

PEMODELAN FAKTOR-FAKTOR YANG MEMPENGARUHI TINGKAT PENGANGGURAN TERBUKA DI INDONESIA DENGAN PENDEKATAN REGRESI SPASIAL

Miranda Agustina*, Makkulau, Bahridin Abapihi, Gusti Ngurah Adhi Wibawa, Ruslan,
Irma Yahya
Program Studi S1 Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas
Halu Oleo
*Email: agustinamiranda313@gmail.com

ABSTRACT

Unemployment is one of the serious problems in development in Indonesia. The indicator used to measure the high unemployment rate is the Open Unemployment Rate (OUR). OUR in an area is estimated to be influenced by OUR in the surrounding area. This may be due to the proximity factor between regions. A regression model that incorporates the spatial relationship between regions into the model, namely the spatial regression model. The purpose of this study was to find out the appropriate model in modeling the factors that affect OUR in Indonesia in 2020 with a spatial regression approach and the factors that influence it. The results of this study indicate that the appropriate model is the Spatial Autoregressive (SAR) model and the influencing factors are the growth rate of Gross Regional Domestic Product (GRDP) based on constant prices (X_3) and the labor force participation rate (X_4).

Keywords: *Open Unemployment Rate, Spatial Effect, Spatial Regression, Spatial Autoregressive*

ABSTRAK

Pengangguran merupakan salah satu masalah yang cukup serius dalam pembangunan di Indonesia. Indikator yang digunakan untuk mengukur tingginya angka pengangguran adalah Tingkat Pengangguran Terbuka (TPT). TPT di suatu daerah diperkirakan dipengaruhi oleh TPT di daerah sekitarnya. Hal ini mungkin terjadi karena adanya faktor kedekatan antardaerah. Suatu model regresi yang memasukkan hubungan spasial antardaerah ke dalam model, yaitu dengan model regresi spasial. Tujuan dari penelitian ini adalah untuk mengetahui model yang sesuai dalam pemodelan faktor-faktor yang mempengaruhi TPT di Indonesia 2020 dengan pendekatan regresi spasial dan faktor-faktor yang mempengaruhinya. Hasil penelitian ini menunjukkan bahwa model yang sesuai adalah model Spatial Autoregressive (SAR) dan faktor-faktor yang mempengaruhinya adalah laju pertumbuhan Produk Domestik Regional Bruto (PDRB) atas dasar harga konstan (X_3) dan tingkat partisipasi angkatan kerja (X_4).

Kata Kunci: *Tingkat Pengangguran Terbuka, efek spasial, regresi spasial, spasial Autoregressive*

PENDAHULUAN

Pengangguran merupakan salah satu masalah yang cukup serius dalam pembangunan di Indonesia. Indikator yang digunakan untuk mengukur tingginya angka pengangguran adalah Tingkat Pengangguran Terbuka atau TPT. Berdasarkan publikasi Badan Pusat Statistik (BPS) Indonesia, menunjukkan bahwa TPT di Indonesia pada bulan Agustus 2020 sebesar 7,07 persen. Angka ini cukup tinggi jika dibandingkan dengan tahun sebelumnya, dimana TPT pada tahun 2019 sebesar 5,23 persen (BPS, 2021). Adanya peningkatan ini menyebabkan angka TPT tahun 2020 melebihi angka yang ditargetkan oleh pemerintah. Adapun angka yang ditargetkan diperkirakan 4,8 sampai 5,1 persen (Kemenkeu, 2019). TPT di suatu daerah diperkirakan dipengaruhi oleh TPT di daerah sekitarnya. Hal ini mungkin terjadi karena adanya faktor kedekatan atau ketetanggaan antar daerah (Wuryandari *et.al*, 2014). Salah satu teori yang menjelaskan adanya keterkaitan antar daerah adalah hukum pertama geografi yang dikemukakan oleh Tobler yang menyatakan bahwa, "Segala sesuatu saling berhubungan satu dengan yang lainnya, tetapi sesuatu yang lebih dekat akan lebih berpengaruh

daripada sesuatu yang jauh” (Djuraidah & Wigena, 2012). Adanya hubungan spasial dalam variabel respon akan menyebabkan pendugaan menjadi tidak tepat karena asumsi keacakan galat dilanggar, sehingga untuk mengatasinya diperlukan suatu model regresi yang memasukkan hubungan spasial antar lokasi ke dalam model, yaitu dengan model regresi spasial (Sanusi *et.al*, 2018).

Model regresi spasial secara resmi ditetapkan dua aspek dari efek spasial, yaitu Efek dependensi spasial yang ditunjukkan dengan adanya kemiripan karakteristik antarlokasi yang saling bertetangga dan efek heterogenitas spasial yang menunjukkan adanya keragaman antarlokasi yang bertetangga (LeSage & Pace, 2009). Jika dalam data memiliki efek dependensi spasial, maka penyelesaian yang dilakukan adalah dengan pendekatan area. Pada pemodelan dependensi spasial terdapat beberapa model yang terbentuk, yaitu Model *Spatial Autoregressive* (SAR) yang memiliki dependensi pada nilai respon antar lokasi, *Spatial Error Models* (SEM) memiliki dependensi pada nilai *error* antar lokasi, dan *Spatial Autoregressive Moving Average* (SARMA) memiliki dependensi pada nilai respon dan nilai *error* antarlokasi. Penelitian TPT dengan pendekatan regresi spasial sebelumnya sudah pernah dilakukan oleh beberapa peneliti, salah satunya diantaranya adalah Wardani *et.al* (2019) tentang pemodelan tingkat pengangguran terbuka provinsi Jawa Tengah menggunakan model *spatial autoregressive moving average*.

Penelitian ini bertujuan untuk mengetahui model yang sesuai dalam pemodelan faktor-faktor yang mempengaruhi TPT di Indonesia tahun 2020 dengan pendekatan regresi spasial dan mengetahui faktor-faktor yang mempengaruhi TPT di Indonesia.

METODE PENELITIAN

Data

Data yang digunakan dalam penelitian ini merupakan jenis data sekunder yang diperoleh Badan Pusat Statistik (BPS) Republik Indonesia melalui (<https://www.bps.go.id>) berupa data dan publikasi tahun 2020. Unit observasi dalam penelitian ini adalah seluruh Provinsi di Indonesia yaitu sebanyak 34 Provinsi. Data terdiri atas 1 variabel respon (Y) dan 4 variabel prediktor (X). Defenisi operasional dari variable respon dan variable prediktor yang digunakan dalam penelitian ini dapat ditampilkan pada table berikut:

Tabel 1 Defenisi operasional variabel penelitian

	Variabel	Defenisi Operasional	Satuan	Referensi
Y	Tingkat Pengangguran Terbuka	Persentase jumlah pengangguran terhadap jumlah angkatan kerja	Persen	
X ₁	Kepadatan Penduduk	Banyaknya penduduk persatuan luas	Jiwa/km ²	Wardani, 2019
X ₂	Persentase penduduk tamat SD Saja	Persentase penduduk 15 tahun keatas berdasarkan pendidikan yang ditamatkan adalah Sekolah Dasar (SD)	Persen	Mariana, 2013
X ₃	Laju pertumbuhan PDRB atas dasar Harga Konstan	Pertumbuhan produksi barang dan jasa di suatu wilayah perekonomian dalam selang waktu tertentu	Persen	Wardiansyah, 2016
X ₄	Tingkat Partisipasi Angkatan Kerja	Persentase penduduk usia 15 tahun keatas yang merupakan angkatan kerja	Persen	Wardani, 2019

Adapun langkah-langkah yang akan dilakukan dalam penelitian ini adalah sebagai berikut:

1. Melakukan Eksplorasi Data terhadap Tingkat Pengangguran Terbuka di Indonesia dan faktor-faktor yang mempengaruhinya.
2. Melakukan penetapan matriks pembobot spasial (W).
3. Melakukan pengujian efek spasial yang terdiri dari uji heterogenitas spasial dan uji dependensi spasial.
4. Melakukan identifikasi awal model spasial yang akan digunakan dengan melihat uji *Lagrange Multiplier* (LM), berdasarkan informasi dari uji LM dilakukan pemodelan spasial berdasarkan parameter yang signifikan.
5. Melakukan pemodelan *Spatial Autoregressive* (SAR).
6. Membuat interpretasi model dan menarik kesimpulan.

Metode

1) Matriks Pembobot Spasial

Dalam pemodelan menggunakan regresi spasial, hal yang pertama dilakukan adalah menetapkan matriks pembobot spasial. Matriks pembobot spasial adalah matriks yang menggambarkan kedekatan suatu lokasi dengan lokasi lainnya. Menurut Kosfeld (2006), informasi lokasi dapat diketahui dari dua sumber yaitu: Hubungan ketetanggaan (*neighborhood*) dan jarak (*distance*).

Pada prinsipnya bobot jarak antara suatu lokasi dengan lokasi lain ditentukan dengan jarak kedua daerah itu. Semakin dekat jarak kedua lokasi tersebut maka bobot yang diberikan semakin besar. Hal ini sesuai dengan hukum geografi pertama Tobler. Jarak di antara lokasi i dan lokasi j umumnya didefinisikan sebagai jarak *Euclidian*. Jarak *Euclidian* dapat diperoleh dengan rumus sebagai berikut :

$$d_{ij} = \sqrt{(u_i - u_j)^2 + (v_i - v_j)^2} \quad (1)$$

dimana, u_i adalah koordinat garis lintang atau *latitude* pada lokasi ke- i ; u_j adalah koordinat garis lintang atau *latitude* pada lokasi ke- j ; v_i adalah koordinat garis bujur atau *longitude* pada lokasi ke- i ; dan v_j adalah koordinat garis bujur atau *longitude* pada lokasi ke- j .

Penelitian ini menggunakan matriks pembobot k - nn sebagai matriks pembobot spasial. Pada cara ini peneliti menentukan sebanyak k lokasi j di sekitar lokasi i yang terdekat dengan lokasi tersebut. Penentuan jarak yang sering digunakan untuk menghitung k tetangga terdekat adalah jarak *Euclidian* atau d_{ij} . Jika nilai $k \in \{1, 2, \dots, n-1\}$, maka k jarak terdekat antara lokasi- i ke semua lokasi- j adalah $N_k \in \{d_{ij[1]}, d_{ij[2]}, \dots, d_{ij[k]}\}$, dengan k himpunan lokasi- j terdekat terhadap lokasi- i adalah $J_k(i) = \{j_{[1]}, j_{[2]}, \dots, j_{[k]}\}$, maka nilai bobot (w_{ij}) didefinisikan sebagai berikut:

$$w_{ij}(k) = \begin{cases} 1, & j \in J_k(i) \\ 0, & \text{lainnya} \end{cases}$$

Perhitungan k tetangga terdekat dilakukan dengan tahapan sebagai berikut: Pertama menghitung jarak *Euclidian* lokasi i ke j ; kedua mengurutkan jarak yang diperoleh, dan ketiga memilih k lokasi dengan jarak terdekat sebagai nilai optimum. Penentuan nilai k pertama kali di dasarkan pada statistik *moran's I* atau indeks Moran, prosesnya dilakukan secara iterasi. Nilai k terpilih berdasarkan nilai *moran's I* terbesar (Jaya *et.al*, 2017).

Menurut Elhorst (2014), matriks pembobot pada umumnya akan dinormalisasikan sehingga tiap barisnya akan berjumlah 1, sehingga menjamin bahwa

Bobot tiap daerah observasi akan berkisar antar 0-1. Normalisasi pembobot berdasarkan baris disebut normalisasi baris (*row standardization*). Nilai dari matriks pembobot spasial terstandarisasi baris ke- i dan kolom ke- j adalah sebagai berikut:

$$W_{ij} = \frac{W_{ij}^*}{\sum_{j=1}^n W_{ij}^*}$$

2) Efek Spasial

2.1 Efek Heterogenitas Spasial

Efek heterogenitas adalah efek yang menunjukkan adanya keragaman antar lokasi. Jadi setiap lokasi mempunyai struktur dan parameter hubungan yang berbeda. Pengujian efek heterogenitas spasial dilakukan dengan uji Breusch-Pagan (BP). Hipotesis yang digunakan pada uji ini yaitu:

$$H_0: \sigma_1^2 = \sigma_2^2 = \dots = \sigma_n^2 \text{ (Homoskedastisitas)}$$

$$H_1: \text{paling tidak ada satu } \sigma_i^2 \neq \sigma_j^2 \text{ dimana } i \neq j \text{ (Heteroskedastisitas)}$$

Statistik uji yang digunakan pada uji Breusch-Pagan sebagai berikut:

$$BP = \frac{1}{2} \mathbf{f}^T \mathbf{Z} (\mathbf{Z}^T \mathbf{Z})^{-1} \mathbf{Z}^T \mathbf{f}$$

dimana:

$$f_i = \left(\frac{e_i^2}{\sigma^2} - 1 \right)$$

dengan e_i^2 adalah *error* untuk observasi ke- i dari hasil estimasi regresi dengan menggunakan ols dan \mathbf{Z} merupakan matriks variabel prediktor berukuran $n \times (k + 1)$. Pengambilan keputusan pada uji ini adalah tolak H_0 jika nilai $BP > \chi^2_{(\alpha, k)}$.

2.2 Efek Dependensi Spasial

Dependensi spasial atau ketergantungan spasial terjadi akibat adanya ketergantungan dalam data wilayah. Adanya dependensi spasial mengindikasikan bahwa nilai atribut pada daerah tertentu terkait oleh nilai atribut daerah lain yang letaknya berdekatan atau bertetangga. Uji yang digunakan untuk mengetahui *spatial dependence* adalah dengan menggunakan statistik *Moran's I*. Hipotesis yang digunakan dalam uji ini adalah sebagai berikut:

$$H_0: I = 0 \text{ (tidak terdapat dependensi spasial)}$$

$$H_1: I \neq 0 \text{ (terdapat dependensi spasial)}$$

Statistik uji yang digunakan yaitu:

$$Z_{hitung} = \frac{I - E(I)}{\sqrt{\text{var}(I)}}$$

dengan:

$$E(I) = I_0 = -\frac{1}{n-1}$$

$$\text{var}(I) = \frac{n^2 S_1 - n S_2 + 3 S_0^2}{(n^2 - 1) S_0^2} - [E(I)]^2$$

dimana indeks *Moran (Morans' I)* dapat dirumuskan sebagai berikut:

$$I = \frac{n \sum_{i=1}^n \sum_{j=1}^n \mathbf{W}_{ij} (y_i - \bar{y})(y_j - \bar{y})}{S_0 \sum_{i=1}^n (y_i - \bar{y})^2}$$

dengan:

$$S_0 = \sum_{i=1}^n \sum_{j=1}^n \mathbf{W}_{ij} ; S_1 = \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n (\mathbf{W}_{ij} + \mathbf{W}_{ji})^2 ;$$

$$S_2 = \sum_{i=1}^n \left(\sum_{j=1}^n \mathbf{W}_{ij} + \sum_{j=1}^n \mathbf{W}_{ji} \right)^2$$

keterangan:

- y_i : data amatan ke- i
- y_j : data amatan ke- j
- \bar{y} : rata-rata data amatan
- \mathbf{W}_{ij} : matriks pembobot spasial terstandarisasi
- n : jumlah wilayah amatan
- I : nilai *Moran's I*
- $E(I)$: nilai rata-rata dari I
- $\text{var}(I)$: nilai variansi dari I
- i, j : 1, 2, ..., n

Dasar pengambilan keputusan pada uji ini adalah tolak H_0 jika $Z_{hitung} > Z_{\alpha/2}$ yang berarti terdapat autokorelasi spasial (Pusdiktasari, 2019).

Moran's I scatterplot merupakan sebuah diagram untuk melihat hubungan antara nilai amatan pada suatu lokasi (distandarisasi) dengan rata-rata nilai amatan dari lokasi-lokasi yang bertetangga dengan lokasi yang bersangkutan. *Moran's scatterplot* dibagi atas empat kuadran yang sesuai untuk empat pola penyebaran nilai dari setiap wilayah yang bertetangga. Pembagian kuadrannya adalah sebagai berikut:

Kuadran II <i>Low-High</i>	Kuadran I <i>High-High</i>
Kuadran III <i>Low-Low</i>	Kuadran IV <i>High-Low</i>

Gambar 1 *Moran's Scatterplot*

Adapun penjelasan dari masing-masing kuadran adalah sebagai berikut: Kuadran I (terletak di kanan atas) disebut *High-High* (HH) yang menunjukkan wilayah yang memiliki pengamatan tinggi yang dikelilingi oleh wilayah yang juga memiliki pengamatan tinggi untuk variabel yang menjadi perhatian. Kuadran II (terletak di kiri atas) disebut *Low-High* (LH) yang menunjukkan wilayah dengan nilai rendah namun dikelilingi wilayah dengan nilai tinggi untuk variabel yang menjadi perhatian. Kuadran III (terletak di kiri bawah) disebut *Low-Low* (LL) yang menunjukkan wilayah dengan nilai pengamatan rendah dan dikelilingi oleh wilayah yang juga mempunyai nilai pengamatan rendah untuk variabel yang menjadi perhatian. Kuadran IV (terletak di kanan bawah) disebut *High-Low* (HL) yang menunjukkan wilayah dengan nilai tinggi

yang dikelilingi wilayah dengan nilai rendah untuk variabel yang menjadi perhatian.

3) Uji *Lagrange Multiplier* (LM)

Uji *Lagrange Multiplier* (LM) digunakan untuk memilih model regresi spasial yang sesuai (LeSage, 1998). Uji LM terdiri dari LM_{lag} dan LM_{error} . Apabila LM_{lag} signifikan maka model yang sesuai adalah SAR, jika LM_{error} signifikan maka model yang sesuai adalah SEM dan jika keduanya signifikan maka model yang sesuai adalah SARMA (*Spatial Autoregressive Moving Average*). Apabila LM_{lag} dan LM_{error} keduanya tidak signifikan, maka model yang sesuai adalah *Ordinary Least Square* (OLS). Uji LM digunakan untuk menentukan kehadiran efek dependensi spasial atau tidak didalam model. Bentuk tes LM, yaitu:

a. *Lagrange Multiplier Lag*

Hipotesis yang digunakan adalah:

$H_0: \rho = 0$ (tidak ada dependensi spasial *lag*)

$H_1: \rho \neq 0$ (ada dependensi spasial *lag*)

Statistik Uji:

$$LM_{lag} = \frac{(\boldsymbol{\varepsilon}^T \mathbf{W}\mathbf{y})^2}{s^2((\mathbf{W}\mathbf{X}\boldsymbol{\beta})^T \mathbf{M}(\mathbf{W}\mathbf{X}\boldsymbol{\beta}) + \mathbf{T}s^2)}$$

dengan,

$$\mathbf{M} = \mathbf{I} - \mathbf{X}(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T, \quad \mathbf{T} = \text{tr}[(\mathbf{W}^T + \mathbf{W})\mathbf{W}], \quad s^2 = \frac{\boldsymbol{\varepsilon}^T \boldsymbol{\varepsilon}}{n}$$

Pengambilan keputusan, tolak H_0 jika $LM_{lag} > \chi^2_{(1,\alpha)}$ atau $p\text{-value} < \alpha$.

b. *Lagrange Multiplier Error*

Hipotesis yang digunakan adalah:

$H_0: \lambda = 0$ (tidak ada dependensi spasial *error*)

$H_1: \lambda \neq 0$ (ada dependensi spasial *error*)

Statistik Uji:

$$LM_{error} = \frac{(\boldsymbol{\varepsilon}^T \mathbf{W}\boldsymbol{\varepsilon} / s^2)^2}{\mathbf{T}}$$

Pengambilan keputusan, tolak H_0 jika $LM_{error} > \chi^2_{(1,\alpha)}$ atau $p\text{-value} < \alpha$.

4) Model *Spatial Autoregressive* (SAR)

Model SAR merupakan model yang mengkombinasikan model regresi sederhana dengan lag spasial pada variabel respon menggunakan data *cross section* (LeSage, 1999). Model ini terbentuk apabila $\rho \neq 0$ dan $\lambda = 0$. Kelebihan dari model SAR adalah model ini tepat digunakan untuk pola spasial dengan pendekatan area. Model umum dari *Spatial Autoregressive Model* atau disebut juga *Spatial lag Model* (SLM) adalah sebagai berikut:

$$\mathbf{y} = \rho \mathbf{W}\mathbf{y} + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \quad \boldsymbol{\varepsilon} \sim N(0, \sigma^2 \mathbf{I})$$

Model persamaan diatas mengasumsikan bahwa proses autoregressive hanya pada variabel respon. Pada persamaan tersebut, variabel respon y dimodelkan sebagai kombinasi linier dari daerah sekitarnya atau daerah yang berimpitan dengan y .

Estimasi parameter model SAR dilakukan dengan menggunakan metode

maximum likelihood atau kemungkinan maksimum. Metode kemungkinan maksimum adalah suatu metode yang digunakan untuk mencari koefisien regresi sehingga probabilitas kejadian dari variabel respon semaksimal mungkin. Metode ini merupakan suatu metode alternatif untuk mengestimasi parameter jika bentuk distribusi *error* diketahui. Estimasi parameter model diperoleh dengan memaksimalkan fungsi kemungkinan yang ekuivalen dengan memaksimalkan logaritma dari fungsi kemungkinan *likelihood* (L) pada persamaan sebagai berikut:

$$\ln(L) = -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln \sigma^2 + \ln |\mathbf{I} - \rho \mathbf{W}| - \frac{1}{2\sigma^2} (\mathbf{y} - \rho \mathbf{W}\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^T (\mathbf{y} - \rho \mathbf{W}\mathbf{y} - \mathbf{X}\boldsymbol{\beta})$$

Estimasi parameter σ^2 , β , dan ρ diperoleh dengan memaksimalkan fungsi log kemungkinan dari persamaan di atas. Parameter pertama yang perlu dilakukan adalah parameter β . Didapatkan estimasi parameter β adalah:

$$\begin{aligned} \hat{\boldsymbol{\beta}} &= (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{A}\mathbf{y} \\ &= (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T (\mathbf{I} - \rho \mathbf{W})\mathbf{y} \end{aligned}$$

Penduga untuk σ^2 diperoleh dengan menurunkan fungsi log *likelihood* di atas terhadap σ^2 sebagai berikut:

$$\hat{\sigma}^2 = \frac{[\mathbf{A}\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}]^T [\mathbf{A}\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}]}{n}$$

Selanjutnya, dilakukan estimasi parameter ρ .

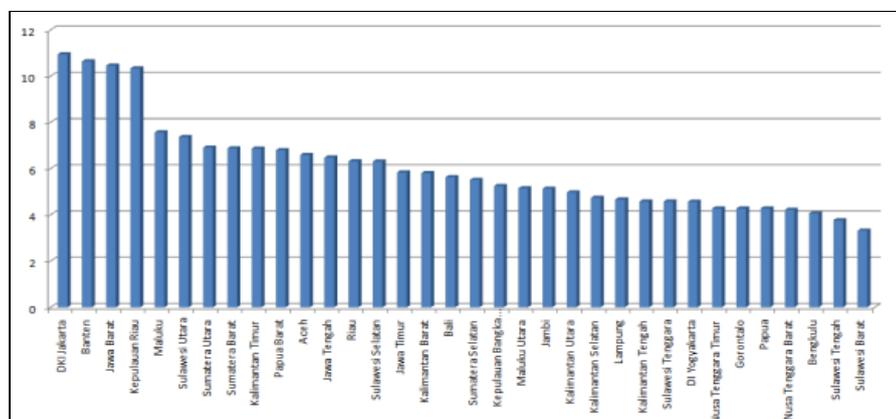
$$f(\rho) = c - \frac{n}{2} \ln(\hat{\mathbf{e}}_0 - \hat{\rho} \hat{\mathbf{e}}_L)^T (\hat{\mathbf{e}}_0 - \hat{\rho} \hat{\mathbf{e}}_L) + \ln |\mathbf{A}|$$

dengan, $c = -\frac{n}{2} \ln(2\pi) + \frac{n}{2} \ln(n)$

HASIL DAN PEMBAHASAN

Eksplorasi Data

Penelitian ini menggunakan objek data TPT menurut provinsi di Indonesia pada tahun 2020. Secara administratif provinsi di Indonesia terdiri dari 34 provinsi, tentunya masing-masing provinsi memiliki karakteristik TPT yang berbeda-beda, sehingga perlu diketahui karakteristik TPT masing-masing provinsi di Indonesia dengan menghitung statistika deskriptif dan secara visual. Berikut adalah TPT menurut Provinsi di Indonesia pada tahun 2020.



Gambar 2 Tingkat Pengangguran Terbuka menurut Provinsi di Indonesia Tahun 2020

Berdasarkan Gambar 2 dapat dilihat bahwa TPT tertinggi terdapat di Provinsi DKI Jakarta sedangkan terendah terdapat di Provinsi Sulawesi Barat.

Selanjutnya, peta sebaran TPT menurut provinsi di Indonesia tahun 2020 dapat dilihat pada gambar 3:



Gambar 3 Peta Sebaran Tingkat Pengangguran Terbuka (TPT) menurut Provinsi di Indonesia Tahun 2020

Berdasarkan Gambar 3 dapat dilihat bahwa sebaran TPT di Indonesia terdiri atas tiga kelompok warna dimana wilayah dengan TPT tinggi ditunjukkan oleh warna yang gelap, wilayah dengan TPT sedang ditunjukkan oleh warna yang sedang, dan wilayah dengan TPT rendah ditunjukkan oleh warna yang lebih terang.

Hasil penelitian analisis statistik deskriptif dapat dilihat pada tabel 2:

Tabel 2 Statistik Deskriptif Variabel Respon dan Variabel Prediktor

Variabel	Rata-rata	Minimum	Maksimum
Y	6,03	3,32	10,95
X ₁	739,24	9	15907
X ₂	22,84	13,10	29,47
X ₃	-1,25	-9,31	4,92
X ₄	68,04	63,40	74,32

Berdasarkan Tabel 2 dapat dilihat bahwa TPT (Y) tertinggi terdapat di Provinsi DKI Jakarta yaitu sebesar 10,95% sedangkan terendah terdapat di Provinsi Sulawesi Barat yaitu sebesar 3,32%. Rata-rata tingkat pengangguran terbuka di Indonesia pada tahun 2020 sebesar 6,03%. Kepadatan penduduk (X₁) tertinggi terdapat di Provinsi DKI Jakarta yaitu sebesar 15907 Jiwa/km² sedangkan terendah terdapat di Provinsi Kalimantan Utara yaitu sebesar 9 Jiwa/km². Rata-rata kepadatan penduduk di Indonesia pada tahun 2020 sebesar 739,24 Jiwa/km². Persentase penduduk tamat SD Saja (X₂) tertinggi terdapat di Provinsi Sulawesi Tengah yaitu sebesar 29,47% sedangkan terendah terdapat di Provinsi DKI Jakarta yaitu sebesar 13,10%. Rata-rata persentase penduduk tamat SD Saja di Indonesia pada tahun 2020 sebesar 22,84%. Laju pertumbuhan PDRB atas dasar harga konstan (X₃) tertinggi terdapat di Provinsi Maluku Utara yaitu sebesar 4,92% sedangkan terendah terdapat di Provinsi Bali yaitu sebesar -9,31%. Rata-rata Laju pertumbuhan PDRB atas dasar harga konstan di Indonesia pada tahun 2020 sebesar -1,25%. Persentase Tingkat Partisipasi Angkatan Kerja (X₄) tertinggi terdapat di Provinsi Bali yaitu sebesar 74,32% sedangkan terendah terdapat di Provinsi Sulawesi Selatan yaitu sebesar 63,40%. Rata-rata tingkat partisipasi angkatan kerja di Indonesia pada tahun 2020 sebesar 68,04%.

Matriks Pembobot Spasial

Matriks pembobot spasial yang digunakan pada penelitian ini adalah menggunakan pembobot spasial k lokasi atau k tetangga terdekat (*k nearest neighbour*). Penentuan nilai *k* yang optimum didasarkan pada statistik *moran's I* atau indeks moran yang paling besar. Perhitungan *k* tetangga terdekat dilakukan dengan

tahapan sebagai berikut: Pertama menghitung jarak *Euclidian* lokasi i ke j pada persamaan 1; kedua mengurutkan jarak yang diperoleh, dan ketiga memilih k lokasi dengan jarak terdekat sebagai nilai optimum. Penentuan nilai k pertama kali di dasarkan pada statistik *moran's I* atau indeks Moran, prosesnya dilakukan secara iterasi. Percobaan penentuan nilai k pada penelitian ini hanya sampai $k = 1, 2, \dots, 7$. Hal ini dikarenakan semakin besar nilai k , maka nilai indeks moran yang didapatkan semakin kecil. Pembentukan nilai k dilakukan dengan menggunakan software R sebagai berikut:

```
library(rgdal)
library(spdep)
#Peta Indonesia
PtIndo=readOGR("F:/Peta Indo/prov.shp")
plot(PtIndo)
coor=coordinates(PtIndo)
nomor=PtIndo$KODE
text(coor,label=nomor,cex=0.6)
#penentuan k tetangga terdekat (knn)
#k=2
k2=knearneigh(coor,k=2,longlat=TRUE)
K_2=knn2nb(k2)
Wkk=nb2listw(K_2)
matbobot2=listw2mat(Wkk)
```

Berdasarkan hasil percobaan penentuan nilai k didapatkan nilai *moran's I* yang paling besar adalah 0,5507 dengan $k = 2$. Ini artinya bahwa berdasarkan identifikasi *Moran's I* diperoleh k yang paling optimum adalah 2. Nilai k yang optimum akan digunakan sebagai jumlah tetangga disetiap lokasi- i , sehingga masing-masing wilayah memiliki tetangga sebanyak 2 tetangga terdekat dengan jarak lokasi terpendek dari lokasi- i .

Efek Spasial

1. Efek Heterogenitas Spasial

Adapun uji yang digunakan pada efek ini yakni menggunakan uji Breusch-Pagan (uji BP). Berikut hasil uji Breusch-Pagan:

Tabel 3 Hasil Uji Heterogenitas Spasial

Breusch-Pagan	DF	<i>p-value</i>	Keterangan
7,5847	4	0,1080	Terima H_0

Berdasarkan Tabel 3 dapat dilihat bahwa *p-value* pada uji BP untuk model ini adalah sebesar 0,1080 dimana nilai ini lebih besar dari 0,05, sehingga keputusan yang diambil adalah menerima H_0 yang berarti tidak terdapat heterogenitas antar wilayah.

2. Efek Dependensi Spasial

Uji untuk mengetahui *spatial dependence* di dalam suatu model adalah statistic *Moran's I*. Hasil yang diperoleh dari uji Indeks Moran dapat dilihat pada tabel berikut:

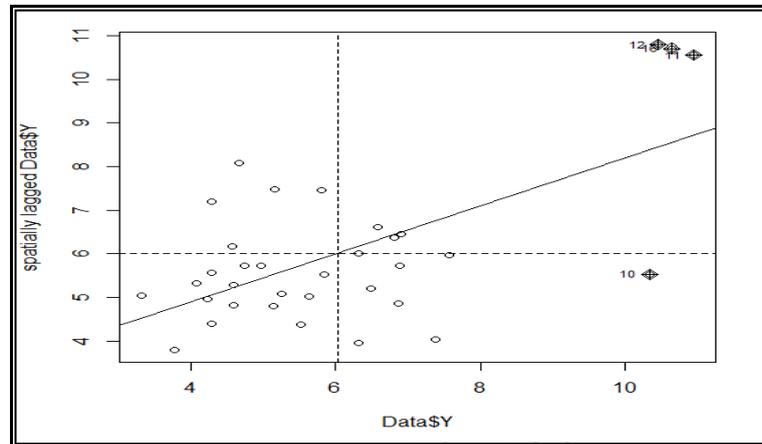
Tabel 4 Hasil Uji Indeks Moran

Morans'I	I_0	Z(I)	<i>-value</i>	Ket
0,5507	0,03	3,83	0,0000	Tolak H_0

Berdasarkan Tabel 4 dapat dilihat bahwa hasil uji Indeks Moran dengan pembobot *k-nn* dengan $k=2$, didapatkan *p-value* sebesar 0,0000 yang lebih kecil dari 0,05 atau nilai Z(I) sebesar 3,8369 yang lebih besar dari $Z_{\alpha/2} = 1,96$. Hal ini menunjukkan bahwa

pada taraf signifikansi 5% H_0 ditolak. Ini artinya bahwa terdapat dependensi spasial antarwilayah provinsi di Indonesia. Nilai yang dihasilkan dalam I atau nilai indeks moran adalah lebih besar dari I_0 , sehingga karakteristik dari pola spasial yang dihasilkan adalah cenderung bergerombol (*clustered*).

Selanjutnya ketergantungan spasial antar wilayah secara visual dapat dilihat menggunakan *moran's scatterplot*. Adapun *moran's scatterplot* dengan menggunakan pembobot *k-nn* ($k = 2$) diberikan pada gambar 4:



Gambar 4 Moran's Scatterplot dengan Pembobot *k-nn* ($k = 2$)

Berdasarkan Gambar 4 dapat dilihat bahwa masing-masing kuadran dapat dijelaskan sebagai berikut:

1. Kuadran I: *High-High* (HH) menunjukkan bahwa wilayah yang mempunyai TPT tinggi dikelilingi oleh wilayah yang mempunyai TPT tinggi pula. Wilayah-wilayah yang berada pada kuadran I, yaitu: Aceh, Sumatera Utara, DKI Jakarta, Jawa Barat, Banten, dan Papua Barat.
2. Kuadran II: *Low-High* (LH) menunjukkan bahwa wilayah yang mempunyai TPT rendah dikelilingi oleh wilayah yang mempunyai TPT tinggi. Wilayah-wilayah yang berada pada kuadran II, yaitu: Lampung, DI Yogyakarta, Kalimantan Barat, Maluku Utara, dan Papua.
3. Kuadran III: *Low-Low* (LL) menunjukkan bahwa wilayah yang mempunyai TPT rendah dikelilingi oleh wilayah yang mempunyai TPT rendah pula. Wilayah-wilayah yang berada pada kuadran III yaitu: Jambi, Sumatera Selatan, Bengkulu, Kepulauan Bangka Belitung, Jawa Timur, Bali, Nusa Tenggara Barat, Nusa Tenggara Timur, Kalimantan Tengah, Kalimantan Selatan, Kalimantan Utara, Sulawesi Tengah, Sulawesi Tenggara, Gorontalo, dan Sulawesi Barat.
4. Kuadran IV: *High-Low* (HL) menunjukkan bahwa wilayah yang mempunyai TPT tinggi dikelilingi oleh wilayah yang mempunyai TPT rendah. Wilayah-wilayah yang berada pada kuadran IV, yaitu: Sumatera Barat, Riau, Kepulauan Riau, Jawa Tengah, Kalimantan Timur, Sulawesi Utara, Sulawesi Selatan, dan Maluku.

Uji *Lagrange Multiplier* (LM)

Hasil pengujian Indeks Moran model TPT di Indonesia menunjukkan bahwa telah terjadi dependensi spasial sehingga perlu dilanjutkan ke uji *Lagrange Multiplier* (LM). Uji ini digunakan untuk melihat model regresi spasial sesuai. Uji LM mendeteksi adanya dependensi spasial secara lebih spesifik yaitu dependensi spasial dalam hal lag, *error*, atau lag dan *error*. Berikut hasil uji *Lagrange Multiplier*:

Tabel 5 Hasil Uji LM

Uji LM	Nilai	<i>p-value</i>	Ket
LM(lag)	8,29	0,0039	Tolak H ₀
LM (error)	0,53	0,4662	Terima H ₀

Berdasarkan Tabel 5 dapat dilihat bahwa *p-value* dari uji LM_{lag} lebih kecil dari 0,05 maka pengambilan keputusannya adalah menolak H₀. Hal ini berarti bahwa terdapat dependensi spasial lag pada tingkat pengangguran terbuka, sehingga perlu dilanjutkan ke pembuatan model SAR.

Model Spatial Autoregressive (SAR)

Berdasarkan hasil uji LM di atas dapat dilihat bahwa model spasial lag model SAR memiliki nilai *p-value* lebih kecil dari nilai α (0,05). Artinya bahwa model yang sesuai untuk analisis dalam penelitian ini adalah model SAR, sehingga dalam penelitian ini menggunakan model SAR.

1. Estimasi Parameter Model SAR 4 Variabel

Hasil estimasi parameter model SAR berdasarkan data TPT di tiap provinsi di Indonesia tahun 2020 disajikan sebagai berikut:

$$\hat{Y}_i = 29,909 + 0,3438 \sum_{j=1, i \neq j}^{34} W_{ij} Y_j + 0,00007 X_1 - 0,0634 X_2 - 0,3093 X_3 - 0,3665 X_4$$

2. Signifikansi Parameter Model SAR 4 Variabel

2.1 Uji Simultan Model

Pengujian simultan model dilakukan untuk melihat apakah model SAR sesuai dalam pemodelan TPT dan melihat apakah variabel prediktor secara bersama-sama dapat mempengaruhi variabel respon. Berikut Hasil Uji simultan Model dapat dilihat sebagai berikut:

Tabel 6 Hasil Uji Simultan Model

Model	F _{hitung}
SAR	21,0050

Berdasarkan Tabel 6 didapatkan nilai F_{hitung} = 21,005 > F_{0,05(4,29)} = 2,70 yang berarti bahwa H₀ ditolak. Maka dari itu dapat disimpulkan bahwa model SAR telah sesuai dan variabel prediktor secara bersama-sama dapat mempengaruhi TPT di Indonesia.

2.2 Uji Parsial Model

Untuk menentukan variabel mana yang memberikan pengaruh pada model SAR ini dapat diuji secara formal dengan menggunakan uji signifikansi parsial:

Tabel 7 Hasil Estimasi dan Pengujian Parameter Model SAR 4 Variabel

Variabel	Koefisien	<i>z-value</i>	<i>p-value</i>
ρ	0.3438	3,1609	0,0033
Constant	29,909	6,3832	0,0000
X ₁	0,00007	0,9337	0,3505
X ₂	-0,0634	-1,4331	0,1518
X ₃	-0,3093	-4,0015	0,0000
X ₄	-0,3665	-5,5874	0,0000
AIC	112,42		

Berdasarkan Tabel 7 diperoleh bahwa variabel X_1 (kepadatan penduduk) dan X_2 (persentase penduduk tamat SD Saja) memiliki nilai p -value lebih besar dari nilai α (0,05). Artinya kepadatan penduduk dan persentase penduduk tamat SD Saja tidak memiliki pengaruh terhadap tingkat pengangguran terbuka di Indonesia tahun 2020, sedangkan variabel X_3 (laju pertumbuhan PDRB atas dasar harga kostan) dan variabel X_4 (tingkat partisipasi angkatan kerja) memiliki nilai p -value lebih kecil dari nilai α (0,05). Artinya laju pertumbuhan PDRB atas dasar harga kostan dan tingkat partisipasi angkatan kerja memiliki pengaruh terhadap TPT di Indonesia tahun 2020.

Tahapan selanjutnya yang dilakukan adalah meregresikan kembali variabel prediktor yang berpengaruh signifikan terhadap TPT menurut provinsi di Indonesia.

3. Estimasi Parameter Model SAR 2 Variabel

Hasil estimasi parameter model SAR berdasarkan data TPT di tiap provinsi di Indonesia tahun 2020 disajikan sebagai berikut:

$$\hat{Y}_i = 29,8265 + 0,3877 \sum_{j=1, i \neq j}^{34} W_{ij} Y_j - 0,3430 X_3 - 0,3903 X_4 \quad (1)$$

3.1 Uji Simultan Model

Hasil uji simultan model SAR 2 variabel disajikan sebagai berikut:

Tabel 8 Hasil Uji Simultan Model

Model	F _{hitung}
SAR	38,98

Berdasarkan Tabel 8 didapatkan nilai $F_{hitung} = 38,98 > F_{0,05(2,31)} = 3,30$ yang berarti bahwa H_0 ditolak. Maka dari itu dapat disimpulkan bahwa model SAR telah sesuai dan variabel prediktor secara bersama-sama dapat mempengaruhi TPT di Indonesia.

3.1 Uji Parsial Model

Untuk menentukan variabel mana yang memberikan pengaruh pada model SAR ini dapat diuji secara formal dengan menggunakan uji signifikansi parsial. sebagai berikut:

Tabel 9 Hasil Estimasi dan Pengujian Parameter Model SAR 2 Variabel

Variabel	Koefisien	z-value	p-value
ρ	0.3877	3,572	0,0006
Constant	29,8265	6,192	0,0000
X_3	-0,3430	-4,377	0,0000
X_4	-0,3903	-5,790	0,0000
AIC	112,43		

Berdasarkan Tabel 9 diperoleh bahwa variabel X_3 (laju pertumbuhan PDRB atas dasar harga kostan) dan variabel X_4 (tingkat partisipasi angkatan kerja) memiliki nilai probabilitas lebih kecil dari nilai α (0,05). Artinya laju pertumbuhan PDRB atas dasar harga kostan dan tingkat partisipasi angkatan kerja memiliki pengaruh terhadap tingkat pengangguran terbuka di Indonesia tahun 2020 dan koefisien ρ yang signifikan mengindikasikan bahwa H_0 pada model SAR ditolak, artinya

ketergantungan lag pada spasial berpengaruh terhadap tingkat pengangguran terbuka suatu provinsi.

Dari model SAR pers. (1) menunjukkan bahwa variabel laju pertumbuhan PDRB atas dasar harga konstan (X_3) dan tingkat partisipasi angkatan kerja suatu provinsi (X_4) memiliki hubungan negatif dengan tingkat pengangguran terbuka provinsi di Indonesia. Hal ini mengindikasikan bahwa ketika laju pertumbuhan PDRB atas dasar harga konstan suatu provinsi naik sebesar satu satuan dengan faktor lain dianggap konstan, maka tingkat pengangguran terbuka cenderung menurun sebesar 0,3430 dan ketika tingkat partisipasi angkatan kerja di suatu provinsi naik sebesar satu satuan dengan faktor lain dianggap konstan, maka tingkat pengangguran terbuka cenderung menurun sebesar 0,3903. Koefisien ρ sebesar 0,3877 menunjukkan bahwa adanya peningkatan pengaruh dari wilayah yang mengelilingi suatu provinsi, maka akan meningkatkan tingkat pengangguran terbuka suatu provinsi.

Nilai koefisien Determinasi (R^2) yang diperoleh dari model ini adalah sebesar 71,54% yang berarti keragaman TPT di Indonesia dapat dijelaskan oleh laju pertumbuhan PDRB atas dasar harga konstan dan tingkat partisipasi angkatan kerja sebesar 71,54%, sedangkan sisanya 28,46% dipengaruhi oleh faktor lain yang tidak dilibatkan dalam model dan nilai *Akaike Information Criterion* (AIC) yang dihasilkan adalah sebesar 112,4200. Nilai AIC yang dihasilkan pada model ini lebih kecil dari nilai AIC sebelum pemodelan berdasarkan variabel yang signifikan yaitu 112,43, sehingga model ini dikatakan layak digunakan untuk memodelkan TPT di Indonesia pada tahun 2020. Hal ini sesuai dengan penelitian yang dilakukan oleh Jaya *et al* (2017) yang mengemukakan bahwa penggunaan pembobot *k-nn* dalam pemodelan spasial dapat meminimumkan nilai AIC.

Berdasarkan model persamaan (1), diperoleh estimasi model SAR untuk setiap provinsi di Indonesia adalah sebagai berikut:

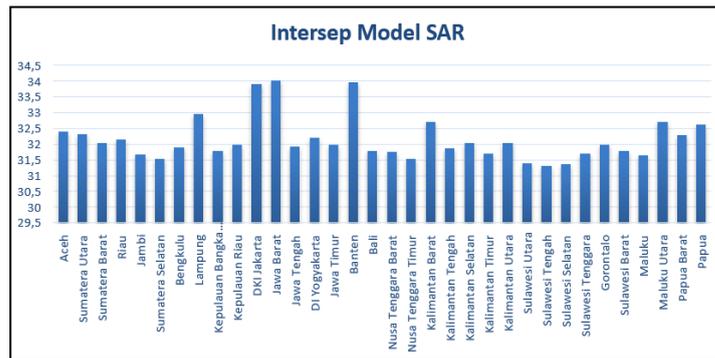
1. Provinsi Aceh

$$\begin{aligned}\hat{Y}_1 &= 29,8265 + 0,3877(W_{ij})(Y_2 + Y_4) - 0,3430X_3 - 0,3903X_4 \\ &= 29,8265 + 0,3877(0.5)(6,91+6,32) - 0,3430X_3 - 0,3903X_4 \\ &= 29,8265 + 2,5646 - 0,3430X_3 - 0,3903X_4 \\ &= 32,3911 - 0,3430X_3 - 0,3903X_4\end{aligned}$$

2. Provinsi Sumatera Utara

$$\begin{aligned}\hat{Y}_2 &= 29,8265 + 0,3877(W_{ij})(Y_1 + Y_4) - 0,3430X_3 - 0,3903X_4 \\ &= 29,8265 + 0,3877(0.5)(6,59+6,32) - 0,3430X_3 - 0,3903X_4 \\ &= 29,8265 + 2,5026 - 0,3430X_3 - 0,3903X_4 \\ &= 32,3291 - 0,3430X_3 - 0,3903X_4\end{aligned}$$

Proses perhitungan model SAR dilakukan berulang pada setiap lokasi, yaitu sampai provinsi Papua, sehingga diperoleh 34 model SAR. Nilai intersep dari masing-masing lokasi dapat dilihat pada gambar 5:



Gambar 5 Diagram Batang Intersep model SAR

Berdasarkan Gambar 5 dapat dilihat bahwa nilai intersep ketiga provinsi yaitu DKI Jakarta, Jawa Barat dan Banten memiliki nilai intersep yang tinggi dan hampir sama dikarenakan ketiga provinsi ini memiliki jarak yang dekat antarsatu samalain. Ketiga provinsi ini daerah yang memiliki intersep tertinggi terdapat di provinsi Jawa Barat yaitu 34,0117. Nilai intersep pada daerah ini tertinggi dikarenakan dipengaruhi oleh daerah tetangganya yang memiliki TPT tinggi pula.

KESIMPULAN

Berdasarkan hasil analisis data dan pembahasan yang telah dijelaskan, maka diambil kesimpulan sebagai berikut: Model yang sesuai berdasarkan hasil uji LM adalah model *Spatial Autoregressive* (SAR) dan faktor-faktor yang mempengaruhi TPT di Indonesia tahun 2020 adalah laju pertumbuhan PDRB atas dasar harga kostan (X_3) dan tingkat partisipasi angkatan kerja (X_4).

DAFTAR PUSTAKA

- BPS. 2021. *Statistik Indonesia 2021*. Jakarta: Badan Pusat Statistik.
- Djuraidah, A., & Wigena, A.H. 2012. Regresi Spasial untuk Menentukan Faktor-faktor Kemiskinan di Provinsi Jawa Timur. *Jurnal Statistika*. Vol.12 No.1 Hal.1.
- Elhorst, J.P. 2014. *Spatial Econometric: From Cross Sectional data to Spatial Panelsi*. Berlin: Springer.
- Jaya, I.G.N.M., Bertho, T., & Zulhanif. 2017. Optimalisasi Matrik Bobot Spasial Berdasarkan K-Nearest Neighbor dalam Spasial Lag Model. *Prosiding Konferensi Nasional Penelitian Matematika dan Pembelajarannya II (KNPMP II)*. Surakarta: 18 Maret 2017.
- Kemenkeu. 2019. *RAPBN 2020 Untuk Akselerasi Daya Saing Melalui Inovasi dan Penguatan Kualitas Sumber Daya Manusia*. Jakarta: Kementerian Keuangan Republik Indonesia
- Kosfeld, R. 2006. *Spatial Econometric*. Germany: University of. Kassel.
- Lesage, J.P. 1998. *Spatial Econometrics*. Departement of Economics, University of Toledo.
- _____. 1999. *The Theory And Practice Of Spasial Econometrics*. Departement of Economics University of Toledo.
- LeSage, J, & Pace, R.K. (2009). *Introduction to Spasial Econometrics*. New York: Springer.
- Mariana. 2013. Pendekatan Regresi Spasial dalam Pemodelan Tingkat Pengangguran Terbuka. *Jurnal Matematika dan Pembelajarannya*. Vol.1 No.1.

- Pusdiktasari, Z. H. 2019. *Pendeteksian Outlier Spasial Menggunakan Average Difference Algorithm (Studi Kasus pada Data Laju PDRB Kota/Kabupaten di Jawa Timur Tahun 2016)*[Skripsi]. Malang: Universitas Brawijaya, Fakultas Matematika dan Ilmu Pengetahuan Alam.
- Sanusi, W., Hisyam, I., & Nur, H.S. 2018. Model Regresi Spasial dan Aplikasinya dalam Menganalisis Angka Putus Sekolah Usia Wajib Belajar di Provinsi Sulawesi Selatan. *Journal of Mathematics, Computations, and Statistics*. Vol.1 No.2 Hal.183.
- Wardani, I.K., Sri, S.H., & Etik, Z. 2019. Pemodelan Tingkat Pengangguran Terbuka Provinsi Jawa Tengah Menggunakan Model Spatial Autoregressive Moving Average. *Prosiding dari Seminar Nasional Matematika dan Pendidikan Matematika (SENDIKA)*. Purworejo.
- Wardiansyah, M., Yulmardi., & Zainul, B. 2016. Analisis Faktor-faktor yang Mempengaruhi Tingkat Pengangguran Terbuka (Studi Kasus Provinsi-provinsi se-Sumatera). *Jurnal Ekonomi Sumberdaya dan Lingkungan*. Vol. 5. No. 1.