

PEMODELAN JUMLAH WISATAWAN DI JAWA TENGAH MENGGUNAKAN METODE GENERALIZED SPACE TIME AUTOREGRESSIVE - SEEMINGLY UNRELATED REGRESSION (GSTAR-SUR)

Innosensia Adella¹, Dwi Ispriyanti², Hasbi Yasin³

^{1,2,3} Departemen Statistika, Fakultas Sains dan Matematika, Universitas Diponegoro

*Email : innosendella26@gmail.com

ABSTRACT

Space-time model is a model that can explain data with spatial and time characteristics. The Generalized Space Time Autoregressive (GSTAR) model is one of the generalized space-time models from the Space Time Autoregressive (STAR) model. The GSTAR model is more flexible when dealing with areas that have heterogeneous characteristics than the STAR model. The GSTAR model models time series data in multiple regions at once. This model can then be used to model data on the number of tourists in four regions in Central Java, namely Semarang, Jepara, Magelang and Semarang district for the 2014 to 2019 period. In Central Java. On the residual model, the Lagrange Multiplier Test is carried out and it is known that there is a correlation between the residuals. The modeling was continued by using the Generalized Space Time Autoregressive – Seemingly Unrelated Regression (GSTAR-SUR) model. GSTAR-SUR is one of the more efficient models used to model GSTAR with correlated residuals. Residual through the white-noise assumption test, it is found that the appropriate model is the GSTAR-SUR(2,1) model. This model can then be used in forecasting data on the number of tourists in Semarang, Jepara, Magelang and Semarang district in the next period.

Keywords : Tourists, Space Time, GSTAR, SUR, White-Noise

1. PENDAHULUAN

Wisata menjadi salah satu kegiatan yang tidak terpisahkan dari masyarakat, wisata merupakan kegiatan perjalanan yang dilakukan oleh seseorang atau sekelompok orang dengan mengunjungi tempat tertentu dalam jangka waktu sementara. Sedangkan wisatawan adalah seseorang yang melakukan kegiatan wisata atau orang yang melakukan perjalanan wisata dan tinggal sementara sekurang-kurangnya 24 jam. Salah satu wilayah di Indonesia yang dikenal memiliki banyak destinasi wisata yang sering dikunjungi baik wisatawan lokal maupun wisatawan mancanegara adalah Jawa Tengah. Menurut data Buku Statistik Pariwisata Jawa Tengah dari Tahun 2015-2019 terdapat empat wilayah yang hampir selalu masuk dalam urutan sepuluh besar wilayah dengan jumlah wisatawan terbanyak di Jawa Tengah setiap tahunnya. Wilayah tersebut diantaranya Kota Semarang, Kabupaten Jepara, Kabupaten Magelang dan Kabupaten Semarang. Jumlah wisatawan yang berkunjung di wilayah tersebut diduga tidak hanya berkaitan dengan jumlah wisatawan periode sebelumnya namun juga lokasi masing-masing wilayah. Model yang berkaitan dengan data runtun waktu dan lokasi adalah model *space time*.

Model GSTAR menggunakan metode OLS dalam mengestimasi parameter. Sedangkan apabila residual antar persamaan dalam model saling berkorelasi maka estimasi dengan metode OLS pada model GSTAR akan menghasilkan estimator yang tidak efisien. Dengan demikian digunakan metode estimasi yang sesuai untuk residual yang saling berkorelasi yaitu *generalized*

least square (GLS) yang digunakan dalam model *seemingly unrelated regression* (SUR). Estimasi model GSTAR dengan metode GLS ini kemudian disebut sebagai model GSTAR-SUR.

2. TINJAUAN PUSTAKA

Stasioneritas memiliki arti bahwa dalam data tidak terdapat unsur musiman dan trend baik trend naik maupun turun. Stasioneritas data dibagi menjadi dua yaitu stasioneritas dalam rata-rata dan stasioneritas dalam varian. Uji stasioneritas dalam mean yang digunakan pada model GSTAR adalah dengan melihat plot *Matrix Autocorrelation Function* (MACF) dan *Matrix Partial Autocorrelation Function* (MPACF). Sedangkan uji stasioner dalam varian tidak dilakukan dan data dianggap stasioner dalam varian.

Matriks fungsi autokorelasi digunakan dalam menentukan orde model Moving Average (MA). Namun semakin besar ukuran dimensi dan vektornya maka akan semakin rumit bentuk grafik dan matriks yang terbentuk, sehingga dapat menyulitkan identifikasi, sehingga Tiao dan Box dalam Wei (2006) memperkenalkan metode untuk meringkas korelasi sampel, yaitu dengan menggunakan simbol (+), (-) dan (.) dalam korelasi ke (i,j). Simbol-simbol tersebut dapat diartikan sebagai berikut :

- Simbol (+) digunakan untuk menunjukkan nilai korelasi yang lebih besar dari 2 kali standard error dari $\hat{\rho}_{ij}(k)$ artinya bahwa komponen (i,j) memiliki hubungan korelasi positif.
- Simbol (-) digunakan untuk menunjukkan nilai korelasi yang lebih kecil dari -2 kali standard error dari $\hat{\rho}_{ij}(k)$ artinya bahwa komponen (i,j) memiliki hubungan korelasi negatif.
- Simbol (.) digunakan untuk menunjukkan nilai korelasi berada di antara 2 dan -2 standard error dari $\hat{\rho}_{ij}(k)$ artinya bahwa komponen (i,j) tidak memiliki hubungan korelasi baik positif maupun negatif.

Untuk memperoleh kesimpulan bahwa data dapat dikatakan stasioner dalam mean maka visual MACF yang diperoleh harus didominasi oleh simbol (.) dari pada simbol (+) dan (-).

MPACF sebagai salah satu uji stasioneritas yang berpengaruh dalam menentukan orde pada model Autoregressive (AR). Sama halnya dengan MACF, MPACF juga disederhanakan dengan menggunakan simbol (.), (+) dan (-). Dengan kesimpulan bahwa data dapat dikatakan stasioner dalam mean maka visual MACF yang diperoleh harus didominasi oleh simbol (.) dari pada simbol (+) dan (-).

Menurut Borovkova (2008) persamaan model *Generalized Space Time Autoregressive* (GSTAR) dapat dinyatakan sebagai berikut :

$$Z_i(t) = \sum_{k=1}^p \sum_{i=1}^N \left[\Phi_{k0}^{(i)} Z_i(t-k) + \sum_{l=1}^{\lambda_k} \Phi_{kl}^{(i)} W^{(l)} Z_i(t-k) \right] + \varepsilon_i(t)$$

Dengan $Z_i(t)$ = nilai observasi pada waktu ke- t dan lokasi ke- i , p = orde autoregressive, λ_k = orde spasial pada autoregressive k , $\Phi_{kl}^{(i)}$ = matriks diagonal parameter pada orde autoregressive k dan orde spasial l di setiap lokasi ke- i , $W^{(l)}$ = matriks bobot, $\varepsilon_i(t)$ = vektor error.

Penentuan orde autoregressive GSTAR dilakukan dengan mengamati MPACF serta melihat nilai AIC setiap orde. Orde dengan kriteria nilai AIC yang terkecil dipilih sebagai orde

autoregressive model GSTAR. Sedangkan orde spasial model GSTAR pada umumnya dibatasi pada orde spasial 1, karena orde yang tinggi akan sulit diinterpretasikan (Wutsqa dkk, 2010).

Pembobot invers jarak atau *Invers Space* adalah metode pembobotan yang diperoleh dengan mengacu pada jarak antar lokasi. Menurut Johntson, pembobot invers jarak antara 2 lokasi id an j dinyatakan sebagai berikut :

$$w_{ij} = \frac{\frac{1}{d_{ij}}}{\sum_j \frac{1}{d_{ij}}}, \text{ jika } i \neq j \text{ dan } w_{ij} = 0, \text{ jika } i = j$$

Estimasi parameter pada model GSTAR dapat dilakukan dengan metode *Ordinary Least Square* (OLS), yaitu dengan meminimumkan jumlah kuadrat residual diperoleh sebagai berikut :

$$\hat{\Phi} = (Z'Z^*)^{-1}Z^*Z$$

Model *Seemingly Unrelated Regression* (SUR) dapat dinyatakan sebagai himpunan N buah persamaan yang berhubungan karena sesatan antar persamaan saling berkorelasi. Metode SUR merupakan pengembangan dari metode *Generalized Least Square* (GLS). Sehingga estimator GSTAR-SUR dengan metode GLS diperoleh sebagai berikut :

$$\hat{\Phi}_{GLS} = (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}Z$$

Dalam model SUR hal yang penting sebelum melakukan estimasi parameter adalah melakukan uji kovarian untuk mengetahui apakah terdapat korelasi antara sesatan persamaan dengan persamaan yang lain dalam struktur tersebut. Salah satu metode yang dapat digunakan adalah Uji Lagrange Multiplier yang dapat dinyatakan sebagai berikut :

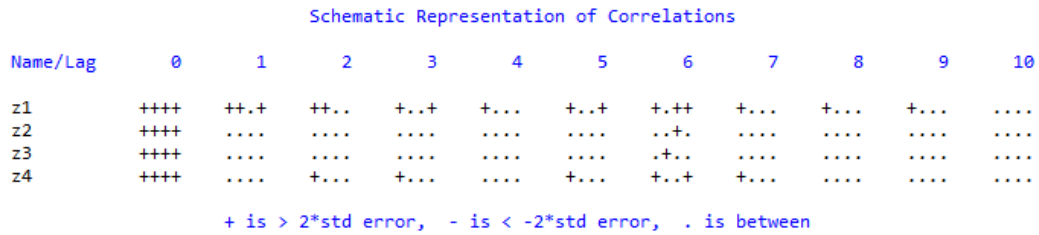
$$\lambda_{LM} = T \sum_{i=2}^N \sum_{j=1}^{i-1} r_{ij}^2$$

Dengan r_{ij} adalah korelasi sesatan antara persamaan ke-i dengan persamaan ke-j dan T adalah banyaknya pengamatan. Terdapat kolerasi sesatan antar persamaan jika $\lambda_{LM} > \chi^2_{\left(\frac{N(N-1)}{2}, \alpha\right)}$.

Model GSTAR dikatakan layak apabila residualnya bersifat *white noise* dan berdistribusi normal. Residual bersifat *white noise* artinya tidak terdapat korelasi antar residual dengan $\mu = 0$ dan σ^2 konstan. Untuk melihat pemenuhan asumsi *white noise* dapat dilakukan dengan uji Ljung Box-Pearce.

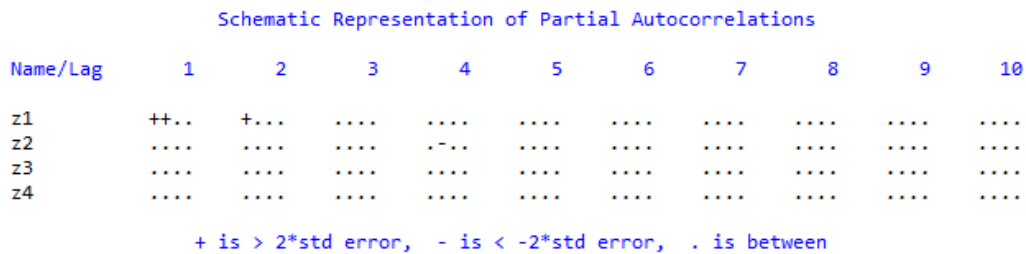
3. HASIL DAN PEMBAHASAN

Pada Gambar 1 di atas terlihat bahwa bahwa simbol (.) muncul lebih banyak dibanding dengan simbol (+) dan (-). Tanda (.) memiliki arti bahwa data jumlah wisatawan pada empat wilayah stasioner secara bersama-sama, dengan demikian tidak perlu dilakukan *differencing*.



Gambar 1. Representation of Correlation

MPACF data jumlah wisatawan di empat wilayah tertera pada **Gambar 2** sebagai berikut :



Gambar 2. Representation of Partial Correlation

Dari **Gambar 2**, diperoleh bahwa lag yang signifikan adalah lag 1, 2 dan 4. Selanjutnya dilakukan pengamatan terhadap nilai AIC pada masing-masing lag tersebut :

Tabel 1. Nilai AIC

Lag	AIC
0	5635,648
1	5621,508*
2	5635,035
3	5642,026

Dari **Tabel 1**, terlihat bahwa nilai AIC terkecil terdapat pada lag 1 dengan nilai AIC sebesar 5621,508. Sehingga orde autoregressive (ρ) yang terpilih adalah orde 1 sedangkan orde spasial yang digunakan yaitu orde 1 karena orde spasial lebih dari 1 akan sulit untuk diinterpretasikan dalam model. Dengan demikian, model GSTAR yang terpilih adalah GSTAR (1,1). Diperoleh matriks pembobot invers jarak sebagai berikut :

$$W = \begin{bmatrix} 0 & 0,291759 & 0,211684 & 0,49557 \\ 0,485472 & 0 & 0,206592 & 0,307936 \\ 0,300817 & 0,176436 & 0 & 0,522746 \\ 0,473148 & 0,176339 & 0,350513 & 0 \end{bmatrix}$$

Diperoleh model GSTAR-OLS (1,1) jumlah wisatawan untuk masing-masing lokasi sebagai berikut :

1. Persamaan Model GSTAR-OLS (1,1) jumlah wisatawan di Kota Semarang

$$\hat{Z}_1(t) = 0,51981Z_1(t-1) + 0,21378Z_2(t-1) + 0,15511Z_3(t-1) + 0,36311Z_4(t-1) + \varepsilon_1(t)$$

2. Persamaan Model GSTAR-OLS (1,1) jumlah wisatawan di Kab. Jepara

$$\hat{Z}_2(t) = 0,21441Z_1(t-1) + 0,09124Z_3(t-1) + 0,136Z_4(t-1) + \varepsilon_2(t)$$

3. Persamaan Model GSTAR-OLS (1,1) jumlah wisatawan di Kab. Magelang

$$\hat{Z}_3(t) = 0,40679Z_3(t-1) + 0,19738Z_1(t-1) + 0,11577Z_2(t-1) + 0,34299Z_4(t-1) + \varepsilon_3(t)$$

4. Persamaan Model GSTAR-OLS (1,1) jumlah wisatawan di Kab. Semarang

$$\hat{Z}_4(t) = 0,19039Z_1(t-1) + 0,07096Z_2(t-1) + 0,14104Z_3(t-1) + \varepsilon_4(t)$$

Uji Lagrange Multiplier dilakukan sebagai berikut :

Hipotesis :

$H_0 : Cov(\varepsilon_{ij}, \varepsilon_{j1}) = 0$ untuk semua $i \neq j$ (tidak ada korelasi sesatan antar lokasi)

$H_1 : Cov(\varepsilon_{ij}, \varepsilon_{j1}) \neq 0$ untuk semua $i \neq j$ (ada korelasi sesatan antar lokasi)

Taraf Signifikansi : $\alpha = 5\%$

Statistik Uji : $\lambda_{LM} = 52,286$

Daerah Kritis :

H_0 ditolak jika nilai $\lambda_{LM} > \chi^2_{\left(\frac{4(4-1)}{2}, 0,05\right)} = 12,59$

Kesimpulan :

H_0 ditolak karena $\lambda_{LM} = 52,286 > 12,59$ sehingga dapat disimpulkan bahwa terdapat korelasi sesatan antar lokasi.

Diperoleh model GSTAR-SUR (1,1) jumlah wisatwan untuk masing-masing lokasi sebagai berikut :

1. Persamaan Model GSTAR-SUR (1,1) jumlah wisatawan di Kota Semarang

$$\hat{Z}_1(t) = 0,51524Z_1(t-1) + 0,21594Z_2(t-1) + 0,15668Z_3(t-1) + 0,36679Z_4(t-1) + \varepsilon_1(t)$$

2. Persamaan Model GSTAR-SUR (1,1) jumlah wisatawan di Kab. Jepara

$$\hat{Z}_2(t) = 0,21381Z_1(t-1) + 0,09099Z_3(t-1) + 0,13562Z_4(t-1) + \varepsilon_2(t)$$

3. Persamaan Model GSTAR-SUR (1,1) jumlah wisatawan di Kab. Magelang

$$\hat{Z}_3(t) = 0,51026Z_3(t-1) + \varepsilon_3(t)$$

4. Persamaan Model GSTAR-SUR (1,1) jumlah wisatawan di Kab. Semarang

$$\hat{Z}_4(t) = 0,21066Z_1(t-1) + 0,07851Z_2(t-1) + 0,15606Z_3(t-1) + \varepsilon_4(t)$$

Berdasarkan Uji Ljung-Box yang dilakukan, maka dapat disimpulkan bahwa residual Model GSTAR-SUR(1,1) untuk lokasi Kabupaten Jepara dan Kabupaten Magelang bersifat white noise. Sedangkan residual Model GSTAR-SUR(1,1) untuk Kota Semarang dan Kabupaten Semarang tidak bersifat white noise.

Dengan demikian diketahui bahwa uji asumsi residual tidak terpenuhi untuk seluruh lokasi. Untuk memperoleh model yang sesuai perlu dilakukan pemodelan ulang dengan menetapkan orde autoregressive yang baru.

Diperoleh model GSTAR-OLS (2,1) jumlah wisatwan untuk masing-masing lokasi sebagai berikut :

1. Persamaan Model GSTAR-OLS (2,1) jumlah wisatawan di Kota Semarang

$$\hat{Z}_1(t) = 0,2292Z_1(t-1) + 0,19591Z_2(t-1) + 0,14134Z_3(t-1) + 0,33089Z_4(t-1) + \varepsilon_1(t)$$

2. Persamaan Model GSTAR-OLS (2,1) jumlah wisatawan di Kab. Jepara

$$\hat{Z}_2(t) = 0,21997Z_1(t-2) + 0,09361Z_3(t-2) + 0,13953Z_4(t-2) + \varepsilon_2(t)$$

3. Persamaan Model GSTAR-OLS (2,1) jumlah wisatawan di Kab. Magelang

$$\hat{Z}_3(t) = 0,402691Z_3(t-1) + \varepsilon_3(t)$$

4. Persamaan Model GSTAR-OLS (2,1) jumlah wisatawan di Kab. Semarang

$$\hat{Z}_4(t) = 0,27013Z_4(t-1) + \varepsilon_4(t)$$

Uji Lagrange Multiplier dilakukan sebagai berikut :

Hipotesis :

$$H_0 : Cov(\varepsilon_{ij}, \varepsilon_{j1}) = 0 \text{ untuk semua } i \neq j \text{ (tidak ada korelasi sesatan antar lokasi)}$$

$$H_1 : Cov(\varepsilon_{ij}, \varepsilon_{j1}) \neq 0 \text{ untuk semua } i \neq j \text{ (ada korelasi sesatan antar lokasi)}$$

Taraf Signifikansi : $\alpha = 5\%$

Statistik Uji : $\lambda_{LM} = 93,38$

Daerah Kritis :

$$H_0 \text{ ditolak jika nilai } \lambda_{LM} > \chi^2_{\left(\frac{4(4-1)}{2}, 0,05\right)} = 12,59$$

Kesimpulan :

H_0 ditolak karena $\lambda_{LM} = 93,38 > 12,59$ sehingga dapat disimpulkan bahwa terdapat korelasi sesatan antar lokasi.

Diperoleh model GSTAR-SUR (2,1) jumlah wisatwan untuk masing-masing lokasi sebagai berikut :

1. Persamaan Model GSTAR-SUR (2,1) jumlah wisatawan di Kota Semarang

$$\hat{Z}_1(t) = 0,2971Z_1(t-1) + 0,17484Z_2(t-1) + 0,12686Z_3(t-1) + 0,29698Z_4(t-1) + \varepsilon_1(t)$$

2. Persamaan Model GSTAR-SUR (2,1) jumlah wisatawan di Kab. Jepara

$$\hat{Z}_2(t) = 0,21997Z_1(t-2) + 0,09361Z_3(t-2) + 0,13953Z_4(t-2) + \varepsilon_2(t)$$

3. Persamaan Model GSTAR-SUR (2,1) jumlah wisatawan di Kab. Magelang

$$\hat{Z}_3(t) = 0,43918Z_3(t-1) + \varepsilon_3(t)$$

4. Persamaan Model GSTAR-SUR (2,1) jumlah wisatawan di Kab. Semarang

$$\hat{Z}_4(t) = 0,10534Z_1(t-2) + 0,03926Z_2(t-2) + 0,10534Z_3(t-2) + \varepsilon_4(t)$$

Uji Ljung-Box masing-masing persamaan adalah sebagai berikut :

Hipotesis :

$$H_0 : \rho_1 = \rho_2 = \dots = \rho_k = 0 \text{ (residual tidak white noise)}$$

$$H_1 : \text{Minimalk ada satu } \rho_k \neq 0, \text{ untuk } k = 1, 2, \dots, K$$

Taraf signifikasnsi : $\alpha = 5\%$

Statistik Uji :

$$Q_k = T(T+2) \sum_{k=1}^K \frac{\hat{\rho}_k^2}{T-k}$$

Daerah Kritis :

$$H_0 \text{ ditolak jika } Q > \chi_{0,05;11} \text{ dengan } \chi_{0,05;11} = 19,675$$

Keputusan :

Melalui perhitungan nilai Ljung-Box masing-masing lokasi diperoleh keputusan sebagai berikut :

Tabel 2. Keputusan Uji Ljung-Box

Lokasi	Qhitung	Keputusan
Kota Semarang	35,279	H_0 ditolak
Kab. Jepara	21,682	H_0 ditolak
Kab. Magelang	53,661	H_0 ditolak
Kab. Semarang	23,762	H_0 ditolak

Kesimpulan :

Berdasarkan Tabel , maka dapat disimpulkan bahwa H_0 ditolak pada model GSTAR-SUR(2,1) semua lokasi, sehingga residual Modal GSTAR-SUR(2,1) untuk semua lokasi bersifat white noise.

4. KESIMPULAN

Berdasarkan hasil analisis dan pembahasan mengenai data Jumlah Wisatawan di Kota Semarang, Kabupaten Jepara, Kabupaten Magelang dan Kabupaten Semarang diperoleh beberapa kesimpulan sebagai berikut:

1. Model GSTAR yang terbentuk dari data Jumlah Wisatawan 4 kota/kabupaten di Jawa Tengah adalah GSTAR dengan orde spasial (ρ) adalah 2 dan orde waktu (λ) adalah 1. Sehingga model yang terbentuk adalah model GSTAR (2,1).
2. Model GSTAR-SUR yang terbentuk adalah GSTAR-SUR (2,1) yang dijabarkan sebagai berikut :
 - Persamaan Model GSTAR-SUR (2,1) jumlah wisatawan di Kota Semarang :
 $\hat{Z}_1(t) = 0,2971Z_1(t-1) + 0,17484Z_2(t-1) + 0,12686Z_3(t-1) + 0,29698Z_4(t-1) + \varepsilon_1(t)$
 - Persamaan Model GSTAR-SUR (2,1) jumlah wisatawan di Kab. Jepara :
 $\hat{Z}_2(t) = 0,21997Z_1(t-2) + 0,09361Z_3(t-2) + 0,13953Z_4(t-2) + \varepsilon_2(t)$
 - Persamaan Model GSTAR-SUR (2,1) jumlah wisatawan di Kab. Magelang :
 $\hat{Z}_3(t) = 0,43918Z_3(t-1) + \varepsilon_3(t)$
 - Persamaan Model GSTAR-SUR (2,1) jumlah wisatawan di Kab. Semarang :
 $\hat{Z}_4(t) = 0,10534Z_1(t-2) + 0,03926Z_2(t-2) + 0,10534Z_3(t-2) + \varepsilon_4(t)$

DAFTAR PUSTAKA

- Badan Pusat Statistik Indonesia. 2018. *Statistik Objek Daya Tarik Wisata*. Jakarta: Badan Pusat Statistik.
- Borovkova, S.A., Lapuhaä, H.P., dan Ruchjana, B.N. 2008. *Consistency and Asymptotic Normality of Least Square Estimators in Generalized STAR Models*. Statistica Neerlandica Vol. 62, No. 4: Hal. 482-508.
- Cliff, A. dan Ord, J.K. 1983. *Spatial Processes: Models and Applications*. London: Pion.

- Halim, S., dan Chandra, Arif. 2011. *Pemodelan Time Series Multivariat secara Automatis*. Jurnal Teknik Industri Vol.13, No. 1: Hal. 19-26.
- Makridakis. 1999. *Metode dan Aplikasi Peramalan Jilid 1*. Jakarta: Erlangga.
- Pfeifer, P.E. dan Deutsch, S.J. 1980. *A Three-Stage Iterative Procedure for Space-Time Modeling*. Technometrics. Vol. 22 No.1: pp 35-47.
- Prastuti, M., dan Suhartono. 2014. *Model GSTAR-SUR Musiman untuk Peramalan Jumlah Wisatawan Mancanegara di Empat Lokasi Wisata di Indonesia*. Tesis. Surabaya: Program Magister Jurusan Statistika FMIPA Institut Teknologi Sepuluh Nopember.
- Wei, W. W. 2006. *Time Series Analysis: Univariate and Multivariate Methods (Second ed.)*. United States of America: Pearson Education, Inc.
- Wutsqa, D.U., Suhartono, dan Sutijo, B. 2010. *Generalized Space-Time Autoregressive Modeling*. Proceedings of the 6th IMT-GT Conference on Mathematics, Statistics and its Applications (ICMSA 2010). Kuala Lumpur: Universiti Tuanku Abdul Rahman.