar yaitu 188,03  $km^2$  selain itu seperempat atau sekitar 28% penduduk tinggal di Jakarta Timur. Disusul Jakarta Barat 24% dan Jakarta Selatan 21% sedangkan wilayah Kabupaten Kepulauan Seribu hanya 0,23% dari total penduduk DKI Jakarta, hal ini disebabkan akses menuju Kepulauan Seribu yang mahal dan sulit sehingga menyebabkan sedikitnya penderita TB Paru (BTA+) di wilayah ini.



**GAMBAR 1. PLOT TIME SERIES PENDERITA TB PARU (BTA+) di DKI JAKARTA**

Berdasarkan Gambar 1 ditunjukkan bahwa plot *time series* pada keenam lokasi penderita TB Paru (BTA+) tidak mengalami pola musiman di masing-masing lokasi. Bila dilihat plot deret waktu Kota Jakarta Timur memiliki jumlah penderita terbanyak kemudian di susul oleh Jakarta Selatan kemudian Jakarta Barat, Jakarta Utara dan Jakarta Pusat, sedangkan untuk Kabupaten Kepulauan Seribu memiliki jumlah penderita TB Paru (BTA+) paling sedikit di antara kelima lokasi.

#### Korelasi Data Banyaknya Penderita TB Paru (BTA+) di DKI Jakarta

Besarnya hubungan satu lokasi terhadap lokasi lain pada suatu waktu dapat dilihat melalui nilai korelasi antar lokasi. Nilai korelasi antar lokasi disajikan pada Tabel 2.

**TABEL 2. KORELASI ANTAR LOKASI DI DKI JAKARTA**

JakPus	1	0,374	0,369	0,401	-0,036	0,142
JakBar	0,374	1	0,695	0,435	0,31	-0,083
JakTim	0,369	0,695	1	0,619	0,543	-0,045
JakUt	0,401	0,435	0,619	1	0,45	0,209
JakSel	-0,036	0,31	0,543	0,45	1	0,28
K. Seribu	0,142	-0,083	-0,045	0,209	0,28	1

Berdasarkan nilai korelasi Tabel 2 korelasi yang tinggi terjadi antara Jakarta Timur dan Jakarta Barat sebesar 0,695 hal ini menunjukkan bahwa Jakarta Timur dan Jakarta Barat heterogen, selanjutnya korelasi yang cukup tinggi terjadi pada Jakarta Timur dan Jakarta Utara sebesar 0,619 sedangkan untuk korelasi antara Kepulauan Seribu dengan kelima Kota lainnya sangat kecil hal ini menunjukkan bahwa data tidak bersifat heterogen. Secara umum data memiliki korelasi yang nyata karena memiliki korelasi dengan signifikansi kurang dari  $\alpha = 0,05$ . Korelasi antara Jakarta Pusat dan Jakarta Selatan yang menunjukkan hubungan tidak nyata dengan signifikansi sebesar 0,820. Namun, lokasi Jakarta Pusat dan Jakarta Selatan tidak harus dihilangkan karena korelasi kedua kota dengan tiga kota yang lain menunjukkan hubungan yang signifikan. Nilai korelasi Kepulauan Seribu dengan kelima Kota lainnya sangat kecil, dan menunjukkan hubungan tidak nyata dengan signifikansi lebih besar dari  $\alpha$ . Lokasi Kepulauan Seribu tidak dimasukkan ke dalam model karena korelasi Kepulauan Seribu dengan lima kota yang lain menunjukkan hubungan yang tidak signifikan.

#### Kestasioneran Data Penderita TB Paru (BTA+)

Syarat utama yang harus terpenuhi dalam deret waktu adalah kestasioneran data. Untuk melihat kesasioneran data dalam *mean* dapat dilakukan uji formal menggunakan uji *Augmented Dickey Fuller* atau (*ADF Test*) Tabel 3.

**TABEL 3. ADF TEST BANYAKNYA PENDERITA TB PARU (BTA+) DI KELIMA LOKASI**

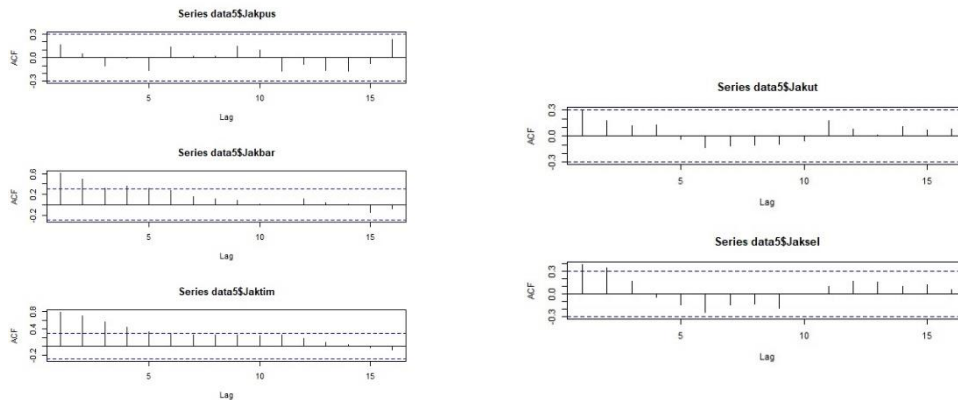
Lokasi	JakPus	JakBar	JakTim	JakUt	JakSel
ADF	-5,7441	-3,6372	-3,9041	-5,3877	-4,5636
p-value	0,01	0,04168	0,02276	0,01	0,01

keterangan:  $\alpha = 0,05$

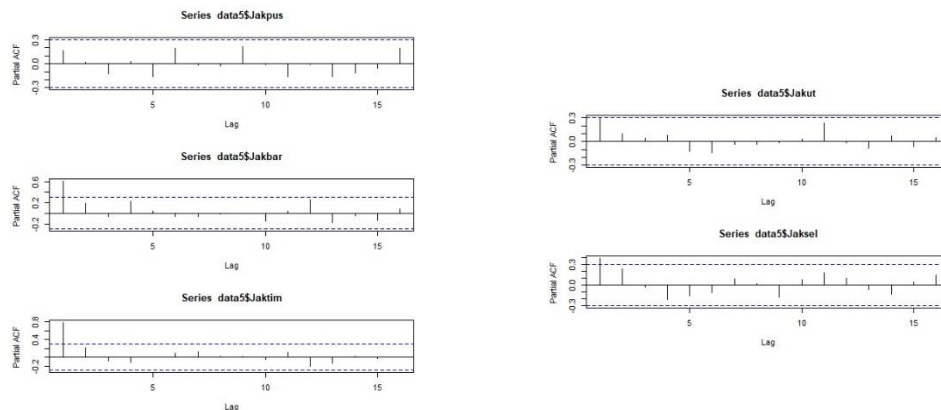
Dari hasil uji *Augmented Dickey Fuller* (ADF) pada Tabel 3 terlihat bahwa untuk data penderita TB Paru (BTA+) di kelima kota di DKI Jakarta, diperoleh nilai ADF di setiap lokasi lebih kecil dari  $\alpha = 0,05$  atau *p-value* < 0,05 sehingga tolak  $H_0$  artinya data sudah stasioner atau tidak mengandung *unit root* pada masing masing lokasi.

### Identifikasi Model GSTAR

Sebelum membangun model GSTAR, identifikasi terhadap data terlebih dahulu guna mendapatkan model yang sesuai dengan data penderita TB Paru (BTA+) di kelima kota DKI Jakarta diperoleh korelogram ACF dan PACF di masing-masing lokasi, yang disajikan pada Gambar 2 dan Gambar 3:



**GAMBAR 2. KORELOGRAM ACF KELIMA LOKASI PENDERITA TB PARU (BTA+)**



**GAMBAR 3. KORELOGRAM PACF KELIMA LOKASI PENDERITA TB PARU (BTA+)**

Selanjutnya dilakukan identifikasi orde model GSTAR untuk data penderita TB Paru (BTA+) di kelima kota di DKI Jakarta. Identifikasi orde spasial model GSTAR pada penelitian ini diberi orde satu, karena lokasi penelitian yang berada pada satu wilayah. Selanjutnya identifikasi orde AR model GSTAR dapat dilakukan dengan melihat korelogram PACF data yang disajikan pada Gambar 3. Berdasarkan Gambar 3 terlihat bahwa pada Kota Jakarta Barat, Jakarta Timur, Jakarta Utara dan Jakarta Selatan memotong di lag ke-1.



Pada model GSTAR penentuan orde *autoregressive* dapat dilakukan dari orde model VAR, dalam hal ini untuk memperoleh model yang sesuai ditentukan dengan panjang lag optimal berdasarkan nilai *Akaike Information Criterion* (AIC) terkecil Tabel 4.

**TABEL 4. NILAI AIC MASING-MASING LAG**

Lag	0	1	2	3	4
AIC	55,02228	54,60588*	55,00238	54,94766	55,23069

Berdasarkan Tabel 4 diperoleh bahwa orde *autoregressive* (p) model GSTAR pada data penderita TB Paru (BTA+) di kelima kota DKI Jakarta adalah 1, karena lag 1 memiliki nilai AIC terkecil sebesar 54,60588. Sehingga model yang sesuai adalah GSTAR (1;1).

**Perhitungan Pembobot Lokasi Invers Jarak**

Matriks bobot lokasi invers jarak diperoleh berdasarkan jarak yang sebenarnya. Pembobot invers jarak memberikan nilai bobot yang besar untuk jarak yang lebih dekat dan memberikan pembobot yang kecil untuk jarak yang jauh, perhitungan pembobot lokasi invers jarak dapat dilihat pada matriks berikut

$$W^1 = \begin{bmatrix} 0 & 0,231 & 0,233 & 0,308 & 0,228 \\ 0,342 & 0 & 0,173 & 0,237 & 0,248 \\ 0,321 & 0,161 & 0 & 0,264 & 0,254 \\ 0,389 & 0,202 & 0,242 & 0 & 0,167 \\ 0,319 & 0,236 & 0,259 & 0,186 & 0 \end{bmatrix}$$

**Perhitungan Pembobot Lokasi Normalisasi Korelasi Silang**

Matriks pembobot lokasi normalisasi korelasi silang diperoleh berdasarkan hasil normalisasi korelasi silang antar lokasi. Perhitungan pembobot lokasi normalisasi korelasi silang dapat dilihat pada matriks berikut

$$W^1 = \begin{bmatrix} 0 & 0,385 & 0,373 & 0,120 & 0,122 \\ 0,127 & 0 & 0,475 & 0,224 & 0,174 \\ 0,162 & 0,381 & 0 & 0,235 & 0,222 \\ 0,077 & 0,302 & 0,393 & 0 & 0,228 \\ 0,101 & 0,263 & 0,351 & 0,285 & 0 \end{bmatrix}$$

**Estimasi Parameter Model GSTAR (1;1) dengan Pembobot Lokasi Invers Jarak dan Pembobot Lokasi Normalisasi Korelasi Silang**

Estimasi parameter model GSTAR(1;1) dengan pembobot lokasi invers jarak dan pembobot lokasi normalisasi korelasi silang dilakukan dengan Metode Kuadrat Terkecil dan disajikan pada Tabel 5 dan Tabel 6.

**TABEL 5. HASIL ESTIMASI PARAMETER GSTAR(1;1) DENGAN PEMBOBOT LOKASI INVERS JARAK**

Parameter	Estimasi	Std. Error	t-value	p-value	Kesimpulan
$\phi_{10}^1$	0,614	0,192	3,193	0,00163	Signifikan
$\phi_{10}^2$	0,566	0,141	4,008	8,62e-05	Signifikan
$\phi_{10}^3$	0,793	0,110	7,208	1,14e-11	Signifikan
$\phi_{10}^4$	-0,044	0,183	-0,239	0,81146	Tdk. Signifikan
$\phi_{10}^5$	0,436	0,130	3,361	0,00093	Signifikan
$\phi_{11}^1$	0,271	0,136	1,990	0,04800	Signifikan
$\phi_{11}^2$	0,497	0,159	3,132	0,00199	Signifikan
$\phi_{11}^3$	0,369	0,188	1,965	0,05076	Tdk. Signifikan
$\phi_{11}^4$	0,681	0,121	5,621	6,32e-08	Signifikan
$\phi_{11}^5$	0,641	0,146	4,395	1,79e-05	Signifikan

**TABEL 6. HASIL ESTIMASI PARAMETER GSTAR(1;1) DENGAN PEMBOBOT LOKASI NORMALISASI KORELASI SILANG**

Parameter	Estimasi	Std. Error	t-value	p-value	Kesimpulan
$\phi_{10}^1$	0,636	0,182	3,496	0,001	Signifikan
$\phi_{10}^2$	0,510	0,142	3,593	0,001	Signifikan
$\phi_{10}^3$	0,705	0,118	6,043	7,29e-09	Signifikan
$\phi_{10}^4$	-0,039	0,177	-0,221	0,825	Tdk. Signifikan
$\phi_{10}^5$	0,493	0,125	3,957	0,000105	Signifikan
$\phi_{11}^1$	0,229	0,115	1,986	0,048	Signifikan
$\phi_{11}^2$	0,474	0,135	3,513	0,000547	Signifikan
$\phi_{11}^3$	0,486	0,186	2,605	0,009867	Signifikan
$\phi_{11}^4$	0,585	0,101	5,780	2,84e-08	Signifikan
$\phi_{11}^5$	0,538	0,131	4,117	5,62e-05	Signifikan

Nilai estimasi parameter yang signifikan dimana  $p$ -value kurang dari  $\alpha = 0,05$ . Sedangkan nilai estimasi parameter dengan pembobot lokasi invers jarak  $\phi_{10}^4$  dan  $\phi_{11}^3$  memiliki  $p$ -value berturut-turut sebesar 0,81146 dan 0,05076, serta  $\phi_{10}^4$  dengan pembobot lokasi normalisasi korelasi silang sebesar 0,825 yang menunjukkan parameter tersebut tidak signifikan. Menurut (Kostenko dan Hyndman, 2008) dalam Rani *et al.* (2013) menjelaskan bahwa variabel yang tidak signifikan dapat digunakan untuk melakukan peramalan. Sehingga dalam penelitian ini parameter yang tidak signifikan akan tetap digunakan.

Persamaan GSTAR(1;1) dengan pembobot lokasi invers jarak sebagai berikut

$$\hat{Z}_1(t) = 0,614Z_1(t-1) + 0,0626Z_2(t-1) + 0,0631(t-1) + 0,0835Z_4(t-1) + 0,0618Z_5(t-1) + e_1(t).$$

$$\hat{Z}_2(t) = 0,1700Z_1(t-1) + 0,566Z_2(t-1) + 0,0860Z_3(t-1) + 0,1178Z_4(t-1) + 0,1233Z_5(t-1) + e_2(t).$$

$$\hat{Z}_3(t) = 0,1184Z_1(t-1) + 0,0594Z_2(t-1) + 0,793Z_3(t-1) + 0,0974Z_4(t-1) + 0,0937Z_5(t-1) + e_3(t).$$

$$\hat{Z}_4(t) = 0,2649Z_1(t-1) + 0,1376Z_2(t-1) + 0,1648Z_3(t-1) - 0,044Z_4(t-1) + 0,1137Z_5(t-1) + e_4(t).$$

$$\hat{Z}_5(t) = 0,2045Z_1(t-1) + 0,1513Z_2(t-1) + 0,1660Z_3(t-1) + 0,1192Z_4(t-1) + 0,436Z_5(t-1) + e_5(t).$$

Persamaan GSTAR(1;1) dengan pembobot lokasi normalisasi korelasi silang sebagai berikut

$$\hat{Z}_1(t) = 0,636Z_1(t-1) + 0,0882Z_2(t-1) + 0,0854Z_3(t-1) + 0,0275Z_4(t-1) + 0,0279Z_5(t-1) + e_1(t).$$

$$\hat{Z}_2(t) = 0,0602Z_1(t-1) + 0,510Z_2(t-1) + 0,2166Z_3(t-1) + 0,1062Z_4(t-1) + 0,0825Z_5(t-1) + e_2(t).$$

$$\hat{Z}_3(t) = 0,0787Z_1(t-1) + 0,1852Z_2(t-1) + 0,705Z_3(t-1) + 0,1142Z_4(t-1) + 0,1079Z_5(t-1) + e_3(t).$$

$$\hat{Z}_4(t) = 0,0450Z_1(t-1) + 0,1767Z_2(t-1) + 0,2299Z_3(t-1) - 0,039Z_4(t-1) + 0,1334Z_5(t-1) + e_4(t).$$

$$\hat{Z}_5(t) = 0,0543Z_1(t-1) + 0,1415Z_2(t-1) + 0,1888Z_3(t-1) + 0,1533Z_4(t-1) + 0,493Z_5(t-1) + e_5(t).$$

**Pengujian Residual (White Noise)**

Pemeriksaan asumsi residual *white noise* ini menggunakan uji *Ljung-Box Pearce* (LB).

Berikut hasil uji LB untuk model GSTAR (1;1) dengan pembobot lokasi invers jarak dan pembobot lokasi normalisasi korelasi silang disajikan pada Tabel 7

**TABEL 7. HASIL PENGUJIAN ASUMSI RESIDUAL WHITE NOISE**

Pembobot Lokasi	Asumsi Residual <i>White Noise</i>
Invers Jarak	<i>White Noise</i>
Normalisasi Korelasi Silang	<i>White Noise</i>

Berdasarkan Tabel 6 pembobot lokasi invers jarak dan pembobot lokasi normalisasi korelasi silang telah memenuhi asumsi residual *white noise* karena nilai  $LB > \chi^2_{1-\alpha, k}$  tabel.

### Pengujian Ketepatan Model

*Root Mean Square Error*(RMSE) digunakan untuk pengukuran kebaikan model dan ketepatan ramalan pada data penderita TB Paru (BTA+) di kelima kota DKI Jakarta di masa yang akan datang. *Mean Absolute Percentage Error* (MAPE) untuk menguji ukuran kesalahan nilai dugaan model yang dinyatakan dalam bentuk rata-rata persentase absolut residual disajikan pada Tabel 8

**TABEL 8. NILAI RMSE DAN MAPE**

Pembobot Lokasi	RMSE	MAPE
Invers Jarak	74,83935	10,6%
Normalisasi Korelasi Silang	73,57728	10,5%

Berdasarkan Tabel 7 didapatkan informasi bahwa model GSTAR(1;1) dengan pembobot lokasi normalisasi korelasi silang memiliki nilai RMSE dan MAPE yang lebih kecil dibandingkan model GSTAR(1;1) dengan pembobot lokasi invers jarak. Maka, model yang akan digunakan untuk meramalkan data penderita TB Paru (BTA+) di kelima kota DKI Jakarta di masa yang akan datang adalah model GSTAR(1;1) dengan pembobot lokasi normalisasi korelasi silang.

### Peramalan

Pada penelitian ini dalam meramalkan banyaknya data penderita TB Paru (BTA+) di kelima kotamadya di DKI Jakarta model yang digunakan adalah GSTAR(1;1) dengan pembobot lokasi normalisasi korelasi silang. Model tersebut digunakan untuk peramalan selama 5 waktu yang akan datang. Hasil ramalan 5 waktu ke depan disajikan pada Tabel 9.

**TABEL 9. RAMALAN 5 WAKTU KE DEPAN DATA BANYAKNYA PENDERITA TB PARU (BTA+) DI KELIMA KOTAMADYA DKI JAKARTA**

Waktu	JakPus	JakBar	JakTim	JakUt	JakSel
Okt-Des 2017	419	662	913	416	665
Jan-Mar 2018	433	659	918	418	680
Apr-Jun 2018	442	662	925	422	690
Jul-Sept 2018	450	666	932	425	697
Okt-Des 2018	455	671	939	428	703

## KESIMPULAN DAN SARAN

### Kesimpulan

Berdasarkan hasil dan pembahasan tersebut, dapat kita ketahui bahwa langkah-langkah sistematis pemodelan data deret waktu dan lokasi menggunakan model *Generalized Space Time Autoregressive* (GSTAR) dengan pembobot lokasi normalisasi korelasi silang dan pembobot lokasi invers jarak yaitu: statistik deskriptif, uji stasioneritas, identifikasi orde *autoregressive* (p), menentukan orde spasial ( $\lambda_k$ ), menghitung pembobot lokasi dengan menggunakan pembobot lokasi normalisasi korelasi silang dan pembobot lokasi invers jarak, estimasi parameter untuk masing-masing pembobot lokasi, pengujian residual *white noise*, uji kelayakan model dan terakhir peramalan.

Model GSTAR(1;1) dengan pembobot lokasi normalisasi korelasi silang merupakan model GSTAR terbaik dengan persamaan sebagai berikut:

$$\hat{Z}_1(t) = 0,636Z_1(t-1) + 0,0882Z_2(t-1) + 0,0854Z_3(t-1) + 0,0275Z_4(t-1) + 0,0279Z_5(t-1) + e_1(t).$$

$$\hat{Z}_2(t) = 0,0602Z_1(t-1) + 0,510Z_2(t-1) + 0,2166Z_3(t-1) + 0,1062Z_4(t-1) + 0,0825Z_5(t-1) + e_2(t).$$

$$\hat{Z}_3(t) = 0,0787Z_1(t-1) + 0,1852Z_2(t-1) + 0,705Z_3(t-1) + 0,1142Z_4(t-1) + 0,1079Z_5(t-1) + e_3(t).$$

$$\hat{Z}_4(t) = 0,0450Z_1(t-1) + 0,1767Z_2(t-1) + 0,2299Z_3(t-1) - 0,039Z_4(t-1) + 0,1334Z_5(t-1) + e_4(t).$$

$$\hat{Z}_5(t) = 0,0543Z_1(t-1) + 0,1415Z_2(t-1) + 0,1888Z_3(t-1) + 0,1533Z_4(t-1) + 0,493Z_5(t-1) + e_5(t).$$

Model GSTAR terbaik yang dapat digunakan untuk peramalan data banyaknya penderita TB Paru (BTA+) di kelima Kota DKI Jakarta adalah model GSTAR(1;1) menggunakan pembobot lokasi normalisasi korelasi silang karena memenuhi asumsi *white noise* dengan RMSE sebesar 73,57728 dan MAPE sebesar 10,5% dibandingkan dengan GSTAR(1;1) menggunakan pembobot lokasi invers jarak.

### Saran

Hal yang disarankan oleh penulis untuk peneliti selanjutnya adalah dapat dikembangkan dengan menerapkan model GSTAR pada kasus lain seperti indeks harga konsumen. Selain itu, penelitian ini dapat dikembangkan dengan pembobot lokasi lain seperti pembobot lokasi biner. Melihat keterbatasan data, maka peneliti lain dapat menambahkan periode maupun menambahkan jumlah lokasi atau wilayah penelitian karena dimungkinkan dengan sampel yang banyak akan lebih bisa menjelaskan nilai sebenarnya sehingga dapat digunakan untuk peramalan di masa mendatang.

### UCAPAN TERIMA KASIH

Terima kasih kepada Ibu Dra. Widyanti Rahayu, M.Si dan Ibu Eti Dwi Wiraningsih, M.Si atas saran dan ketersediaan waktu dalam membimbing penulis, sehingga tulisan ini berhasil selesai dengan baik serta P2PML Kementerian Kesehatan atas saran dan ketersediaan data dalam menunjang penelitian ini.

### REFERENSI

- Anggraeni, D. et al. Aplikasi Generalized Space Time Autoregressive (GSTAR) pada Pemodelan Volume Kendaraan Masuk Tol Semarang. Jurnal Media Statistik. Vol. 6, No. 2, pp. 71-80, 2013.
- Borovkova, S.A. et al. Consistency and Asymptotic Normality of Least Squares Estimators in Generalized STAR Models. Statistica Neerlandica, 2008.
- Borovkova, S.A. et al. Least Squares Estimation of Generalized Space Time Autoregressive (GSTAR) Model and Its Properties. Statistica Neerlandica, 2008.
- Departemen Kesehatan RI. 2016. Profil Kesehatan Provinsi DKI Jakarta 2016.
- Dhoriva Urwatul Wutsqa et al. 2010. Generalized Space Time Autoregressive Modeling. The 6th IMT-GT Conference on Mathematics, Statistics, and its Applications (ICMSA): Proceeding. pp(752-261). Kuala Lumpur: Universitas Tunku Abdul Rahman.
- Gujarati. 2006. Dasar-Dasar Ekonometrika = Basic Econometrics. Edisi Ketiga. Erlangga.
- Kementrian Kesehatan RI. 2016. Infodatin Tuberkulosis.
- Kristien Margi S. et al. Analisa dan Penerapan Metode Single Exponential Smoothing untuk Prediksi Penjualan Pada Periode Tertentu.. Prosiding SNATIF Ke-2, 2015.

- Lutkepohl, H. 2005. *New Introduction to Multiple Time Series Analysis*. Springer.
- Makridakis, S. et al. 1995. *Metode dan Aplikasi Peramalan*. Edisi Kedua. Erlangga.
- Ruchjana B.N. 2002. *Suatu Model Generalized Space Time Autoregressive dan Penerapannya pada Produksi Minyak Bumi*. Disertasi tidak diterbitkan. Bandung: Program Doktor ITB.
- Rani S.A.P. et al. 2013. *Pemodelan Generalized Space Time Autoregressive (GSTAR(p;1)) (Penerapan pada Data Angka Kesakitan Penyakit ISPA di Kota Malang)*. Jurnal. Universitas Brawijaya Malang.
- Suhartono dan Subanar. *The Optimal Determination of Space Weight in GSTAR by Using Cross-Correlation Inference*. *Journal of Quantitative Methods*. Vol. 2, No. 2, December 2006.
- Susanti, S. et al. *Model Generalized Space Time Autoregressive Integrated dengan Pembobot Normalisasi Korelasi Silang pada Perkembangan Aset BPR di Provinsi Jawa Barat, Jawa Tengah, dan Jawa Timur*. *Jurnal Statistik*, 2017.
- Wei, W.W.S. 2006. *Time Series Univariate and Multivariate Methods*. Addison Wesley Publishing Company. Inc: Canada.
- Wijayanti Lumbanraja. 2015. *Penerapan Model Generalized Space Time pada Harga Bawang Merah di Pulau Jawa*. Skripsi. Institut Pertanian Bogor.