Analisis Metode *Geographically Weighted Generalized Poisson Regression* untuk Pemodelan Faktor yang Mempengaruhi Jumlah Kematian Anak di Provinsi Jawa Timur

Muhamad Adryanta dan Purhadi
Departemen Statistika, Fakultas Matematika, Komputasi, dan Sains Data,
Institut Teknologi Sepuluh Nopember (ITS)
e-mail: purhadi@statistika.its.ac.id

***Abstrak*—Jumlah kematian anak merupakan bagian permasalahan dari salah satu tujuan dalam *Sustainable Development Goals* (SDGs), yaitu menjamin hidup sehat dan kehidupan yang lebih baik untuk seluruh penduduk dunia di segala umur. Jumlah kematian anak berdasarkan Profil Kesehatan Provinsi Jawa Timur periode 2014/2015 diketahui sebesar 428 per 1000 anak 1-4 tahun, jauh dari target global yaitu kurang dari 25 per kematian per 1000 kelahiran hidup. Penelitian ini berfokus pada analisis jumlah kematian anak di Provinsi Jawa Timur pada tahun 2017 beserta faktor-faktor yang diduga mempengaruhinya dengan menggunakan metode *Geographically Weighted Generalized Poisson Regression* (GWGPR) dikarenakan keberadaan faktor spasial yaitu antar kabupaten/kota. Jumlah kematian anak di Jawa Timur pada tahun 2017 memiliki persebaran terbesar berada di wilayah tapal kuda, bagian pusat, dan bagian barat laut wilayah Jawa Timur. Pemodelan GWGPR tanpa menggunakan *exposure* diketahui variabel yang dominan signifikan terhadap masing-masing kabupaten/kota yaitu variabel persentase rumah tangga ber-PHBS, persentase rumah sehat, persentase penduduk dengan akses terhadap fasilitas sanitasi yang layak (jamban sehat), persentase desa/kelurahan yang melaksanakan sanitasi total berbasis masyarakat, penduduk dengan akses berkelanjutan terhadap air minum berkualitas (layak), dan persentase penduduk miskin yang terbagi dalam lima kelompok kabupaten/kota. Pemodelan GWGPR menggunakan *exposure* diketahui tidak ada faktor atau variabel yang berpengaruh secara signifikan terhadap masing-masing kabupaten /kota. Kriteria kesesuaian model menggunakan AICc menunjukkan bahwa *Generalized Poisson Regression* secara umum memiliki nilai AICc terkecil dibandingkan metode lainnya.**

***Kata Kunci*—AICc, Exposure, GPR, GWGPR, Jawa Timur, Kematian Anak**

I. PENDAHULUAN

JAMINAN hidup yang sehat dan baik bagi seluruh penduduk dunia di segala umur merupakan salah satu tujuan dari adanya *Sustainable Development Goals* (SDGs) yang dijadikan kerangka pembangunan dan perundingan seluruh negara secara berkelanjutan hingga tahun 2030 [1]. Masalah kematian khususnya kematian di bawah lima tahun (*under five mortality*) diharapkan secara global mencapai target kurang dari 25 per kematian per 1000 kelahiran hidup [2]. Jumlah kematian anak berdasarkan Profil Kesehatan Provinsi Jawa Timur periode 2014/2015 diketahui sebesar 428 per 1000 anak 1-4 tahun. Analisis jumlah kematian anak menjadi penting dikarenakan kematian anak di beberapa negara adalah masalah yang lebih besar daripada kematian bayi dan laporan statistik kematian anak terdapat korelasi kuat antara penelitian atau riset dengan kebijakan pemerintah [3].

Kondisi wilayah di Jawa Timur yang beragam menyebabkan jumlah kejadian antar wilayah berbeda. Metode *Geographically Weighted Generalized Poisson Regression* (GWGPR) merupakan metode yang memperhitungkan faktor spasial atau karakteristik wilayah dan memodelkan korelasi antar variabel respon yang berupa data *count*. Metode tersebut bentuk dari pengembangan dari *Generalized Poisson Regression* dengan lokasi pengambilan data diasumsikan berdistribusi poisson. Penggunaan *Generalized Poisson Regression* dikarenakan data jumlah kematian anak di Jawa Timur memiliki nilai varians yang lebih besar dari nilai mean (overdispersi). Metode ini adalah menggunakan syarat regresi kernel yang menggunakan fungsi pembobotan spasial untuk memperkirakan variasi spasial dalam parameter *Generalized Poisson Regression*.

Penelitian sebelumnya yang menggunakan faktor spa-sial yaitu tentang pemodelan angka kematian ibu di Indonesia dengan metode *Geographically Weighted Poisson Regression* (GWPR). Kesimpulannya adalah lebih baik menggunakan model GWPR dibandingkan model Regresi Poisson dengan alasan rendahnya nilai AIC dan *deviance* model GWPR [4]. Penelitian sejenis yaitu menganalisis jumlah kematian bayi dan jumlah kematian ibu di Jawa Timur tahun 2013 [5]. Penelitian tersebut melakukan pemodelan data yang berdistribusi poisson tergeneralisasi dengan memperhitungkan variasi spasial dalam hubungan antar variabel respon bivariat dan variabel-variabel prediktornya dengan hasil akhir parameter model yang berbeda-beda pada setiap lokasi.

Penelitian kali ini melakukan analisis jumlah kematian anak di provinsi Jawa Timur pada tahun 2017 beserta faktor-faktor yang mempengaruhinya dengan menggunakan metode *Geographically Weighted Generalized Poisson Regression* (GWGPR) secara univariat. Alasannya variabel yang digunakan merupakan peristiwa yang diduga mengikuti distribusi *Generalized Poisson* dan bersifat random serta terdapat perbedaan karakteristik di setiap kabupaten/kota atau memperhatikan faktor spasial. Asumsi overdispersi bisa diatasi pula menggunakan metode tersebut. Beberapa faktor-faktor yang signifikan terhadap jumlah kematian anak dan model terbaik pada setiap kabupaten/kota di provinsi Jawa Timur bisa diketahui menggunakan metode GWGPR.

II. TINJAUAN PUSTAKA

A. *Regresi Poisson Univariat*

Regresi poisson adalah model regresi yang dapat digunakan pada data yang variabel responnya berdistribusi tidak normal dan berjenis diskrit, yaitu berdistribusi Poisson sebagai syarat utamanya. Distribusi Poisson memberikan suatu model yang realistis untuk berbagai macam fenomena acak selama nilai dari variabel acak Poisson berupa bilangan bulat non-negatif. Variabel random diskrit Y berdistribusi Poisson dengan parameter λ_q jika dan hanya jika memiliki fungsi peluang sebagai berikut [6].

$$P(Y=y|\lambda(q))=\begin{cases}\dfrac{e^{-\lambda(q)}\left(\lambda(q)\right)^y}{y!}, & y=0,1,2,...;\lambda(q)>0\\ 0 & ,y \text{ yang lain}\end{cases} \quad (1)$$

Keterangan: q merupakan *exposure*

Jika diberikan sampel random $Y_i \sim P\left(\lambda(q_i)\right), i=1,2,\dots,n$ dengan q_i merupakan *exposure* dengan definisi jumlah anak usia 1−4 tahun di kabupaten/kota di Jawa Timur pada unit ke-*i*. Model regresi Poisson berupa logaritma natural dari nilai ekspektasi Y_i yang proporsional terhadap jumlah q_i (apabila model tanpa *exposure* maka tidak menggunakan variabel q_i) dan bergantung pada variabel bebas $\boldsymbol{x}_i$, bisa dituliskan persamaannya sebagai berikut.

$$\ln\left(\frac{E(Y_i)}{q_i}\right)=\boldsymbol{x}_i^T\boldsymbol{\beta}$$

$$\lambda(q_i,\boldsymbol{x}_i)=q_i\,e^{\boldsymbol{x}_i^T\boldsymbol{\beta}} \quad (2)$$

B. *Multikolinieritas*

Multikolinearitas adalah terdapatnya hubungan atau korelasi secara signifikan (nilainya tinggi) antara variabel prediktor satu dengan variabel prediktor lain. Pendeteksi keberadaan multikolinearitas bisa menggunakan nilai *Variance Inflating Factor* (VIF) yang dinyatakan pada persamaan berikut [7],

$$VIF=\frac{1}{1-R_j^2} \quad (2)$$

Nilai VIF yang lebih besar dari 10 menunjukkan adanya multikolinearitas.

C. *Overdispersi*

Overdispersi merupakan sebuah kondisi dimana nilai *varians* yang lebih besar dari nilai *mean*. Dampak adanya *overdispersi* dalam pemodelan adalah suatu variabel bisa saja muncul sebagai parameter yang signifikan padahal sebenarnya tidak signifikan [8]. Pemeriksaan overdispersi dapat menggunakan perbandingan rasio sebagai berikut [9].

$$\theta=\frac{D(\hat{\boldsymbol{\beta}})}{df}=\frac{\left(-2\ln(\Lambda)\right)}{n-k-1}=\frac{\left(-2\ln\left(\dfrac{L(\hat{\omega})}{L(\hat{\Omega})}\right)\right)}{n-k-1} \quad (3)$$

Terdapat syarat tertentu untuk parameter dispersi θ dalam hal pengaruh terhadap penentuan model yang akan dianalisis sebagai berikut.

$$\theta=\begin{cases}0 & \text{, model regresi poisson}\\ >1 & \text{, model GPR overdispersi}\\ <0 & \text{, model GPR underdispersi}\end{cases}$$

D. *Generalized Poisson Regression (GPR)*

Model GPR adalah salah satu metode statistika yang merupakan model yang digunakan untuk data *count* dalam rangka penanganan pelanggaran asumsi e*quidispersi* yang terjadi pada model regresi poisson, dengan asumsi variabel random merupakan berdistribusi Poison tergeneralisasi [10]. Persamaan model GPR dapat ditulis sebagai berikut (model GPR dengan *exposure* menggunakan variabel q_i).

$$\ln\left(\frac{\mu(q_i,\boldsymbol{x}_i)}{q_i}\right)=\boldsymbol{x}_i^T\boldsymbol{\beta}=\beta_0+\beta_1x_{1i}+\beta_2x_{2i}+\ldots+\beta_kx_{ki} \quad (4)$$

$$\mu_i=\mu(\boldsymbol{x}_i)=e^{\boldsymbol{x}_i^T\boldsymbol{\beta}} \quad (5)$$

$$\mu_i=\mu(q_i,\boldsymbol{x}_i)=q_ie^{\boldsymbol{x}_i^T\boldsymbol{\beta}} \quad (6)$$

Bentuk logaritma model GPR tanpa *exposure* sebagai berikut.

$$\begin{aligned}\ln\left(f(y_i)\right)&=y_i\left(\ln\left(e^{\boldsymbol{x}_i^T\boldsymbol{\beta}}\right)-\ln\left(1+\theta e^{\boldsymbol{x}_i^T\boldsymbol{\beta}}\right)\right)+y_i\ln(1+\theta y_i)\\&-\ln(1+\theta y_i)-\ln(y_i!)-\frac{\mu_i(1+\theta y_i)}{1+\theta e^{\boldsymbol{x}_i^T\boldsymbol{\beta}}}\end{aligned} \quad (7)$$

Bentuk logaritma model GPR dengan *exposure* sebagai berikut.

$$\begin{aligned}\ln\left(f(y_i)\right)&=y_i\left(\ln\left(q_ie^{\boldsymbol{x}_i^T\boldsymbol{\beta}}\right)-\ln\left(1+\theta q_ie^{\boldsymbol{x}_i^T\boldsymbol{\beta}}\right)\right)+y_i\ln(1+\theta y_i)\\&-\ln(1+\theta y_i)-\ln(y_i!)-\frac{\mu_i(1+\theta y_i)}{1+\theta q_ie^{\boldsymbol{x}_i^T\boldsymbol{\beta}}}\end{aligned} \quad (8)$$

Penaksiran parameter dari model GPR menggunakan metode *Maximum Likelihood Estimation* (MLE) dengan melakukan iterasi Newton-Rhapson. Fungsi *likelihood* model GPR untuk yang tanpa *exposure* dan dengan *exposure* masing-masing dituliskan sebagai berikut.

$$L(\boldsymbol{\beta},\theta)=\prod_{i=1}^{n}\left(\frac{e^{\boldsymbol{x}_i^T\boldsymbol{\beta}}}{1+\theta e^{\boldsymbol{x}_i^T\boldsymbol{\beta}}}\right)^{y_i}\frac{(1+\theta y_i)^{y_i-1}}{y_i!}\exp\left(-\frac{e^{\boldsymbol{x}_i^T\boldsymbol{\beta}}(1+\theta y_i)}{1+\theta e^{\boldsymbol{x}_i^T\boldsymbol{\beta}}}\right) \quad (9)$$

$$L(\boldsymbol{\beta},\theta)=\prod_{i=1}^{n}\left(\frac{q_ie^{\boldsymbol{x}_i^T\boldsymbol{\beta}}}{1+\theta q_ie^{\boldsymbol{x}_i^T\boldsymbol{\beta}}}\right)^{y_i}\frac{(1+\theta y_i)^{y_i-1}}{y_i!}\exp\left(-\frac{q_ie^{\boldsymbol{x}_i^T\boldsymbol{\beta}}(1+\theta y_i)}{1+\theta q_ie^{\boldsymbol{x}_i^T\boldsymbol{\beta}}}\right) \quad (10)$$

Langkah berikutnya adalah membentuk fungsi ln-*likelihood* model GPR yang ditunjukkan sebagai persamaan berikut.

$$\begin{aligned}\ln\left(L(\boldsymbol{\beta},\theta)\right)&=\sum_{i=1}^{n}y_i\left(\ln\left(q_ie^{\boldsymbol{x}_i^T\boldsymbol{\beta}}\right)-\ln\left(1+\theta q_ie^{\boldsymbol{x}_i^T\boldsymbol{\beta}}\right)\right)+y_i\ln(1+\theta y_i)\\&-\ln(1+\theta y_i)-\ln(y_i!)-\frac{q_ie^{\boldsymbol{x}_i^T\boldsymbol{\beta}}(1+\theta y_i)}{1+\theta q_ie^{\boldsymbol{x}_i^T\boldsymbol{\beta}}}\end{aligned} \quad (11)$$

Selanjutnya adalah dilakukan iterasi Newton-Rhapson dalam rangka mendapatkan penaksir parameter model GPR yang *close form* dengan cara melakukan repetisi iterasi yang dimulai dari suku ke m hingga $m=m+1$ pada persamaan $\hat{\boldsymbol{\beta}}_{(m+1)}=\hat{\boldsymbol{\beta}}_{(m)}-\boldsymbol{H}^{-1}\left(\hat{\boldsymbol{\beta}}_{(m)}\right)\boldsymbol{g}\left(\hat{\boldsymbol{\beta}}_{(m)}\right)$ hingga nilai $\hat{\boldsymbol{\beta}}_{(m)}$ konvergen, yaitu $\left\|\hat{\boldsymbol{\beta}}_{(m+1)}-\hat{\boldsymbol{\beta}}_{(m)}\right\|<\varepsilon$ dimana ε lebih besar dari nol dan sangat kecil

Pengujian model secara serentak dapat menggunakan *Maximum Likelihood Ratio Test* (MLRT) untuk menguji signifikansi dari faktor geografis atau lokasi. Hipotesis dapat dituliskan sebagai berikut [11].

$H_0:\beta_1=\beta_2=\ldots=\beta_j=0$

$H_1:$ minimal ada satu $\beta_j\neq 0$, $j=1,2,\ldots,k$

Langkah kedua menentukan statistik rasio *likelihood* sebagai statistik uji yang ditulis dalam persamaan berikut.

$$D(\hat{\boldsymbol{\beta}}_j)=-2\ln(\Lambda)=-2\ln\left(\frac{L(\hat{\omega})}{L(\hat{\Omega})}\right) \quad (12)$$

Daerah penolakan H_0 adalah jika nilai $D(\hat{\boldsymbol{\beta}}_j)$ lebih besar dibandingkan nilai kritis $\chi^2_{(\alpha,k)}$, yang bermakna minimal terdapat satu variabel prediktor dalam model GPR yang berpengaruh signifikan terhadap variabel respon. Kemudian dilakukan pengujian parameter model secara parsial dalam rangka mengetahui parameter mana saja yang signifikan mempengaruhi variabel responnya. Berikut adalah bentuk dari hipotesis yang digunakan.

$H_0 : \beta_j = 0$

$H_1 : \beta_j \neq 0$, $j = 1,2,\ldots,k$

Statistik uji yang digunakan adalah sebagai berikut.

$$Z = \frac{\hat{\beta}_j}{SE(\hat{\beta}_j)} \qquad (13)$$

Daerah penolakan H_0 adalah jika nilai statistik uji $|Z|$ lebih besar dibandingkan nilai kritis $Z_{\left(\frac{\alpha}{2}\right)}$, dengan makna variabel prediktor tersebut berpengaruh signifikan terhadap variabel respon pada setiap lokasi dalam model GPR.

E. *Heterogenitas Spasial*

Heterogenitas spasial merupakan gambaran peristiwa a-tau proses yang terdistribusi pada wilayah yang tidak merata. Pengujian heterogenitas spasial dilakukan untuk mengetahui apakah terdapat karakteristik atau keunikan di setiap lokasi pengamatan. Kejadian heterogenitas spasial dapat menghasilkan parameter regresi yang berbeda-beda di setiap lokasi pengamatan. Heterogenitas spasial dapat diuji menggunakan uji *Breusch-Pagan* (BP) yang mempunyai hipotesis sebagai berikut:

$H_0 : \sigma_1^2 = \sigma_2^2 = \ldots = \sigma_n^2 = \sigma^2$ (*Varians* antar lokasi sama)

H_1 : minimal ada satu $\sigma_i^2 \neq \sigma^2; i = 1,2,...,k$ (*Varians* antar lokasi berbeda)

Persamaan statistik uji yang digunakan adalah sebagai berikut.

$$BP = \left(\frac{1}{2}\right) \boldsymbol{f}^T \boldsymbol{Z}\left(\boldsymbol{Z}^T\boldsymbol{Z}\right)^{-1} \boldsymbol{Z}^T \boldsymbol{f} \qquad (14)$$

dimana elemen vektor $\boldsymbol{f}$ adalah $f_i = \frac{e_i^2}{\hat{\sigma}^2} - 1$ dengan $e_i = y_i - \hat{y}_i$, $\boldsymbol{Z}$ merupakan matriks berukuran $n \times (k+1)$ yang berisi vektor yang sudah dinormalstandarkan untuk setiap observasi. Keputusan menolak H_0 diperoleh jika nilai statistik uji *Breusch-Pagan* lebih besar dari nilai kritis $\chi^2_{(\alpha,k)}$ atau nilai *p-value* kurang dari taraf signifikansi α, dengan makna varians antar lokasi berbeda alias terjadi heteroskedasititas dalam model.

F. *Matrik Pembobot Spasial*

Keragaman spasial antara lokasi satu dengan lokasi lain ditunjukkan dengan adanya pembentukan matriks pembobot **W** dengan menggunakan fungsi kernel yang entri-entrinya merupakan fungsi kontinu dari jarak *Euclidian* antar lokasi. Matriks pembobot yang dapat digunakan salah satunya yaitu *Adaptive Kernel Bisquare,* yaitu suatu fungsi kernel dengan nilai *bandwidth* yang berbanding terbalik pada setiap lokasi pengamatan terhadap besarnya kepadatan data pada wilayah analisis. Berikut persamaan fungsi dari kernel *Adaptive Kernel Bisquare*.

$$w_{ij} = \begin{cases} \left(1 - \left(\frac{d_{ij}}{b_i}\right)^2\right)^2 \\ 0 \end{cases} \qquad (15)$$

d_{ij} merupakan jarak *Euclidian* antar lokasi ke-*i* dan lokasi ke-*j* dengan rumus $d_{ij} = \sqrt{(u_i - u_j)^2 + (v_i - v_j)^2}$ Kemudian b_i adalah *bandwith* lokasi ke-*i*, yang merupakan radius suatu lingkaran dimana titik data yang berada di dalamnya diperhitungkan dalam membentuk model lokasi ke-*i*. Sementara u_i dan v_i masing-masing merupakan lokasi geografis lintang (*latitude*) dan lokasi geografis bujur (*longitude*) kabupaten/kota ke-*i*. Pemilihan *bandwidth* optimum memiliki kepentingan dalam hal pembentukan matriks pembobot, pengaruh ketepatan model terhadap data, dan mengatur besar kecilnya *varians* dan bias penaksir yang dihasilkan [12]. Penggunaan metode *Cross Validation* (CV) dapat dilakukan untuk menentukan *bandwidth* optimum. *Bandwidth* yang optimal ditunjukkan dengan nilai CV minimum. Persamaan yang digunakan adalah berikut [14].

$$CV(b) = \sum_{i=1}^{n} \left(y_i - \hat{y}_{\neq i}(b_i)\right)^2 \qquad (16)$$

dimana $\hat{y}_{\neq i}(b_i)$ merupakan nilai estimasi y_i dengan syarat pengamatan lokasi (u_i, v_i) dihilangkan dari proses estimasi dan n adalah jumlah sampel lokasi pengamatan.

G. *Geographically Weighted Generalized Poisson Regression (GWGPR)*

Model GWGPR adalah salah satu metode statistika yang merupakan pengembangan dari *Generalized Poisson Regression* dengan penaksir parameter model bersifat lokal untuk setiap titik atau lokasi pengamatan (titik garis lintang dan titik garis bujur). Persamaan model GWGPR untuk lokasi ke-*i* dengan menggunakan fungsi *log-link* dapat dituliskan sebagai berikut.

$$\eta_i = g(\mu_i) = \ln(\mu_i) = \boldsymbol{x}_i^T \boldsymbol{\beta}(u_i, v_i)$$

$$\mu_i = g^{-1}(\eta_i) = g^{-1}\left(\boldsymbol{x}_i^T \boldsymbol{\beta}(u_i, v_i)\right)$$

$$\mu_i = \exp\left(\boldsymbol{x}_i^T \boldsymbol{\beta}(u_i, v_i)\right) \qquad (17)$$

$$\mu_i = q_i \exp\left(\boldsymbol{x}_i^T \boldsymbol{\beta}(u_i, v_i)\right) \qquad (18)$$

Penaksiran parameter dari model GWGPR menggunakan metode *Maximum Likelihood Estimation* (MLE) dengan cara memaksimumkan fungsi *likelihood* [12]. Langkah pertama membentuk fungsi *likelihood* dengan persamaan:

$$L(\boldsymbol{\beta}(u_i, v_i), i = 1,2,\ldots,n) = \prod_{i=1}^{n} \left(\frac{\mu_i}{1+\theta\mu_i}\right)^{y_i} \frac{(1+\theta y_i)^{y_i - 1}}{y_i!} \exp\left(-\frac{\mu_i(1+\theta y_i)}{1+\theta\mu_i}\right)$$

Langkah kedua adalah merubah fungsi *likelihood* menjadi bentuk logaritma natural (ln-*likelihood*) sebagai berikut.

$$\ln\left(L(\boldsymbol{\beta}(u_i, v_i), i = 1,2,\ldots,n)\right) = \ln\left(\prod_{i=1}^{n} f(y_i)\right) = \sum_{i=1}^{n} \ln(f(y_i))$$

$$= \sum_{i=1}^{n} \ln\left[\prod_{i=1}^{n} \left(\frac{\mu_i}{1+\theta\mu_i}\right)^{y_i} \frac{(1+\theta y_i)^{y_i-1}}{y_i!} \exp\left(-\frac{\mu_i(1+\theta y_i)}{1+\theta\mu_i}\right)\right]$$

$$= \sum_{i=1}^{n} \left[y_i \left(\ln(\mu_i) - \ln(1+\theta\mu_i) \right) + (y_i - 1)\ln\left((1+\theta y_i)\right) \right]$$

$$+ \sum_{i=1}^{n} \left[-\ln(y_i!) - \frac{\mu_i(1+\theta y_i)}{1+\theta\mu_i} \right]$$

$$= \sum_{i=1}^{n} \left[y_i \left(\ln\left(\exp\left(q_i \boldsymbol{x}_i^T \boldsymbol{\beta}(u_i,v_i)\right)\right) - \ln\left(1+\theta\exp\left(q_i \boldsymbol{x}_i^T \boldsymbol{\beta}(u_i,v_i)\right)\right)\right) \right]$$

$$+ \sum_{i=1}^{n} \left[(y_i - 1)\ln\left((1+\theta y_i)\right) - \ln(y_i!) \right]$$

$$+ \sum_{i=1}^{n} \left[-\frac{\exp\left(q_i \boldsymbol{x}_i^T \boldsymbol{\beta}(u_i,v_i)\right)(1+\theta y_i)}{1+\theta\exp\left(q_i \boldsymbol{x}_i^T \boldsymbol{\beta}(u_i,v_i)\right)} \right]$$

Langkah ketiga adalah memaksimumkan fungsi ln-*likelihood* model GWGPR dengan menambahkan pembobot geografis dari masing-masing lokasi yang ditunjukkan sebagai persamaan berikut

$$\ln\left(L^*\left(\boldsymbol{\beta}\left(u_{i^*},v_{i^*}\right)\right)\right) = \sum_{i=1}^{n} w_{ii^*} \left[y_i \left(q_i \boldsymbol{x}_i^T \boldsymbol{\beta}(u_i,v_i) \right) \right]$$

$$+ \sum_{i=1}^{n} w_{ii^*} \left[-\ln\left(1+\theta\exp\left(q_i \boldsymbol{x}_i^T \boldsymbol{\beta}(u_i,v_i)\right)\right) \right]$$

$$+ \sum_{i=1}^{n} w_{ii^*} \left[(y_i - 1)\ln\left((1+\theta y_i)\right) - \ln(y_i!) \right]$$

$$+ \sum_{i=1}^{n} w_{ii^*} \left[-\frac{\exp\left(q_i \boldsymbol{x}_i^T \boldsymbol{\beta}(u_i,v_i)\right)(1+\theta y_i)}{1+\theta\exp\left(q_i \boldsymbol{x}_i^T \boldsymbol{\beta}(u_i,v_i)\right)} \right]$$

Langkah keempat adalah dilakukan iterasi Newton-Rhapson dalam rangka mendapatkan penaksir parameter model GWGPR yang *close form* dengan cara melakukan repetisi iterasi yang dimulai dari suku ke m hingga $m = m + 1$ pada persamaan $\hat{\boldsymbol{\beta}}_{(m+1)} = \hat{\boldsymbol{\beta}}_{(m)} - \boldsymbol{H}^{-1}\left(\hat{\boldsymbol{\beta}}_{(m)}\right)\boldsymbol{g}\left(\hat{\boldsymbol{\beta}}_{(m)}\right)$ hingga nilai $\hat{\boldsymbol{\beta}}_{(m)}$ konvergen, yaitu $\left\|\hat{\boldsymbol{\beta}}_{(m+1)} - \hat{\boldsymbol{\beta}}_{(m)}\right\| < \varepsilon$ dimana ε lebih besar dari nol dan sangat kecil.

Pengujian model secara serentak dapat menggunakan *Maximum Likelihood Ratio Test* (MLRT) untuk menguji signifikansi dari faktor geografis atau lokasi. Hipotesis dapat dituliskan sebagai berikut [11].

$H_0 : \beta_1(u_i,v_i) = \beta_2(u_i,v_i) = \ldots = \beta_k(u_i,v_i) = 0 \; ; i = 1,2,\ldots,n$

$H_1 :$ minimal ada satu $\beta_j(u_i,v_i) \neq 0 \; ; i = 1,2,\ldots,n ; \; j = 1,2,\ldots,k$

Langkah kedua menentukan statistik rasio *likelihood* sebagai statistik uji yang ditulis dalam persamaan berikut.

$$D\left(\hat{\boldsymbol{\beta}}_j(u_i,v_i)\right) = -2\ln\Lambda = -2\ln\left[\frac{L(\hat{w})}{L(\hat{\Omega})}\right] \qquad (19)$$

Daerah penolakan H_0 adalah jika nilai $D\left(\hat{\boldsymbol{\beta}}_j(u_i,v_i)\right)$ lebih besar dibandingkan nilai kritis $\chi^2_{(a,k)}$, dengan makna minimal terdapat satu variabel prediktor dalam model GWGPR yang berpengaruh signifikan terhadap variabel respon.

Kemudian dilakukan pengujian parameter model se-cara parsial dalam rangka mengetahui parameter mana saja yang signifikan mempengaruhi variabel responnya. Berikut adalah bentuk dari hipotesis yang digunakan.

$H_0 : \beta_j(u_i,v_i) = 0$ (parameter $\beta_j(u_i,v_i)$ tidak berpengaruh signifikan pada lokasi (u_i,v_i))

$H_1 : \beta_j(u_i,v_i) \neq 0$ (parameter $\beta_j(u_i,v_i)$ berpengaruh signifikan pada lokasi (u_i,v_i))

Statistik uji yang digunakan adalah sebagai berikut.

$$Z = \frac{\hat{\beta}_j(u_i,v_i)}{SE\left(\hat{\beta}_j(u_i,v_i)\right)} = \frac{\hat{\beta}_j(u_i,v_i)}{\sqrt{var\left(\hat{\beta}_j(u_i,v_i)\right)}} \qquad (19)$$

Daerah penolakan H_0 adalah jika nilai statistik uji $|Z|$ lebih besar dari nilai kritis $Z_{(\alpha/2)}$ dengan makna variabel prediktor tersebut berpengaruh signifikan terhadap variabel respon pada setiap lokasi dalam model GWGPR.

H. Kriteria Kebaikan Model

Pemilihan model terbaik antara model *Generalized Poisson Regression* dan *Geographically Weighted Generalized Poisson Regression* dengan *exposure* bisa ditinjau kriteria kesesuaian model secara statistik menggunakan metode *Akaike's Information Criteriation Correction* (AICc) dengan sampel kecil. AICc digunakan untuk mengatasi potensi *overfitting* pada *Akaike's Information Criteriation* (AIC), yaitu terjadinya kondisi lebih banyak jumlah parameter dibandingkan dengan jumlah pengamatan yang mengakibatkan bias terhadap penarikan kesimpulan [13]. AICc memiliki formula sebagai berikut.

$$AICc = AIC + 2\left(\frac{K(K+1)}{n-K-1}\right)$$

$$AICc = -2\ln L(\hat{\beta}) + 2K + 2\left(\frac{K(K+1)}{n-K-1}\right) \qquad (19)$$

$L(\hat{\beta})$ merupakan nilai maksimum dari fungsi likelihood parameter model dan K merupakan jumlah parameter dalam model. adalah nilai maksimum dari fungsi likelihood parameter model, n adalah jumlah pengamatan, dan K merupakan jumlah parameter dalam model. Model terbaik yang dipilih adalah model dengan nilai AICc paling kecil.

Tabel 1.
Variabel Penelitian

Variabel	Deskripsi
Y	Jumlah kematian anak
X_1	Persentase rumah tangga ber-PHBS
X_2	Persentase rumah sehat
X_3	Rasio Puskesmas per 30.000 penduduk
X_4	Persentase penduduk dengan akses terhadap fasilitas sanitasi yang layak (jamban sehat)
X_5	Persentase desa/kelurahan yang melaksanakan sanitasi total berbasis masyarakat
X_6	Penduduk dengan akses berkelanjutan terhadap air minum berkualitas (layak)
X_7	Persentase Penduduk Miskin

I. Angka Kematian Anak

Angka kematian anak didefinisikan sebagai kematian anak yang tergolong berusia 1-4 tahun dalam kurun satu tahun tertentu per 1000 anak umur yang sama pada pertengahan tahun itu yang dimana tidak termasuk kematian bayi [14]. Kematian anak pada suatu wilayah tertentu mencerminkan kondisi kesehatan lingkungan yang langsung mempengaruhi tingkat kesehatan anak.

III. METODOLOGI PENELITIAN

A. *Sumber Data*

Penelitian ini menggunakan data sekunder yang berasal dari BPS Jawa Timur dan Profil Kesehatan Jawa Timur yang dikeluarkan oleh Dinas Kesehatan Provinsi Jawa Timur. Unit pengamatan yang digunakan merupakan kabupaten/kota di Jawa Timur yang berjumlah 38 buah, yang terdiri atas 29 kabupaten dan 9 kota.

B. *Variabel Penelitian*

Variabel-variabel yang digunakan dalam penelitian ini ditabulasikan pada Tabel 1.

Keterangan: q_i merupakan *exposure* dengan definisi jumlah anak dengan rentang usia 0-4 tahun.

C. *Langkah Analisis*

Langkah analisis dalam penelitian ini yang digunakan adalah sebagai berikut.

1. Mengumpulkan data-data yang digunakan sebagai variabel penelitian
2. Melakukan pengolahan data berupa analisis karakteristik jumlah kematian anak dan faktor-faktor penyebabnya menggunakan statistik deskriptif
3. Melakukan pewarnaan yang sama pada peta kabupaten/kota di Jawa Timur berdasarkan beberapa *range* tertentu
4. Melakukan identifikasi korelasi antar variabel penelitian (variabel respon dengan variabel prediktor) menggunakan korelasi Pearson dan pengecekan *scatterplot*
5. Melakukan pengujian multikolinieritas antar variabel prediktor menggunakan nilai *Variance in Factor* (VIF)
6. Melakukan analisis model *Generalized Poisson Regression* (GPR) tanpa *exposure*
7. Pengujian overdispersi model *Generalized Poisson Regression* (GPR) tanpa *exposure*
8. Melakukan analisis model *Generalized Poisson Regression* (GPR) dengan *exposure*
9. Menguji signifikansi parameter model *Generalized Poisson Regression* (GPR) dengan *exposure* secara serentak dan secara parsial
10. Pengujian heterogenitas spasial
11. Pembentukan matrik pembobot spasial menggunakan *Adaptive Kernel Bisquare*
12. Penentuan pembobot atau *bandwith* optimum
13. Melakukan analisis model *Geographically Weighted Generalized Poisson Regression* (GWGPR) dengan cara membentuk model dan membentuk penaksir parameter model tersebut
14. Menguji signifikansi parameter model *Geographically Weighted Generalized Poisson Regression* (GWGPR) dengan *exposure* secara serentak dan secara parsial
15. Membandingkan kebaikan antar model *Generalized Poisson Regression* (GPR) dan *Geographically Weighted Generalized Poisson Regression* (GWGPR) dengan *exposure* menggunakan nilai *Akaike's Information Criterion Correction* (AICc)
16. Mengelompokkan berbagai wilayah berdasarkan variabel-variabel yang signifikan
17. Melakukan pewarnaan pada peta kabupaten/kota di Jawa Timur berdasarkan variabel-variabel yang signifikan
18. Menarik kesimpulan dan saran

IV. ANALISIS DAN PEMBAHASAN

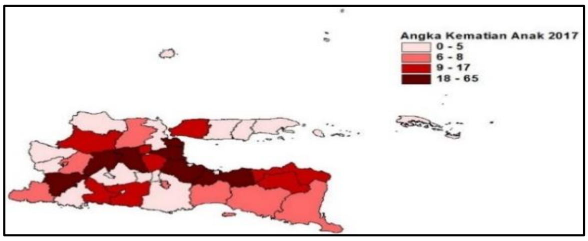
A. *Karakteristik Data Jumlah Kematian Anak*

Karakteristik jumlah kematian anak di Jawa Timur pada tahun 2017 beserta faktor-faktor yang mempengaruhinya dapat ditinjau menggunakan hasil analisis statistika deskriptif yang ditabulasikan sebagai berikut.

Berdasarkan Tabel 2 tersebut jumlah kematian anak di Jawa Timur tahun 2017 rata-rata di setiap kabupaten/kota sebesar 10,92 atau sekitar 11 kejadian. Jumlah kasus tertinggi yaitu pada Kota Sidoarjo sejumlah 65 kasus, sementara jumlah kasus terendah terjadi di Kota Blitar dan Kabupaten Pamekasan sejumlah 0 kasus. Persebaran jumlah kematian anak di Jawa Timur berdasarkan pengelompokan wilayah disajikan dalam Gambar 1.

Tabel 2.
Statistika Deskriptif Variabel Penelitian

Variabel	Mean	Varians	Min	Max
Y	10,92	218,301	0	65
X_1	54,0061	230,350	24,22	100,00
X_2	70,5258	269,010	25,08	96,84
X_3	0,8116	0,044	0,36	1,30
X_4	80,0713	298,040	25,15	100,16
X_5	80,0653	473,188	14,29	100,00
X_6	78,6834	367,442	17,00	100,00
X_7	11,6255	22,278	4,17	23,56



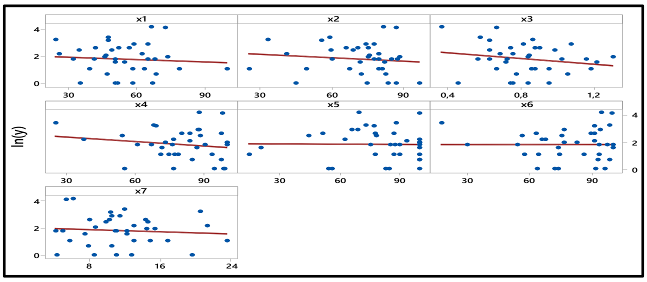
Gambar 1. Peta Persebaran Jumlah Kematian Anak di Jawa Timur Tahun 2017

Gambar 1 menunjukkan persebaran jumlah kematian anak pada tahun 2017 terbesar berada di wilayah tapal kuda (terutama Kabupaten Probolinggo, dan Kabupaten Pasuruan), bagian pusat (terutama Kota Surabaya, Kabupaten Sidoarjo, dan Kabupaten Jombang), dan bagian barat laut (terutama Kabupaten Nganjuk) wilayah Jawa Timur. Terdapat 6 kabupaten/kota dengan jumlah kematian anak diatas kuartil 3 (sekitar 18-65 jiwa) dan 16 kabupaten/kota dengan jumlah kematian anak di bawah kuartil 1 (sekitar 0-5 jiwa).

B. *Identifikasi Hubungan Antar Variabel Penelitian*

Hubungan antara variabel dapat diketahui dengan meninjau visualisasi persebaran titik-titik data yang digunakan dengan menggunakan *scatterplot*. *Scatterplot* tersebut bisa menunjukkan karakter hubungan antar variabel penelitian, yaitu hubungan positif, negatif, ataupun tidak memiliki hubungan. Berikut *scatterplot* antara variabel respon (jumlah kematian anak pada kabupaten/kota di Jawa Timur) dengan berbagai variabel prediktor (faktor-faktor yang diduga mempengaruhi).

Berdasarkan *scatterplot* pada Gambar 2 diketahui memiliki kecenderungan data yang menyebar acak secara luas dan tidak berpola, sehingga secara visual kecenderungan hubungan antara variabel respon dengan masing-masing variabel prediktor (sejumlah tujuh variabel) yaitu hubungan yang lemah. Untuk mengetahui arah dan kekuatan hubungan antara variabel respon terhadap variabel prediktor, bisa ditinjau berdasarkan nilai korelasi Pearson yang dihasilkan, pada Tabel 3.



Gambar 2. *Scatterplot* Variabel Respon terhadap Berbagai Variabel Prediktor

Tabel 3.
Matrik Nilai Korelasi Variabel Respon Terhadap Variabel Prediktor

Variabel	X_1	X_2	X_3	X_4
Y	0,080	0,040	-0,344	-0,058
	0,633	0,809	0,034	0,729
Variabel	X_5	X_6	X_7	
Y	-0,022	0,109	-0,206	
	0,896	0,517	0,215	

Berdasarkan Tabel 3 hubungan antara jumlah kematian anak pada kabupaten/kota di Jawa Timur dengan tujuh variabel prediktor secara umum adalah lemah dikarenakan nilai korelasi yang dominan kecil (mendekati nilai 0). Variabel persentase rumah tangga ber-PHBS (X_1), persentase rumah sehat (X_2), dan penduduk dengan akses berkelanjutan terhadap air minum berkualitas (layak) (X_6) memiliki kecenderungan hubungan (korelasi) positif terhadap jumlah kematian anak pada kabupaten/kota di Jawa Timur. Sementara variabel rasio puskesmas per 30.000 penduduk (X_3), persentase penduduk dengan akses terhadap fasilitas sanitasi yang layak (jamban sehat) (X_4), persentase desa/kelurahan yang melaksanakan sanitasi total berbasis masyarakat (X_5), dan persentase penduduk miskin (X_7) memilki kecenderungan hubungan (korelasi) yang negatif dengan jumlah kematian anak pada ka-bupaten/kota di Jawa Timur. Variabel rasio puskesmas per 30.000 penduduk (X_3) memiliki hubungan yang signifikan dikarenakan nilai *p-value* dibawah nilai taraf signifikansi 0,05. Arah hubungan-nya adalah korelasi negatif maka dapat dikatakan peningkatan nilai rasio puskesmas per 30.000 penduduk berbanding terbalik dengan peningkatan jumlah kematian anak.

C. Pengujian Multikolinieritas

Pemeriksaan kasus multikolinieritas untuk variabel prediktor antar tahunnya ditinjau berdasarkan nilai VIF yang bersyarat dibawah angka 10, dengan tabulasi sebagai berikut.

Tabel 4.
Nilai VIF Variabel Prediktor

Respon	Prediktor	VIF
X_1	X_2,X_3,X_4,X_5,X_6,X_7	1,39
X_2	X_1,X_3,X_4,X_5,X_6,X_7	2,48
X_3	X_1,X_2,X_4,X_5,X_6,X_7	1,15
X_4	X_1,X_2,X_3,X_5,X_6,X_7	3,08
X_5	X_1,X_2,X_3,X_4,X_6,X_7	1,23
X_6	X_1,X_2,X_3,X_4,X_5,X_7	2,35
X_7	X_1,X_2,X_3,X_4,X_5,X_6	2,21

Hasil pemeriksaan multikolinieritas masing-masing variabel prediktor diketahui semua memiliki nilai VIF yang kurang dari 10 sehingga tidak terjadi kasus multikolinieritas dan bisa digunakan dalam analisis pemodelan selanjutnya.

D. Pemodelan Jumlah Kematian Anak di Jawa Timur Menggunakan Generalized Poisson Regression Tanpa Exposure

Pengujian parameter model secara serentak diperlukan untuk mengetahui adanya minimal satu variabel prediktor dalam model yang berpengaruh signifikan terhadap variabel respon dengan hipotesis sebagai berikut.

$H_0 : \beta_1 = \beta_2 = \ldots = \beta_j = 0$

$H_1 :$ minimal ada satu $\beta_j \neq 0$, $j = 1,2,\ldots,k$

Statistik uji berupa nilai devians $D(\hat{\boldsymbol{\beta}})$ sebesar 248,5 yang kemudian dibandingkan dengan nilai kritis $\chi^2_{(0,05,7)}$ sebesar 14,067. Didapatkan keputusan untuk tolak H_0 yang bermakna seluruh parameter secara serentak mempunyai pengaruh dalam model. Pengujian parameter model secara parsial dilakukan untuk memilih parameter yang berpengaruh signifikan terhadap variabel respon dengan hipotesis sebagai berikut.

$H_0 : \beta_j = 0$

$H_1 : \beta_j \neq 0$, $j = 1,2,\ldots,7$

Statistik uji berupa nilai $|Z|$ setiap parameter variabel prediktor yang dibandingkan dengan nilai kritis $Z_{(\alpha/2)} = Z_{(0,025)}$ sebesar 1,96. Bisa pula ditinjau dengan menggunakan nilai *p-value* yang dibandingkan dengan taraf signifikansi 0,05 atau 5%. Diketahui bahwa variabel atau faktor yang berpengaruh secara signifikan terhadap jumlah kematian anak pada tahun 2017 yaitu persentase penduduk dengan akses terhadap fasilitas sanitasi yang layak (jamban sehat) (X_4) dan penduduk dengan akses berkelanjutan terhadap air minum berkualitas (layak) (X_6). Sehingga persamaan model yang diperoleh adalah sebagai berikut.

$$\hat{\mu} = \exp\left[5,0208 + 0,02493X_1 + 0,003335X_2 - 0,5157X_3\right]$$
$$\times \exp\left[-0,05182X_4 - 0,01952X_5 + 0,03164X_6 - 0,05566X_7\right]$$

Nilai devians $D(\hat{\boldsymbol{\beta}})$ juga dipakai sebagai pemeriksaan overdispersi dengan cara melakukan pembagian $D(\hat{\boldsymbol{\beta}})$ dengan nilai derajat bebas $n-k-1$ sebesar 30. Dihasilkan nilai sebesar 8,283 yang bernilai lebih dari satu. Sehingga data jumlah kematian anak di Jawa Timur pada tahun 2017 terdapat kasus overdispersi sehingga diperlukan pengolahan dengan metode lain.

E. Pemodelan Jumlah Kematian Anak di Jawa Timur Menggunakan Generalized Poisson Regression Dengan Exposure

Pengujian parameter model secara serentak diperlukan untuk mengetahui adanya minimal satu variabel prediktor dalam model yang berpengaruh signifikan terhadap variabel respon dengan hipotesis sebagai berikut.

$H_0 : \beta_1 = \beta_2 = \ldots = \beta_j = 0$

$H_1 :$ minimal ada satu $\beta_j \neq 0$, $j = 1,2,\ldots,7$

Statistik uji berupa nilai devians $D(\hat{\boldsymbol{\beta}})$ sebesar 232,1 yang kemudian dibandingkan dengan nilai kritis $\chi^2_{(0,05,7)}$ sebesar 14,067. Didapatkan keputusan untuk tolak H_0 yang bermakna seluruh parameter secara serentak mempunyai pengaruh dalam model. Pengujian parameter model secara parsial dilakukan untuk memilih parameter yang berpengaruh signifikan terhadap variabel respon dengan hipotesis sebagai berikut.

$H_0 : \beta_j = 0$

$H_1 : \beta_j \neq 0$, $j = 1, 2, \ldots, k$

Statistik uji berupa nilai $|z|$ setiap parameter variabel prediktor yang dibandingkan dengan nilai kritis $Z_{(\alpha/2)} = Z_{(0,025)}$ sebesar 1,96. Bisa pula ditinjau dengan menggunakan nilai *p-value* yang dibandingkan dengan taraf signifikansi 0,05 atau 5%. Diketahui bahwa variabel atau faktor yang berpengaruh secara signifikan terhadap jumlah kematian anak pada tahun 2017 yaitu persentase penduduk dengan akses terhadap fasilitas sanitasi yang layak (jamban sehat) (X_4), dikarenakan variabel tersebut memiliki nilai *p-value* kurang dari taraf signifikansi sebesar 0,05. Sehingga persamaan model yang diperoleh adalah sebagai berikut.

$$\hat{\mu} = \exp\left[-8,1232 + 0,001432X_1 + 0,008092X_2 + 1,0588X_3\right]$$
$$\times \exp\left[-0,03673X_4 - 0,0045X_5 + 0,02156X_6 - 0,06183X_7\right]$$

Nilai devians $D(\hat{\boldsymbol{\beta}})$ juga dipakai sebagai pemeriksaan overdispersi dengan cara melakukan pembagian $D(\hat{\boldsymbol{\beta}})$ dengan nilai derajat bebas $n-k-1$ sebesar 30. Dihasilkan nilai sebesar 7,73667 yang bernilai lebih dari satu. Sehingga data jumlah kematian anak di Jawa Timur pada tahun 2017 terdapat kasus overdispersi sehingga diperlukan pengolahan dengan metode lain.

F. Pengujian Heterogenitas Spasial Pemodelan Jumlah Kematian Anak di Jawa Timur

Pengujian heterogenitas spasial dilakukan untuk mengetahui apakah terdapat kekhasan pada setiap lokasi pengamatan. Tinjauannya dengan menggunakan statistik uji *Breusch-Pagan* dengan hipotesis sebagai berikut.

$H_0 : \sigma_1^2 = \sigma_2^2 = \ldots = \sigma_n^2 = \sigma^2$ (*Varians* antar lokasi sama)

H_1 : minimal ada satu $\sigma_i^2 \neq \sigma^2; i = 1, 2, ..., k$ (*Varians* antar lokasi berbeda)

Nilai statistik uji *Breusch-Pagan* untuk pemodelan jum-lah kematian anak pada tahun 2017 baik tanpa *exposure* mau-pun menggunakan *exposure* didapatkan masing-masing sebe-sar 16,276 dan 14,268. Kemudian dibandingkan dengan nilai kritis $\chi^2_{(\alpha,k)} = \chi^2_{(0,05,7)}$ sebesar 14,067, masing-masing diperoleh keputusan tolak H_0 yang bermakna *varians* antar lokasi berbeda. Bisa disimpulkan bahwa terdapat perbedaan karakteristik pada suatu kejadian di setiap lokasi pengamatan (kabupaten/kota di Jawa Timur), baik data yang tanpa melibatkan *exposure* maupun yang melibatkan *exposure*.

G.Pemodelan Jumlah Kematian Anak di Jawa Timur Menggunakan Geographically Weighted Generalized Poisson Regression Tanpa Exposure

Pengujian parameter model secara serentak diperlukan untuk mengetahui adanya minimal satu variabel prediktor dalam model yang berpengaruh signifikan terhadap variabel respon dengan hipotesis sebagai berikut.

$H_0 : \beta_1(u_i, v_i) = \beta_2(u_i, v_i) = \ldots = \beta_k(u_i, v_i) = 0$

H_1 : minimal ada satu $\beta_j(u_i, v_i) \neq 0$; $i = 1, 2, \ldots, 38; j = 1, 2, \ldots, 7$

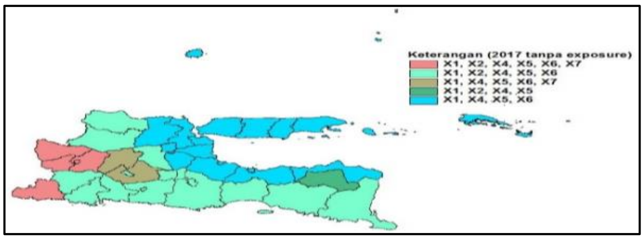
Statistik uji berupa nilai devians $D(\hat{\boldsymbol{\beta}}_j(u_i, v_i))$ sebesar -15882,41 yang dibandingkan dengan nilai kritis $\chi^2_{(0,05,7)}$ sebesar 14,067. Didapatkan keputusan untuk gagal tolak H_0 yang berarti seluruh parameter secara serentak tidak mempunyai pengaruh dalam model.

Pengujian parameter model secara parsial dilakukan untuk memilih parameter yang berpengaruh signifikan terhadap variabel respon dengan hipotesis sebagai berikut.

$H_0 : \beta_j(u_i, v_i) = 0$

$H_0 : \beta_j(u_i, v_i) \neq 0$

Statistik uji berupa nilai $|z|$ setiap parameter variabel prediktor yang dibandingkan dengan nilai kritis $Z_{(\alpha/2)} = Z_{(0,025)}$ sebesar 1,96. Pemodelan jumlah kematian anak pada tahun 2017 diketahui variabel persentase rumah tangga ber-PHBS (X_1), persentase rumah sehat (X_2), persentase penduduk dengan akses terhadap fasilitas sanitasi yang layak (jamban sehat) (X_4), persentase desa/kelurahan yang melaksanakan sanitasi total berbasis masyarakat (X_5), penduduk dengan akses berkelanjutan terhadap air minum berkualitas (layak) (X_6), dan persentase penduduk miskin (X_7) merupakan faktor yang dominan berpengaruh secara signifikan terhadap masing-masing kabupaten/kota. Pemetaan pengelompokan kabupaten/kota berdasarkan variabel yang signifikan dapat digam-barkan sebagai berikut.



Gambar 3. Peta Pengelompokan Untuk Jumlah Kematian Anak Menggunakan Metode GWGPR Tanpa *Exposure*.

Ditunjukkan pemodelan metode GWGPR tanpa *exposure* pada lokasi ke-37 yaitu Kota Surabaya diketahui sebagai berikut.

$$\hat{\mu} = \exp\left[4,103978 + 0,041932X_1 - 0,03906X_2 - 0,36571X_3\right]$$
$$\times \exp\left[-0,00606X_4 - 0,00426X_5 - 0,00528X_6 - 0,00832X_7\right]$$

Model tersebut bisa diinterpretasikan salah satunya setiap penambahan 10% persentase rumah tangga ber-PHBS maka akan meningkatkan jumlah kematian anak di Kota Surabaya sebesar $\exp[0.041932 * 10] = 0,41$ kali dengan asumsi variabel konstan.

H.Pemodelan Jumlah Kematian Anak di Jawa Timur Menggunakan Geographically Weighted Generalized Poisson Regression Dengan Exposure

Pengujian parameter model secara serentak diperlukan untuk mengetahui adanya minimal satu variabel prediktor dalam model yang berpengaruh signifikan terhadap variabel respon dengan hipotesis sebagai berikut.

$H_0 : \beta_1(u_i, v_i) = \beta_2(u_i, v_i) = \ldots = \beta_k(u_i, v_i) = 0$; $i = 1, 2, \ldots, n$

H_1 : minimal ada satu $\beta_j(u_i, v_i) \neq 0$; $i = 1, 2, \ldots, 38; j = 1, 2, \ldots, 7$

Statistik uji berupa nilai devians $D\left(\hat{\boldsymbol{\beta}}_j\left(u_i,v_i\right)\right)$ sebesar 23,3316 yang dibandingkan dengan nilai kritis $\chi^2_{(0,05,7)}$ sebesar 14,067. Didapatkan keputusan untuk tolak H_o yang berarti seluruh parameter secara serentak mempunyai pengaruh dalam model. Pengujian parameter model secara parsial dilakukan untuk memilih parameter yang berpengaruh signifikan terhadap variabel respon dengan hipotesis sebagai berikut.

$H_0 : \beta_j\left(u_i,v_i\right)=0$

$H_1 : \beta_j\left(u_i,v_i\right)\neq 0$

Statistik uji berupa nilai $|Z|$ setiap parameter variabel prediktor yang dibandingkan dengan nilai kritis $Z_{(\alpha/2)}=Z_{(0,025)}$ sebesar 1,96. Pemodelan jumlah kematian anak pada tahun 2017 diketahui tidak ada faktor yang berpengaruh secara signifikan terhadap masing-masing kabupaten/kota. Ditunjukkan pemodelan metode GWGPR dengan *exposure* pada lokasi ke-21 yaitu Kabupaten Ngawi diketahui sebagai berikut.

$$\hat{\mu}=\exp\left[7,87\cdot 10^{-5}+0,000936X_1+0,001922X_2+0,000744X_3\right]$$
$$\times\exp\left[0,000422X_4+0,00021X_5-0,00049X_6+0,003775X_7\right]$$

Model tersebut bisa diinterpretasikan salah satunya setiap penambahan 10% persentase rumah tangga ber-PHBS maka jumlah kematian anak di Kabupaten Ngawi meningkat sebesar $\exp\left[0,000936*10\right]=0,009$ kali dengan asumsi variabel konstan. Setiap penambahan 10% faktor penduduk dengan akses berkelanjutan terhadap air minum berkualitas (layak) maka akan menurunkan jumlah kematian anak di Kabupaten Ngawi dalam jumlah $\exp\left[0,00049*10\right]=0,0049$ kali dengan asumsi variabel konstan.

I. Pemilihan Model Terbaik

Kriteria AICc mampu mengetahui kesesuaian model secara statistik, dengan tujuan mengidentifikasi faktor yang berpengaruh signifikan terhadap model. Perhitungan AICc untuk masing-masing pemodelan jumlah kematian anak pada tahun 2017 adalah sebagai berikut.

Tabel 5.
Nilai AICc Setiap Metode Pemodelan

Metode	Nilai AICc
Generalized Poisson Regression tanpa *Exposure*	266,5
Generalized Poisson Regression dengan *Exposure*	256,6
Geographically Weighted Generalized Poisson Regression tanpa *Exposure*	16310,2
Geographically Weighted Generalized Poisson Regression dengan *Exposure*	752,7

Nilai AICc untuk metode *Generalized Poisson Regres-sion* dengan *exposure* secara umum memiliki nilai AICc ter-kecil dibandingkan metode lainnya sehingga metode tersebut diutamakan dalam hal menganalisis jumlah kematian anak di Jawa Timur.

V. KESIMPULAN DAN SARAN

A. Kesimpulan

Berdasarkan analisis dan pembahasan yang telah dilakukan pada bab IV, maka diperoleh kesimpulan sebagai berikut. Jumlah kematian anak di Jawa Timur tahun 2017 memiliki persebaran terbesar berada di wilayah tapal kuda, bagian pusat, dan bagian barat laut Jawa Timur. Faktor-faktor yang diduga mempengaruhi jumlah kematian anak di Jawa Timur memiliki karakteristik persebaran yang berbeda-beda antar faktor, namun dibandingkan antar tahunnya tidak berbeda secara signifikan. Hasil penaksiran parameter pemodelan jumlah kematian anak menggunakan *Generalized Poisson Regression* baik yang menggunakan atau yang tanpa menggunakan *exposure* seluruh parameter secara serentak mempunyai pengaruh dalam model, namun secara parsial hanya beberapa variabel saja yang terpilih dan semuanya mengalami kasus overdispersi. Pemodelan GWGPR tanpa menggunakan *exposure* diketahui variabel yang dominan signifikan terhadap masing-masing kabupaten/kota yaitu variabel persentase rumah tangga ber-PHBS, persentase rumah sehat, persentase penduduk dengan akses terhadap fasilitas sanitasi yang layak (jamban sehat), persentase desa/kelurahan yang melaksanakan sanitasi total berbasis masyarakat, penduduk dengan akses berkelanjutan terhadap air minum berkualitas (layak), dan persentase penduduk miskin, yang terbagi dalam lima kelompok. Sementara pemodelan GWGPR menggunakan *exposure* diketahui tidak ada faktor atau variabel yang berpengaruh secara signifikan terhadap masing-masing kabupaten/kota. Kriteria kesesuaian model menggu-nakan AICc menunjukkan bahwa metode *Generalized Pois-son Regression* secara umum memiliki nilai AICc terkecil dibandingkan metode lainnya.

B. Saran

Saran untuk penelitian selanjutnya yaitu menambah atau mengganti faktor terkait sebagai variabel prediktor dalam penelitian yang sesuai dengan kerangka berpikir kematian anak secara umum agar model yang dihasilkan lebih baik dan bisa tersebar merata signifikansi variabel prediktornya pada berbagai wilayah kabupaten/kota di Jawa Timur.

DAFTAR PUSTAKA

[1] United Nations, *Transforming Our World: The 2030 Agenda for Sustainable*. New York: United Nations, 2015.

[2] United Nations Development Programme, “A Social Determinants Approach to Maternal Health, Discussion Paper,” New York, 2011.

[3] M. Seipel, “Child Mortality: A Preventable Tragedy,” *J. Multicult. Soc. Work*, vol. 4, no. 1, pp. 63–74, 1996.

[4] R. Destyanugraha and R. Kurniawan, “Pemodelan Angka Kematian Ibu di Indonesia dengan Pendekatan Geographically Weighted Poisson Regression,” *J. Mat. Sains, dan Teknol.*, vol. 18, no. 2, pp. 77–95, 2017.

[5] D. Setiawan, Purhadi, and Sutikno., “Parameter Estimation of Geographically Weighted Bivariate Generalized Poisson Regression Model,” in *International Conference on Theoretical and Applied Statistics*, 2016.

[6] A. C. Cameron and P. K. Trivedi, *Regression Analysis of Count Data*, 2nd ed. USA: Cambridge University Press, 2013.

[7] R. Hocking, *Method and Applications of Linear Models*. New York: John Wiley & Sons, Inc, 1996.

[8] J. Hilbe, *Modelling Count Data*. New York: Cambridge University Press, 2014.

[9] N. Ismail and A. Jemain, “Handling Overdispersion with Negative Binomial and Generalized Poisson Regression Models,” Virginia, 2007.

[10] A. Melliana, Y. Setyorini, H. Eko, S. Rosi, and Purhadi., “The Comparison Of Generalized Poisson Regression and Negative Binomial Regression Methods In Overcoming Overdispersion,” *Int. J. Sci. Technol.*, vol. 8, no. 2, pp. 255–258, 2013.

[11] R. Caraka and H. Yasin, *Geographically Weighted Regreession (GWR) Sebuah Pendekatan Regresi Geografis*. Yogyakarta: MOBIUS, 2017.

[12] T. Nakaya, A. S. Fotheringham, C. Brudson, and M. Charlton, “Geographically Weighted Poisson Regression for Disease Association Mapping,” *Stat. Med.*, pp. 2695–2717, 2005.

[13] A. McQuarrie and C. Tsai, *Regression and Time Series Model Selection*. Singapore: World Scientific Publishing, 1998.

[14] United Nations, *Step-By-Step Guide To The Estimation Of Child Mortality*. New York: United Nations, 1990.