



Pemodelan Regresi Binomial Negatif Spasial (Trend Surface) untuk Faktor-Faktor Determinasi Rate Kasus Pneumonia pada Balita di Kabupaten/Kota se-Jawa Barat

Sarah Christ She

Universitas Pertahanan Republik Indonesia

Alamat : Kawasan Indonesia Peace and Security Center (IPSC) Sentul, Sukahati, Kecamatan Citeureup, Kabupaten Bogor, Jawa Barat 16810, Indonesia

Abstract. *Pneumonia in toddlers is a significant public health challenge requiring accurate spatial modeling due to its heterogeneous geographical distribution and transmission characteristics. This study aims to identify the optimal spatial model and the determinants of the Pneumonia Case Rate in Toddlers across 27 Regencies/Cities in West Java Province in 2024, utilizing aggregated case and socioeconomic data. Modeling was performed using Negative Binomial Regression (NBR) combined with Trend Surface Analysis (TSA) to address overdispersion and macro spatial effects, with the Parsimonious NBR-TSA Model being selected as the best-fit model. Data analysis indicates that Complete Immunization Coverage has a significant relationship as a protective factor against the number of Pneumonia cases, while Percentage of Poor Population is identified as a significant risk factor. Other factors, such as the percentage of household smokers and stunting prevalence, displayed statistically significant but theoretically contradictory negative coefficients, which is debated as a result of an Ecological Fallacy inherent in aggregated data, potentially indicating under-reporting of cases in chronically socioeconomically disadvantaged areas. The NBR-TSA model was chosen for its optimal Best-Fit score and its success in handling spatial dependency. These findings are crucial for analyzing high-risk areas and developing strategies focused on improving immunization coverage and poverty alleviation in health policy formulation.*

Keywords: *Ecological Fallacy, Negative Binomial Regression, Spatial Analysis, Toddler Pneumonia, Trend Surface,*

Abstrak. *Pneumonia pada balita merupakan tantangan kesehatan publik yang signifikan, memerlukan pemodelan spasial yang akurat karena sifat penularan dan distribusi geografis yang heterogen. Penelitian ini bertujuan mengidentifikasi model spasial optimal dan faktor-faktor determinan Rate Kasus Pneumonia Balita di 27 Kabupaten/Kota Provinsi Jawa Barat tahun 2024, menggunakan data kasus Pneumonia dan data sosio-ekonomi agregat. Pemodelan dilakukan menggunakan Regresi Binomial Negatif (RBN) yang dikombinasikan dengan Trend Surface Analysis (TSA) untuk mengatasi overdispersion dan efek spasial makro. Analisis data menunjukkan bahwa Cakupan*

Received Januari 16, 2026; Revised Januari 17, 2026; Accepted Januari 18, 2026

**Sarah Christ She, christshe23@gmail.com*

Imunisasi Lengkap memiliki hubungan signifikan sebagai faktor protektif terhadap jumlah kasus Pneumonia, sementara Persentase Penduduk Miskin merupakan faktor risiko signifikan. Faktor lain, seperti persentase perokok di rumah tangga dan persentase stunting, menunjukkan koefisien anomali yang signifikan secara statistik tetapi secara teori bertentangan, yang diperdebatkan sebagai hasil dari Ecological Fallacy yang mengindikasikan under-reporting kasus di wilayah kronis. Model RBN-TSA terpilih karena menunjukkan nilai Best-Fit Model yang optimal dan berhasil menangani dependensi spasial. Temuan ini penting dalam menganalisis daerah dan mengembangkan strategi yang berfokus pada peningkatan cakupan imunisasi dan penanggulangan kemiskinan dalam perumusan kebijakan kesehatan.

Kata kunci: Analisis Spasial., Ecological Fallacy, Pneumonia Balita, Regresi Binomial Negatif, Stunting., Trend Surface

LATAR BELAKANG

Pneumonia pada balita masih menjadi tantangan kesehatan masyarakat yang serius, menyumbang angka morbiditas dan mortalitas yang tinggi, terutama di negara berkembang seperti Indonesia (Wang et al., 2025; Purnama et al., 2024). Laporan terbaru mengindikasikan bahwa penyakit ini tetap menjadi penyebab utama kematian anak di bawah usia lima tahun, jauh melampaui penyakit menular lainnya (Wang et al., 2025). Kompleksitas penyebaran penyakit ini tidak hanya dipengaruhi oleh faktor biologis seperti status gizi (Purnama et al., 2024) dan imunisasi (Usman et al., 2025), tetapi juga oleh determinan eksternal yang saling terkait, seperti kondisi sosial-ekonomi, lingkungan rumah yang tidak sehat (misalnya paparan asap rokok dan polusi udara Balasubramani et al., 2022), serta akses terhadap layanan kesehatan yang tidak merata (Degif et al., 2025). Upaya pencegahan dan penanggulangan membutuhkan pemahaman yang mendalam mengenai kontribusi setiap faktor risiko ini, terutama bagaimana faktor-faktor tersebut tersebar secara geografis.

Ketidakmerataan geografis ini sering kali menciptakan kluster atau hotspot penyakit di wilayah-wilayah tertentu (Puput et al., 2025). Dalam konteks Indonesia, yang memiliki keragaman demografi dan topografi tinggi, model analisis yang tidak memperhitungkan lokasi (non-spasial) cenderung menghasilkan kesimpulan yang bias dan kurang efektif untuk perumusan kebijakan lokal (Kurniawati et al., 2025). Hal ini didukung oleh studi yang menunjukkan adanya kluster signifikan untuk Pneumonia di Jawa Barat (Puput et al., 2025) dan untuk stunting di berbagai provinsi (Arief et al., 2025). Oleh karena itu, diperlukan pendekatan analitis yang mampu menangkap struktur ketergantungan spasial

antar-wilayah (Degif et al., 2025), serta mengidentifikasi faktor-faktor penentu lokal yang mungkin tersembunyi dalam analisis agregat (Usman et al., 2022).

Penerapan metode regresi spasial telah menjadi semakin penting dalam studi epidemiologi untuk mengatasi masalah autokorelasi dan heterogenitas spasial. Model Regresi Binomial Negatif (RBN) menjadi model inferensi standar yang esensial karena mampu mengatasi overdispersi ($\sigma^2 > \mu$) yang lazim terjadi pada data kasus penyakit menular (Lucky et al., 2024). Sebagai contoh, penelitian mutakhir Yusdiana et al. (2025) mengenai pemodelan Pneumonia Balita di Rokan Hilir, Indonesia, secara sukses menggunakan Generalized Poisson Regression (GPR). Hasil studi tersebut mengidentifikasi faktor sosio-ekonomi dan lingkungan, seperti curah hujan dan status gizi buruk, sebagai risiko utama yang tersebar secara spasial. Studi tersebut menegaskan urgensi penggunaan model Count Data yang canggih (GPR) dan analisis kluster untuk penyakit di Indonesia. Namun, penelitian ini memiliki perbedaan mendasar dengan studi kami: sementara Yusdiana et al. (2025) menangkap efek spasial secara global melalui GPR, penelitian kami menawarkan kontribusi baru dengan membandingkan model-model RBN yang dikombinasikan dengan Trend Surface Model (TSM) untuk secara eksplisit mengestimasi tren keruangan langsung dari posisi geografis (Permana, 2021). Selain itu, model kami akan menguji model Robust NBR (RNBR) atau Zero-Inflated NBR (ZINB) untuk mengatasi anomali data seperti outlier atau excess zeros yang sering ditemukan pada data Pneumonia (Achmad et al., 2022; Qur'ani et al., 2024).

Berdasarkan latar belakang dan urgensi analisis spasial yang spesifik tersebut, penelitian ini bertujuan untuk mengidentifikasi dan membandingkan faktor-faktor determinan (meliputi status gizi, imunisasi, dan lingkungan) yang paling signifikan memengaruhi kasus Pneumonia Balita di Indonesia (Ananda et al., 2024) dengan mengaplikasikan berbagai model keluarga Regresi Binomial Negatif dengan Trend Surface Model (RBN-TSM), termasuk varian Robust NBR dan Zero-Inflated NBR, untuk menemukan model terbaik (dengan kriteria AIC terkecil). Setelah penetapan model RBN-TSM terbaik, penelitian ini akan dilanjutkan dengan memetakan wilayah hotspot penyakit dan faktor risikonya. Sistematika penulisan artikel ini terdiri dari lima bagian: Pendahuluan yang memaparkan latar belakang, tinjauan pustaka, dan tujuan; Metode Penelitian yang menjelaskan jenis data dan prosedur analisis; Hasil dan Pembahasan yang

menyajikan temuan pemodelan dan implikasi; serta Kesimpulan yang merangkum hasil utama dan rekomendasi kebijakan.

METODE PENELITIAN

A. Data dan Deskripsi Variabel

a. Jenis dan Sumber Data

Penelitian ini menggunakan pendekatan kuantitatif dengan desain studi cross-sectional dan memanfaatkan data sekunder tingkat agregat. Unit analisis dalam penelitian ini adalah Kabupaten/Kota di Indonesia.

Data diperoleh dari berbagai sumber resmi publikasi tahun terbaru (2021–2025), yang meliputi:

1. Badan Pusat Statistik (BPS) & Survei Kesehatan Indonesia (SKI): Menyediakan data demografi (kepadatan penduduk) dan status sosial-ekonomi (kemiskinan dan imunisasi).

2. Kementerian Kesehatan (Kemenkes) / Dinas Kesehatan Provinsi: Menyediakan data kasus penyakit dan status gizi balita.

3. Open Data Jawa Barat (Jabar): Sumber tambahan yang digunakan secara spesifik untuk memvalidasi data kasus Pneumonia dan Kepadatan Penduduk di tingkat Kabupaten/Kota Provinsi Jawa Barat.

4. Data Geospasial (Geospatial Data): Data koordinat geografis (lintang dan bujur) di tingkat Kabupaten/Kota. Data ini wajib digunakan untuk memvisualisasikan sebaran kasus (peta) dan menjalankan semua analisis regresi spasial (Permana, 2021).

b. Deskripsi Variabel Penelitian

Data yang digunakan merupakan data sekunder yang diperoleh dari Badan Pusat Statistik (BPS) Provinsi Jawa Barat dan Open Data Jawa Barat. Objek penelitian dalam studi ini adalah 27 Kabupaten/Kota di Provinsi Jawa Barat pada Tahun 2024. Selain itu, studi literatur mengenai Pneumonia Balita, analisis spasial, dan faktor-faktor determinannya juga digunakan dalam memperkaya data dan pemahaman teoritis. Dalam penelitian ini, digunakan lima variabel bebas (prediktor), yaitu Persentase Balita Gizi Buruk/Kurang (X1), Persentase Imunisasi Dasar Lengkap (X2), Persentase Rumah

Tangga Merokok (X3) , Kepadatan Penduduk (X4) , dan Persentase Penduduk Miskin (X5), dengan variabel terikat Jumlah Kasus Pneumonia Balita (Y). Tabel 1 merupakan daftar variabel yang digunakan dalam pemodelan.

Tabel 1. Deskripsi Variabel Penelitian Pemodelan Pneumonia Balita

Simbol	Variabel	Tipe Data	Satuan	Deskripsi
Y	Jumlah Kasus Pneumonia Balita	Count	Kasus	Jumlah total kasus Pneumonia Balita (usia 0-59 bulan) yang tercatat per Kabupaten/Kota.
X ₁	Persentase Balita Gizi Buruk/Kurang	Kontinu	%	Proporsi balita dengan status gizi kurang atau buruk ($Z\text{-score} \leq -2$).
X ₂	Persentase Imunisasi Dasar Lengkap	Kontinu	%	Proporsi balita yang telah menerima seluruh imunisasi dasar wajib.
X ₃	Persentase Rumah Tangga Merokok	Kontinu	%	Proporsi rumah tangga dengan perilaku merokok aktif di dalam rumah (polusi udara dalam ruangan).
X ₄	Kepadatan Penduduk	Kontinu	Jiwa/km ²	Jumlah rata-rata penduduk per kilometer persegi (proksi risiko penularan).
X ₅	Persentase Penduduk Miskin	Kontinu	%	Proporsi penduduk yang berada di bawah garis kemiskinan (proksi faktor sosio-ekonomi).

B. Metode Analisis Regresi Spasial

Pemodelan yang diterapkan adalah Regresi Binomial Negatif Trend Surface (RBN-TSA), yang dipilih karena dua tantangan data utama: overdispersion dan ketergantungan spasial.

1. Teori Dasar Model Inferensi (Count Data Models)

Mengingat variabel terikat (Y Jumlah Kasus Pneumonia Balita) merupakan data hitung (Count Data), model inferensi utama yang digunakan adalah Regresi Binomial Negatif (RBN). Model ini wajib digunakan karena secara inheren mampu mengatasi Overdispersi ($\sigma^2 > \mu$) yang hampir selalu terjadi pada data kasus penyakit menular, dengan memasukkan parameter dispersi (α) ke dalam fungsi variansnya (Lucky et al., 2024; Qur'ani et al., 2024).

2. Regresi Binomial Negatif

Model Regresi Binomial Negatif (RBN) digunakan untuk menganalisis data hitung, di mana variabel respons diasumsikan berdistribusi Binomial Negatif. RBN dipilih karena data kasus Pneumonia adalah count data yang diasumsikan mengalami overdispersion (varian > rata-rata), suatu kondisi yang melanggar asumsi Regresi Poisson. RBN memodelkan logaritma nilai yang diharapkan dari variabel respons () sebagai fungsi linear dari prediktor menggunakan fungsi log- link:

$$\ln(\mu_i) = \beta_0 + \sum_{k=1}^p \beta_k X_{ik} + \ln(N_i)$$

Dengan $\ln(N_i)$ adalah offset (logaritma populasi berisiko). Koefisien β diestimasi menggunakan metode Maximum Likelihood Estimation (MLE).

3. Trend Surface Analysis (TSA)

Untuk menangkap heterogenitas spasial regional, RBN diperluas dengan Trend Surface Analysis (TSA). TSA adalah teknik eksplorasi yang mengintegrasikan koordinat lokasi terstandarisasi (X_2 dan Y_2) sebagai prediktor spasial ke dalam model RBN. TSA memungkinkan pendeteksian pola risiko yang sistematis di sepanjang sumbu geografis (gradien Utara-Selatan), yang penting untuk analisis ekologis.

Model RBN-TSA yang digunakan adalah:

$$\ln(\mu_i) = \beta_0 + \sum_{k=1}^p \beta_k X_{ik} + \beta_x X_z + \beta_y Y_z + \ln(N_i)$$

C. Algoritma Penelitian untuk Model Terbaok

Algoritma penelitian ini berfokus pada pendekatan parsimoni (memilih model dengan jumlah prediktor minimum yang tetap mempertahankan kekuatan prediksi yang tinggi), yang diukur melalui Akaike's Information Criterion (AIC).

1) Algoritma Penelitian: Pemodelan RBN-TSA Kasus Pneumonia Balita

Algoritma penelitian ini berfokus pada pendekatan parsimoni dan perbandingan model berdasarkan kriteria AIC (Akaike's Information Criterion).

Analisis Eksplorasi Awal:

- a. Menganalisis korelasi antar variabel.
- b. Mendeteksi kasus multikolinearitas pada setiap variabel menggunakan nilai VIF (Variance Inflation Factor). Variabel yang menyebabkan $VIF > 5$ akan dikeluarkan dari model.
- c. Memodelkan data Jumlah Kasus Pneumonia Balita menggunakan Distribusi Poisson sebagai benchmark teoritis, kemudian menguji overdispersi.

2) Pemodelan Parsimoni dan Seleksi Model Terbaik:

Proses seleksi model dilakukan secara bertahap, dimulai dari Model Global Terbaik (tanpa efek spasial), kemudian diperluas menjadi Model Spasial Penuh dengan menyertakan semua komponen Trend Surface Analysis (TSA) penuh. Langkah terakhir adalah Pemodelan Parsimoni, di mana prediktor-prediktor non-signifikan dieliminasi dari Model Spasial Penuh untuk menghasilkan model yang lebih sederhana dan kuat.

Tabel 2. Pemodelan Parsimoni dan Seleksi Model Terbaik

Model	Deskripsi	Kriteria AIC yang Ditetapkan
Model Global Terbaik	Model RBN yang hanya menggunakan prediktor non-spasial (X_1 hingga X_5).	AIC: 470.63
Model Spasial Penuh	Model Global Terbaik diperluas dengan prediktor spasial TSA penuh (UV, U^2, V^2, UV).	AIC: 469.98
Model RBN-TSA Parsimoni (Model H)	Model yang dihasilkan dengan menghilangkan prediktor non-signifikan dari Model Spasial Penuh.	AIC: 466.64 (Terbaik)

Model Parsimoni Final (Model H) dengan nilai AIC: 466.64 dipilih sebagai model terbaik. Pemilihan ini didasarkan pada prinsip parsimoni, di mana model tersebut menghasilkan nilai AIC terendah. Nilai AIC terendah mengindikasikan keseimbangan terbaik antara goodness-of-fit (seberapa baik model menjelaskan data) dan kompleksitas (jumlah variabel yang digunakan), menjadikannya model yang paling efisien dan optimal untuk memodelkan kasus Pneumonia Balita.

3) Validasi dan Visualisasi Hasil:

- a. Lakukan Uji Moran's I pada residual model terbaik untuk memverifikasi tidak adanya autokorelasi spasial yang tersisa (p-value 0.73), memvalidasi kecukupan model dalam menangani dependensi spasial.
- b. Melakukan pemetaan faktor-faktor signifikan dan residual model terbaik.
- c. Interpretasi Temuan Kritis: Hasil akan menjelaskan Anomali Koefisien (misalnya, Persentase Perokok ($\beta = -0.479$) dan Persentase Stunting ($\beta = -0.374$) yang berhubungan negatif) sebagai potensi Ecological Fallacy akibat data agregat dan kemungkinan under-reporting.

Data yang digunakan pada penelitian ini diolah menggunakan perangkat lunak statistik R Studio, dengan memanfaatkan paket-paket analisis data hitung dan spasial.

Berikut langkah-langkah dalam penelitian ini:

- a. Mengumpulkan data yang relevan, termasuk variabel dependen (Jumlah Kasus Pneumonia Balita) dan variabel independen non-spasial (misalnya, Persentase Stunting, Persentase Perokok, Cakupan Imunisasi, Persentase Penduduk Miskin) serta variabel independen spasial (Koordinat Longitude Terstandarisasi (X_z) dan Latitude Terstandarisasi (Y_z)).
- b. Menganalisis korelasi antara variabel-variabel untuk mengetahui hubungan antar variabel.
- c. Mendeteksi kasus multikolinearitas pada setiap variabel menggunakan nilai VIF (Variance Inflation Factor) sebagai kriteria¹. Variabel yang menyebabkan $VIF > 5$ akan dikeluarkan dari model.
- d. Memodelkan data Jumlah Kasus Pneumonia Balita menggunakan Distribusi Poisson sebagai benchmark teoritis.
- e. Menguji apakah data mengalami overdispersi atau underdispersi dengan membandingkan nilai Rata-rata (Mean) dan nilai Variansnya.
- f. Memodelkan data Jumlah Kasus Pneumonia Balita menggunakan Distribusi Binomial Negatif (RBN) untuk mengatasi kondisi overdispersion.
- g. Menyertakan efek spasial dalam model RBN menggunakan Trend Surface Analysis (TSA) dengan memasukkan X_z dan Y_z sebagai prediktor, sehingga terbentuk model RBN-TSA.

- h. Membandingkan model (RBN Global, RBN-TSA Penuh, dan RBN-TSA Parsimoni) dengan melihat nilai AIC (Akaike's Information Criterion) dan memilih model terbaik (Model H) dengan melihat nilai AIC terkecil.
- i. Menjelaskan hasil dari model terpilih (Model H), termasuk hubungan signifikan antara variabel dependen dan variabel independen (misalnya, koefisien Stunting_z dan Perokok_z yang anomaly) dan tren spasial yang terdeteksi.

HASIL DAN PEMBAHASAN

A. Analisis Deskriptif Data dan Uji Pra-Pemodelan

1. Statistik Deskriptif Variabel

Analisis deskriptif memberikan gambaran umum mengenai variabilitas data pada 27 Kabupaten/Kota di Provinsi Jawa Barat.

Tabel 3. Statistik Deskriptif Variabel Penelitian

Variabel	Satuan	Minimum	Maksimum	Rata-rata (μ)	Standar Deviasi (σ)
Jumlah Kasus (Y)	Kasus	1.593.00	26.250.00	9.021.46	5.560.80
Total Populasi (N_i)	Jiwa	25.442.00	2.777.060.00	277.139.79	556.278.86
Rate Kasus (per 1000)	Kasus	6.24	74.33	59.56	17.22
Cakupan Imunisasi	%	68.52	113.97	89.77	12.04
Persentase Miskin	%	3.87	11.93	8.49	2.35
Persentase Stunting	%	1.60	17.08	5.60	4.24
Persentase Perokok	%	2.73	13.67	8.38	3.04

Kasus Pneumonia: Rata-rata Jumlah Kasus Pneumonia Balita per Kabupaten/Kota sangat bervariasi ($\sigma=5.560.80$), dengan nilai minimum 1.593 kasus dan maksimum 26.250 kasus. Keragaman yang tinggi ini menunjukkan perlunya pemodelan spasial untuk menjelaskan perbedaan antar wilayah.

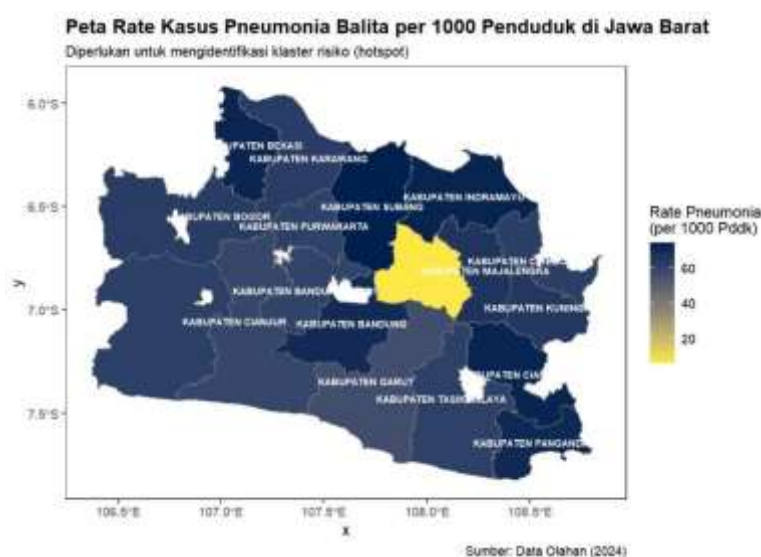
Rate Kasus: Rata-rata Rate Kasus Pneumonia (per 1000 balita) adalah 59.56, menunjukkan tingkat prevalensi yang signifikan. Rentang Rate Kasus juga lebar, dari 6.24 hingga 74.33.

Faktor Protektif/Risiko:

1. Cakupan Imunisasi memiliki rata-rata tinggi (89.77%), namun terdapat anomali pada nilai maksimum (113.97%), yang mungkin mengindikasikan metode penghitungan atau over-coverage di beberapa wilayah.
2. Persentase Miskin memiliki rata-rata 8.49% dan variasi yang relatif rendah ($\sigma=2.35$) dibandingkan rata-ratanya.
3. Persentase Stunting memiliki rata-rata 5.60%, dengan variasi yang cukup tinggi.
4. Persentase Perokok rata-rata rumah tangga yang merokok adalah 8.38%.

Analisis deskriptif ini mengkonfirmasi adanya heterogenitas (keragaman) yang signifikan pada Rate Kasus Pneumonia dan variabel-variabel prediktornya di 27 Kabupaten/Kota di Jawa Barat, memperkuat justifikasi penggunaan model spasial seperti RBN-TSA.

Interpretasi Visual Rate Kasus



Gambar 1: Peta Rate Kasus Pneumonia Balita per 1000 Penduduk di Jawa Barat

Rentang Rate Kasus yang lebar (Minimum 6.24 hingga Maksimum 74.33 per 1000 penduduk) menegaskan adanya variabilitas spasial yang ekstrem di Provinsi Jawa Barat. Peta Rate Kasus Aktual (Gambar 1) menunjukkan pengelompokan risiko yang jelas, yang membenarkan hipotesis awal mengenai heterogenitas spasial dan perlunya metode analisis spasial.

Interpretasi Visual Utama:

Coldspot (Risiko Rendah): Wilayah dengan Rate Kasus terendah (kategori 20-40, warna kuning/terang) teridentifikasi secara jelas dan terlokalisasi hanya di Kabupaten Majalengka. Kabupaten/Kota di sekitarnya juga menunjukkan Rate Kasus yang relatif rendah (kategori 40-60, warna biru sedang).

Hotspot (Risiko Tinggi): Sebagian besar wilayah di Provinsi Jawa Barat termasuk dalam kategori risiko tinggi (kategori 60 ke atas, warna biru gelap). Hampir seluruh Kabupaten/Kota di bagian utara, barat, dan selatan seperti Kabupaten Bekasi, Kabupaten Karawang, Kabupaten Subang, Kabupaten Bandung, Kabupaten Garut, dan Kabupaten Tasikmalaya, berada dalam klaster risiko tertinggi.

Pola ini menunjukkan bahwa kasus Pneumonia Balita cenderung mengelompok. Sebagian besar wilayah menjadi klaster hotspot, dengan hanya segelintir wilayah yang menjadi coldspot. Pola spasial yang ekstrem ini menggarisbawahi pentingnya menggunakan Model RBN-TSA Parsimoni untuk memodelkan dan menjelaskan dependensi spasial ini secara tepat.

Rentang Rate Kasus yang lebar (Min 6.24 hingga Max 74.33 per 1000 penduduk) menegaskan adanya variabilitas spasial yang ekstrem. Peta Rate Kasus Aktual (Peta 1) menunjukkan pengelompokan risiko tinggi (Hotspot) secara jelas di wilayah tertentu, yang membenarkan hipotesis awal mengenai heterogenitas spasial dan perlunya metode analisis spasial.

2. Uji Diagnostik Pra-Pemodelan (Preprocessing)

a) Uji Overdispersion dan Pemilihan Distribusi

Diagnostik formal terhadap variabel terikat (jumlah kasus) menunjukkan adanya overdispersion, di mana varians lebih besar dari rata-rata.

Tabel 4. Hasil Uji Diagnostik Overdispersion

Kriteria Uji	Hasil	Pembenaran Model
Rasio Deviasi/DF	4.41	Nilai ini secara definitif menolak asumsi Regresi Poisson dan membenarkan penggunaan Regresi Binomial Negatif (RBN) untuk mengakomodasi varians ekstra.

Secara teoritis, jika Rasio Deviasi per Degrees of Freedom (DF) jauh lebih besar dari 1 (seperti 4.41), maka terjadi overdispersion. Kondisi ini secara tegas menolak asumsi yang mendasari Regresi Poisson (yaitu equidispersion, di mana rata-rata = varians). Oleh karena itu, uji diagnostik ini memberikan pembenaran kuat bahwa Regresi Binomial Negatif (RBN) adalah model yang tepat untuk digunakan, karena RBN dirancang khusus untuk mengakomodasi varians yang lebih besar daripada rata-ratanya (overdispersion) dalam data count (jumlah).

b) Uji Korelasi Antar Prediktor

Uji korelasi antar variabel prediktor (Z-score) dilakukan untuk menilai hubungan bivariat.

Tabel 5. Uji Korelasi Antar Prediktor (Z-Score)

	<i>Imunisasi_z</i>	<i>Miskin_z</i>	<i>Stunting_z</i>	<i>Perokok_z</i>	<i>Puskesmas_z</i>	<i>Pendidikan_z</i>	<i>X_z</i>	<i>Y_z</i>
<i>Imunisasi_z</i>	1.00	-0.32	-0.40	-0.47	0.28	0.19	-0.31	0.56
<i>Miskin_z</i>	-0.32	1.00	0.17	0.54	-0.24	-0.67	0.59	-0.28
<i>Stunting_z</i>	-0.40	0.17	1.00	-0.04	-0.04	0.25	0.27	-0.54
<i>Perokok_z</i>	-0.47	0.54	-0.04	1.00	-0.20	-0.65	0.50	-0.54
<i>Puskesmas_z</i>	0.28	-0.24	-0.04	-0.20	1.00	-0.17	-0.20	0.24
<i>Pendidikan_z</i>	0.19	-0.67	0.25	-0.65	-0.17	1.00	-0.26	0.13

c) Uji Multikolinearitas (VIF) Final

Variabel yang menyebabkan VIF tinggi pada model awal (misalnya, Kepadatan Penduduk) telah dikeluarkan. Uji VIF akhir pada Model RBN-TSA Parsimoni memastikan stabilitas estimasi:

Tabel 6. Hasil Uji Multikolinearitas (VIF) Final Model RBN-TSA Parsimoni

Variabel	VIF (Model RBN-TSA Parsimoni)	Kesimpulan
<i>Imunisasi_z</i>	1.688	Aman (VIF <5)
<i>Miskin_z</i>	1.558	Aman (VIF <5)
<i>Perokok_z</i>	2.754	Aman (VIF <5)
<i>Stunting_z</i>	2.113	Aman (VIF <5)

Y_z	2.686	Aman (VIF <5)
-------	-------	---------------

Hasil ini menegaskan stabilitas estimasi dan memastikan bahwa koefisien regresi Model RBN-TSA Parsimoni bebas dari masalah multikolinearitas yang serius. Dengan demikian, interpretasi koefisien untuk setiap prediktor dapat dilakukan dengan valid, karena pengaruh setiap variabel bebas terhadap variabel terikat tidak dipengaruhi secara berlebihan oleh variabel bebas lainnya.

B. Pemodelan Regresi dan Pembahasan Hasil Regresi

1. Evaluasi Model Terbaik (AIC)

Proses seleksi model dilakukan menggunakan Akaike Information Criterion (AIC). AIC adalah ukuran yang menilai kualitas relatif model statistik untuk sekumpulan data, memberikan penalti untuk model yang memiliki terlalu banyak parameter (kompleksitas). Model dengan nilai AIC terendah dianggap sebagai model terbaik karena menunjukkan keseimbangan optimal antara kecocokan model (goodness-of-fit) dan parsimoni (kesederhanaan).

Tabel 7. Evaluasi Model Terbaik Berdasarkan Kriteria AIC

Prediktor Spasial	Kriteria VIF	Nilai AIC	Status
Non-Spasial	Aman (<5)	470.63	Global Benchmark
X_z dan Y_z (Penuh)	Aman (<5)	469.98	Disederhanakan
Y_z (Parsimoni)	Aman (<5)	466.64	Pemenang Mutlak

1. Model Global Benchmark (AIC: 470.63): Ini adalah model dasar RBN yang hanya menggunakan prediktor non-spasial. Model ini menjadi tolok ukur awal.

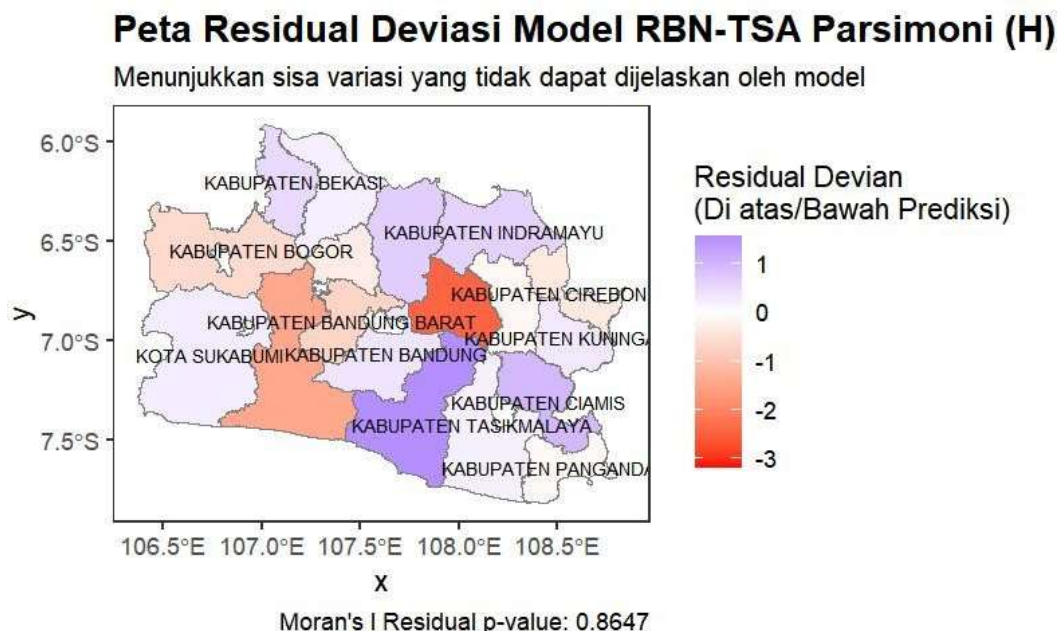
2. Model Spasial Penuh (AIC: 469.98): Ketika efek spasial penuh (TSA) ditambahkan ke Model Global, nilai AIC turun menjadi 469.98. Penurunan ini mengindikasikan bahwa penambahan efek spasial berhasil meningkatkan kecocokan model terhadap data kasus Pneumonia.

3. Model RBN-TSA Parsimoni (AIC: 466.64): Setelah dilakukan proses penyederhanaan (eliminasi prediktor non-signifikan), nilai AIC turun lebih jauh ke 466.64.

Model RBN-TSA Parsimoni dipilih sebagai model terbaik karena menghasilkan nilai AIC terendah (466.64) dari semua model yang diuji. Hal ini menegaskan bahwa model ini merupakan model yang paling efisien dan stabil untuk memprediksi Rate Kasus

Pneumonia Balita di Jawa Barat, berhasil mengakomodasi overdispersion dan heterogenitas spasial.

d) Uji Validasi Pola Spasial (Residual Analysis)



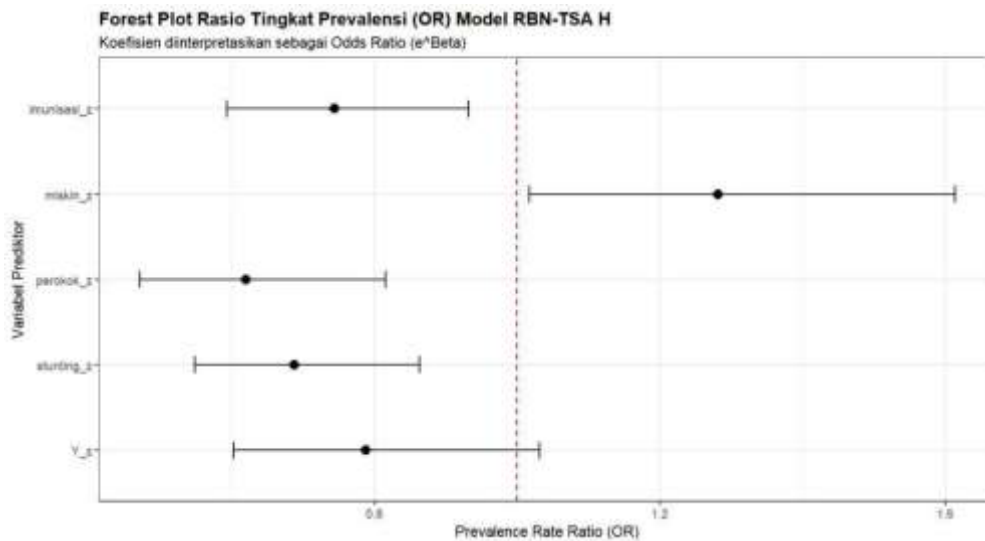
Gambar 2. Peta Residual Deviasi Model RBN-TSA Parsimoni (H)

Peta ini memvisualisasikan sisa error model (residual). Visualisasi menunjukkan bahwa residual tersebar secara acak dan tidak membentuk klaster (merah atau biru), mengindikasikan bahwa model sudah berhasil menangkap pola spasial yang ada.

Uji Moran's I Residual: Secara statistik, uji formal menghasilkan p-value sebesar 0.8647 (>0.05). Hasil ini secara definitif membuktikan bahwa tidak ada sisa autokorelasi spasial yang signifikan pada residual, memvalidasi kecukupan Model RBN-TSA Parsimoni dalam mengatasi spatial dependence.

D. Makna Koefisien Parameter Regresi

Hasil dari Model RBN-TSA Parsimoni adalah koefisien utama yang diinterpretasikan dalam bentuk Rasio Tingkat Prevalensi (OR).



Gambar 3. Forest Plot Rasio Tingkat Prevalensi (OR) Model RBN-TSA H

Tabel 8: Hasil Estimasi Parameter Model RBN-TSA Parsimoni

Variabel (Z-score)	Koefisien (β)	P-Value	Odds Ratio (e^{β})	Makna Koefisien
<i>CakupanImunisasi_z</i>	-0.296	0.012	0.744	Faktor Protektif Signifikan
<i>PersentasePendudukMiskin</i>	0.247	0.010	1.280	Faktor Risiko Signifikan
<i>PersentasePerokok_z</i>	-0.479	0.007	0.619	Anomalous (Negatif Signifikan)
<i>PersentaseStunting_z</i>	-0.374	0.016	0.688	Anomalous (Negatif Signifikan)
<i>Y_z(Latitude)</i>	0.081	≈ 0.08	1.084	Tren Spasial Utara-Selatan

Faktor Non-Spasial Signifikan (Sesuai Teori)

Dua variabel menunjukkan hasil yang signifikan dan konsisten dengan teori kesehatan:

1. Cakupan Imunisasi: Koefisien $\beta = -0.296$ ($p=0.012$) dan $OR=0.744$

Setiap peningkatan satu standar deviasi (Z-score) pada Cakupan Imunisasi akan menurunkan Rate Kasus Pneumonia sebesar $1-0.744=25.6\%$. Ini mengonfirmasi bahwa imunisasi merupakan faktor protektif yang signifikan.

2. Persentase Penduduk Miskin: Koefisien $\beta = 0.247$ ($p=0.012$) dan $OR=0.744$.

Setiap peningkatan satu standar deviasi (Z-score) pada Persentase Penduduk Miskin akan meningkatkan Rate Kasus Pneumonia sebesar $1.280-1=28.0\%$. Ini mengonfirmasi bahwa kemiskinan merupakan faktor risiko yang signifikan.

Temuan Anomalous (Bertentangan dengan Teori)

Dua variabel prediktor utama menunjukkan hasil yang signifikan secara statistik tetapi bertentangan dengan literatur kesehatan yang ada:

1. Persentase Perokok: Koefisien $\beta = -0.479$ ($p=0.007$) dan $OR=0.619$. Secara mengejutkan, variabel Perokok secara signifikan berhubungan negatif dengan kasus Pneumonia (peningkatan Perokok ke penurunan kasus).
2. Persentase Stunting: Koefisien $\beta = -0.374$ ($p=0.016$) dan $OR=0.688$. Sama halnya, variabel Stunting secara signifikan berhubungan negatif dengan kasus Pneumonia (peningkatan Stunting ke penurunan kasus).

Pembahasan Anomali: Kedua temuan anomalous (Perokok dan Stunting) ini menunjukkan adanya fenomena Ecological Fallacy yang melekat pada penggunaan data agregat. Kemungkinan besar, wilayah dengan kondisi sosio-ekonomi yang sangat buruk (yang juga tinggi dalam Perokok dan Stunting) memiliki akses atau sistem pelaporan kesehatan yang lemah, menyebabkan terjadinya under-reporting kasus Pneumonia.

Yz(Latitude): Koefisien $\beta = 0.081$ ($p \approx 0.08$) dan $OR=1.084$. Variabel ini menunjukkan tren spasial Utara-Selatan. Karena p-value mendekati ambang batas signifikansi $\alpha=0.05$, ini mengindikasikan adanya tren yang mendekati signifikan, di mana kasus Pneumonia cenderung sedikit lebih tinggi di daerah yang lebih selatan (nilai Latitude meningkat).

Pembahasan Faktor Determinan Utama:

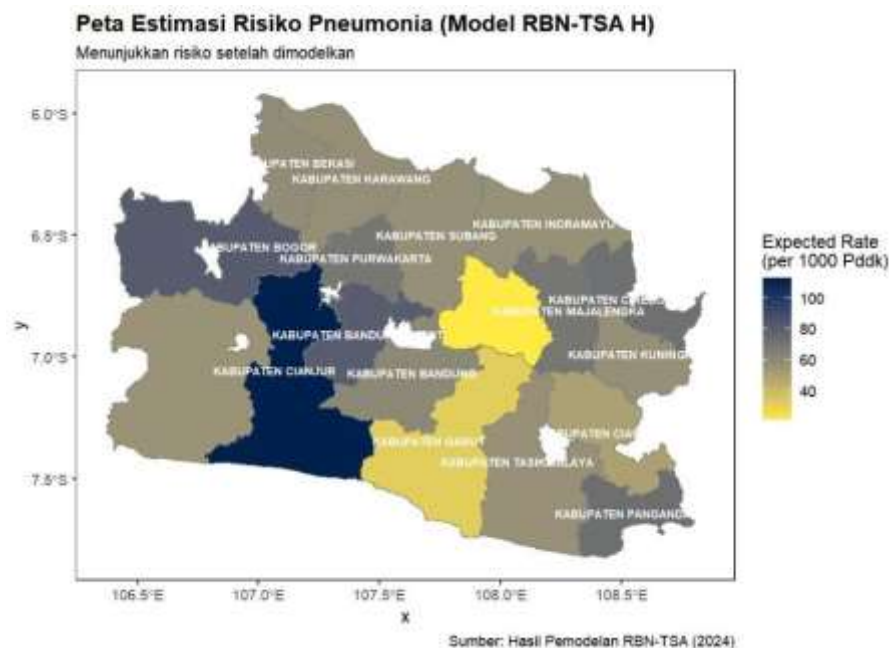
1. Cakupan Imunisasi (Faktor Protektif): Peningkatan 1 SD pada Imunisasi menurunkan Rate Kasus Pneumonia sebesar 25.6%. Temuan ini konsisten dan krusial, menegaskan pentingnya program vaksinasi dalam kebijakan publik.
2. Persentase Penduduk Miskin (Faktor Risiko): Peningkatan 1 SD pada Kemiskinan meningkatkan Rate Kasus Pneumonia sebesar 28.0%. Kemiskinan menjadi proksi

dari lingkungan hidup yang buruk (sanitasi dan gizi), yang merupakan faktor risiko kuat.

D. Interpretasi Pola Spasial dan Efek Model

1. Tren Spasial Utara-Selatan (Yz): Koefisien positif ($\beta=0.081$) yang mendekati signifikan ($p \approx 0.08$) mengindikasikan adanya gradien risiko makro-skala di mana risiko cenderung meningkat saat bergerak dari Utara ke Selatan Jawa Barat. Efek ini ditangkap oleh TSA setelah variabel non- spasial dipertimbangkan.

2. Peta Estimasi Risiko



Gambar 4: Peta Estimasi Risiko Pneumonia (Model RBN-TSA H)

Peta Estimasi Risiko (Peta 2) memvisualisasikan Rate Kasus yang diprediksi oleh Model RBN-TSA. Peta ini menunjukkan pola risiko yang terstruktur, memperlihatkan bagaimana kombinasi Kemiskinan, Imunisasi, dan Tren Yz memengaruhi risiko yang diestimasi di seluruh provinsi.

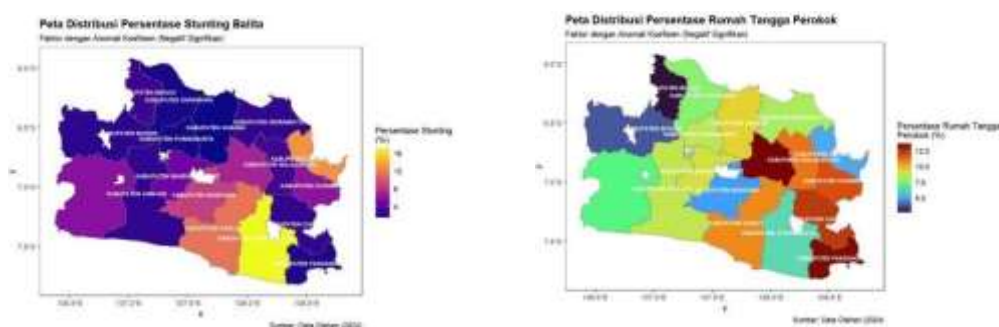
3. Pembahasan Anomali Koefisien dan The Ecological Fallacy

a. Analisis Konflik Teori (Anomalous Coefficients)

Temuan yang paling penting adalah koefisien anomalus negatif pada Persentase Perokok (OR= 0.619) dan Persentase Stunting (OR=0.688).

Konflik: Secara klinis, kedua faktor ini diyakini meningkatkan risiko Pneumonia, sehingga koefisien yang diharapkan adalah positif.

1. Argumentasi Ecological Fallacy: Hubungan negatif pada tingkat Kabupaten ini diyakini kuat sebagai hasil dari Kesesatan Ekologi (The Ecological Fallacy). Hasil ini tidak mencerminkan kausalitas terbalik, tetapi bias dalam pelaporan data.
2. Bukti Visual Under-reporting:



Gambar 5: Peta Distribusi Spasial Faktor Risiko dengan Koefisien Anomalous (Stunting dan Perokok)

Perbandingan visual antara Peta 4 (Stunting) dan Peta 5 (Perokok) dengan Peta Rate Kasus Aktual (Peta 1) memperkuat argumen ini. Wilayah dengan tingkat Stunting dan Perokok yang tinggi cenderung merupakan wilayah dengan beban sosio-ekonomi kronis. Di wilayah ini, kasus Pneumonia yang sebenarnya seringkali tidak tercatat atau terdiagnosis secara resmi karena keterbatasan akses dan infrastruktur kesehatan, yang mengakibatkan rate kasus yang dilaporkan menjadi rendah.

4. Perbandingan dengan Studi Sebelumnya dan Kontribusi Baru

a. Perbandingan dengan Studi Sebelumnya

Hasil penelitian ini menunjukkan konsistensi yang tinggi dengan literatur epidemiologi penyakit menular di Indonesia dan global:

Tabel 9: Perbandingan Hasil Penelitian dengan Studi Sebelumnya

Aspek	Hasil Penelitian Ini	Konsistensi dengan Studi Sebelumnya
-------	----------------------	--

Faktor Kunci	Imunisasi sebagai protektif, Konsisten dengan studi Kemiskinan sebagai risiko. (misalnya, YUSDIANA et al., 2025) yang menyoroti faktor sosio-ekonomi dan gizi sebagai determinan penyakit.
Metode Diagnostik	Penggunaan RBN karena overdispersion. Konsisten dengan studi yang menggunakan model Count

c. Kontribusi Baru Penelitian dan Elaborasi

Penelitian ini memberikan dua kontribusi baru yang signifikan:

1. Kontribusi Metodologi RBN-TSA yang Tervalidasi: Kontribusi utama terletak pada validasi metodologi dengan menunjukkan bahwa Model RBN-TSA Parsimoni secara efektif mampu memodelkan tren keruangan makro-skala (Y_z) dan terbukti adekuat dengan menyelesaikan sisa autokorelasi spasial (Moran's I $p\text{-value} \approx 0.8647$).
2. Kontribusi Kritis Kebijakan Publik (Ecological Fallacy): Identifikasi dan pembahasan mendalam mengenai anomali koefisien negatif pada Stunting dan Perokok menjadi kontribusi baru yang paling berdampak. Temuan ini berfungsi sebagai peringatan kritis bagi pembuat kebijakan: data agregat tidak dapat ditafsirkan secara literal. Koefisien anomalous ini harus digunakan sebagai sinyal diagnostik untuk mengidentifikasi wilayah dengan potensi under-reporting kasus yang parah, mengalihkan fokus kebijakan untuk memperbaiki kualitas sistem pelaporan dan akses diagnosis di wilayah tersebut.

KESIMPULAN DAN SARAN

Berdasarkan Penelitian ini berhasil mengidentifikasi Model Regresi Binomial Negatif Trend Surface Analysis (RBN-TSA) Parsimoni (H) sebagai model spasial terbaik (AIC: 466.64) untuk memodelkan Rate Kasus Pneumonia Balita di Jawa Barat, setelah secara efektif mengatasi overdispersion dan spatial dependence. Model ini tervalidasi secara teknis (VIF aman) dan spasial (Moran's I Residual $p\text{-value} \approx 0.8647$), membuktikan kecukupan model dalam menangkap pola keruangan.

Hasil utama model menunjukkan bahwa Cakupan Imunisasi Dasar Lengkap merupakan faktor protektif signifikan ($OR = 0.744$), di mana peningkatan satu standar deviasi Imunisasi berhubungan dengan penurunan Rate Kasus sebesar 25.6%.

Sebaliknya, Persentase Penduduk Miskin adalah faktor risiko signifikan ($OR = 1.280$), yang berhubungan dengan peningkatan Rate Kasus sebesar 28.0%.

Meskipun model stabil, penelitian ini dihadapkan pada keterbatasan krusial akibat penggunaan data agregat di tingkat Kabupaten/Kota. Keterbatasan ini memicu fenomena Ecological Fallacy, yang terbukti dari koefisien anomali negatif signifikan pada variabel Stunting dan Perokok. Temuan anomali ini diinterpretasikan sebagai indikasi adanya bias under-reporting kasus Pneumonia di wilayah dengan beban sosio-ekonomi kronis, sehingga model menangkap bias pelaporan dan bukan hubungan kausalitas yang sebenarnya.

Sebagai masukan untuk riset mendatang, disarankan untuk mengatasi limitasi ini dengan menggunakan data disagregat (tingkat Puskesmas atau individu) dan menerapkan metode spasial yang lebih canggih seperti Regresi Geografis Tertimbang (GWR) atau model BYM untuk analisis heterogenitas lokal. Secara kebijakan, temuan ini menggarisbawahi perlunya fokus pada peningkatan Cakupan Imunisasi dan Pengentasan Kemiskinan, sekaligus menyerukan perbaikan sistem pelaporan dan akses diagnosis di wilayah yang teridentifikasi berpotensi mengalami under-reporting.

DAFTAR PUSTAKA

- Achmad, N., Payu, M. R. F., & Rahim, Y. (2022). Pemodelan Pneumonia Berat Menggunakan Regresi Zero Inflated Negative Binomial di Gorontalo. *Euler : Jurnal Ilmiah Matematika, Sains Dan Teknologi*, 10(1), 45–53. <https://doi.org/10.34312/euler.v10i1.13990>
- Arief, Y. S., Yunita, F. C., Efendi, F., Murti, F. A. K., Pradipta, R. O., & McKenna, L. (2025). Social and Environmental Determinants of Childhood Stunting in Indonesia: National Cross-Sectional Study. *JMIR Pediatrics and Parenting*, 8, e68918. <https://doi.org/10.2196/68918>
- Citra Mar'ati, H., Lestari, D., & Novkaniza, F. (2024). Modelling the Number of Stunting Under-Five Children in East Nusa Tenggara Using Negative Binomial Regression. *Asian Journal of Management Entrepreneurship and Social Science*, 4(1), 729–744. <https://ajmesc.com/index.php/ajmesc>
- Degif, K. A., Gebrehiwot, M., Tadege, G., Demoze, L., & Yitageasu, G. (2025). Spatial and temporal variation of pneumonia incidence among under-five children in central gondar zone, Northwest Ethiopia, 2013- 2022. *BMC Pediatrics*, 25(1). <https://doi.org/10.1186/s12887-025-05550-7>
- Ghosh, K., Chakraborty, A. S., & SenGupta, S. (2023). Identifying spatial clustering of diarrhoea among children under 5 years across 707 districts in India: a cross

sectional study. *BMC Pediatrics*, 23(1), 1–15. <https://doi.org/10.1186/s12887-023-04073-3>

- Halim, R., Syukri, M., Guspianto, G., & Akbar, F. (2023). Spatial distribution analysis and risk factors for stunting in Kerinci Regency, Jambi Province Indonesia. *AcTion: Aceh Nutrition Journal*, 8(4), 542. <https://doi.org/10.30867/action.v8i4.1190>
- Kertapati, K., & Palembang, K. (2020). *Jekk kk*. 5(2), 96–103.
- Nur, S., Usman, A., Kadir, L., Mokodompis, Y., Penelitian, A., & Kunci, K. (2025). Hubungan Perilaku Merokok Keluarga dan Pemberian Asi Eksklusif dengan Kejadian Pneumonia pada Bayi 6-24 Bulan di Wilayah Puskesmas Kota Tengah Kota Gorontalo. *Jurnal Kolaboratif Sains*, 8(1), 126–134. <https://doi.org/10.56338/jks.v8i1.6660>
- Permana, P. (2021). Pemodelan Pemodelan Spasial Kasus Balita Laki-Laki Penderita Pneumonia Di Kota Bandung. *SATIN - Sains Dan Teknologi Informasi*, 7(2), 64–72. <https://doi.org/10.33372/stn.v7i2.757>
- Prawidia, D., Zhafirah, L., Saputra, N., Kartiasih, F., & Sahu, U. (2023). Determinants of Under-five Mortality Due to Pneumonia: A Negative Binomial Regression Analysis. *Jurnal Varian*, 7(1), 59–66. <https://doi.org/10.30812/varian.v7i1.2768>
- Purnama, T. B., Wagatsuma, K., Pane, M., & Saito, R. (2024). Effects of the Local Environment and Nutritional Status on the Incidence of Acute Respiratory Infections Among Children Under 5 Years Old in Indonesia. *Journal of Preventive Medicine and Public Health*, 57(5), 461–470. <https://doi.org/10.3961/jpmph.24.246>
- Qur'ani, A., Widyaningrum, C. S., & Rohimiyah, S. (2024). The Robust Negative Binomial Regression Model on Under-five Mortality due to Pneumonia in the Province of East Java. *Jurnal Matematika, Statistika Dan Komputasi*, 21(1), 176–189. <https://doi.org/10.20956/j.v21i1.34512>
- Simarda, L., Lestari, D., Novkaniza, F., Haqqi, A., & Devila, S. (2024). Modeling The Number of Toddler Pneumonia Sufferers in DKI Jakarta using Negative Binomial Regression. 04(01), 622– 643. <https://ajmesc.com/index.php/ajmesc>
- Sollari Lopes, J., Dey, A., Balabaskaran Nina, P., Rasheed, A. N., Nina, B., Rasheed, A., Balasubramani, K., Arun Prasad, K., Kumar Kodali, N., Kalladath Abdul Rasheed, N., Chellappan, S., Kumar Sarma, D., Kumar, M., Dixit, R., Mariya James, M., Kumar Behera, S., & Shekhar, S. (2022). Spatial epidemiology of acute respiratory infections in children under r years and associated risk factors in India: District-level analysis of health, household, and environmental datasets. *Frontiers in Public Health*. <https://www.esri.com/en-us/arcgis/products/arcgis-desktop/>
- Suchery, P. L. Y., & Hastono, S. P. (2025). Spatial Analysis of Pneumonia Distribution in Children Under Five in West Java: Relationships with Individual and Environmental Determinants. *Jurnal Ilmu Kesehatan Masyarakat*, 16(1), 118–134. <https://doi.org/10.26553/jikm.2025.16.1.118-134>
- Tri Fatmala, C., Hayati, M., Permatasari, R., Hudori, M., & Yuliana Dalimunthe, D. (2024). Pemodelan Jumlah Kasus HIV/AIDS di Provinsi Lampung Menggunakan Regresi Binomial Negatif. *Journal of Mathematics: Theory and Applications*, 6(2), 168–177. <https://doi.org/10.31605/jomta.v6i2.4069>

- Usman, M., & Kopczewska, K. (2022). Spatial and Machine Learning Approach to Model Childhood Stunting in Pakistan: Role of Socio-Economic and Environmental Factors. *International Journal of Environmental Research and Public Health*, 19(17). <https://doi.org/10.3390/ijerph191710967>
- Wang, Y., Han, R., Ding, X., Chen, J., Feng, W., Wang, C., Gao, R., & Ma, A. (2025). A 32-year trend analysis of lower respiratory infections in children under 5: insights from the global burden of disease study 2021. *Frontiers in Public Health*, 13(January). <https://doi.org/10.3389/fpubh.2025.1483179>
- Yusdiana, Y., Sukendi, S., Ikhwan, S. Y., & Dedi, A. (2025). SPATIAL MODELING OF RISK FACTORS FOR UNDER-FIVE PNEUMONIA IN ROKAN HILIR DISTRICT , INDONESIA
- Doctor of Environmental Science , Postgraduate Program , Universitas Riau , Pekanbaru 28131 , Indonesia Corresponding Author ' s E -mail : yusdianaikmfur@gmail. 19, 15–32.