



# ANALISIS POLA PENYEBARAN PENYAKIT CAMPAK DI KABUPATEN SUMENEP DENGAN ALGORITMA RANDOM FOREST BERBASIS GEOSPASIAL

Anandito Wisnu Widya Pratama<sup>1)</sup>, I Kadek Dwi Nuryana<sup>2)</sup>

<sup>1)</sup>Sistem Informasi, Fakultas Teknik, Universitas Negeri Surabaya, Surabaya, Indonesia

Email : [anandito.22147@mhs.unesa.ac.id](mailto:anandito.22147@mhs.unesa.ac.id)

<sup>2)</sup>Sistem Informasi, Fakultas Teknik, Universitas Negeri Surabaya, Surabaya, Indonesia

Email : [dwinuryana@unesa.ac.id](mailto:dwinuryana@unesa.ac.id)

## Abstract

Measles remains a Highly contagious disease and poses a surveillance challenge in Sumenep Regency, whose administrative area consists of mainland and island subdistricts. This study aimed to analyze the spatial distribution of measles, forecast monthly cases, and implement the results in a WebGIS dashboard. The study used measles records from 2022–2025 obtained from the Sumenep Regency Health Office and subdistrict boundaries in GeoJSON format. Annual records for 2022–2024 were temporally disaggregated using the monthly distribution pattern of the detailed 2025 data, while individual 2025 records were aggregated by subdistrict and month. The resulting panel comprised 27 subdistricts, 48 monthly periods, and 1,296 observations. Spatial analysis compared Queen Contiguity and K-Nearest Neighbor weights, followed by Global Moran's I and Local Indicators of Spatial Association with 999 permutations. Random Forest Regression used Lag-1 to Lag-4, month, and year, with a chronological split and five-fold expanding-window RandomizedSearchCV. KNN with  $k=5$  produced a connected spatial network and significant positive autocorrelation (Moran's  $I=0.3384$ ;  $p=0.004$ ). The tuned model achieved  $MAE=3.6021$ ,  $RMSE=5.6325$ , and  $R^2=0.3025$ , outperforming the Lag-1 baseline. Lag-1 was the dominant predictor. The integrated WebGIS presents case maps, LISA clusters, forecasts, and operational priority areas. The approach supports data-driven surveillance, although predictions remain moderate and tend to underestimate extreme increases.

**Keywords:** Measles, Moran's I, LISA, Random Forest Regression; WebGIS.

## Abstrak

Campak merupakan penyakit yang sangat menular dan menjadi tantangan surveilans di Kabupaten Sumenep yang memiliki wilayah daratan dan kepulauan. Penelitian ini bertujuan menganalisis pola spasial kasus campak, memprediksi jumlah kasus bulanan, dan mengimplementasikan hasilnya dalam WebGIS Dashboard. Data yang digunakan meliputi kasus campak tahun 2022–2025 dari Dinas Kesehatan Kabupaten Sumenep dan batas administrasi kecamatan berformat GeoJSON. Data tahunan 2022–2024 didisagregasi secara temporal menggunakan pola distribusi bulanan data rinci 2025, sedangkan data individu 2025 diagregasi berdasarkan kecamatan dan bulan. Dataset akhir terdiri atas 27 kecamatan, 48 periode bulanan, dan 1.296 observasi. Analisis spasial membandingkan pembobot Queen Contiguity dan K-Nearest Neighbor, kemudian menghitung Global Moran's I dan Local Indicators of Spatial Association dengan 999 permutasi. Random Forest Regression menggunakan Lag-1 sampai Lag-4, Bulan, dan Tahun, dengan pembagian kronologis serta RandomizedSearchCV lima fold berbasis expanding window. KNN  $k=5$  menghasilkan jaringan yang terhubung dan autokorelasi spasial positif signifikan (Moran's  $I=0,3384$ ;  $p=0,004$ ). Model hasil tuning memperoleh  $MAE=3,6021$ ,  $RMSE=5,6325$ , dan  $R^2=0,3025$  serta lebih baik daripada baseline Lag-1. Lag-1 menjadi prediktor paling dominan. WebGIS mengintegrasikan peta kasus, kluster LISA, hasil prediksi, dan wilayah prioritas operasional. Pendekatan ini mendukung pemantauan berbasis data, tetapi prediksi masih moderat dan cenderung meremehkan lonjakan kasus ekstrem.

**Kata Kunci:** Campak, Moran's I, LISA, Random Forest Regression, WebGIS.



## PENDAHULUAN

Campak masih menjadi masalah kesehatan masyarakat karena tingkat penularannya tinggi dan dapat menimbulkan komplikasi, terutama pada anak-anak. World Health Organization melaporkan sekitar 10,3 juta kasus campak terjadi secara global pada 2023, meningkat 20% dibandingkan tahun sebelumnya. Di Indonesia, kebangkitan kasus juga terlihat pada 2022 ketika tercatat 4.845 kasus terkonfirmasi laboratorium yang tersebar di 32 provinsi. Kondisi tersebut memperlihatkan bahwa pemantauan berbasis wilayah dan waktu tetap diperlukan untuk mendukung deteksi dini serta penentuan prioritas intervensi.

Penyebaran penyakit menular tidak selalu bersifat acak. Perbedaan cakupan imunisasi, kepadatan penduduk, mobilitas, dan akses pelayanan kesehatan dapat membentuk pola spasial dan temporal yang berbeda antarwilayah. Analisis spasial membantu mengidentifikasi apakah wilayah dengan nilai kasus serupa cenderung berdekatan, sedangkan analisis lokal dapat menunjukkan lokasi kluster dan pencilaan secara lebih rinci. Kajian global juga menunjukkan bahwa kebangkitan campak membentuk pola spasio-temporal dan hotspot yang bervariasi antarnegara.

Kabupaten Sumenep memiliki karakteristik geografis yang khas karena terdiri atas wilayah daratan dan kepulauan. Kondisi ini menyebabkan hubungan ketetanggaan tidak selalu dapat direpresentasikan hanya melalui persinggungan batas administrasi. Queen Contiguity sesuai untuk wilayah yang berbagi sisi atau titik sudut, sedangkan K-Nearest Neighbor (KNN) membentuk hubungan berdasarkan kedekatan centroid dan dapat menjaga keterhubungan wilayah kepulauan. Pemilihan matriks bobot perlu diuji karena dapat memengaruhi nilai autokorelasi dan kluster yang terbentuk.

Selain mengetahui pola spasial, informasi prediksi diperlukan agar peningkatan kasus dapat diantisipasi. Random Forest Regression mampu menangkap hubungan nonlinier dan telah digunakan dalam prediksi wabah campak berdasarkan data historis. Namun, banyak penelitian masih memisahkan analisis spasial dan prediksi temporal. Penelitian ini mengintegrasikan Global Moran's I dan LISA dengan Random Forest Regression berbasis lag feature, kemudian menyajikan hasilnya melalui WebGIS. Tujuannya adalah mengidentifikasi pola dan kluster penyebaran campak, memprediksi jumlah kasus periode berikutnya, serta menghasilkan informasi wilayah prioritas pada tingkat kecamatan.

## METODE PENELITIAN

### 3.1 Data dan Kerangka Analisis

Penelitian mengikuti tahapan CRISP-DM yang mencakup pemahaman kebutuhan, pemahaman data, persiapan data, pemodelan, evaluasi, dan deployment. Data utama berupa rekap jumlah kasus campak 2022–2024, data individu kasus 2025, dan batas administrasi 27 kecamatan dalam format GeoJSON. GeoJSON digunakan untuk menyimpan geometri dan atribut wilayah sehingga dapat langsung diintegrasikan dengan pemetaan berbasis web.

**Tabel 1.** Ringkasan Dataset dan Pembagian Data

| Komponen         | Hasil                        |
|------------------|------------------------------|
| Wilayah          | 27 kecamatan                 |
| Periode          | Jan 2022–Des 2025            |
| Jumlah periode   | 48 bulan                     |
| Jumlah observasi | 1.296                        |
| Training         | Mei 2022–Des 2024 (864, 73%) |
| Testing          | Jan–Des 2025 (324, 27%)      |
| Prediktor        | Lag-1–Lag-4, Bulan, Tahun    |
| Target           | Jumlah kasus bulanan         |

*Sumber : Hasil Pengolahan Data Peneliti (2026)*

### 3.2 Persiapan Data

Data individu 2025 dibersihkan melalui pemeriksaan missing value, duplikasi, format tanggal, dan konsistensi nama kecamatan, kemudian diagregasi menjadi jumlah kasus per kecamatan dan bulan. Data 2022–2024 tersedia dalam bentuk rekap tahunan, sehingga dilakukan disagregasi temporal menggunakan proporsi bulanan yang diperoleh dari data rinci 2025. Jumlah hasil distribusi tetap disesuaikan agar sama dengan total tahunan. Data hasil disagregasi diperlakukan sebagai estimasi pembelajaran dan bukan pengganti observasi aktual.

Setelah resolusi waktu diseragamkan, seluruh data digabungkan menjadi balanced panel 27 kecamatan untuk 48 bulan. Pembentukan lag dilakukan secara terpisah pada setiap kecamatan. Lag-1, Lag-2, Lag-3, dan Lag-4 merepresentasikan jumlah kasus satu sampai empat bulan sebelumnya. Empat periode awal tidak digunakan dalam pemodelan karena belum memiliki seluruh lag, sehingga data pemodelan dimulai pada Mei 2022.

### 3.3 Analisis Spasial

Matriks bobot spasial dibentuk menggunakan Queen Contiguity dan KNN. Uji sensitivitas KNN dilakukan pada  $k=3$  sampai  $k=7$ . Bobot ditransformasi dengan row standardization. Global Moran's I digunakan untuk menguji pola autokorelasi secara keseluruhan, sedangkan LISA digunakan untuk mengidentifikasi kluster High-High (HH), Low-Low (LL), High-Low (HL), Low-High (LH), dan wilayah tidak signifikan. Signifikansi ditentukan melalui 999 permutasi dengan taraf 5%. Pendekatan LISA digunakan karena mampu menunjukkan lokasi kluster yang tidak dapat dijelaskan hanya melalui statistik global.

### 3.4 Pemodelan Random Forest dan Evaluasi

Data dibagi secara kronologis agar informasi masa depan tidak masuk ke proses pelatihan. Data Mei 2022–Desember 2024 digunakan sebagai training, sedangkan seluruh bulan tahun 2025 digunakan sebagai testing. Optimasi hyperparameter dilakukan hanya pada training menggunakan RandomizedSearchCV dan lima fold expanding-window time-series cross-validation. Skema ini mempertahankan urutan temporal dan mengurangi risiko data leakage. Random search dipilih karena efisien untuk mengevaluasi ruang parameter yang luas. Parameter yang dicari meliputi jumlah pohon, kedalaman maksimum, minimum sampel split, minimum sampel leaf, jumlah fitur,

dan bootstrap, parameter tersebut berpengaruh terhadap kompleksitas dan generalisasi Random Forest.

Kinerja model diukur menggunakan Mean Absolute Error (MAE), Root Mean Square Error (RMSE), dan koefisien determinasi ( $R^2$ ). Sebagai pembandingan digunakan baseline naive Lag-1, yaitu nilai periode sebelumnya dijadikan prediksi periode saat ini. Setelah evaluasi, model terbaik dilatih kembali menggunakan data hingga desember 2025 untuk memprediksi januari 2026. Hasil prediksi kemudian dianalisis kembali dengan LISA dan dikombinasikan dengan kuartil prediksi untuk membentuk kategori prioritas operasional.

### 3.5 Implementasi WebGIS

Hasil analisis diimplementasikan dalam WebGIS Dashboard berbasis Flask, Leaflet, dan GeoJSON. Backend mengelola data serta hasil model, sedangkan Leaflet menampilkan peta interaktif. Integrasi peta dan dashboard dapat membantu penyajian tren, klaster, prediksi, dan ringkasan surveilans dalam satu antarmuka. Pengujian fungsional dilakukan menggunakan black box testing.

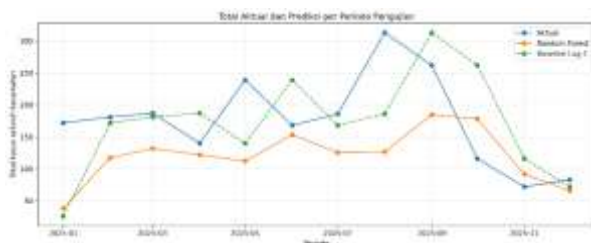


**Gambar 1.** Tampilan Dashboard Overview WebGIS  
*Sumber : Hasil Pengolahan Data Peneliti (2026)*

## HASIL DAN PEMBAHASAN

### 4.1 Karakteristik Data

Dataset akhir berisi 1.296 observasi dengan total 3.967 kasus. Jumlah kasus per tahun adalah 168 pada 2022, 1.368 pada 2023, 312 pada 2024, dan 2.119 pada 2025. Kasus bulanan tertinggi terjadi pada Agustus 2025 sebanyak 313 kasus. Perbedaan tingkat kasus antarperiode menunjukkan pola yang fluktuatif dan meningkatkan kesulitan prediksi, terutama ketika data pengujian memiliki lonjakan yang jauh lebih tinggi daripada sebagian besar data pelatihan.



**Gambar 2.** Perbandingan Aktual, Prediksi Random Forest, dan Baseline per Periode  
*Sumber : Hasil Pengolahan Data Peneliti (2026)*

### 4.2 Autokorelasi Spasial dan Klaster LISA

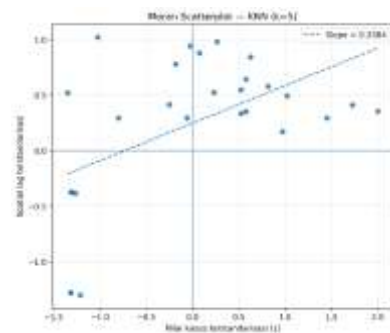
**Tabel 2.** Perbandingan Global Moran's I

| SWM   | I    | Z     | p    | Komp. | Island |
|-------|------|-------|------|-------|--------|
| Queen | 0,44 | 2,727 | 0,00 | 8     | 5      |
| KNN   | 0,33 | 3,906 | 0,00 | 1     | 0      |
| k=5   | 84   | 5     | 4    |       |        |

*Sumber : Hasil Pengolahan Data Peneliti (2026)*

Kedua pembobot menghasilkan Moran's I positif dan signifikan, sehingga persebaran kasus tidak bersifat acak dan cenderung mengelompok. Queen Contiguity menghasilkan nilai I yang lebih tinggi, tetapi membentuk delapan komponen dan lima island. KNN k=5 menghasilkan satu jaringan yang terhubung tanpa island.

Oleh karena itu, KNN k=5 digunakan sebagai pembobot utama pada analisis operasional karena seluruh kecamatan, termasuk wilayah kepulauan, tetap memiliki hubungan spasial. Pemilihan ini tidak menunjukkan bahwa hubungan melalui laut merupakan bukti penularan langsung, KNN hanya menjadi representasi matematis kedekatan wilayah.



**Gambar 3.** Moran Scatterplot KNN k=5  
*Sumber : Hasil Pengolahan Data Peneliti (2026)*

LISA KNN k=5 mengidentifikasi klaster HH pada Ambunten, Batuputih, dan Gapura; klaster LL pada Arjasa, Kangayan, dan Sapeken, serta klaster LH pada Batuan, Guluk Guluk, dan Talango. Delapan belas kecamatan lainnya tidak signifikan pada taraf 5%. Klaster HH menunjukkan konsentrasi wilayah bernilai tinggi yang berdekatan dan dapat digunakan sebagai penyaringan awal untuk surveilans. Kategori tidak signifikan tetap dapat memiliki jumlah kasus tinggi, tetapi pola lokalnya belum cukup kuat secara statistik.

### 4.3 Kinerja Random Forest Regression

**Tabel 3.** Evaluasi Model pada Data Testing 2025

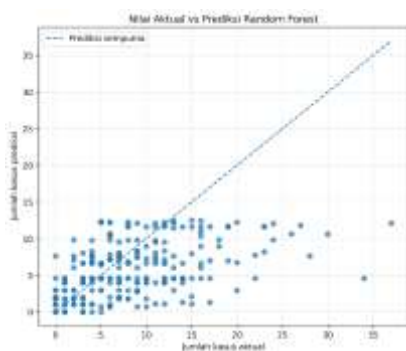
| Model               | MAE    | RMSE   | $R^2$  |
|---------------------|--------|--------|--------|
| Random Forest awal  | 3,6673 | 5,7991 | 0,2606 |
| Random Forest tuned | 3,6021 | 5,6325 | 0,3025 |
| Baseline Lag-1      | 4,0154 | 6,0952 | 0,1832 |

*Sumber : Hasil Pengolahan Data Peneliti (2026)*

Model terbaik menggunakan 500 pohon, max\_depth=5, min\_samples\_split=5, min\_samples\_leaf=2,

max\_features=1, dan bootstrap=True. Rata-rata RMSE cross-validation adalah 1,5694 dengan standar deviasi 1,7010. Perbedaan error antarfold menunjukkan bahwa tingkat kesulitan prediksi berubah pada setiap rentang waktu. Pada testing 2025, model tuned menghasilkan MAE 3,6021, yang berarti rata-rata selisih absolut sekitar 3,6 kasus. RMSE 5,6325 lebih rendah daripada baseline 6,0952, sedangkan  $R^2$  0,3025 lebih tinggi daripada baseline 0,1832. Dengan demikian, Random Forest memberi peningkatan dibandingkan aturan sederhana Lag-1, tetapi baru menjelaskan sekitar 30,25% variasi jumlah kasus.

Visualisasi aktual dan prediksi pada skripsi menunjukkan bahwa model mampu mengikuti arah umum perubahan, namun cenderung menghasilkan nilai lebih rendah saat terjadi lonjakan kasus. Kondisi ini konsisten dengan RMSE yang lebih tinggi daripada MAE karena beberapa error besar memperoleh penalti lebih kuat. Hasil tersebut menunjukkan bahwa prediksi sebaiknya digunakan sebagai informasi pendukung, bukan pengganti data surveilans aktual. Kajian spasio-temporal campak juga menekankan bahwa prediksi wabah dipengaruhi oleh perubahan pola epidemiologi dan kualitas data historis.



**Gambar 4.** Nilai Aktual dan Prediksi Random Forest  
*Sumber : Hasil Pengolahan Data Peneliti (2026)*

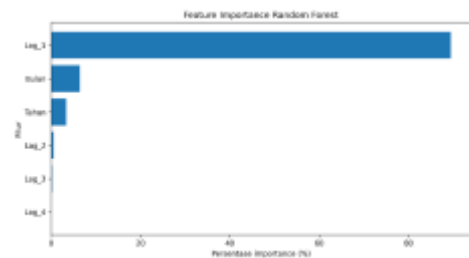
#### 4.4 Kontribusi Fitur

**Tabel 4.** Feature importance Random Forest

| Fitur | Impurity (%) | Permutation Mean |
|-------|--------------|------------------|
| Lag-1 | 89,4870      | 2,3929           |
| Bulan | 6,3928       | 0,1036           |
| Tahun | 3,3496       | 0,0000           |
| Lag-2 | 0,5455       | 0,0126           |
| Lag-3 | 0,1495       | 0,0080           |
| Lag-4 | 0,0756       | 0,0002           |

*Sumber : Hasil Pengolahan Data Peneliti (2026)*

Lag-1 menjadi fitur paling dominan dengan impurity importance 89,4870% dan permutation mean 2,3929. Artinya, jumlah kasus satu bulan sebelumnya merupakan informasi yang paling banyak digunakan model untuk membentuk prediksi. Bulan dan Tahun memberikan kontribusi tambahan, sedangkan Lag-2 sampai Lag-4 relatif rendah. Nilai Feature importance menunjukkan kontribusi terhadap keputusan model dan tidak dapat diartikan sebagai hubungan sebab-akibat epidemiologis.



**Gambar 5.** Persentase Feature importance Random Forest  
*Sumber : Hasil Pengolahan Data Peneliti (2026)*

#### 4.5 Prediksi Januari 2026 dan Prioritas Operasional

**Tabel 5.** Lima Prediksi Tertinggi Januari 2026

| Peringkat | Kecamatan    | Prediksi |
|-----------|--------------|----------|
| 1         | Rubaru       | 11,6647  |
| 2         | Saronggi     | 11,5975  |
| 3         | Pasongsongan | 11,5928  |
| 4         | Ambunten     | 11,5014  |
| 5         | Pragaan      | 11,4204  |

*Sumber : Hasil Pengolahan Data Peneliti (2026)*

Prediksi Januari 2026 menunjukkan nilai tertinggi pada Rubaru, Saronggi, Pasongsongan, Ambunten, dan Pragaan. Analisis Moran's I terhadap permukaan prediksi menghasilkan  $I=0,4533$ ,  $z=5,1801$ , dan  $p=0,002$ , sehingga nilai prediksi juga membentuk autokorelasi spasial positif. LISA prediksi mengidentifikasi HH pada Pasongsongan, Pragaan, Bluto, Lenteng, dan Guluk Guluk. Prioritas tinggi diberikan kepada wilayah HH atau wilayah dengan prediksi minimal kuartil ketiga. Aturan ini menggabungkan besarnya prediksi temporal dengan konteks spasial, sehingga wilayah yang tidak signifikan secara LISA tetap dapat masuk prioritas tinggi apabila prediksinya besar.



**Gambar 6.** Peta Kategori Prioritas Spasial  
*Sumber : Hasil Pengolahan Data Peneliti (2026)*

Prediksi periode berikutnya merupakan hasil deployment dan belum dapat digunakan untuk menghitung metrik evaluasi karena nilai aktual januari 2026 belum tersedia. Oleh sebab itu, kategori prioritas berfungsi sebagai alat penyaringan operasional yang perlu dikonfirmasi melalui data surveilans, cakupan imunisasi, kondisi pelayanan kesehatan, dan investigasi epidemiologi.

#### 4.6 Implementasi dan Pengujian WebGIS

WebGIS menampilkan peta kasus, peta kluster LISA, ringkasan Moran's I, grafik time series, hasil prediksi

Random Forest, Feature importance, tabel prioritas, dan laporan integrasi. Pengujian black box terhadap fungsi peta, grafik, tabel, prediksi, dan navigasi menunjukkan seluruh komponen berhasil berjalan sesuai masukan dan keluaran yang diharapkan. Penyajian terintegrasi memudahkan pengguna membandingkan kondisi aktual, pola lokal, hasil prediksi, dan tingkat prioritas tanpa berpindah ke perangkat analisis terpisah. Pendekatan dashboard spasial serupa telah digunakan untuk mendukung pemantauan penyakit dan penyajian informasi wilayah berisiko



**Gambar 7.** Halaman Integrasi Analisis Spasial dan Prediksi

*Sumber : Hasil Pengolahan Data Peneliti (2026)*

#### 4.7 Implikasi Operasional

Hasil integrasi dapat digunakan sebagai tahapan penyaringan awal bagi Dinas Kesehatan Kabupaten Sumenep. Wilayah prioritas tinggi perlu memperoleh verifikasi data kasus secara lebih cepat, penelusuran kontak, pemeriksaan cakupan imunisasi, dan kesiapan logistik vaksin. Kecamatan yang berada pada kluster High-High memerlukan koordinasi lintaswilayah karena nilai tinggi tidak berdiri sendiri, tetapi berdekatan dengan wilayah yang juga memiliki nilai tinggi. Sementara itu, kecamatan dengan prediksi tinggi tetapi kategori LISA tidak signifikan tetap perlu diperhatikan karena besarnya estimasi temporal dapat menunjukkan beban kasus yang tinggi meskipun belum membentuk kluster lokal secara statistik.

Wilayah prioritas sedang dapat diarahkan pada pemantauan berkala dan validasi terhadap kemungkinan perubahan pola. Kategori Low-High atau High-Low perlu ditelaah sebagai pencilaan spasial karena kondisi kecamatan berbeda dengan lingkungan sekitarnya. Perbedaan tersebut dapat berkaitan dengan mobilitas penduduk, akses pelayanan, kualitas pencatatan, atau faktor epidemiologis lain yang tidak tercakup dalam model. Wilayah prioritas rendah tetap memerlukan surveilans rutin dan kegiatan pencegahan karena kategori rendah tidak berarti bebas risiko.

Hubungan KNN antara kecamatan daratan dan kepulauan merupakan hubungan kedekatan matematis antar-centroid, bukan pernyataan bahwa laut menjadi media penyebaran campak. Penularan tetap terjadi melalui interaksi antarmanusia. Oleh karena itu, interpretasi peta perlu dilengkapi dengan informasi rute transportasi, mobilitas antarpulau, cakupan imunisasi, dan hasil investigasi lapangan. WebGIS berfungsi menyatukan hasil analisis tersebut agar pengguna dapat menelusuri wilayah

prioritas, tetapi keputusan intervensi tetap memerlukan konfirmasi dari surveilans epidemiologi.

#### 4.8 Keterbatasan Penelitian

Keterbatasan utama terletak pada data historis 2022–2024 yang tersedia dalam resolusi tahunan sehingga harus didisagregasi menjadi estimasi bulanan menggunakan pola 2025. Asumsi tersebut dapat membuat pola temporal historis terlihat lebih teratur daripada kondisi sebenarnya. Selain itu, model hanya menggunakan lag dan fitur kalender tanpa cakupan imunisasi, kepadatan penduduk, mobilitas, cuaca, maupun akses layanan kesehatan.  $R^2$  yang moderat dan kecenderungan underprediction pada lonjakan menunjukkan perlunya data lebih panjang dan variabel eksternal pada penelitian selanjutnya.

#### KESIMPULAN

Analisis spasial menunjukkan autokorelasi positif yang signifikan pada persebaran kasus campak di Kabupaten Sumenep. Queen Contiguity dan KNN menghasilkan arah kesimpulan yang sama, sedangkan KNN  $k=5$  dipakai sebagai pembobot utama karena membentuk satu jaringan tanpa island. LISA mengidentifikasi kluster HH, LL, dan LH yang dapat digunakan sebagai dasar awal penentuan wilayah surveilans dan pengendalian.

Random Forest Regression berbasis Lag-1 sampai Lag-4, Bulan, dan Tahun menghasilkan MAE 3,6021, RMSE 5,6325, dan  $R^2$  0,3025 serta mengungguli baseline Lag-1. Lag-1 menjadi fitur paling dominan. Integrasi prediksi, Moran's I, LISA, dan kategori prioritas dalam WebGIS menghasilkan informasi spasial-temporal yang lebih mudah digunakan, namun prediksi tetap perlu ditafsirkan secara hati-hati karena keterbatasan data disagregasi dan kemampuan model dalam menangkap lonjakan ekstrem.

#### DAFTAR PUSTAKA

- Ahn, E., Liu, N., Parekh, T., Patel, R., Baldacchino, T., Mullavey, T., Robinson, A., & Kim, J. (2021). A Mobile App and Dashboard for Early Detection of Infectious Disease Outbreaks: Development Study. *JMIR Public Health and Surveillance*, 7(3), e14837. <https://doi.org/10.2196/14837>
- Arifin, O., & Supriyatna, A. R. (2023). Sistem Informasi Geografis Untuk Pemetaan Lahan Kakao Menggunakan Leaflet Js Dan Geojson. *Jurnal Teknoinfo*, 17(1), 364. <https://doi.org/10.33365/jti.v17i1.2397>
- Badan Pusat Statistik Kabupaten Sumenep. (2025, February 4). Luas Daerah dan Jumlah Pulau Menurut Kecamatan di Kabupaten Sumenep, 2024. <https://Sumenepkab.Bps.Go.Id/Id/Statistics-Table/3/VUZwV01tSlpPVlpsWIRKbmMxcFhhSGhEVjFoUFFUMDkjMw==/Luas-Daerah-Dan-Jumlah-Pulau-Menurut-Kecamatan-Di-Kabupaten-Sumenep.Html?Year=2024>
- Baskara, Z. W., Septiami, A., & Hadijati, M. (2024). Analisis Penyebaran Gizi Buruk Pada Balita Di Nusa Tenggara Barat (Ntb) Menggunakan Indeks Moran (Moran's I) Dan Local Indicator Of Spatial



- Association (LISA). *Jurnal Gaussian*, 13(2), 462–471. <https://doi.org/10.14710/j.gauss.13.2.462-471>
- Butler, H., Daly, M., Doyle, A., Gillies, S., Hagen, S., & Schaub, T. (2016). The GeoJSON Format. <https://doi.org/10.17487/RFC7946>
- Cerqueira, V., Torgo, L., & Mozetič, I. (2020). Evaluating time series forecasting models: an empirical study on performance estimation methods. *Machine Learning*, 109(11), 1997–2028. <https://doi.org/10.1007/s10994-020-05910-7>
- Feurer, M., & Hutter, F. (2019). Hyperparameter Optimization (pp. 3–33). [https://doi.org/10.1007/978-3-030-05318-5\\_1](https://doi.org/10.1007/978-3-030-05318-5_1)
- Isnan, S., Bin Abdullah, A. F., Shariff, A. R., Ishak, I., Syed Ismail, S. N., & Appanan, M. R. (2025). Moran's I and Geary's C: investigation of the effects of spatial weight matrices for assessing the distribution of infectious diseases. *Geospatial Health*, 20(1). <https://doi.org/10.4081/gh.2025.1277>
- Lin, C.-H., & Wen, T.-H. (2022). How Spatial Epidemiology Helps Understand Infectious Human Disease Transmission. *Tropical Medicine and Infectious Disease*, 7(8), 164. <https://doi.org/10.3390/tropicalmed7080164>
- Luo, Z., Shen, Y., Xu, X., Li, J., & Wang, X. (2025). Spatial-temporal pattern and drivers associated with measles resurgence from 2018 to 2023: a global perspective from 192 countries. *BMJ Public Health*, 3(1), e001912. <https://doi.org/10.1136/bmjph-2024-001912>
- Probst, P., Wright, M. N., & Boulesteix, A. (2019). Hyperparameters and tuning strategies for Random Forest. *WIREs Data Mining and Knowledge Discovery*, 9(3). <https://doi.org/10.1002/widm.1301>
- Ru, B., Kujawski, S., Lee Afanador, N., Baumgartner, R., Pawaskar, M., & Das, A. (2023). Predicting Measles Outbreaks in the United States: Evaluation of Machine Learning Approaches. *JMIR Formative Research*, 7, e42832. <https://doi.org/10.2196/42832>
- Sakinah, N., Saputra, W., Nurfitra, N., Satriani, S., & Junaidi, J. (2022). Analisis Spasial Penyebaran Penyakit Schistosomiasis Menggunakan Indeks Moran Untuk Mendukung Eradikasi Schistosomiasis Di Provinsi Sulawesi Tengah Berbasis Web Dashboard. *Jambura Journal of Probability and Statistics*, 3(2), 120–127. <https://doi.org/10.34312/jjps.v3i2.16580>
- Stoian, V.-I., Stănescu, A., Debita, M., Ignat, M. D., Barbu, R. E., Matei, M. N., Baltă, A. A., Ștefania, Bulza, V., Baroiu, L., & Condratovici, C. P. (2025). Spatio-Temporal Modelling and Forecasting of the Prolonged Measles Outbreak in Romania: Insights and Challenges. *Healthcare*, 13(18), 2364. <https://doi.org/10.3390/healthcare13182364>
- World Health Organization. (2023, April 28). Disease Outbreak News: Measles - Indonesia. <https://www.who.int/emergencies/disease-outbreak-news/item/2023-DON462>
- World Health Organization. (2024, November 14). Measles cases surge worldwide, infecting 10.3 million people in 2023. <https://www.who.int/news/item/14-11-2024-measles-cases-surge-worldwide--infecting-10.3-million-people-in-2023>