

Penerapan *Geographically Weighted Zero-Inflated Poisson Regression* pada Jumlah Kasus Tuberkulosis di Provinsi Sumatera Utara

Khairun Nikmah^{1*}, Riri Syafitri Lubis², Ismail Husein³

^{1,2,3}Matematika, Fakultas Sains dan Teknologi, Universitas Islam Negeri Sumatera Utara, Indonesia

e-mail: khairungemini785@gmail.com

Abstrak

Tuberkulosis adalah penyakit menular yang disebabkan oleh infeksi dari bakteri *Mycobacterium Tuberculosis* yang masuk kedalam paru-paru dan mengakibatkan penderita mengalami gangguan sesak napas yang disertai dengan batuk terus menerus. Perbedaan karakteristik geografis, ekonomi, sosial dan budaya antar lokasi menyebabkan penyebaran covid-19 semakin luas. Penelitian ini bertujuan untuk melihat faktor pengaruh jumlah kasus tuberkulosis di setiap kabupaten/kota menggunakan metode GWZIPR, dimana merupakan pengembangan dari metode GWR dan ZIP. Hasil penelitian menunjukkan bahwa faktor-faktor yang berpengaruh terhadap kasus tuberkulosis di setiap kab/kota terbagi menjadi 1 kelompok berdasarkan variabel signifikan pada model log (poisson state) dan logit (zero state) yaitu semua variabel signifikan pada model log dan logit tersebar di 33 kab/kota yaitu variabel persentase kepadatan penduduk, persentase jumlah puskesmas, persentase jumlah dokter, persentase penduduk miskin dan persentase tenaga perawat.

Kata kunci— Tuberkulosis, Sumatera Utara, GWZIPR

1. PENDAHULUAN

Tuberkulosis adalah penyakit menular yang disebabkan oleh infeksi dari bakteri *Mycobacterium Tuberculosis* yang masuk kedalam paru-paru dan mengakibatkan penderita mengalami gangguan sesak napas yang disertai dengan batuk terus menerus. Indonesia berada pada posisi kedua dengan jumlah kasus TBC didunia setelah India dan diikuti China. TBC di Indonesia diperkirakan sebanyak 969.000 kasus TBC (terdapat satu orang tiap 33 detik). Adapun gejala-gejala yang ditimbulkan seperti batuk, demam, berkeringat pada malam hari atau penurunan berat badan dapat terasa ringan saat seseorang mengalami penyakit aktif.

Tuberkulosis ini masih merupakan masalah kesehatan masyarakat yang menjadi tantangan global. Indonesia merupakan salah satu negara yang mempunyai beban tuberkulosis yang terbesar diantara 5 negara yaitu India, Indonesia, China, Philippina dan Pakistan. Pada tahun 2020 tuberkulosis merupakan penyakit menular dengan angka kematian tertinggi nomor dua setelah covid-19. Berdasarkan laporan kemenkes RI tahun 2020 penderita kasus tuberkulosis mencapai angka 351,936 kasus. Salah satu provinsi yang memiliki kasus tertinggi di Indonesia termasuk Sumatera Utara dengan jumlah kasus 33.779.

Selain merugikan secara ekonomis, TB juga memberikan dampak buruk lainnya secara sosial, stigma dan bahkan dikucilkan oleh masyarakat. Faktor-faktor yang mempengaruhi jumlah kasus tuberkulosis di provinsi sumatera utara yaitu kepadatan penduduk, jumlah rumah sakit, jumlah puskesmas, jumlah tenaga perawat dan persentase penduduk miskin. Adapun tujuan dari penelitian ini adalah untuk melihat faktor-faktor yang mempengaruhi jumlah kasus tuberkulosis di kabupaten/kota Provinsi Sumatera Utara. Untuk melihat faktor pengaruh jumlah tuberkulosis di Sumatera Utara menggunakan analisis regresi, dimana jumlah Covid-19 merupakan variabel respon berupa data diskrit. Analisis regresi merupakan regresi yang digunakan untuk melihat hubungan antara dua variabel.

Model regresi Poisson memiliki asumsi equidispersi yaitu mean dan varians bernilai sama terhadap variabel respon. Terdapat proporsi nilai nol pada variabel respon membuat pelanggaran asumsi regresi poisson, menyebabkan varians lebih besar dari mean disebut sebagai overdispersi. Regresi ZIP akan menghasilkan parameter bersifat global untuk seluruh lokasi data. Data diambil dari setiap lokasi mewakili kondisi yang berbeda dari masing-masing lokasi observasi. Hal ini, dipengaruhi oleh karakteristik masyarakat (sosial,

budaya), kondisi geografis dan perekonomian antara lokasi satu dan lokasi lain, sehingga menyebabkan heterogenitas spasial. Dengan adanya hal tersebut, dilakukan pengembangan dari regresi ZIP yang melibatkan faktor spasial, yakni dengan metode Geographically Weighted Zero-Inflated Poisson Regression.

Penelitian kasus tuberkulosis telah dilakukan oleh Verra dkk (2022) menggunakan model Geographically Weighted Regression (GWR) bahwa faktor yang mempengaruhi. Selanjutnya, penelitian menggunakan model Geographically Weighted Zero-Inflated Poisson Regression (GWZIPR) telah dilakukan oleh Ismah dkk (2020) tentang pada kasus difteri dan Rahmat (2021) tentang Jumlah Kasus Malaria dan Amaliana (2018) mengenai kasus filiarisis.

Berdasarkan pemaparan yang telah dijelaskan diatas, maka penelitian ini , diperlukan suatu model untuk mengetahui faktor-faktor yang mempengaruhi jumlah kasus tuberkulosis di Sumatera Utara yaitu dengan menerapkan Geographically Weighted Zero-Inflated Poisson Regression (GWZIPR).

2. METODE PENELITIAN

Penelitian ini dilakukan di Dinas Kesehatan Provinsi Sumatera Utara (Dinkes Sumut) yang beralamat di Jl. Prof. HM. Yamin Sh No.41AA, Perintis, Kec.Medan Timur, Kota Medan, Sumatera Utara 20232. Adapun variabel yang digunakan pada penelitian ini adalah variabel respon (Y) dan variabel prediktor (X). Dimana variabel Y adalah jumlah kasus Tuberkulosis di Provinsi Sumatera Utara.

Adapun tahapan-tahapan yang akan dilakukan pada penelitian ini adalah, sebagai berikut :

1. Deskripsi variabel respon dan variabel prediktor jumlah kasus Covid-19 di Sumatera Utara.
2. Melakukan pemeriksaan multikolinieritas pada variabel prediktor.
3. Melakukan pemeriksaan Overdispersi pada variabel respon.
4. Melakukan model Zero Inflated Poisson (ZIP).
5. Melakukan model Geographically Weighted Zero-Inflated Poisson Regression dengan langkah berikut :

- a. Menghitung jarak Euclidean antar lokasi pengamatan berdasarkan titik koordinat.

$$d_{ij} = \sqrt{(u_i - u_l)^2 + (v_i - v_l)^2}$$
- b. Menentukan bandwidth optimum untuk setiap lokasi pengamatan menggunakan cross-validation (CV).
- c. Menghitung pembobot dengan fungsi kernel gaussian seperti persamaan.
- d. Melakukan pengujian dependensi spasial menggunakan uji Moran's I dengan persamaan.
- e. Melakukan uji heterogenitas menggunakan uji Breusch Pagan.
- f. Pendugaan parameter menggunakan metode Maximum Likelihood Estimation pada model GWZIPR.

3. HASIL DAN PEMBAHASAN

3.1 Analisis Data

Berdasarkan dari data statistik deskriptif penelitian yang dilakukan didapatkan bahwa nilai varian lebih besar dari mean pada variabel respon dan menyebabkan data menjadi overdispersi. Dimana overdispersi keadaan dimana nilai ragam lebih besar dari rata-rata untuk variabel respon.

Tabel 1. Deskripsi statistik data

Variabel	Rata-rata	Ragam	Deviasi	db	Deviasi/db
Y	523,45	278,51	316,32	27	11,7155

Berdasarkan tabel diatas menunjukkan bahwa nilai rata-rata dari variabel respon (Y) adalah 523,45 dan nilai ragam dari variabel respon (Y) adalah 29873,506. Nilai varians lebih besar dari nilai mean pada variabel respon (Y), sehingga disimpulkan data pada kasus tuberkulosis di sumatera utara mengalami *overdispersi*.

Untuk mengetahui hubungan antara variabel respon dapat dilakukan uji multikolinieritas. Mendeteksi adanya multikolinieritas dilakukan dengan melihat nilai *Varian Inflation Factor* (VIF) pada variabel prediktor. ditunjukkan pada tabel 2 berikut ini:

Tabel 2. Uji Multikolinieritas

Variabel	Nilai VIF
X_1	1,557
X_2	1,032
X_3	2,720
X_4	1,229
X_5	3,086

3.2 Regresi ZIP

Model ZIP adalah model untuk mengatasi masalah pada data kasus yang mengalami *overdispersi* .hal tersebut dikarenakan data mengalami banyak nilai nol. Estimasi parameter pada regresi ZIP dihitung menggunakan bantuan software R yang dirangkum pada tabel 3 berikut ini dengan nilai $\chi^2_{(0,05,1)} = 3,481$:

Tabel 3. Estimasi Parameter *Zero-Inflated Poisson*

Parameter	Estimate	Std.Error	Uji Wald
β_0	2,324	0,345	45,37682
β_1	0,042	0,0027	241,9753
β_2	0,032	0,0015	455,1111
β_3	0,024	0,00451	28,31844
β_4	0,0112	0,00117	91,63562
β_5	0,153	0,0168	82,94005
γ_0	3,258	0,254	164,5261
γ_1	0,117	0,0127	84,87197
γ_2	0,052	0,0151	11,85913
γ_3	0,0423	0,1926	0,048236
γ_4	0,019	0,0078	5,933596
γ_5	0,174	0,0273	40,62311

Berdasarkan hasil dari uji signifikan menggunakan uji *Wald* bahwa pada *poisson state* semua variabel prediktor berpengaruh signifikan terhadap kasus tuberkulosis, sedangkan pada *zero state* hanya variabel $\gamma_1, \gamma_2, \gamma_4, \gamma_5$ yang berpengaruh signifikan terhadap jumlah kasus tuberkulosis. Berdasarkan tabel.3 diatas dan setelah dilakukan uji signifikansi parameter, Sehingga didapatkan model log dan logit sebagai berikut :

$$\ln(\mu_i) = 2,324 + 0,042X_1 + 0,032X_2 + 0,023X_3 + 0,0112X_4 + 0,153X_5$$

$$\text{logit}(\omega_i) = 3,258 + 0,117X_1 + 0,052X_2 + +0,019X_4 + 0,174X_5$$

3.3 Penggunaan Geographically Weighted Zero-Inflated Poisson Regression Pada kasus Tuberkulosis

3.3.1 Uji Heterogenitas Spasial dan Dependensi Spasial

Berdasarkan hasil dari uji dependensi spasial , yang diperoleh dari nilai Moran's I adalah 1,579 lebih kecil dari nilai $Z_{0,025} = 1,960$. Selain itu, nilai *p-value* adalah 0,0598 dan $\alpha = 0,05$. Sehingga, dpt diambil keputusan bahwa gagal tolak H_0 atau tidak ada dependensi spasial. Begitu juga,berdasarkan uji heterogenitas spasial, diperoleh dari nilai *Breusch Pagan* sebesar 11,177 lebih $\chi^2_{(0,05,5)} = 11,0705$. Sehingga dapat diambil kesimpulan bahwa terdapat heterogenitas spasial pada data penelitian.

3.3.2 Estimasi Parameter Model GWZIPR

Model GWZIPR adalah model yang digunakan untuk menduga parameter yang terdapat pengaruh faktor lokasi pada data *overdispersi*. Pada model GWZIPR ini, menggunakan matriks pembobot spasial dengan fungsi *Kernel Gaussian* dan matriks pembobotnya diperoleh dari hasil perhitungan nilai *bandwidth* optimum dan jarak *Euclidean*.

Berdasarkan hasil dari pengujian parsial estimasi parameter model GWZIPR di setiap kabupaten/kota Sumatera Utara. Sebagai contoh, pengujian parameter akan dilakukan di Kabupaten Labuhan Batu seperti berikut ini hasil perhitungan estimasi parameter model GWZIPR menggunakan *software* R. Pengujian ini dilakukan dengan uji *Wald*, kriteria tolak H_0 apabila $W_i > Z_{(0,05/2)}^2 = 1,960$,

Tabel 4. Estimasi Parameter Kab. Labuhanbatu

Model	Parameter	Estimasi	Wald	Kriteria
	β_1	-1,8053	-66,8635	Tolak H_0
<i>Poisson</i>	β_2	-1,8071	-1204,73	Tolak H_0
<i>State</i>	β_3	-1,743	-386,571	Tolak H_0
	β_4	-2,279	-1947,98	Tolak H_0
	β_5	-1,321	-78,633	Tolak H_0
	γ_1	-1,723	-135,679	Tolak H_0
<i>Zero</i>	γ_2	-2,473	-163,82	Tolak H_0
<i>State</i>	γ_3	-2,354	-12,226	Tolak H_0
	γ_4	-1,506	-193,17	Tolak H_0
	γ_5	-0,864	-31,658	Tolak H_0

Berdasarkan hasil yang diperoleh, dapat disimpulkan bahwa tolak didapatkan bahwa tolak H_0 pada semua variabel prediktor yaitu X_1, X_2, X_3, X_4, X_5 untuk model *poisson state* dan variabel X_1, X_2, X_3, X_4, X_5 untuk model *zero state* di Kabupaten Labuhanbatu, memberikan pengaruh pada variabel respon.

Selanjutnya, setelah dilakukan uji parameter parsial GWZIPR, selanjutnya membentuk GWZIPR dengan model *poisson state* dan *zero state* sebagai berikut :

$$\ln(\mu_7) = -187,4563 - 1,8053X_1 - 1,8071X_2 - 1,743X_3 - 2,279X_4 - 1,321X_5$$

$$\text{logit}(\omega_7) = -1703,83 - 1,723X_1 - 2,473X_2 - 2,354X_3 - 1,506X_4 - 0,864X_5$$

Pada model *Poisson State*, apabila terjadi kenaikan 1% akan mengalami peningkatan kasus tuberkulosis pada persentase kepadatan penduduk sebesar $e^{-1,8053} = 0,1644$ $e^{-1,8071} = 0,164$. $e^{-1,743} = 0,17499$. $e^{-2,279} = 0,102$. $e^{-1,321} = 0,266$ Untuk *Zero State*, setiap kenaikan 1% akan memberikan peluang penurunan kasus tuberkulosis seperti $e^{-1,723} = 0,178$.

Berdasarkan perhitungan estimasi parameter dan uji parameter yang dilakukan untuk tiap kabupaten/kota Provinsi Sumatera Utara, dapat disimpulkan bahwa pada dalam penerapan metode GWZIPR yang memiliki 2 model yaitu *Poisson State* dan *Zero State* dan dilakukan uji signifikansi menggunakan uji *Wald* menghasilkan semua variabel signifikan atau memiliki pengaruh terhadap jumlah kasus tuberkulosis di 33 Kabupaten/Kota di Provinsi Sumatera Utara yaitu variabel persentase kepadatan penduduk (X_1), jumlah Puskesmas (X_2), Jumlah Dokter (X_3), persentase Penduduk Miskin (X_4), Jumlah Tenaga Perawat (X_5).

4. KESIMPULAN

Berdasarkan hasil penelitian diatas, dapat disimpulkan bahwa faktor-faktor signifikan yaitu semua variabel signifikan pada model log dan logit tersebar di 33 kab/kot Provinsi Sumatera Utara. Faktor-faktor yang berpengaruh terhadap jumlah kasus tuberkulosis di setiap kab/kota terbagi menjadi 1 kelompok berdasarkan variabel signifikan pada model *Poisson State* (log) dan *Zero State* (logit). Penyebaran ini memiliki letak geografis yang berdekatan antara satu sama lain. Sebagai contoh model GWZIPR pada lokasi pengamatan di Kabupaten Labuhanbatu yaitu sebagai berikut :

$$\ln(\mu_7) = -187,4563 - 1,8053X_1 - 1,8071X_2 - 1,743X_3 - 2,279X_4 - 1,321X_5$$

$$\text{logit}(\omega_7) = -1703,83 - 1,723X_1 - 2,473X_2 - 2,354X_3 - 1,506X_4 - 0,864X_5$$

DAFTAR PUSTAKA

- Adeliana, “Estimation Parameter Geographically Weighted Zero-Inflated Poisson Regression (GWZIPR) With Fixed Bisquare Kernel”, *Proceeding of Conference on Green Technology*, 2018.
- Alfisyahrina Hapsery. D. T, Aplikasi *Geographically Weighted Regression* (GWR) Untuk Pemetaan Faktor Yang Mempengaruhi Indeks Aktivitas Literasi Membaca Di Indonesia. *Riset dan Aplikasi Matematika*, 2022, Vol.5 No.2.
- Amalia, D. F. “Pemodelan Jumlah Kematian Bayi Di Kota Bandung Dengan Regresi Zero-Inflated Poisson”, *Matematika UNAND*, 2022 , Vol.11 No.1.
- Amaliana. F. S, “Comparison of Two Weighting Functions in Geographically Weighted Zero-Inflated Poisson Regression on Filariasis Data”. *Statistics Department Brawijaya University* . 2018
- Adeliana. “Estimation Parameter Geographically Weighted Zero-Inflated Poisson Regression (GWZIPR) With Fixed Bisquare Kernel”, *Proceeding of Conference on Green Technology* . 2018
- Badriyah. L, “Estimasi Parameter Model Regresi Poisson Di Perumun Dengan Metode Maximum Likelihood”, *UIN Maulana Malik Ibrahim* .2019
- BPS Provinsi Sumatera Utara, “Provinsi Sumatera Utara Dalam Angka 2022”, Badan Pusat Statistik Sumatera Utara, 2022.
- Charlton. F, *Geographically Weighted Regression White Paper*. Ireland: Maynooth University, 2009
- Dyah. R. D, “Analisis Jumlah Kasus Malaria di Wilayah Sumatera Menggunakan Geographically Weighted Zero-Inflated Poisson Regression”, *Indonesian journal of Statistics and its Applications* (eISSN:2599-0802), 2020
- Faricha. M, “Pemodelan Geographically Weighted Zero-Inflated Poisson Regression (GWZIPR) dengan Pembobot Adaptive Gaussian Kernel dan Adaptive Bisquare Kernel”, *Universitas Brawijaya*, 2016.
- Greis. L, “Analisis Regresi Zero-Inflated Poisson Untuk Mengatasi Overdispersi Pada Data Kecelakaan Lalu Lintas Di Kota Medan”, *Institusi Universitas Sumatera Utara*, 2017.
- Heinrich Rakuasa. M. R, “Analisis sebaran spasial tingkat kejadian kasus Covid-19 dengan Metode Kernel Density di Kota Ambon”. *Geografi Universitas Negeri Semarang*. 2021.
- Ismah. M. A, “Pendekatan *Geographically Weighted Zero-Inflated Poisson Regression* (GWZIPR) dengan Pembobot *Fixed Bisquare* Pada Kasus Difteri di Indonesia”, *Ilmu Matematika dan Terapan*, 2020.
- Pristianing, N, “Metode Regresi Poisson dengan Spasial Lag Pada Pemdolan HIV/Aids di Kabupaten dan Kota di Provinsi Jawa Timur Tahun 2017”, *Unibversitas Brawijaya*. 2019
- Qorry, V. A. “Geographically Weighted Regression dalam Menganalisis Faktor-Faktor yang Mempengaruhi Kasus Tuberkulosis di Sumatera Utara”, *Sainmatika*, 2022, vol 9 No 2.
- Rahmat. D. D, “Analisis Jumlah Kasus Malaria di Wilayah Sumatera Menggunakan GWZIPR”. *Indonesian Journal of Statistics and Aplications*, 2020, Vol 4 No 4.
- Samson, S. H. “Perbandingan Regresi Zero-Inflated Poisson (ZIP) dan Regresi Zero-Inflated Negative Binomial (ZINB) Pada Data Overdispersi”, *Buletin Ilmiah Mat. Stat. dan Terapannya (Bimaster)*, 2022.
- Tb Indonesia, “Ayo Akhiri TBC”. Diambil, 16 Juni 2023, 2022, <https://tbindonesia.or.id/ayobersamaakhiritbc-indonesia>.