

KAJIAN ANALISIS SPASIAL PADA DATA CACAH BIRESPO DENGAN OVERDISPERSI DAN NOL BERLEBIH

YUFAN PUTRI ASTRINI



PROGRAM STUDI MAGISTER STATISTIKA DAN SAINS DATA
SEKOLAH SAINS DATA, MATEMATIKA, DAN INFORMATIKA
INSTITUT PERTANIAN BOGOR
BOGOR
2026



@Hak cipta milik IPB University

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



PERNYATAAN MENGENAI TESIS DAN SUMBER INFORMASI SERTA PELIMPAHAN HAK CIPTA

Dengan ini saya menyatakan bahwa tesis dengan judul “Kajian Analisis Spasial pada Data Cacah Birespon dengan Overdispersi dan Nol Berlebih” adalah karya saya dengan arahan dari dosen pembimbing dan belum diajukan dalam bentuk apa pun kepada perguruan tinggi mana pun. Sumber informasi yang berasal atau dikutip dari karya yang diterbitkan maupun tidak diterbitkan dari penulis lain telah disebutkan dalam teks dan dicantumkan dalam Daftar Pustaka di bagian akhir tesis ini.

Dengan ini saya melimpahkan hak cipta dari karya tulis saya kepada Institut Pertanian Bogor.

Bogor, Juli 2026

Yufan Putri Astrini
G1501222051



@Hak cipta milik IPB University

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



@Hak cipta milik IPB University

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



RINGKASAN

YUFAN PUTRI ASTRINI. Kajian Analisis Spasial pada Data Cacah Birespon dengan Overdispersi dan Nol Berlebih. Dibimbing oleh BUDI SUSETYO dan KUSMAN SADIK.

Data cacah sering digunakan untuk merepresentasikan jumlah kejadian pada berbagai bidang, seperti kesehatan, pertanian, demografi, dan lain sebagainya. Pemodelan data cacah biasanya menggunakan regresi Poisson, tetapi model tersebut kurang sesuai ketika data mengalami overdispersi, yaitu kondisi ketika ragam lebih besar daripada rataannya. Permasalahan menjadi lebih kompleks apabila data juga mengandung nilai nol berlebih, terdiri atas dua peubah respon yang saling berkaitan, serta memiliki keragaman spasial. *Bivariate Zero-Inflated Generalized Poisson Regression* (BZIGPR) dapat digunakan untuk memodelkan data cacah birespon yang mengalami overdispersi dan nol berlebih, tetapi model ini menghasilkan parameter yang berlaku secara global. *Geographically Weighted Bivariate Zero-Inflated Generalized Poisson Regression* (GWBZIGPR) mengembangkan model tersebut dengan membentuk parameter lokal pada setiap lokasi sehingga mampu mengakomodasi keragaman spasial. Fleksibilitas model lokal berpotensi meningkatkan ketepatan prediksi, tetapi juga menyebabkan kompleksitas model yang lebih tinggi.

Penelitian ini bertujuan mengevaluasi kinerja BZIGPR dan GWBZIGPR berdasarkan *Akaike Information Criterion Corrected* (AICc) dan *Root Mean Square Error* (RMSE) pada data yang memiliki keragaman spasial. Penelitian dilakukan melalui kajian simulasi dengan tiga ukuran contoh, yaitu 50, 100, dan 200, serta tiga skenario proporsi nol teramati birespon, yaitu (0,3; 0,3), (0,3; 0,6), dan (0,6; 0,6). Setiap skenario dilakukan sebanyak 30 ulangan. Parameter dispersi ditetapkan sebesar 0,3 dan parameter ketergantungan sebesar 0,6. Koordinat spasial dibangkitkan secara acak dan digunakan untuk membentuk parameter regresi lokal, sehingga peubah penjelas dan peubah respon berbeda menurut lokasi. Ketergantungan antara kedua respon dibentuk menggunakan *Gaussian Copula*. Data hasil simulasi divalidasi berdasarkan kesesuaian proporsi nol, tingkat overdispersi, korelasi antar respon, dan keragaman spasial.

Pendugaan parameter BZIGPR dan GWBZIGPR dilakukan menggunakan metode *Maximum Likelihood Estimation* dengan algoritma optimasi *Limited-memory Broyden-Fletcher-Goldfarb-Shanno with Bound constraints* (L-BFGS-B). Pada GWBZIGPR, pendugaan dilakukan secara lokal menggunakan pembobot *Adaptive Gaussian Kernel*, sedangkan *bandwidth* optimum ditentukan berdasarkan *Generalized Cross Validation* (GCV). Evaluasi model dilakukan berdasarkan *Akaike Information Criterion Corrected* (AICc) dan *Root Mean Square Error* (RMSE). AICc digunakan untuk menilai kecocokan model dengan mempertimbangkan kompleksitas, sedangkan RMSE digunakan untuk mengukur ketepatan prediksi respon.

Hasil penelitian menunjukkan bahwa berdasarkan AICc, BZIGPR menghasilkan nilai lebih rendah daripada GWBZIGPR. Pada $n = 50$, selisih AICc berkisar antara 36,91-42,48, sedangkan pada $n = 200$ mengecil menjadi 3,47-8,50. Hasil uji ART ANOVA menyimpulkan bahwa model, ukuran contoh, dan proporsi nol masing-masing berbeda nyata terhadap AICc ($p < 0,001$) Interaksi antara



model dan ukuran contoh juga nyata dengan $F = 11,84$ dan $p = 0,01$, tetapi interaksi model dan proporsi nol tidak nyata ($p = 0,91$). Berdasarkan RMSE gabungan, GWBZIGPR memberikan hasil akurasi yang lebih baik dengan median 4,60 dibandingkan dengan BZIGPR sebesar 4,97. Perbedaan ini signifikan dengan nilai $p < 0,001$. Model ($F=33,23$), ukuran contoh ($F=7,76$), dan proporsi nol ($F=678,79$) berbeda nyata terhadap nilai RMSE, tetapi interaksinya tidak berbeda nyata. Evaluasi kinerja menurut AICc menghasilkan kesimpulan bahwa BZIGPR memberikan performa yang lebih baik berdasarkan keseimbangan kecocokan dan kompleksitas model. Namun, GWBZIGPR lebih unggul daripada BZIGPR dari sisi meningkatkan ketepatan prediksi pada data yang memiliki keragaman spasial.

Kata kunci: keragaman spasial, nol berlebih, cacah birespon, BZIGPR, GWBZIGPR.



SUMMARY

YUFAN PUTRI ASTRINI. Spatial Study of Overdispersed Bivariate Count Data with Excess Zeros. Supervised by BUDI SUSETYO and KUSMAN SADIK.

Count data are widely used to represent the number of events occurring in various fields, including health, agriculture, and demography. Such data are commonly analyzed using Poisson regression; however, this model is inappropriate when overdispersion is present, that is, when the variance exceeds the mean. The modeling problem becomes more complex when the data also contain excess zeros, involve two related response variables, and exhibit spatial heterogeneity. Bivariate Zero-Inflated Generalized Poisson Regression (BZIGPR) can accommodate overdispersed bivariate count data with excess zeros, but it assumes that the regression parameters are constant across all locations. Geographically Weighted Bivariate Zero-Inflated Generalized Poisson Regression (GWBZIGPR) extends this model by estimating location-specific parameters, thereby allowing spatial heterogeneity to be incorporated. Although this local modeling framework may improve predictive accuracy, it also increases model complexity.

This study aimed to evaluate the performance of BZIGPR and GWBZIGPR using the corrected Akaike information criterion (AICc) and root mean square error (RMSE) for data with spatial heterogeneity. A simulation study was conducted using three sample sizes, namely 50, 100, and 200, and three scenarios of observed zero proportions for the two responses: (0.3, 0.3), (0.3, 0.6), and (0.6, 0.6). Each scenario was replicated 30 times. The dispersion parameter was fixed at 0.3, while the dependence parameter was set at 0.6. Spatial coordinates were generated randomly and used to construct local regression parameters, allowing the relationships between the explanatory variables and the response variables to vary across locations. Dependence between the two responses was generated using a Gaussian copula. The simulated data were validated in terms of the observed zero proportions, degree of overdispersion, correlation between responses, and spatial heterogeneity.

The parameters of both BZIGPR and GWBZIGPR were estimated using maximum likelihood estimation with the Limited-memory Broyden–Fletcher–Goldfarb–Shanno algorithm with bound constraints (L-BFGS-B). For GWBZIGPR, local estimation was performed using an adaptive Gaussian kernel, and the optimal bandwidth was selected using generalized cross-validation (GCV). Model performance was assessed using AICc and RMSE. AICc was used to evaluate model fit while accounting for model complexity, whereas RMSE was used to measure the predictive accuracy of the response estimates.

The results showed that BZIGPR consistently produced lower AICc values than GWBZIGPR. At $n=50$, the differences in AICc ranged from 36.91 to 42.48, whereas at $n=200$, they decreased to between 3.47 and 8.50. The ART ANOVA results indicated that model type, sample size, and zero proportion each had a significant effect on AICc ($p<0.001$). The interaction between model type and sample size was also significant ($F=11.84$, $p=0.01$), whereas the interaction between model type and zero proportion was not significant ($p=0.91$). In terms of combined RMSE, GWBZIGPR achieved greater predictive accuracy, with a median RMSE of 4.60 compared with 4.97 for BZIGPR. This difference was

@Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

statistically significant ($p < 0.001$). Model type ($F = 33.23$), sample size ($F = 7.76$), and zero proportion ($F = 678.79$) each had a significant effect on RMSE, while their interaction effects were not significant. Overall, BZIGPR performed better according to AICc because it provided a more favorable balance between model fit and complexity. In contrast, GWBZIGPR was superior in terms of predictive accuracy for data exhibiting spatial heterogeneity.

Keywords: spatial heterogeneity, excess zeros, bivariate count data, BZIGPR, GWBZIGPR





Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

© Hak Cipta milik IPB, tahun 2026
Hak Cipta dilindungi Undang-Undang

Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan atau menyebutkan sumbernya. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik, atau tinjauan suatu masalah, dan pengutipan tersebut tidak merugikan kepentingan IPB.

Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apa pun tanpa izin IPB.

KAJIAN ANALISIS SPASIAL PADA DATA CACAH BIRES PON DENGAN OVERDISPERSI DAN NOL BERLEBIH

YUFAN PUTRI ASTRINI

Tesis
sebagai salah satu syarat untuk memperoleh gelar
Magister pada
Program Studi Statistika dan Sains Data

**PROGRAM STUDI MAGISTER STATISTIKA DAN SAINS DATA
SEKOLAH SAINS DATA, MATEMATIKA, DAN INFORMATIKA
INSTITUT PERTANIAN BOGOR
BOGOR
2026**

@Hak cipta milik IPB University

IPB University





@Hak cipta milik IPB University

IPB University

Penguji Luar Komisi Pembimbing pada Ujian Tesis:
Prof. Dr. Ir. Muhammad Nur Aidi, M.S.



IPB University
— Bogor Indonesia —

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

Perpustakaan IPB University



Judul Tesis : Kajian Analisis Spasial pada Data Cacah Birespon dengan
Overdispersi dan Nol Berlebih
Nama : Yufan Putri Astrini
NIM : G1501222051

Disetujui oleh

Pembimbing 1:
Dr. Ir. Budi Susetyo, M.S.

Pembimbing 2:
Dr. Kusman Sadik, S.Si., M.Si.

Diketahui oleh

Ketua Program Studi:
Dr. Agus Mohamad Soleh, S.Si., M.T.
NIP 197503151999031004

Dekan Sekolah Sains Data, Matematika, dan Informatika
Prof. Dr. Ir. Agus Buono, M.Si., M. Kom.
NIP 196607021993021001

Tanggal Ujian: 27 Juli 2026

Tanggal Lulus:



@Hak cipta milik IPB University

IPB University



IPB University
— Bogor Indonesia —

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

Perpustakaan IPB University

PRAKATA

Puji syukur penulis panjatkan ke hadirat Tuhan Yang Maha Esa atas segala rahmat dan karunia-Nya sehingga tesis berjudul “Kajian Analisis Spasial pada Data Cacah Birespon dengan Overdispersi dan Nol Berlebih” dapat diselesaikan dengan baik. Penyusunan tesis ini tidak terlepas dari bantuan, dukungan, dan arahan dari berbagai pihak. Oleh karena itu, penulis mengucapkan terima kasih kepada:

- 1 Bapak Dr. Ir. Budi Susetyo, M.S. dan Dr. Kusman Sadik, S.Si., M.Si. selaku komisi pembimbing tesis yang telah memberikan arahan, masukan, dan bimbingan selama proses penyusunan tesis ini.
- 2 Dr. Ir. Indahwati, M.Si. selaku pemimpin ujian tesis dan Prof. Dr. Ir. Muhammad Nur Aidi, M.S. selaku penguji luar komisi pembimbing, yang telah membantu lancarnya proses penyelesaian tugas akhir penulis.
- 3 Seluruh dosen dan tenaga kependidikan Program Studi Statistika dan Sains Data atas ilmu, pengalaman, dan dukungan yang diberikan selama masa perkuliahan.
- 4 Orang tua, suami, dan keluarga penulis yang senantiasa memberikan doa, dukungan, dan motivasi.
- 5 Teman-teman Magister Statistika dan Sains Data angkatan 2022 atas kebersamaan, diskusi, dan dukungan selama proses studi hingga penyusunan tesis ini.
- 6 Semua pihak yang telah membantu, baik secara langsung maupun tidak langsung, dalam penyelesaian tesis ini.

Semoga tesis ini dapat memberikan manfaat bagi pengembangan ilmu pengetahuan dan bagi pihak-pihak yang membutuhkan.

Bogor, Juli 2026

Yufan Putri Astrini



@Hak cipta milik IPB University

IPB University



IPB University
— Bogor Indonesia —

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

Perpustakaan IPB University



DAFTAR ISI

DAFTAR TABEL	iv
DAFTAR GAMBAR	iv
DAFTAR LAMPIRAN	iv
I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Tujuan	3
II TINJAUAN PUSTAKA	4
2.1 Overdispersi	4
2.2 Keragaman Spasial	4
2.3 Model <i>Zero-Inflated</i>	5
2.4 Model <i>Bivariate Zero-inflated Generalized Poisson</i>	5
2.5 Matriks Pembobot Spasial	9
2.6 Regresi <i>Geographically Weighted Bivariate Zero-Inflated Generalized Poisson</i>	10
2.7 <i>Gaussian Copula</i>	11
III METODE	13
3.1 Pembangkitan Data Simulasi	13
3.2 Analisis Data	16
IV HASIL DAN PEMBAHASAN	20
4.1 Validasi Karakteristik Data Simulasi	20
4.2 Evaluasi Model BZIGPR dan GWBZIGPR	22
V SIMPULAN DAN SARAN	29
5.1 Simpulan	29
5.2 Saran	29
DAFTAR PUSTAKA	30
LAMPIRAN	33
RIWAYAT HIDUP	37



DAFTAR TABEL

7	Skenario simulasi	13
8	Rata-rata AICc model BZIGPR dan GWBZIGPR pada sembilan skenario simulasi	22
9	ART ANOVA dari AICc berdasarkan model, ukuran contoh, dan proporsi nol	23
10	Hasil uji lanjut ART-C pada interaksi model dan ukuran contoh berdasarkan AICc	24
11	Rata-rata RMSE model BZIGPR dan GWBZIGPR pada sembilan skenario simulasi	25
12	ART ANOVA dari RMSE gabungan berdasarkan model, ukuran contoh, dan proporsi nol	26
13	Hasil uji <i>Wilcoxon Signed-Rank</i> rata-rata RMSE gabungan model BZIGPR dan GWBZIGPR	27

DAFTAR GAMBAR

14	Diagram alir tahapan data simulasi	19
15	Proporsi nol teramati dan target pada setiap skenario simulasi	20
16	Sebaran <i>Variance to Mean Ratio</i> (VMR) pada setiap skenario	20
17	Sebaran korelasi antara Y1 dan Y2	21
18	Hasil uji keragaman spasial <i>Koenker-Basset</i>	22
19	Perbandingan rata-rata AICc pada model BZIGPR dan GWBZIGPR	23
20	Plot interaksi antara model dan ukuran contoh pada AICc	24
21	Sebaran rata-rata RMSE Gabungan (A), RMSE Y1 (B), dan RMSE Y2 (C) pada sembilan skenario simulasi	26
22	Plot interaksi proporsi pada model BZIGPR dan GWBZIGPR	27
23	Plot interaksi ukuran contoh pada model BZIGPR dan GWBZIGPR	28

DAFTAR LAMPIRAN

24	Hasil uji skor nol berlebih	34
25	Hasil uji overdispersi	34
26	Hasil uji keragaman spasial <i>Koenker-Basset</i>	35
27	Ringkasan pemilihan <i>bandwidth</i> adaptif model GWBZIGPR	35