

## PERBANDINGAN METODE *REGRESI SPATIAL AUTO REGRESSIVE MODEL* DAN *SPATIAL ERROR MODEL* UNTUK KEJADIAN DIARE DI PROVINSI BALI

**Luh Putu Safitri Pratiwi\***

ITB STIKOM Bali, Denpasar, Bali, Indonesia, 80111

**Abstrak.** Untuk mengendalikan laju pertumbuhan kasus diare, langkah yang dapat diambil adalah menganalisis faktor-faktor yang berkaitan dengan diare yang memiliki potensi untuk meningkatkan kasus diare dari perspektif wilayah. Pemahaman tentang pengaruh daerah atau pengaruh geografis dalam pemodelan dikenal sebagai pemodelan spasial. Teridentifikasi bahwa terdapat dua varian umum dari model spasial yang sering digunakan, meliputi *Spatial Autoregressive Model (SAR)* dan *Spatial Error Model (SEM)*. Tujuan dari penelitian ini adalah untuk melihat model terbaik tingkat kejadian diare di wilayah Bali dengan menerapkan metode analisis spasial seperti *SAR* dan *SEM*. Hasil penelitian menunjukkan bahwa data yang digunakan dalam analisis memiliki elemen spasial yang signifikan, sehingga pemodelan yang mengintegrasikan aspek spasial menjadi penting. Dalam hal ini, metode *SEM* terbukti sebagai model terbaik, jika dibandingkan dengan *OLS* dan *SAR*. Nilai  $R^2$  Model *SEM* sebesar 96,01% dan nilai *AIC* sebesar 267,558. Sehingga hasil ini menunjukkan bahwa model *SEM* lebih tepat dalam menjelaskan hubungan antara variabel-variabel yang terlibat dalam penelitian ini.

**Kata Kunci:** metode regresi, spasial auto regresif, spasial error model

**Abstract.** To control the growth rate of diarrheal cases, one of the steps that can be taken is to analyze factors related to diarrhea and their potential to contribute to an increase in diarrhea cases from a regional perspective. Understanding the influence of regions or geographic factors in modeling is known as spatial modeling. It has been identified that there are two common variants of spatial models often used, which include the *Spatial Autoregressive Model (SAR)* and the *Spatial Error Model (SEM)*. The objective of this research is to determine the best model for the incidence of diarrhea in the Bali region by applying spatial analysis methods such as *SAR* and *SEM*. The research results indicate that the data used in the analysis contain significant spatial elements, highlighting the importance of integrating spatial aspects into modeling. In this context, the *SEM* method has proven to be the most suitable model when compared to *Ordinary Least Squares (OLS)* and *SAR*. The  $R$ -squared value for the *SEM* model is 96.01%, and it achieves the lowest *AIC* value of 267.558. Therefore, these findings demonstrate that the *SEM* model is more accurate in explaining the relationships among the variables involved in this study.

**Keywords:** regression method, spatial auto regressive, spatial error model

Sitasi: Pratiwi, L.P.S. 2023. Perbandingan Metode Regresi Spasial Auto Regressive Model dan Spatial Error Model Untuk Kejadian Diare di Provinsi Bali. *MES (Journal of Mathematics Education and Science, 9(1): 67-72.*

Submit: 10 Oktober 2023	Revisi: 19 Oktober 2023	Publish: 21 Oktober 2023
----------------------------	----------------------------	-----------------------------

### PENDAHULUAN

Pemahaman tentang pengaruh daerah atau pengaruh geografis dalam pemodelan dikenal sebagai pemodelan spasial. Teridentifikasi bahwa terdapat dua varian umum dari model spasial yang sering digunakan, meliputi *Spatial Autoregressive Model (SAR)*, *Spatial Error Model (SEM)*, serta kombinasi model antara *SAR* dan *SEM*. Dalam merancang pemodelan *SAR* dan *SEM*, dibutuhkan data yang merujuk pada posisi, objek, dan interaksi di dalam ruang geografis,

\*Corresponding Author: [putu\\_safitri@stikom-bali.ac.id](mailto:putu_safitri@stikom-bali.ac.id)

dikenal sebagai data spasial. Aspek yang paling esensial dari data spasial adalah matriks pembobot, yang memperlihatkan relasi antara berbagai wilayah di dalamnya (Grasa,1989).

Langkah yang digunakan dalam menghasilkan matriks pembobot melibatkan pendekatan berdasarkan kedekatan geografis antar wilayah (LeSage, 1999) dan penyesuaian bobot berdasarkan variabel lain (customized), sebagaimana dilakukan dalam studi yang dilakukan oleh (Pratiwi et al. 2020).

Penelitian sebelumnya yang dilakukan oleh (Rahmawati & Bimanto, 2021) telah menginvestigasi pemodelan SAR dan SEM. Hasil dari penelitian tersebut mengindikasikan adanya ketergantungan dalam bentuk dependensi lag dan error, diukur melalui uji Lagrange Multiplier. Selain itu, ditemukan bahwa model SEM lebih cocok untuk menganalisis Indeks Pembangunan Manusia (IPM) di Provinsi Jawa Timur dibandingkan dengan model SAR dan regresi OLS. Hal ini terbukti dari nilai  $R^2$  tertinggi mencapai 0,998 dan nilai AIC terendah sebesar 2,345

SAR dan SEM dapat diaplikasikan pada bidang kesehatan salah satunya mengenai persebaran kejadian diare pada suatu daerah. Diare/buang air adalah penyakit yang ditandai dengan buang air besar berbentuk tinja encer atau cair, kadang-kadang bercampur darah atau lendir, yang umumnya terjadi 3 kali atau lebih dalam 24 jam. Diare dapat disertai dengan muntah-muntah, maupun penurunan kesadaran. Istilah lainnya adalah mencret atau bocor (BPS,2020). Menurut Profil Kesehatan Provinsi Bali Tahun 2021, angka kesakitan diare di Provinsi Bali dilaporkan sebanyak 102.197 jiwa pada semua umur, tertinggi yaitu di Kota Denpasar sebesar 26.932 dan terendah di Kabupaten Buleleng sebesar 1.951. Sedangkan, persentase cakupan pelayanan yaitu Kabupaten Tabanan persentasenya sebesar (33,60%), Kabupaten Kelungkung sebesar (74,47%), Kabupaten Bangli sebesar (28,70%), Kabupaten Gianyar sebesar (34,20%), Kabupaten Jembrana (63,8%), Kabupaten Karangasem sebesar (55,6%), Kabupaten Buleleng sebesar (32,50%), Kabupaten Jembrana sebesar (23,80%), kabupaten Karangasem sebesar (26,60%), dan terendah adalah di Kota Denpasar dan Kabupaten Badung yaitu hanya sebesar (18,70%) (Dinkes Bali, 2021).

Untuk mengendalikan laju pertumbuhan kasus diare, langkah yang dapat diambil adalah menganalisis faktor-faktor yang berkaitan dengan diare dan memiliki potensi untuk meningkatkan kasus diare dari perspektif wilayah. Beberapa faktor yang berpengaruh pada jumlah kasus diare di Provinsi Bali meliputi tingkat kepadatan penduduk dan ketersediaan sarana sanitasi di setiap kecamatan di wilayah tersebut (Sulasih et al. 2021), pelayanan kesehatan yang tersedia, tingkat akses rumah tangga terhadap sanitasi yang memadai, ketersediaan sumber air minum yang bersih, fasilitas tempat buang air besar bersama di rumah tangga, dan masalah timbunan sampah (Mardhatillah, 2020).

Penyebaran kasus diare sesuai untuk dianalisis menggunakan metode spasial karena kejadian diare di suatu wilayah dapat dipengaruhi oleh wilayah-wilayah tetangga. Dalam hal ini, jarak antara setiap wilayah memiliki dampak signifikan pada interaksi antar wilayah tersebut. Oleh karena itu, kita perlu untuk mengidentifikasi faktor-faktor dari wilayah-wilayah terdekat yang berkontribusi pada penurunan jumlah kasus diare di suatu daerah tertentu dan mengukur sejauh mana pengaruhnya. Untuk tujuan ini, diperlukan pendekatan pemodelan spasial berbasis area seperti SAR dan SEM untuk menganalisis faktor-faktor yang berperan signifikan dalam mengurangi kasus diare di Provinsi Bali.

## **METODE**

Pelaksanaan penelitian ini dilakukan di wilayah provinsi Bali yang mencakup keseluruhan 56 Kecamatan. Variable penelitian yang digunakan sebagai berikut: Variabel dependen pada penelitian ini yaitu jumlah kasus diare disetiap kecamatan di Provinsi Bali (Y).

Variabel independen pada penelitian ini adalah kepadatan penduduk ( $X_1$ ), jumlah sarana air minum ( $X_2$ ), jumlah sarana sanitasi ( $X_3$ ).

Adapun tahapan analisis data yang akan dilakukan dalam penelitian ini adalah sebagai berikut:

1. Melakukan deskripsi sebagai gambaran awal untuk mendeskriptifkan data kejadian diare di Provinsi Bali
2. Melakukan Pemodelan variabel kejadian diare di Provinsi Bali dan faktor-faktor memengaruhinya.
3. Melakukan pemodelan *Ordinary Least Square* (OLS), terdiri dari:
4. Memeriksa aspek spasial pada data yang digunakan
5. Pemodelan Regresi Spasial kejadian diare digunakan untuk mengidentifikasi dependensi spasial yang lebih spesifik
6. Selanjutnya, pemodelan SAR dan SEM dilakukan untuk menghasilkan model regresi terbaik dengan kriteria AIC dalam analisis kejadian diare.

### 1. Regresi Spasial

Anselin (1988) mengemukakan bahwa model regresi yang memperhitungkan dampak dari relasi spasial dikenal sebagai model regresi spasial. Salah satu bentuk relasi spasial adalah autokorelasi spasial, di mana keberadaan autokorelasi spasial menghasilkan pembentukan parameter spasial autoregresif dan *moving average*. Dengan demikian, terbentuklah sebuah proses spasial seperti yang diuraikan berikut.

$$y = \rho W_1 y + X\beta + u \quad (1)$$

$$u_t = \lambda W_2 u_{t-1} + \varepsilon \quad (2)$$

$$\varepsilon \sim N(0, \sigma^2 I)$$

sehingga model umum yang terbentuk adalah:

$$y = \rho W_1 y + X\beta + \lambda W_2 u + \varepsilon \quad (3)$$

### 2. Pemilihan Model Terbaik

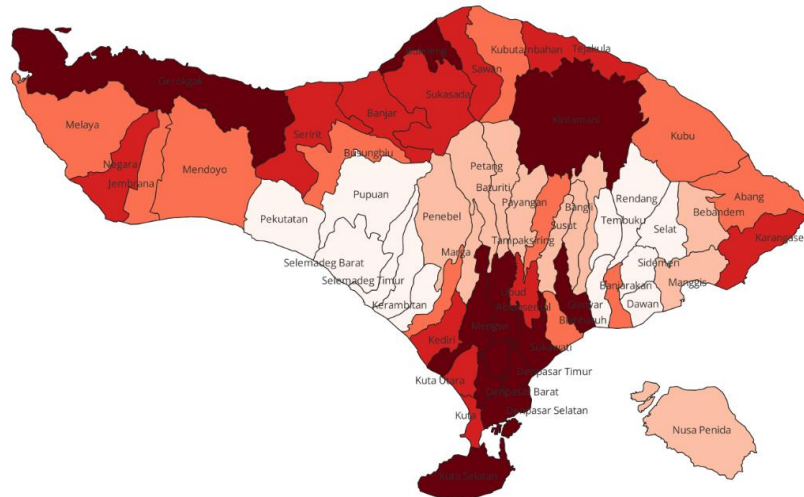
Dalam memilih model yang optimal untuk mendukung penelitian ini, dilakukan evaluasi berdasarkan kriteria yang sesuai. Salah satu kriteria yang digunakan adalah Kriteria Informasi *Akaike's Information Criterion* (AIC) yang disesuaikan. Persamaan yang merepresentasikan AIC adalah:

$$AIC(k^*) = -2 \frac{\ln L(k^*)}{n} + 2 \frac{k^*}{n} \quad (4)$$

Di mana  $L$  ( $L(k^*)$ ) mewakili nilai maksimum likelihood yang melibatkan  $k$  variabel independen,  $k$  adalah jumlah parameter, dengan  $k^* = 1, 2, \dots, k$ , dan  $n$  adalah ukuran sampel. Model yang dipilih berdasarkan kriteria AIC adalah model dengan nilai AIC terendah (Burnham & Anderson, 2002).

## HASIL DAN PEMBAHASAN

Hasil penelitian mengindikasikan bahwa daerah-daerah dengan angka kejadian diare tertinggi terletak di Gianyar, Gerogak, Kintamani, Sukawati, Abiansemal, Buleleng, Kuta Selatan, Mengwi, serta berbagai bagian Timur, Utara, Barat, dan Selatan di Denpasar. Di sisi lain, wilayah-wilayah seperti Selemadeg Barat, Selemadeg, Selemadeg Timur, Pekutatan, Tembuku, Sidemen, Dawan, Kerambitan, Pupuan, Selat, Banjarangkan, dan Rendang menunjukkan kategori angka kejadian yang sangat rendah.



**Gambar 1.** Persebaran Kejadian Diare

Pembentukan model regresi klasik OLS dengan meregresikan variabel Y dengan X<sub>1</sub>, X<sub>2</sub>, X<sub>3</sub>. Model regresi klasik (OLS) yang diperoleh adalah:

$$\hat{y} = -3,3 + 0,027X_1 + 0,001X_2 - 0,002X_3$$

Model ini menjelaskan bahwa terdapat peningkatan sekitar 0,027 persen dalam kejadian diare ketika terjadi pertambahan satu persen pada variabel kepadatan penduduk (X<sub>1</sub>), dengan asumsi variabel prediktor lainnya tetap konstan. Selain itu, peningkatan sekitar 0,001 persen dalam kejadian diare dapat terjadi dengan peningkatan satu persen pada jumlah sarana air minum (X<sub>2</sub>), dengan syarat variabel prediktor lain tetap konstan. Interpretasi yang serupa berlaku untuk setiap variabel dalam model regresi ini. Begitu pula, terdapat penurunan sekitar 0,002 persen dalam kejadian diare jika terjadi peningkatan satu persen pada variabel jumlah sarana sanitasi (X<sub>3</sub>).

**Tabel 1.** Estimasi Parameter, Nilai Statistik Uji, dan VIF dengan Metode OLS

Parameter	Estimasi	p-value	VIF
$\beta_0$	-3,3	0,964	
$\beta_1$	0,027	0,000*	1,27
$\beta_2$	0,001	0,690	3,22
$\beta_3$	0,002	0,565	3,56
R <sup>2</sup>		98,02%	
SSerror		1668,46	
F <sub>hitung</sub>		876,09	
AIC		267,558	
p-value		0,000	

Ket: \*) signifikan pada  $\alpha = 5\%$

Selanjutnya dilakukan identifikasi awal (diagnosis) adanya efek spasial. Tujuan dari diagnosis ini adalah untuk mengidentifikasi adanya variasi spasial yang berbeda dan ketergantungan spasial. Hal ini penting dilakukan untuk menentukan tindakan selanjutnya, yaitu menentukan model spasial manakah yang akan digunakan untuk memodelkan kejadian diare. Berikut diagnostik untuk spatial dependence (Tabel 2).

**Tabel 2.** Hasil Diagnostik Dependensi Spasial

Uji Spatial Dependence	<i>P-value</i>
Moran's I ( <i>error</i> )	0,9082709
LM (lag)	0,2008444
Robust LM (lag)	0,0456321*
LM ( <i>error</i> )	0,7236528
Robust LM ( <i>error</i> )	0,0424754*

Ket : \*) signifikan pada  $\alpha = 5\%$

Nilai *p-value* Robust LM (lag) sebesar 0,0456321 dan Robust LM (*error*) 0,0424754. Hal ini menunjukkan bahwa ada dependensi spasial dalam *error* regresi. LM dapat mendeteksi dependensi spasial secara lebih spesifik yaitu dependensi spasial dalam hal lag, *error*, atau keduanya (lag dan *error*).

Pemodelan dengan metode spasial berbasis area yang digunakan adalah model SEM, dengan estimasi parameter sebagai berikut

**Tabel 3.** Estimasi Parameter dengan Metode SEM

Parameter	Coeff	<i>p-value</i>
$\beta_0$	80,05	0,132
$\beta_1$	0,534	0,103
$\beta_2$	0,567	0,011*
$\beta_3$	0,059	0,049*
P	-0,63	0,033*
$R^2$	96,01%	
AIC	355.597	

Ket: \*) signifikan pada  $\alpha = 5\%$ ;  $Z_{0,05;30} = 1,645$

Dengan tingkat signifikansi  $\alpha=5\%$  maka menghasilkan keputusan tolak  $H_0$  yaitu jumlah sarana air minum ( $X_2$ ) dan jumlah sarana sanitasi ( $X_3$ ). Sehingga model dari metode SEM yang terbentuk adalah:

$$\hat{y}_i = -0,63 \sum_{j=1}^n w_{ij} y_j + 80,05 + 0,534 X_1 - 0,567 X_2 + 0,059 X_3$$

Pemilihan model terbaik berdasarkan nilai *R-Squared* terbesar dan nilai AIC terkecil. Berikut tampilan perbandingan ketiga model:

**Tabel 4.** Pemilihan Model Terbaik

METODE	$R^2$	AIC
OLS	88,02%	355.597
SAR	95,34%	350.561
SEM	96,01%	267,558

Pada tabel 4 dapat dilihat, bahwa model terbaik yang diperoleh berdasarkan ketiga metode dengan melihat  $R^2$  yang terbesar dan nilai AIC terkecil yaitu model SEM, dimana nilai  $R^2$  pada metode SEM yaitu sebesar 96,01% dan nilai AIC 267,558. Model akhir yang diperoleh dalam yaitu metode SEM yaitu :

$$\hat{y}_i = -0,63 \sum_{j=1}^n w_{ij} y_j + 80,05 + 0,534 X_1 - 0,567 X_2 + 0,059 X_3$$

Model tersebut menjelaskan bahwa apabila kepadatan penduduk ( $X_1$ ) bertambah sebanyak satu satuan kejadian diare akan bertambah sebanyak 0,543 satuan dengan asumsi variabel yang lain konstan. Begitu juga interpretasi pada masing-masing variabel prediktor



yang lain. Nilai  $p$ -value untuk  $\rho$  sebesar 0,63 yang berarti bahwa ada dependensi spasial lag pada variabel kejadian diare dan  $\sum_{j=1}^n w_{ij}$  dalam hal lain memiliki arti ada pengaruh letak kecamatan yang saling berdekatan( $i$ ) dengan kecamatan yang diamati ( $j$ ) terhadap kejadian diare. Nilai  $R^2$  yang dihasilkan sebesar 96,01% nilai tersebut memiliki arti bahwa besaran variansi dari kejadian diare yang dapat dijelaskan oleh model.

## KESIMPULAN

Data yang digunakan dalam analisis memiliki unsur spasial yang signifikan, sehingga dilakukan pemodelan berbasis spasial wilayah. Hasil pemodelan menunjukkan bahwa metode SEM lebih unggul dibandingkan dengan OLS dan SAR. Hal ini terbukti dari tingginya nilai  $R^2$  sebesar 96,01% pada model SEM, yang signifikan lebih tinggi dibandingkan dengan model-model lainnya. Selain itu, model SEM juga memiliki nilai AIC yang lebih rendah sebesar 267,558, menandakan bahwa model ini lebih cocok dalam menggambarkan hubungan antara variabel-variabel yang ada dalam data.

## UCAPAN TERIMA KASIH

Terima kasih kami ucapkan kepada semua pihak yang telah membantu dalam penyelesaian penelitian ini sehingga dapat terselesaikan dengan baik, serta kami ucapkan terimakasih juga kepada ITB Stikom Bali yang telah mendanai penelitian ini melalui dana hibah Penelitian Internal Tahun-Penelitian Dosen Individual 2023.

## DAFTAR PUSTAKA

- Anselin, L. (1988). *Spatial Econometrics: Methods and Models*. Netherlands: Kluwer Academic Publishers.
- BPS. (2020). *Sirusa Badan Pusat Statistik Republik Indonesia*. Retrieved from Badan Pusat Statistik Republik Indonesia: <http://sirusa.bps.go.id>
- Burnham, K., & Anderson, D. (2002). *Model Selection and Multimodel Inference: a Practical Information-Theoretic Approach*. New York: Springer-Verlag.
- Dinas Kesehatan Provinsi Bali. (2021). *Profil Kesehatan Provinsi Bali Tahun 2020*. Bali, Dinas Kesehatan Provinsi Bali.
- Grasa, A.A. (1989). Econometric Model Selection: A New Approach. Kluwer. *Jurnal Informatika Mulawarman*, Vol. 4(3).
- Le-Sage, J.P. (1999). *The Theory and Practice of Spatial Econometrics*, Toledo: Department of Economics University of Toledo.
- Mardhatillah, M. (2020). *Faktor-Faktor yang Mempengaruhi Kasus Penyakit Diare di Provinsi Sumatera Barat Menggunakan Analisis Regresi Spasial*. Skripsi (tidak diterbitkan). Padang: Universitas Negeri Padang.
- Pratiwi, L.P.S., Hendayanti, N.P.N., dan Suniantara, I.K.P. (2020). Perbandingan Pembobotan *Seemingly Unrelated Regression-Spatial Durbin Model* untuk Faktor Kemiskinan dan Pengangguran, *Jurnal Varian*, Vol. 3(2): 51-64.
- Rahmawati, D. dan Bimanto, H. (2021). Perbandingan *Spatial Autoregressive Model* dan *Spatial Error Model* dalam Pemodelan Indeks Pembangunan Manusia di Provinsi Jawa Timur. *Jurnal Statistika dan Aplikasinya*. Vol. 5(1), June 2021.
- Sulasih, I.G.A.D., Susilawati, M., dan Suciptawati, N.L.P. (2021). Pemodelan Kasus Diare Di Provinsi Bali Dengan Metode Analisis Regresi Spasial, *E-Jurnal Matematika* Vol. 10(2): pp. 95-102