

Pemodelan *generalized additive mixed model* pada data jumlah kasus kusta di Sulawesi Tengah

Siti Azizah Karim^{1*}, Lailany Yahya², La Ode Nashar²

¹ Mahasiswa Statistika, FMIPA, Universitas Negeri Gorontalo, Gorontalo

² Dosen Matematika, FMIPA, Universitas Negeri Gorontalo, Gorontalo

azizahnursin821@gmail.com

Diterima: 16-06-2026; Direvisi: 30-06-2026; Dipublikasi: 30-06-2026

Abstract

Leprosy is a contagious disease caused by the bacterium Mycobacterium leprae and remains a concern in Indonesia, especially in Central Sulawesi Province, which is classified as a high endemic area. Nationally, in 2022 Central Sulawesi ranked 8th in terms of new leprosy cases. This study aims to model the number of leprosy cases and identify significant factors using the Generalized Additive Mixed Model (GAMM) method. The data used are the number of leprosy cases in 2021–2023 in Central Sulawesi Province. The independent variables used included population density, percentage of poor population, access to proper sanitation, health complaints, number of medical personnel, male population, health services for the elderly, and access to proper drinking water sources. The GAMM model was applied with a P spline smoothing function and parameter estimation using Restricted Maximum Likelihood (REML). The results showed that male population was significantly and positively associated with the number of leprosy cases. The random effect of district/city also indicated variation among regions that was not fully explained by the fixed effects in the model. These findings indicate the presence of statistical association between demographic characteristics and leprosy case counts, as well as heterogeneity among districts/cities, without implying a causal relationship.

Keywords: leprosy; central sulawesi; generalized additive mixed model; p-spline; negative binomial

Abstrak

Kusta adalah salah satu penyakit menular yang disebabkan oleh bakteri Mycobacterium Leprae dan masih menjadi perhatian di Indonesia, terutama di Provinsi Sulawesi Tengah yang tergolong daerah high endemic. Secara nasional, pada tahun 2022 Sulawesi Tengah berada di peringkat ke-8 dalam hal kasus baru kusta. Penelitian ini bertujuan untuk memodelkan jumlah kasus kusta dan mengetahui faktor-faktor yang signifikan menggunakan metode Generalized Additive Mixed Model (GAMM). Data yang digunakan ialah data jumlah kasus kusta tahun 2021–2023 di Provinsi Sulawesi Tengah. Variabel independen yang digunakan mencakup kepadatan penduduk, persentase penduduk miskin, akses sanitasi layak, keluhan kesehatan, jumlah tenaga medis, jumlah penduduk laki-laki, pelayanan kesehatan lansia, dan akses sumber air minum layak. Model GAMM diterapkan dengan fungsi penghalus P-spline dan estimasi parameter menggunakan Restricted Maximum Likelihood (REML). Hasil penelitian menunjukkan bahwa jumlah penduduk laki-laki memiliki hubungan yang signifikan dan positif dengan jumlah kasus kusta. Efek acak kabupaten/kota juga menunjukkan adanya variasi antarwilayah yang belum sepenuhnya dijelaskan oleh efek tetap dalam model. Temuan tersebut menunjukkan adanya hubungan statistik antara karakteristik demografi dan jumlah kasus kusta serta heterogenitas antar kabupaten/kota, tanpa menginterpretasikannya sebagai hubungan sebab-akibat.

Kata Kunci: kusta; sulawesi tengah; generalized additive mixed model; p-spline; binomial negatif

1. PENDAHULUAN

Kesehatan merupakan parameter utama kualitas hidup manusia yang berdampak langsung pada produktivitas sosial dan ekonomi. Indonesia, melalui rencana aksi *Sustainable Development Goals* (SDGs), masih berupaya keras mengeliminasi penyakit menular tropis, salah satunya adalah kusta (*Leprosy*) (Bappenas, 2020). Penyakit yang disebabkan oleh bakteri *Mycobacterium Leprae* ini menyerang kulit, saraf tepi, hingga organ tubuh lainnya, dan jika terlambat dideteksi dapat menyebabkan disabilitas permanen. Penderita kusta seringkali mendapatkan stigma negatif dari masyarakat dikarenakan dianggap penyakit kutukan, keturunan, maupun penyakit aib. Adanya perlakuan diskriminatif masyarakat tentunya memiliki dampak buruk pada penderita kusta yang harus menanggung beban fisik serta moral dalam waktu yang berkepanjangan.

Secara global, Indonesia menempati peringkat ketiga kasus kusta terbanyak setelah India dan Brasil. Di tingkat nasional, Provinsi Sulawesi Tengah menjadi perhatian khusus karena termasuk wilayah *high endemic*. Pada tahun 2022, angka penemuan kasus baru di wilayah ini mencapai 9,7 per 100.000 penduduk, menempatkannya di peringkat ke-8 nasional (Kemenkes, 2023). Meskipun status eliminasi secara administratif telah dicapai, fluktuasi kasus masih terjadi, di mana pada akhir 2023 angka insidensi kembali merangkak naik menjadi 0,98 per 10.000 penduduk.

Beberapa penelitian mengenai pemodelan penyakit kusta telah dilakukan dengan berbagai pendekatan statistik. Andreas et al., (2022) memodelkan prevalensi kusta di Jawa Timur menggunakan *geographically weighted regression* (GWR) untuk mengakomodasi variasi hubungan antarvariabel menurut lokasi. Aipassa et al., (2023) menganalisis faktor-faktor yang memengaruhi jumlah penyakit kusta di Provinsi Maluku menggunakan regresi binomial negatif untuk menangani karakteristik data cacah yang mengalami *overdispersi*. Sementara itu, Azwarini & Purhadi, (2021) menggunakan regresi *zero-inflated Poisson inverse Gaussian* (ZIPIG) untuk memodelkan jumlah kasus kusta yang memiliki banyak nilai nol. Pendekatan-pendekatan tersebut menunjukkan bahwa karakteristik data kusta dapat ditangani dengan model yang berbeda sesuai dengan permasalahan statistik yang ditemukan.

Meskipun demikian, penelitian-penelitian tersebut belum secara simultan mengakomodasi kemungkinan hubungan nonlinear antara variabel penjelas dengan jumlah kasus, karakteristik data cacah dengan distribusi binomial negatif, serta adanya pengamatan berulang pada wilayah yang sama selama beberapa tahun. Regresi binomial negatif dapat mengakomodasi *overdispersi* pada data cacah, tetapi hubungan antara prediktor dan respons pada bentuk regresi standar umumnya dinyatakan secara parametrik. Sementara itu, GWR berfokus pada variasi hubungan menurut lokasi, sedangkan model *zero-inflated* terutama ditujukan untuk menangani kelebihan nilai nol.

Oleh karena itu, diperlukan pendekatan yang dapat mengakomodasi lebih dari satu karakteristik tersebut dalam satu kerangka pemodelan.

Untuk mengatasi keterbatasan tersebut, penelitian ini menerapkan pendekatan Generalized Additive Mixed Model (GAMM) yang merupakan pengembangan dari Generalized Additive Model (GAM). GAM merupakan pengembangan dari generalized linear model yang memungkinkan hubungan antara variabel respons dan prediktor dimodelkan secara fleksibel melalui fungsi aditif dan fungsi penghalus (Hastie & Tibshirani, 1990). Dalam penerapannya, hubungan nonlinear dalam GAM dapat direpresentasikan menggunakan fungsi penghalus yang fleksibel, dengan tingkat kehalusan dikendalikan melalui penalisasi sehingga model dapat menyesuaikan kompleksitas kurva berdasarkan informasi dalam data (Wood, 2017). Salah satu basis yang dapat digunakan dalam fungsi penghalus adalah P-spline, yang menggabungkan basis B-spline dengan penalti terhadap koefisien untuk mengendalikan tingkat kehalusan kurva (Eilers & Marx, 1996). Komponen mixed model memungkinkan dimasukkannya efek acak untuk menangkap variasi antar kelompok pengamatan (Bolker et al., 2009).

Penggunaan efek acak kabupaten/kota dalam penelitian ini perlu dibedakan dari efek spasial. Efek acak wilayah tidak dimaksudkan untuk memodelkan ketergantungan geografis antarwilayah, melainkan untuk menangkap variasi atau heterogenitas antar kabupaten/kota yang tidak sepenuhnya dapat dijelaskan oleh variabel penjelas yang dimasukkan dalam model. Pemodelan efek spasial secara eksplisit memerlukan struktur yang merepresentasikan hubungan atau ketergantungan geografis antarwilayah. Dengan demikian, penggunaan random effect kabupaten/kota dalam penelitian ini tidak dapat diartikan sebagai pemodelan spasial secara langsung.

Research gap penelitian ini terletak pada keterbatasan penelitian terdahulu dalam menggabungkan fleksibilitas hubungan nonlinear dengan struktur data berkelompok pada pemodelan jumlah kasus kusta tingkat kabupaten/kota. Kebaruan penelitian ini terletak pada penerapan GAMM dengan distribusi binomial negatif, fungsi penghalus P-spline, serta random effect kabupaten/kota untuk memodelkan jumlah kasus kusta di 13 kabupaten/kota di Sulawesi Tengah selama periode 2021–2023. Pendekatan tersebut digunakan untuk memperoleh model yang dapat mempertimbangkan karakteristik distribusi data, kemungkinan ketidaklinearan hubungan, dan variasi antarwilayah secara bersamaan. Berdasarkan penelitian sebelumnya terdapat beberapa faktor yang mempengaruhi kejadian penyakit kusta, seperti kepadatan penduduk, penduduk dengan keluhan kesehatan, dan penduduk miskin (Dhahari & Sofro, 2024). Kepadatan penduduk dan penggunaan air bersih juga signifikan terhadap kasus kusta (Darsyah & Ramadhan, 2022).

Penelitian ini bertujuan untuk memodelkan jumlah kasus penyakit kusta di Sulawesi Tengah menggunakan Generalized Additive Mixed Model dan mengidentifikasi variabel yang memiliki hubungan signifikan dengan jumlah kasus kusta. Hasil penelitian

diharapkan dapat memberikan informasi statistik mengenai faktor-faktor yang berkaitan dengan variasi jumlah kasus kusta serta memberikan alternatif pendekatan pemodelan untuk data kesehatan yang memiliki struktur berkelompok dan kemungkinan hubungan nonlinear.

2. METODE PELAKSANAAN

Penelitian ini menggunakan data sekunder yang diperoleh dari situs resmi Dinas Kesehatan Provinsi Sulawesi Tengah dan Badan Pusat Statistik Provinsi Sulawesi Tengah. Data penelitian terdiri dari 13 kabupaten/kota di wilayah Provinsi Sulawesi Tengah dengan periode observasi selama tiga tahun, yakni 2021-2023. Penelitian ini menggunakan variabel dependen berupa jumlah kasus kusta (Y), serta delapan variabel independen, yaitu kepadatan penduduk (X_1), penduduk miskin (X_2), rumah tangga belum memiliki akses sanitasi layak (X_3), penduduk dengan keluhan kesehatan dan pernah berobat jalan (X_4), jumlah tenaga medis (X_5), jumlah penduduk laki-laki (X_6), pelayanan kesehatan usia lanjut (X_7), dan rumah tangga memiliki akses sumber air minum layak (X_8).

Analisis data dilakukan menggunakan software Rstudio. Adapun langkah-langkah yang dilakukan untuk analisis data dalam penelitian ini adalah sebagai berikut:

1. Identifikasi distribusi variabel dependen: menguji asumsi normalitas terhadap variabel dependen menggunakan uji *shapiro-wilk*. Jika asumsi normalitas ditolak, dilakukan uji kesesuaian distribusi menggunakan uji *chi-square goodness of fit*.
2. Uji korelasi: melihat hubungan awal antara variabel independen dan dependen menggunakan korelasi *rank spearman*.
3. Uji multikolinearitas: menghitung nilai *variance inflation factor* (VIF) pada variabel terpilih.
4. Uji linearitas: mendeteksi bentuk hubungan fungsional antara variabel menggunakan *ramsey RESET test*. Variabel yang terbukti tidak linear akan dimodelkan dengan menambahkan fungsi penghalus *P-Spline*.
5. Estimasi parameter GAMM: membentuk model GAMM
6. Uji signifikansi yang dilakukan secara simultan dan parsial
7. Uji diagnostik yang meliputi uji normalitas, pemeriksaan heteroskedastisitas, serta uji independensi terhadap residual model

3. HASIL DAN PEMBAHASAN

3.1 Identifikasi Distribusi Variabel Dependen

Salah satu syarat menggunakan Generalized Additive Mixed Model adalah distribusi variabel dependennya merupakan distribusi keluarga eksponensial. Berikut hasil identifikasi distribusi normal menggunakan uji *shapiro-wilk*:

Tabel 1. Hasil Uji Distribusi Normal Variabel Dependen

Variabel	P-Value
Y	0,0168

Berdasarkan hasil uji Shapiro-Wilk, diperoleh p-value sebesar 0,0168 sehingga hipotesis nol ditolak dan variabel dependen tidak berdistribusi normal. Karena variabel dependen berupa data cacah, selanjutnya dipertimbangkan distribusi Poisson dan binomial negatif sebagai distribusi alternatif. Distribusi Poisson merupakan salah satu distribusi dasar yang umum digunakan untuk memodelkan data cacah, dengan asumsi bahwa nilai harapan dan varians bersyarat memiliki nilai yang sama (Cameron & Trivedi, 2013). Sementara itu, distribusi binomial negatif dapat digunakan sebagai alternatif ketika data cacah menunjukkan dispersi yang lebih besar dibandingkan yang dapat dijelaskan oleh distribusi Poisson (Hilbe, 2011). Untuk memastikan distribusi yang lebih sesuai dengan karakteristik data penelitian, selanjutnya dilakukan uji kesesuaian distribusi menggunakan chi-square goodness-of-fit. Hasil pengujian distribusi Poisson dan binomial negatif disajikan pada Tabel 2.

Tabel 2. Hasil Perbandingan Distribusi

Distribusi	P-Value
Poisson	$< 2.2e-16$
Negative Binomial	0,4752

Berdasarkan tabel tersebut distribusi poisson memiliki nilai $p - value < 0,05$ yang berarti tolak H_0 , sedangkan distribusi negative binomial memiliki $p - value > 0,05$ yang berarti gagal tolak H_0 . Dengan demikian hasil uji goodness of fit menunjukkan bahwa tidak terdapat bukti yang cukup untuk menolak kesesuaian data dengan distribusi binomial negatif. Pemilihan distribusi Binomial Negatif dalam penelitian ini sejalan dengan Budiharti & Sunendiari, (2021) yang menunjukkan bahwa model Binomial Negatif lebih sesuai dibandingkan regresi Poisson dalam pemodelan jumlah penderita kusta.

3.2 Uji Korelasi

Uji korelasi dilakukan untuk melihat hubungan antar variabel. Pada penelitian ini uji korelasi dilakukan menggunakan *rank spearman* dengan hasil sebagai berikut:

Tabel 3. Hasil Uji Korelasi

Variabel	r_s	P - Value
X_1	0,3801	0,0170
X_2	0,3548	0,02669
X_3	0,2684	0,0986
X_4	-0,0835	0,6134
X_5	0,3386	0,0350
X_6	0,7305	$1,299 \times 10^{-7}$
X_7	0,2825	0,0814
X_8	-0,0371	0,8227

Pada tabel diatas hasil uji korelasi menunjukkan bahwa variabel X_1, X_2, X_5 dan X_6 memiliki nilai $p - value < 0,05$ yang berarti tolak H_0 sehingga dapat disimpulkan bahwa keempat variabel tersebut memiliki korelasi signifikan dengan variabel dependen. Sedangkan, X_3, X_4, X_7 dan X_8 tidak memiliki korelasi yang signifikan dengan variabel dependen.

3.3 Uji Multikolinearitas

Uji multikolinearitas dilakukan untuk melihat apakah terdapat hubungan yang kuat antar variabel independen dengan cara menghitung nilai VIF dengan hasil sebagai berikut:

Tabel 4. Hasil Uji Multikolinearitas

	X_1	X_2	X_5	X_6
X_1	Inf	1,638513	3,441109	1,163543
X_2	1,638513	Inf	1,598526	1,130261
X_5	3,441109	1,598526	Inf	1,348209
X_6	1,163543	1,130261	1,348209	Inf

Pada tabel menunjukkan hasil perhitungan multikolinearitas antar variabel independen memenuhi kriteria, yaitu seluruh variabel memiliki nilai $VIF < 10$. Maka dapat disimpulkan bahwa seluruh variabel independen tidak menunjukkan adanya multikolinearitas, dengan VIF digunakan sebagai salah satu ukuran untuk mengidentifikasi korelasi antarvariabel prediktor (Jamco et al., 2023).

3.4 Uji Linearitas

Uji linearitas dilakukan untuk mengetahui apakah terdapat hubungan yang linear antara variabel dependen dan variabel independen. Uji ini dilakukan menggunakan uji *ramsey RESET* dengan hasil sebagai berikut:

Tabel 5. Hasil Uji Linearitas

Variabel	<i>P-Value</i>
X_1 dan Y	0,249
X_2 dan Y	0,0033
X_5 dan Y	0,3905
X_6 dan Y	0,8482

Pada tabel tersebut menunjukkan bahwa variabel X_1, X_5 dan X_6 memiliki $p - value > 0,05$ yang berarti ketiga variabel tersebut memiliki hubungan linear dengan variabel dependen. Sementara itu, variabel X_2 dengan $p - value < 0,05$ yang berarti bahwa variabel X_2 tidak linear dengan variabel dependen.

3.5 Estimasi Parameter Model GAMM

Setelah melakukan uji asumsi terhadap semua variabel, maka selanjutnya dilakukan pemodelan GAMM. Berikut ini adalah hasil estimasi parameter dari model GAMM:

Tabel 6. Hasil Estimasi Parameter

Parameter	Estimasi Parameter
<i>Intercept</i>	1,701
X_1	0,00006
X_5	-0,00138
X_6	$9,791 \times 10^{-6}$

Tabel tersebut menampilkan hasil estimasi parameter model GAMM. Nilai intercept sebesar 1,701 menunjukkan nilai log rata-rata jumlah kasus kusta ketika seluruh variabel independen bernilai nol. Adapun variabel yang memiliki koefisien positif mengindikasikan bahwa peningkatan nilai variabel-variabel tersebut cenderung meningkatkan jumlah kasus kusta dengan asumsi variabel lain konstan. Sebaliknya, variabel dengan koefisien negatif menunjukkan kecenderungan penurunan variabel tersebut akan diikuti dengan peningkatan variabel dependen yakni jumlah kasus kusta.

Tabel 7. Hasil Estimasi Komponen Penghalus

	edf
$s(X_2)$	1

Tabel diatas menyajikan hasil estimasi komponen penghalus pada model GAMM, khususnya untuk variabel $s(X_2)$ yang dimodelkan dengan fungsi penghalus. Nilai effective degrees of freedom (edf) sebesar 1 menunjukkan bahwa hubungan antara variabel independen X_2 dan variabel dependen mendekati linear.

3.6 Uji Signifikansi Parameter

Pengujian signifikansi parameter dilakukan secara simultan dan parsial dengan hasil sebagai berikut:

Tabel 8. Hasil Uji Simultan

G	<i>P-Value</i>
13,082	0,0226

Pada tabel menunjukkan hasil uji simultan yang dilakukan menggunakan uji rasio likelihood yang menghasilkan nilai *p-value* sebesar 0,0226 yang mana nilai tersebut kurang dari $\alpha = 0,05$ maka H_0 ditolak. Artinya, minimal ada satu parameter yang signifikan terhadap model dan terdapat variasi antar wilayah yang tidak dapat dijelaskan sepenuhnya oleh variabel independen, sehingga pemodelan dengan memasukkan efek acak wilayah pada model GAMM layak digunakan.

Tabel 9. Hasil Uji Parsial

Parameter	<i>P-Value</i>
X_1	0,39116
X_5	0,26648
X_6	0,00188
$s(X_2)$	0,673

Berdasarkan tabel hasil uji parsial model GAMM menunjukkan bahwa parameter variabel jumlah penduduk laki-laki (X_6) memiliki $p - value < 0,05$ ini membuktikan bahwa variabel X_6 berpengaruh signifikan terhadap jumlah kasus kusta. Sedangkan variabel X_1, X_5 dan X_2 memiliki $p - value > 0,05$ yang berarti ketiga variabel tersebut tidak berpengaruh signifikan terhadap variabel dependen yakni jumlah kasus kusta

3.7 Uji Diagnostik Model

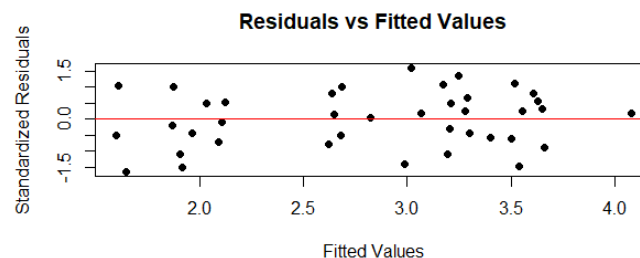
Pada penelitian ini uji diagnostik terhadap residual model digunakan dengan melakukan pemeriksaan asumsi normalitas, pemeriksaan heteroskedastisitas, serta uji independensi terhadap residual. Adapun hasil analisis uji diagnostik sebagai berikut:

Tabel 10. Uji Normalitas Residual

	<i>P-Value</i>
Residual	0,4394

Hasil uji normalitas residual pada tabel 10 menunjukkan nilai $p - value > 0,05$ yang berarti gagal tolak H_0 . Artinya residual berdistribusi normal dan model dinilai cukup baik dalam merepresentasikan variasi data jumlah kasus kusta.

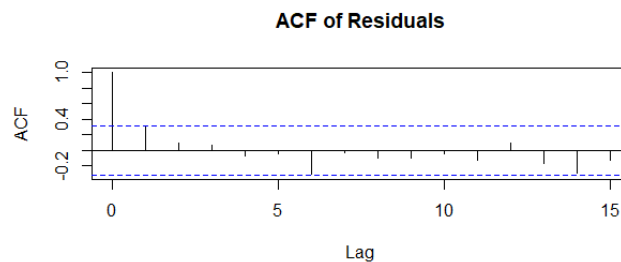
Uji heteroskedastisitas bertujuan untuk melihat varians residual bersifat konstan (homokedastisitas) atau berubah-ubah. Berikut adalah hasil pemeriksaan heteroskedastisitas:



Gambar 1. Plot Residual vs Fitted Values

Gambar 1 menampilkan plot residual terhadap nilai prediksi model. Dapat dilihat pada plot tersebut tidak terdapat struktur yang jelas pada titik-titik residu yang tersebar secara acak disekitar garis nol. Ini menunjukkan bahwa tidak terjadi masalah heteroskedastisitas.

Uji independensi residual bertujuan untuk mengetahui ada atau tidaknya korelasi serial atau autokorelasi antarresidual. Pemeriksaan dilakukan menggunakan plot autocorrelation function (ACF)



Gambar 2. Plot ACF of Residual

Gambar 2. menampilkan plot ACF dari residual model. Dapat dilihat pada lag 1 sampai 15 seluruh nilai ACF berada dalam batas kepercayaan, sehingga tidak ditemukan autokorelasi yang signifikan antar residual. Dengan demikian, dapat dikatakan bahwa asumsi independensi residual telah terpenuhi.

3.8 Pembahasan

Model yang diperoleh pada pemodelan GAMM adalah sebagai berikut:

$$Y = 1,701 + 9,791 \times 10^{-6}X_6$$

Hasil pemodelan Generalized Additive Mixed Model menunjukkan bahwa jumlah penduduk laki-laki merupakan variabel yang memiliki hubungan signifikan dengan jumlah kasus kusta. Koefisien positif sebesar 0,000009791 menunjukkan bahwa peningkatan jumlah penduduk laki-laki berkaitan dengan peningkatan rata-rata jumlah kasus kusta ketika variabel lain dalam model dianggap konstan. Dengan menggunakan interpretasi pada skala log, peningkatan sebesar 1.000 penduduk laki-laki berkaitan dengan perubahan rata-rata jumlah kasus kusta sebesar sekitar 0,98%. Interpretasi tersebut merupakan hubungan statistik berdasarkan model dan tidak dimaksudkan sebagai hubungan sebab-akibat.

Hubungan positif tersebut perlu dipahami dengan mempertimbangkan bahwa variabel respons dalam penelitian ini merupakan jumlah kasus kusta, bukan angka kejadian yang telah distandardisasi terhadap jumlah penduduk. Oleh karena itu, wilayah dengan jumlah penduduk laki-laki yang lebih besar secara demografis juga memiliki jumlah penduduk yang lebih besar sebagai populasi yang berpotensi mengalami kasus kusta. Dengan demikian, hubungan positif antara jumlah penduduk laki-laki dan jumlah kasus dapat mencerminkan perbedaan ukuran populasi antarwilayah selain kemungkinan adanya perbedaan karakteristik sosial, ekonomi, mobilitas, maupun akses terhadap pelayanan kesehatan. Penjelasan tersebut merupakan interpretasi substantif yang perlu dipandang sebagai kemungkinan dan bukan sebagai bukti bahwa jenis kelamin laki-laki secara kausal menyebabkan terjadinya kusta.

Temuan mengenai jumlah penduduk laki-laki juga dapat dibandingkan dengan penelitian sebelumnya. Aipassa et al. (2023) menunjukkan bahwa karakteristik demografis merupakan salah satu aspek yang perlu dipertimbangkan dalam pemodelan jumlah kasus kusta menggunakan regresi binomial negatif. Hasil penelitian ini memberikan tambahan perspektif bahwa ketika data dianalisis pada tingkat kabupaten/kota dan respons yang digunakan berupa jumlah kasus, ukuran populasi suatu kelompok demografis dapat berkaitan dengan variasi jumlah kasus antarwilayah. Perbedaan hasil antarpelitian dapat dipengaruhi oleh karakteristik wilayah, periode pengamatan, definisi variabel, ukuran populasi, serta metode statistik yang digunakan.

Sementara itu, variabel independen lainnya yang tidak menunjukkan signifikansi statistik dalam model tidak dapat langsung diartikan bahwa variabel tersebut tidak memiliki hubungan dengan kejadian kusta. Ketidaksignifikanan statistik dalam penelitian ini menunjukkan bahwa berdasarkan data dan spesifikasi model yang digunakan, belum terdapat bukti yang cukup untuk menyatakan adanya hubungan yang signifikan antara variabel tersebut dan jumlah kasus kusta pada taraf signifikansi yang digunakan. Kondisi tersebut dapat berkaitan dengan variasi antarwilayah yang relatif terbatas, ukuran sampel yang kecil, keterkaitan antarvariabel penjelas, maupun adanya faktor lain yang tidak dimasukkan ke dalam model. Oleh karena itu, interpretasi terhadap variabel yang tidak signifikan perlu dibatasi pada konteks data penelitian ini.

Penggunaan fungsi penghalus P-spline dalam model dilakukan untuk mengakomodasi kemungkinan hubungan nonlinear yang teridentifikasi pada tahap pengujian sebelumnya. Hasil estimasi komponen penghalus menunjukkan nilai effective degrees of freedom (edf) sebesar 1. Nilai tersebut menunjukkan bahwa bentuk hubungan yang diestimasi oleh fungsi penghalus cenderung mendekati hubungan linear. Temuan ini menunjukkan bahwa fleksibilitas GAMM tetap dapat digunakan untuk mengevaluasi kemungkinan nonlinearitas, meskipun hasil akhir fungsi penghalus pada data penelitian ini tidak menunjukkan pola nonlinear yang kuat.

Selain efek tetap, model juga memasukkan random effect kabupaten/kota. Nilai standar deviasi efek acak sebesar 0,3103 menunjukkan adanya variasi antar kabupaten/kota setelah mempertimbangkan variabel-variabel tetap dalam model. Hasil uji simultan yang menghasilkan p-value sebesar 0,0226 juga menunjukkan bahwa terdapat bukti statistik bahwa komponen model secara keseluruhan memberikan kontribusi terhadap variasi respons. Keberadaan random effect menunjukkan bahwa karakteristik masing-masing kabupaten/kota masih memiliki variasi yang belum sepenuhnya dijelaskan oleh variabel independen yang digunakan.

Variasi tersebut tidak serta-merta menunjukkan adanya efek spasial. Random effect kabupaten/kota dalam penelitian ini merepresentasikan heterogenitas antar kelompok wilayah, sedangkan efek spasial secara khusus berkaitan dengan ketergantungan atau

hubungan berdasarkan kedekatan geografis antarwilayah. Oleh karena itu, hasil penelitian ini lebih tepat diinterpretasikan sebagai adanya heterogenitas antar kabupaten/kota, bukan sebagai bukti adanya autokorelasi spasial. Analisis spasial secara eksplisit dapat menjadi pengembangan penelitian selanjutnya dengan memasukkan informasi lokasi dan struktur ketetanggaan antarwilayah.

Secara keseluruhan, hasil penelitian menunjukkan bahwa pemodelan GAMM memberikan kerangka yang memungkinkan karakteristik data jumlah kasus, kemungkinan hubungan nonlinear, dan variasi antarwilayah dipertimbangkan secara bersamaan. Namun, hasil penelitian tetap perlu dipahami dalam keterbatasan jumlah observasi yang hanya mencakup 13 kabupaten/kota selama tiga tahun. Selain itu, karena penelitian menggunakan data observasional tingkat wilayah, hasil yang diperoleh menggambarkan hubungan statistik pada tingkat kabupaten/kota dan tidak dapat digunakan untuk menyimpulkan hubungan kausal pada tingkat individu.

4. SIMPULAN

Berdasarkan hasil pemodelan Generalized Additive Mixed Model pada jumlah kasus kusta di Sulawesi Tengah tahun 2021–2023, diperoleh model dengan distribusi binomial negatif, fungsi penghalus P-spline, dan random effect kabupaten/kota. Hasil pengujian parsial menunjukkan bahwa jumlah penduduk laki-laki merupakan variabel yang memiliki hubungan signifikan dan positif dengan jumlah kasus kusta. Koefisien positif menunjukkan bahwa peningkatan jumlah penduduk laki-laki berkaitan dengan peningkatan rata-rata jumlah kasus kusta dalam konteks data dan variabel yang digunakan dalam penelitian ini.

Komponen random effect kabupaten/kota menunjukkan adanya variasi antarwilayah yang belum sepenuhnya dijelaskan oleh variabel independen dalam model. Variasi tersebut menunjukkan heterogenitas antar kabupaten/kota dan tidak secara langsung dapat diinterpretasikan sebagai efek spasial atau ketergantungan geografis. Oleh karena itu, hasil penelitian ini menunjukkan adanya hubungan statistik antara variabel jumlah penduduk laki-laki dan jumlah kasus kusta serta adanya variasi antarwilayah, tetapi tidak dimaksudkan untuk menyimpulkan hubungan sebab-akibat.

5. UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada semua pihak yang telah memberikan kontribusi dalam penelitian ini. Terima kasih juga disampaikan kepada pihak Dinas Kesehatan Provinsi Sulawesi Tengah dan Badan Pusat Statistik yang telah menyediakan data sekunder secara transparan bagi publik, sehingga analisis data dalam penelitian ini dapat diselesaikan dengan baik..

6. REKOMENDASI

Penelitian selanjutnya disarankan untuk menggunakan pendekatan *Spatio-Temporal* GAMM guna mengakomodasi variasi waktu dan lokasi secara bersamaan, terutama mengingat penyebaran kasus kusta cenderung bersifat geografis. Peneliti dapat membandingkan performa akurasi model GAMM ini dengan metode berbasis *machine learning* lainnya seperti *Random Forest* atau *Gradient Boosting* untuk mendapatkan performa prediksi terbaik.

7. REFERENSI

- Aipassa, A. D., Wattimena, A. Z., & Haumahu, G. (2023). Analisis Faktor-Faktor yang Mempengaruhi Jumlah Penyakit Kusta Di Provinsi Maluku dengan Menggunakan Regresi Binomial Negatif. *PARAMETER*, 02(02), 87–100.
- Andreas, C., Horidah, H., Sulistiana, R., Venosia, D., & Chamidah, N. (2022). Pemodelan Prevalensi Penyakit Kusta Di Jawa Timur Dengan Pendekatan Geographically Weighted Regression. *Aplikasi Statistika & Komputasi Statistika*, 14(1), 33–48.
- Azwarini, R., & Purhadi. (2021). Pemodelan Jumlah Kasus Kusta di Kabupaten Mojokerto dan Kabupaten Jombang Tahun 2019 Menggunakan Regresi Zero-Inflated Poisson Inverse Gaussian. *INFERENSI*, 4(2), 85–97.
- Bappenas. (2020). *Pedoman Teknis Penyusunan Rencana Aksi: Tujuan Pembangunan Berkelanjutan (TPB)/Sustainable Development Goals (SDGs) (II)*. BAPPENAS. <https://sdgs.jakarta.go.id/storage/01JYNJGJQ0V4860BF4FH9GNF4Z.pdf>
- Bolker, B. M., Brooks, M. E., Clark, C. J., Geange, S. W., Poulsen, J. R., Stevens, M. H. H., & White, J.-S. S. (2009). Generalized linear mixed models: A practical guide for ecology and evolution. *Trends in Ecology & Evolution*, 24(3), 127–135. <https://doi.org/10.1016/j.tree.2008.10.008>
- Budiharti, L. M., & Sunendiari, S. (2021). Pemodelan dan Pemetaan Jumlah Penderita Kusta di Jawa Barat dengan Regresi Binomial Negatif dan Flexibly Shaped Spatial Scan Statistic. *Riset Statistika*, 1(2), 99–106. <https://api.semanticscholar.org/CorpusID:238782417>
- Cameron, A. C., & Trivedi, P. K. (2013). *Regression Analysis of Count Data (2nd ed.)*. Cambridge University Press.
- Darsyah, M. Y., & Ramadhan, M. N. (2022). Pemodelan Jumlah Kasus Penyakit Kusta Di Provinsi Sulawesi Tenggara Menggunakan Metode Regresi Poisson Inverse Gaussian. *Jurnal Litbang Edusaintech*, 3(1), 11–25. <https://doi.org/10.51402/jle.v3i1.9>
- Dhahari, N. M., & Sofro, A. (2024). Analisis Regresi Binomial Negatif Untuk Pemodelan Angka Positif Penyakit Kusta Di Jawa Timur. *Math Unesa*, 12(01), 49–56.
- Eilers, P. H. C., & Marx, B. D. (1996). Flexible Smoothing with B-Splines and Penalties.

- Statistical Science*, 11(2), 89–121. <https://doi.org/10.1214/ss/1038425655>
- Hastie, T., & Tibshirani, R. (1990). *Generalized Additive Models*. Chapman and Hall.
- Hilbe, J. M. (2011). *Negative Binomial Regression (2nd ed.)* (Second). Cambridge University Press.
- Jamco, J. C., Kondolembang, F., & Noya Van Delsen, M. S. (2023). Penanganan multikolinearitas pada regresi linier berganda menggunakan regresi LASSO (Studi kasus: Distribusi persentase Produk Domestik Regional Bruto di Provinsi Maluku tahun 1999–2021). *Parameter: Jurnal Matematika, Statistika Dan Terapannya*, 2(2), 145–154.
- Kemenkes, R. (2023). *Profil Kesehatan Indonesia 2023*.
- Wood, S. N. (2017). *Generalized Additive Models: An Introduction with R (2nd ed.)* (Second). Chapman and Hall.