Geospatial Health 2025; volume 20:1379

A Bayesian spatiotemporal Poisson conditional autoregressive model for dengue haemorrhagic fever in Indonesia integrating satellite-generated environmental data

Sukarna,[1,2] Hari Wijayanto,[1] Yenni Angraini,[1] Anang Kurnia[1]

[1]Statistics and Data Science Study Program, School of Data Science, Mathematics, and Informatics, IPB University, Dramaga, Bogor; [2]Faculty of Mathematics and Natural Science, Universitas Negeri Makassar, Makassar, Indonesia

Correspondence: Hari Wijayanto, Statistics and Data Science Study Program - School of Data Science, Mathematics, and Informatics - IPB University, Jl. Meranti Wing 22 Level 4, Dramaga, Bogor 16680, Indonesia. E-mail: hari@apps.ipb.ac.id

Key words: hierarchical Bayesian, spatio-temporal, DHF, proxy data, Sentinel-2 satellite imagery, Indonesia.

Funding: this research was funded by a doctoral dissertation grant (*Penelitian Disertasi Doktor,* PDD) from the Directorate General of Higher Education, Research, and Technology, Ministry of Education, Culture, Research, and Technology, Indonesia. This research was conducted under the Research Program Implementation Contract 2024, ID: 027/E5/PG.02.00.PL/2024, dated June 11, 2024. Additional financial support was provided by the Indonesian Education Scholarship (*Beasiswa Pendidikan Indonesia*, BPI) and IPB University under their research funding scheme. The authors declare no competing financial or non-financial interests related to this work.

Conflict of interest: the authors declare no potential conflict of interest, and all authors confirm accuracy.

Availability of data and materials: the data and R code will be made available upon request.

Acknowledgment: this research was made possible by the generous support of the Department of Statistics and Data Science at IPB University, the Faculty of Mathematics and Sciences at *Universitas Negeri Makassar,* the *Puskesmas* in Makassar City, the *Dinas Kesehatan Kota Makassar* of South Sulawesi Province, the Social Security Agency (*BPJS*), the Indonesian Education Scholarship (*BPI*), and the Directorate of Higher Education, Research, and Technology (Ministry of Education, Culture, Research, and Technology).

Received: 8 February 2025.
Accepted: 23 April 2025.

Geospatial Health 2025; 20:1379
doi:10.4081/gh.2025.1379

Abstract

In association with cases of Dengue Haemorrhagic Fever (DHF), Indonesia's Breteau Index has consistently fallen below the national standard of 95% over the past 12 years (2007–2019). Currently, the country relies on survey methods to map DHF spread, but these methods are costly and require substantial resource support since monitoring DHF cases necessitates considering both spatial and temporal aspects. As an alternative, we proposed a pilot study utilizing a localized version of the hierarchical Bayesian spatiotemporal conditional autoregressive model (LHBSTCARM) to predict the DHF cases in Makassar City, Indonesia. Using this approach, we examined the relationship between DHF and the normalized difference built-up index (NDBI), the Normalized Difference Vegetation Index (NDVI), and the Normalized Difference Water Index (NDWI) that were downloaded from the Sentinel-2 satellite. Based on these datasets, we identified an optimal LHBSTCARM model that classified areas in Makassar City into distinct spatial risk groups based on the likelihood of dengue occurrence. Specifically, the model identified four districts with low relative risk, one with high relative risk and the remaining districts with moderate relative risk. Incorporating covariates, the model also revealed that NDVI and NDWI were significant predictors for dengue outbreaks, whereas NDBI was not. Both significant covariates showed negative effects, with a one-unit increase in NDVI and NDWI associated with reductions in DHF cases by 84.5% and 81.5%, respectively. Thus, NDVI and NDWI are the environmental variables of choice for the prediction of DHF incidence.

Introduction

Aedes mosquitoes, particularly *Ae. aegypti*, serve as the primary vector for the transmission of the dengue virus, which causes Dengue Haemorrhagic Fever (DHF) in humans (Martheswaran *et al.,* 2022). This arbovirus is widely recognized for its significant public health impact, characterized by high mortality rates and severe clinical outcomes (Martheswaran *et al.,* 2022). Globally, dengue infections account for an estimated 2.35 million cases annually (Baharom *et al.,* 2022) resulting in over 25,000 deaths (Martheswaran *et al.,* 2022). While dengue has traditionally been endemic to tropical and subtropical regions, recent evidence indicates its emergence also in temperate zones, including the continental United States (Martheswaran *et al.,* 2022). According to the World Health Organisation (WHO) and scientific articles, more than 40% of the global population is currently at risk of this infection, with approximately 50-100 million new cases reported each

year across more than 100 countries (Baharom *et al.,* 2022; Martheswaran *et al.,* 2022). The European Centre for Disease Prevention and Control (EPDPC) reported about DHF in 2024 (WHO, 2024), stating that more than 13 million people worldwide are infected (as of September 2024), with more than 8,500 deaths.

The majority of cases are in Asia (Khaidir *et al.,* 2022), with Indonesia reporting the highest number of cases from 1968 to 2010 (Dian Maya Sari *et al.,* 2018; Khaidir *et al.,* 2022) and 2022 (Supangat *et al.,* 2023). Over the past two decades (2004–2024), Indonesia has significantly contributed to the regional surge in DHF cases. According to the Indonesian Ministry of Health, there were 58,301 dengue cases with 658 deaths in 2004 (WHO, 2004), compared to 88,593 confirmed cases with 621 deaths in April 2024 (WHO, 2024). Notably, the number of cases surged to 119,709 by June 2024, representing an increase of more than 200% compared to April 2004. These figures highlight the escalating burden of DHF in Indonesia over time.

Consistent rapid and proper epidemic detection will be helpful to mitigate the threat of an infectious disease outbreak before it reaches an even more severe level. One reason for plausible infectious disease prevention is the dependence of contagious, vector-borne diseases on environmental conditions (Ziemann *et al.,* 2018) recognized through estimation and prediction based on the encouraging use of satellite data (Ziemann *et al.,* 2018). Satellite-generated data can serve as the proper proxy for the environment (Ziemann *et al.,* 2018) as they reliably determine and monitor land conditions (Ardiansyah *et al.,* 2023; Kurniawati *et al.,* 2023; Kusumaningrum *et al.,* 2024), e.g., paddy growing.

Ae. aegypti mosquitoes necessitate specific environmental conditions for breeding and survival. These mosquitoes thrive within a temperature range of 4 to 37 °C, with an average lifespan of 14 to 28 days. During their lifespan, females oviposit four times, producing approximately 10 to 100 eggs per oviposition event (Islam *et al.,* 2021). The complete metamorphosis from egg to adult takes 8–10 days under optimal conditions (Susanti & Suharyo, 2017). The dengue virus, transmitted by *Ae. aegypti*, undergoes an incubation period of 8 to 12 days within the mosquito vector, while in humans, clinical symptoms typically manifest between 3 and 14 days post-infection (Aswi *et al.,* 2020a; Islam *et al.,* 2021). The larval stages develop in aquatic environments, progressing from eggs to pupae before emerging as adult mosquitoes that disperse into various habitats. Consequently, stagnant water is a critical breeding site for *Aedes* larvae (Susanti & Suharyo, 2017; Ziemann *et al.,* 2018). The abundance of larvae can be quantified using indices such as the "Breteau Index" (Moreno-Madriñán *et al.,* 2014). In addition, remote sensing technologies have proven effective in identifying potential breeding sites by detecting standing water (Muhsoni, 2015; Ziemann *et al.,* 2018), such as the normalized difference vegetation index (NDVI), the Normalized Difference Water Index (NDWI), the Normalized Difference Moisture Index (NDMI), or the Normalized Difference Built-Up Index (NDBI).

To the best of our knowledge, satellite data have not been used to monitor the DHF incidences in Makassar City, Indonesia. Existing research has predominantly modeled DHF cases using covariates derived from three primary categories: i) climate variables, such as rainfall, precipitation, humidity and temperature (Rasjid *et al.,* 2019; Aswi *et al.,* 2020a; Aswi *et al.,* 2021; Faridah *et al.,* 2022; Solís-Navarro *et al.,* 2022); ii) clinical variables, including White Blood Cell (WBC) count, Red Blood Cell (RBC) count, Haemoglobin (HGB) levels, Haematocrit (HCT) and Platelet (PLT) count (Aswi *et al.,* 2020b; Anggraeni & Mahmudah 2021; Silitonga *et al.,* 2021); and iii) demographic and socioeconomic variables, such as age, gender, population density, and educational attainment (Aswi *et al.,* 2020b; Silitonga *et al.,* 2021). However, acquiring these variables by terrestrial-based research entails huge financial costs or labor-intensive efforts, which may impact their scalability and practicality for larger-scale monitoring.

DHF cases are usually reported as count data, and their distributions are often modelled by assumption as binomial (Sukarna *et al.,* 2023), Poisson (Fitri *et al.,* 2024) or gamma (Hii *et al.,* 2012). These distributions are from the basis of generalized linear mixed models (GLMM), including spatio-temporal models. Bayesian spatio-temporal (BST) modelling has been widely applied across diverse research domains, such as the analysis of sparse data (Mukhopadhyay *et al.,* 2019), forecasting road traffic congestion (Alghamdi *et al.,* 2021; Lian *et al.,* 2023), investigating the impact of climate factors on dengue fever transmission (Aswi *et al.,* 2021), analysing temperature trends (García *et al.,* 2023), predicting tuberculosis incidence (Sukarna *et al.,* 2021; Chen *et al.,* 2023) and assessing ecosystem dynamics (Song *et al.,* 2024). The versatility of BST modelling enables its application across a broad spectrum of contexts, demonstrating its utility in addressing complex spatio-temporal phenomena.

This study identified the optimal model for predicting DHF incidence by leveraging Sentinel-2 satellite imagery-derived indices capturing three critical environmental dimensions, namely vegetation density (utilizing the NDVI), surface wetness (using the NDWI), and built area density (employing the NDBI). The proposed model is a localized hierarchical Bayesian spatiotemporal autoregressive model (LHBSTCARM), a statistical model used to analyze data that vary across both space and time, while accounting for uncertainty and local variations in the data structure (Lee & Lawson, 2016).

Materials and Methods

Study area

The study area was Makassar and surrounding districts. Makassar City, located in the district of Makassar, is the capital of South Sulawesi Province, located in the central region of Indonesia. Therefore, it might be regarded as an ideal representation of the Indonesian climate. Makassar (Figure 1) is the sixth most populous city in Indonesia and encompasses an area of 175.79 km², comprises 14 mainland districts and had a population of 1,459,412 people in 2024 (BPS-Makassar 2024). The districts of Makassar, Mariso, and Mamajang are classified as urban, while Biringkanaya, Tamalanrea, and Manggala are designated as rural, and the others are categorized as suburban (Figure 1). Tamalanrea and Makassar districts exhibit the lowest and highest population densities, respectively. Tamalate, Rappocini, and Biringkanaya districts are the most extensive districts in Makassar.

Spatiotemporal data

The DHF data utilized in this study were obtained through a tripartite collaboration involving the Social Security Agency (*Badang Penyelenggara Jaminan Sosial*, *BPJS*), the Health Office of Makassar City and the community health centers (*Puskesmas* or *PKM*). While BPJS maintains insurance claim records encompassing DHF cases, its utility as a comprehensive data source is con-

strained by incomplete coverage of healthcare users. The Makassar Health Office consolidates epidemiological surveillance data from various healthcare facilities, including hospitals, clinics and PKM; however, the reporting frequency exhibits temporal inconsistencies. The PKMs, as primary healthcare providers at the community level, offer granular, locality-specific data, though their records are incomplete as not all individuals affected by DHF seek treatment at there. This study prioritized data from the source exhibiting the most consistent reporting cadence and comprehensive spatial coverage across the study period.

The retrospective analysis spanned 32 months (January 2022 to August 2024), encompassing 14 administrative districts in Makassar. This timeframe was strategically selected to avoid confounding from the COVID-19 pandemic (March 2020 to December 2021), which disrupted routine disease surveillance systems and healthcare-seeking behaviours. The post-pandemic period allows a more precise assessment of baseline DHF transmission patterns.

Satellite-generated data

The Copernicus Sentinel-2 satellite mission, launched by the European Space Agency over the period 2015-2024, consists of three polar-orbiting satellites aims to monitor changes in land surface conditions. The satellites have a 290 km swath width a spatial resolution ranging from 10 to 60 meters. This resolution is higher than that of other widely used satellites, such as the Terra/Aqua pair with their Moderate Resolution Imaging Spectroradiometer (MODIS) and Landsat (Bergquist & Manda 2019), making it particularly suitable for studying the limited geographical extent and diverse topography of Makassar City. As a result, Sentinel-2 is recognized as a valuable resource for acquiring environmental data, especially about vegetation, water bodies and built areas. Satellite-derived NDBI, NDVI and NDWI indices were utilized to capture these environmental features. These indices are calculated using four spectral bands as defined in previous studies (Muhsoni 2015; Suhet 2015) as follows:

$$NDBI = \frac{B11-B8}{B11+B8}, NDVI = \frac{B8-B4}{B8+B4}, NDWI = \frac{B3-B8}{B3+B8} \quad \text{Eq. 1}$$

where: B3 represents the green part of the spectrum; B4 the red part; B8 the near infrared (NIR) wavelengths; and B11 the shortwave infrared ones (SWIR-1).

The NDBI combines SWIR and NIR providing insight into built areas with index values ranging from -1 (no buildings) to 1 (densely built-up), the NDVI is a vegetation index that is calculated by combining the red wavelengths (B-4) and NIR to estimate

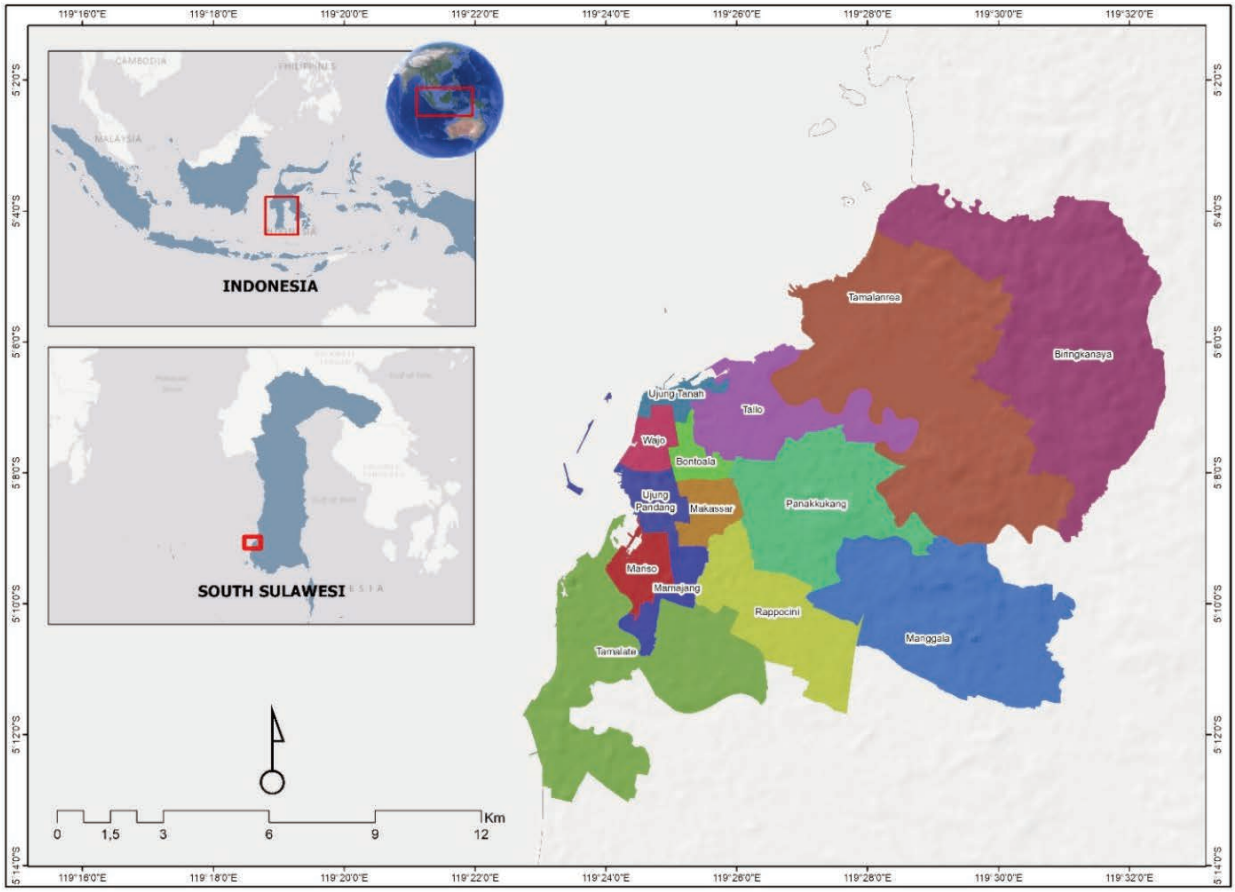

Figure 1. Mainland of Makassar City, South Sulawesi, Indonesia, with 14 districts.

the density or cover of vegetation on the land, while the NDWI is a wetness index that is calculated by combining the green wavelengths (B-3) and NIR, which is utilized to determine the land cover of an area with respect to the presence of water.

Model formulation

The LHBSTCARM demonstrates the capacity to detect spatial clusters characterized by distinct response patterns attributable to their specific geographic and temporal contexts (Lee & Lawson, 2016). Unlike conventional approaches constrained by assumptions of spatial homogeneity, the model does not presuppose that neighbouring regions must exhibit analogous responses after accounting for specified covariates, thereby preserving flexibility in capturing localized heterogeneity.

Localized indicates that the model allows parameters to vary across space instead of assuming one global model for all locations, while hierarchical, accommodates different levels of variation (e.g., across regions or time) and the conditional autoregressive (CAR) part means that it is model is a type of spatial random effect model where the value at one location depends on the values at neighbouring locations, defined through a spatial adjacency matrix. Briefly, the LHBSTCARM approach: i) captures both spatial and temporal dependencies; ii) allows model parameters to vary across locations, with increasing local accuracy; and iii) quantifies uncertainty and can incorporate prior information by using Bayesian inference.

The crude dengue Standardized Incidence Ratio (SIR), also referred to as Relative Risk (RR), was computed as the ratio of the observed to the expected number of dengue cases in a given region and time. The expected number of dengue cases, $e_{s,t}$, is calculated based on the overall incidence rate of dengue in the Makassar region and the population at risk in each respective region and time (Eq. 2) as follow:

$$e_{s,t} = \sum_s \sum_t \frac{Y_{s,t}}{\sum_s \sum_t pop_{s,t}} . pop_{s,t} \quad \text{Eq. 2}$$

The model assumes a Poisson distribution for the reaction and is defined as follows:

$$Y_{s,t} \sim \text{Poisson}(e_{s,t}\theta_{s,t}) \quad \text{Eq. 3}$$

$$\log(\theta_{s,t}) = X_{s,t}^T \beta + u_{s,t} + \lambda_{z_{s,t}},$$

where $Y_{s,t}$ represent the observed dengue cases in the s^{th} region during the t^{th} time, where s ranges from 1 to I and t from 1 to J; $e_{s,t}$ the expected number of dengue cases; $\theta_{s,t}$ the relative risk of dengue in region s at time t, $X^T_{s,t}\beta$ the linear combination of the chosen covariates (fixed effects); $u_{s,t}$ and $\lambda_{z_{s,t}}$, serve as the latent, smoothing components (representing spatially and temporally correlated variation and the constant group component or intercept, respectively); β the vector of covariate regression parameters; and a multivariate Gaussian prior with mean zero and variance 100,000 is assumed. These priors were chosen based on the defaults of the CARBayesST package.

Data points that are spatially and temporally adjacent ($Y_{s,t}$, $Y_{k,l}$) will exhibit automatic correlation when $\lambda_{z_{s,t}} = \lambda_{z_{k,l}}$ (indicating identical intercepts), whereas they will display significantly different values when $\lambda_{z_{s,t}} \neq \lambda_{z_{k,l}}$. In the context of dengue fever incidence, $\lambda_{z_{s,t}} = \lambda_{z_{k,l}}$ permits spatially and temporally adjacent regions to exhibit significantly different probabilities of dengue fever occurrence. The intercept component of $\lambda_{z_{s,t}}$ identifies distinct groups of the areas characterized by high or low likelihood of dengue disease occurrence. It encompasses a maximum of G distinct groups organized according to priors: $\lambda_1 < \lambda_2 < \lambda_3 < \ldots < \lambda_G$.

$$\lambda_k \sim \text{Uniform}(\lambda_{k-1}, \lambda_{k+1}) \text{ for } k = 1, 2, \ldots, G \quad \text{Eq. 4}$$

where λ_0 equals $-\infty$ and $\lambda_{G+1} + \infty$. An observation in region s at time t is assigned to one of the G intercepts, denoted by $Z_{s,t} \in \{1,2,3, \ldots G\}$. The model specifies the value of G, which is advised to be assigned a small odd number (Lee & Sarran, 2015). The case study presented examines the values of G = 2 and G = 3. Higher values were excluded because of the limited number of regions. The value of $Z_{s,t}$ is adjusted towards the central intercept value G^*, defined as G*: $G^* = \frac{G+1}{2}$ for odd G and $G^*: G^* = \frac{G}{2}$ for even G, via the penalty function $f(Z_{s,t})$:

$$f\left(Z_{s,t} \middle| Z_{s,t-1}\right) = \frac{\exp\left(-\delta\left[\left(Z_{s,t}-Z_{s,t-1}\right)^2 + \left(Z_{s,t}-G^*\right)^2\right]\right)}{\sum_{r=1}^{G} \exp\left(-\delta\left[\left(r-Z_{s,t-1}\right)^2 + \left(r-G^*\right)^2\right]\right)}, \text{ for } t = 2, \ldots, J \quad \text{Eq. 5}$$

$$f(Z_{s,1}) = \frac{\exp\left(-\delta\left(Z_{s,1}-G^*\right)^2\right)}{\sum_{r=1}^{G} \exp\left(-\delta\left(r-G^*\right)^2\right)},$$

where, δ represents the penalty parameter following a uniform distribution defined by the bounds L_δ and U_δ. In the following case study, the hyperparameters L_δ and U_δ are established at 1 and 100, respectively, following the recommendations by Lee and Lawson (2016). The smoothing component $u_{s,t}$ is modelled using a multivariate autoregressive process characterized by spatial autocorrelation ρ_S=1, which aligns with the Intrinsically Conditionally Autoregressive (ICAR) model (the first level of the hierarchy)

$$u_s \sim N(0, \tau^2 Q(W, \rho_S)^{-1}), (u_s | u_{t-1}) \sim N(\rho_T u_{t-1}, \tau^2 Q(W, \rho_S)^{-1}), t = 2, \ldots, J. \quad \text{Eq. 6}$$

The variance component τ^2 follows an Inverse Gamma (IG) distribution, representing the second level of the hierarchical model. Four hyperpriors are considered for the variance component τ^2: IG(0.25, 0.0005), IG(0.50, 0.0005), IG(0.75, 0.0005), and IG(1.00, 0.0005), $\rho_T \sim U(01,)$. Parameter values were determined under empirical guidelines derived from methodological precedents established in prior research, ensuring alignment with rigorously validated practices (Lee & Lawson 2016). The value ρ_s=1 enforces significant spatial smoothing on μ_t, thereby regulating each distinct value (gradual change) on the surface $\lambda_{z_{s,t}}$.

Results

Monthly DHF cases

Figures 2 and 3 depict that each district exhibits a comparable trend (Figure 2) alongside significant variability (Figure 3). The temporal patterns are evident in all domains indicating the presence of a temporal component in the data. Consequently, a spatio-temporal modelling approach was deemed appropriate for analyz-

ing this dataset, as it effectively captures the spatial and temporal dynamics inherent in the observed patterns.

Figures 2 and 3 generally illustrate an upward trend in cases, beginning in January and peaking in May for each year. Subsequently, the trend declines steadily until June, reaching its lowest point between September and November. The highest incidence was approximately 70 cases reported in March 2024, primarily in the districts of Rappocini and Tamalate. In addition, the districts of Rappocini, Tamalate, Panakukang, and Biringkanaya record between 58 and 71 monthly cases. In contrast, no cases were recorded in October and November for each year.

Evaluation of the model without covariates

The LHBSTCARM was evaluated for component configurations G=2 and G=3. Model selection was conducted using the Watanabe–Akaike Information Criterion (WAIC) and the Mean Squared Error (MSE) to identify the optimal configuration. As illustrated in Table 1.

Based on Table 1, the WAIC and MSE values for G=2 were 2483 and 6.508, respectively, whereas those for G=3 were 2460 and 6.169, respectively. The result demonstrates that G=3 yields lower WAIC and MSE values compared to G=2, indicating a superior model fit and predictive accuracy. Consequently, the LHBSTCARM with G=3 was identified as the most statistically optimal configuration.

The main output of the model without covariates is the district-level RR, which serves as the basis for quantifying spatial risk variation. As shown in Table 2, the LHBSTCARM configuration with G=2 delineated four districts as high-risk, classifying the remaining districts as low-risk. In contrast, the G=3 configuration identified four districts as low-risk, one district as high-risk and assigned the remaining districts as moderate-risk classification. These results underscore the enhanced granularity of the G=3 model, which introduces a moderate-risk category that is absent in the G=2 configuration, thereby enabling a more nuanced spatial risk assessment. Based on Table 2, a total of 418 months were classified into the medium-level group under the G = 3 clustering scenario, while 27 months were categorized as low-level and only 3 months as high-level. Ujungpandang was the only district with three months classified as high. The table displays a clustering result with a predominant classification rate of 93%.

Monthly indices

The index plot is informed by using descriptive plots, which are spatial combined with time-series plots (or spatial and temporal plots), histograms, and boxplots. Two different types of spatial and temporal plots are presented as individual and descriptive statistics. The NDBI of Makassar (Figure 4) displays a proclivity towards thin vegetation and open land with minimal buildings (median NDBI = 0.06492), distributed from index -0.38195 (minimum or no built-up area) to 0.20358 (maximum or medium density built-up area). Figure 5 illustrates a fragile vegetation index (median NDVI = 0.17456) distributed across a considerable range, from -0.05259 (minimum, non-vegetation) to 0.57786 (maximum,

Table 1. Localized hierarchical Bayesian spatiotemporal conditional autoregressive model: comparative outcome without covariates.

	Lambda1 Outcome (CI)	Lambda2 Outcome (CI)	Lambda3 Outcome (CI)	WAIC	MSE
M1 (G=2)	-0.1534 (-0.3744, -0.0410)	0.8060 (0.4143, 1.3232)	-	2483.020	6.508
M2 (G=3)	-0.9040 (-1.0060, -0.8048)	-0.0133 (-0.0803, 0.0546)	0.5502 (0.4773,0.6360)	2460.622	6.169

Lambda quantifies how much the value of a variable at a given location depends on the values at neighbouring locations; M1, Model 1; M2, Model 2; G, the number of clusters: CI, credible interval; WAIC, Watanabe–Akaike information criterion; MSE, mean squared error.

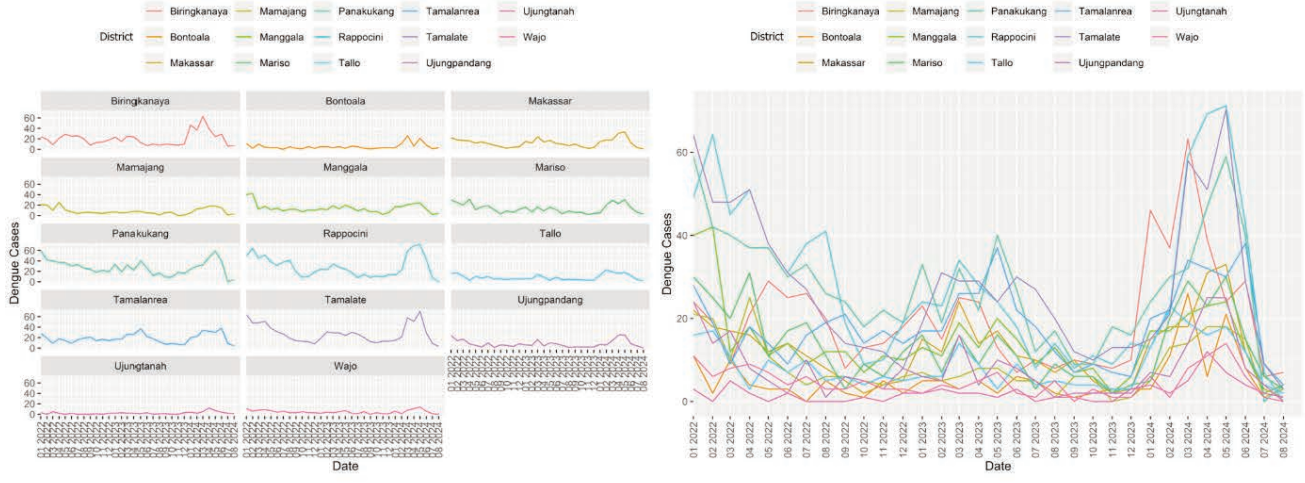

Figure 2. Number of dengue cases across the14 districts of Makassar, Jan 2022 – Aug 2024.

Table 2. Recapitulation of the number of months each district was classified into each group under G=2 and G=3 cluster scenarios (Total of 32 months).

District	G=2		G=3		
	1	2	1	2	3
Biringkanaya	32		3	29	
Bontoala	31	1	7	25	
Makassar	32			32	
Mamajang	32			32	
Manggala	32			32	
Mariso	32			32	
Panakukang	20	12		32	
Rappocini	21	11		32	
Tallo	32		8	24	
Tamalanrea	32			32	
Tamalate	32			32	
Ujungpandang	10	22		29	3
Ujungtanah	32		9	23	
Wajo	32			32	

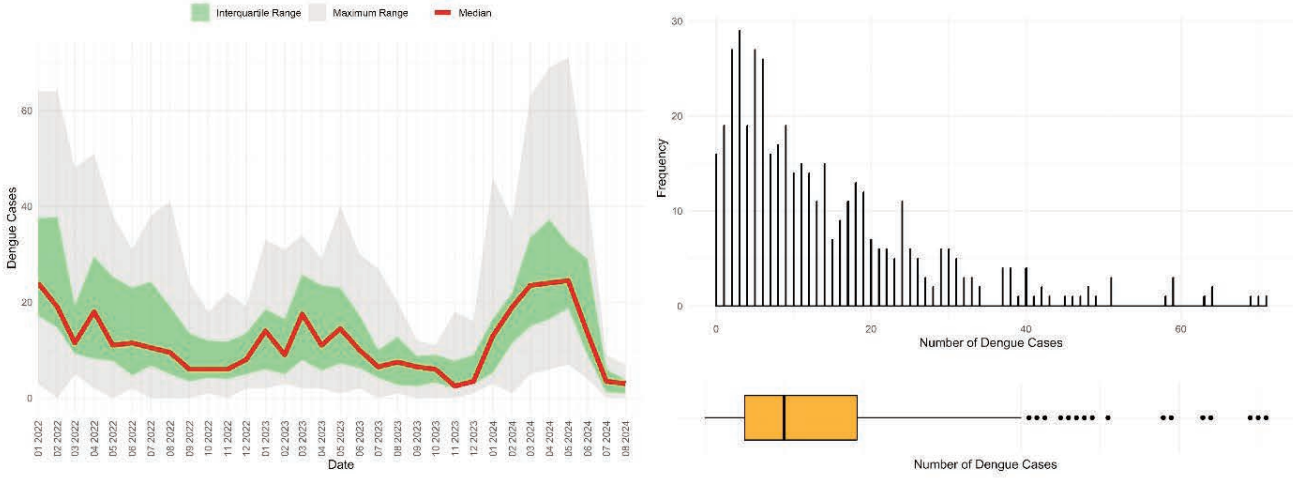

Figure 3. Descriptive analysis for dengue cases in Makassar, Jan 2022 – Aug 2024.

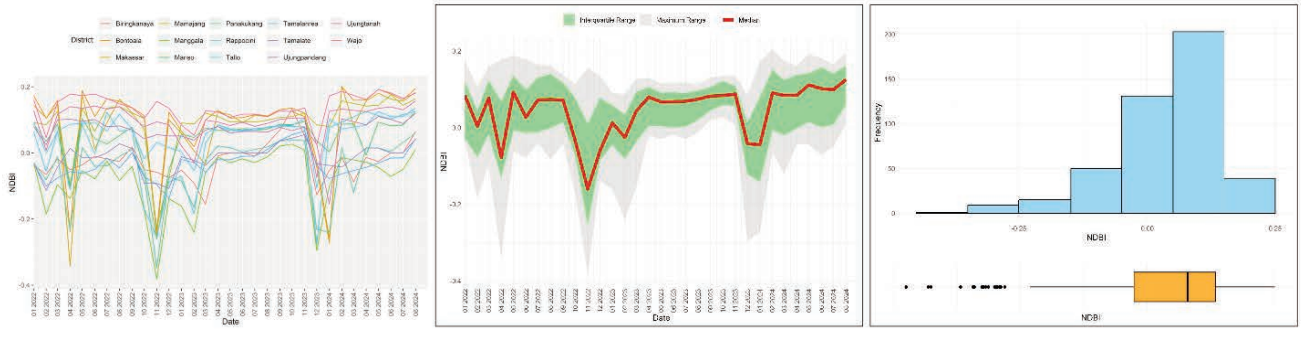

Figure 4. Descriptive plot of NDBI in Makassar, January 2022 – August 2024.

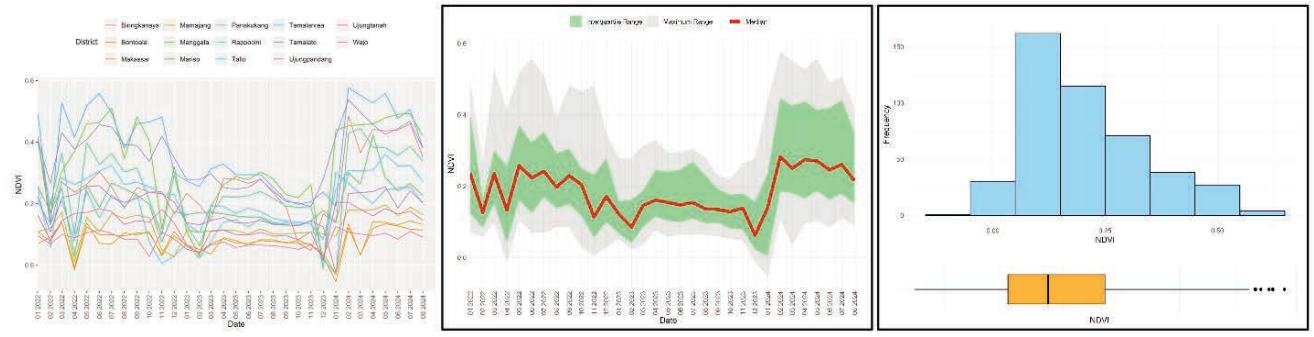

Figure 5. Descriptive plot of NDVI in Makassar. January 2022 – August 2024.

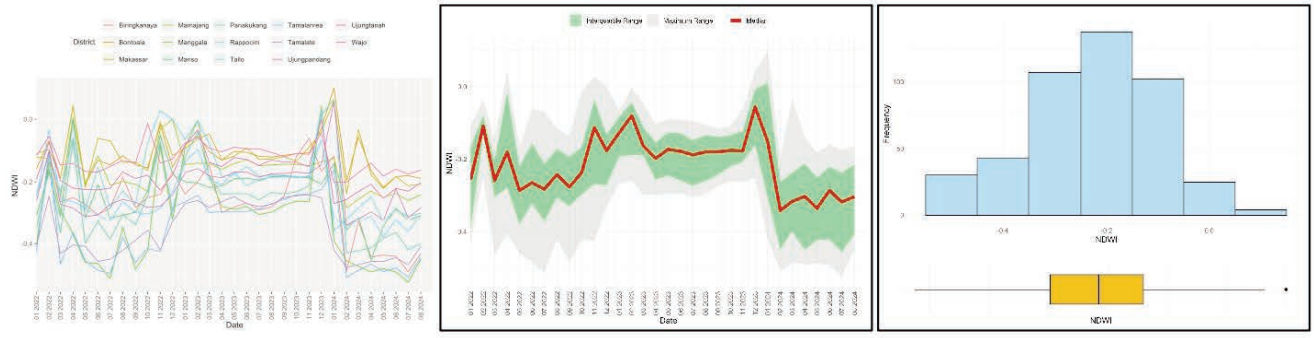

Figure 6. Descriptive plot of NDWI in Makassar, January 2022 – August 2024.

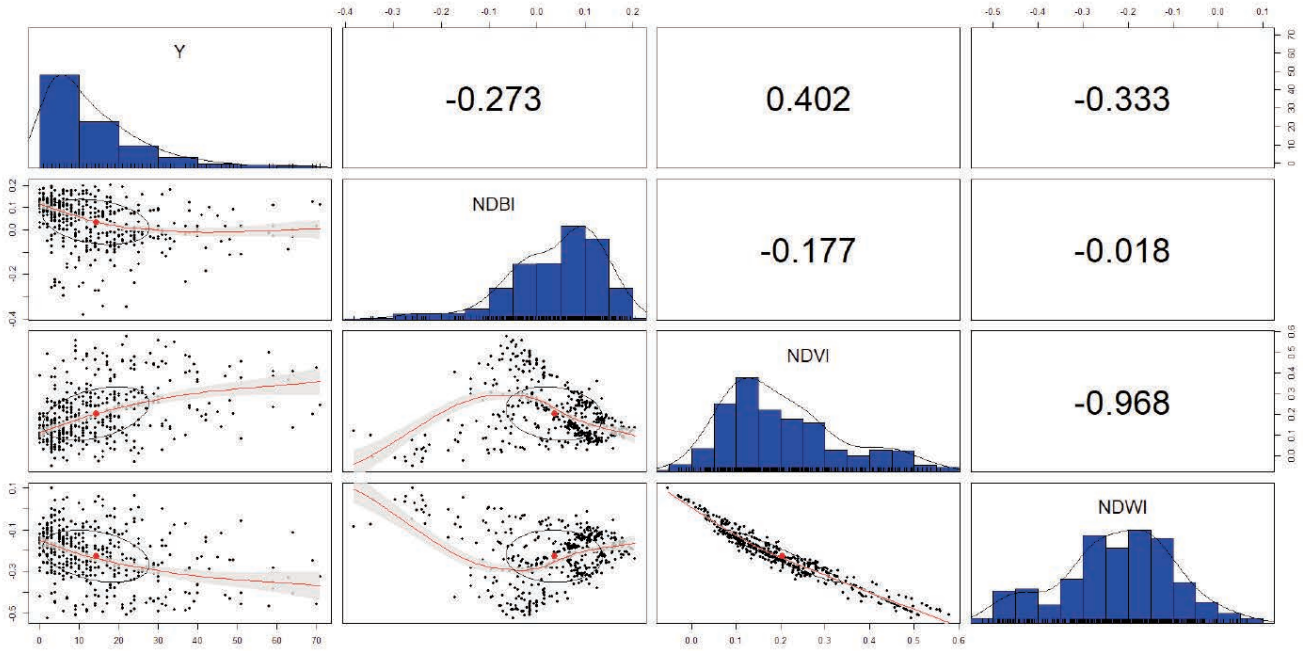

Figure 7. Correlation matrix and pairs plot of matrix correlation (upper right), histogram (diagonally), and pairs scutter (lower left).

healthy, moderate vegetation). The observed range is substantial (0.63039), indicating that the vegetation levels between districts vary considerably, from no vegetation to thin vegetation. Figure 6 illustrates that the city of Makassar is characterized by dry conditions, with a median value of -0.21538. These conditions span a considerable range, from a minimum index of -0.52384 (indicating very dry non-water) to a maximum index of 0.09987 (indicating slightly moist soil). The range is notable, spanning 0.62371 from dry to moderately moist soils. Consequently, the city's districts exhibit a general dryness with varying soil moisture levels.

Relationship between DHF and the remotely sensed indices

The Pearson correlation was applied to detect the relationship between all study variables, particularly between dengue cases and indices derived from the Sentinel-2-generated data (Table 3, Figure 7)

The significant negative correlation between DHF on the one hand and NDBI and NDWI on the other reveals that an increase in DHF cases can be observed concurrently with a decline in these indices, including building density and land covered by water. Nevertheless, the strength of the relationship remained relatively low ($-0.5 < \rho < 0.5$). Conversely, the positive correlation between DHF and NDVI was also only weakly significant, suggesting that an increase in the vegetation index would likely precede an increase in DHF incidence. Figure 7 presents a scatter diagram in the lower triangle, which visualizes the relationship between the variables and demonstrates both linear and nonlinear relationships. A linear relationship was evident between the incidence of DHF

Table 3. Correlation between DHF and covariates by month.

Pearson's correlation	Dengue	NDBI	NDVI
NDBI	-0.273 (4.303e-09)		
NDVI	0.402 (0.000e+00)	-0.177 (1.650e-04)	
NDWI	-0.333 (4.341e-13)	-0.018 (7.068e-01)	-0.968 (0.000)

Cell contents: Pearson's correlation (p-value).

Table 4. Difference between the model's predicted values and the actual values for various mean priors at G=3.

Model	Covariate	Mean	2.50%	97.50%	DIC	WAIC	RMSE	Hyperprior
1	NDBI + NDVI + NDWI	-0.303 -2.315 -2.220	-1.082 -3.995 -4.047	0.385 -0.427 -0.156	2345.095	2460.562	2.499	IG(0.25,0.0005)
2	NDBI + NDVI + NDWI	-0.348 -2.331 -2.220	-1.181 -4.053 -3.990	0.481 -0.480 -0.290	2334.540	2461.996	2.537	IG(0.50,0.0005)
3	NDBI + NDVI + NDWI	-0.085 -1.863 -1.684	-0.862 -4.096 -3.834	0.551 -0.362 -0.041	2351.077	2457.207	2.581	IG(0.75,0.0005)
4	NDBI + NDVI + NDWI	-0.284 -2.289 -2.213	-1.333 -4.185 -4.107	0.400 -0.476 -0.298	2346.963	2466.959	2.599	IG(1.00,0.0005)

WAIC, Watanabe–Akaike information criterion; RMSE, root mean squared error; DIC, deviance information criterion; IG, inverse gamma.

Table 5. The result of HBSTCARL G=3.

	$\hat{\beta}$	2.5%	97.5%	DIC	WAIC	MSE	RMSE
NDBI	-0.0849	-0.8618	0.5508	2351.077	2457.207	6.664	2.581
NDVI	-1.8625	-4.0963	-0.3619				
NDWI	-1.6840	-3.8340	-0.0409				
λ_1	-0.8825	-0.9944	-0.7730				
λ_2	0.0049	-0.0531	0.0651				
λ_3	0.5716	0.5025	0.6383				
δ	1.1329	1.0075	1.3519				
$\tau 2$	0.0185	0.0086	0.0337				
ρT	0.6205	0.3475	0.8057				

NDBI, normalized difference built-up index; NDVI, normalized difference vegetation index; NDWI, normalized difference water index; DIC, deviance information criterion; WAIC, Watanabe–Akaike information criterion; MSE, mean squared error; RMSE, root mean squared error; λ, the intercept for the appropriate cluster, where $\lambda_1 < \lambda_2 < \lambda_3$; δ, the penalty parameter; τ, the variance parameter associated with the spatial random effects u_1; ρ_T = temporal autocorrelation parameter.

and the covariates, and a linear trend was also observed between the NDVI and NDWI indices. The NDBI displayed a nonlinear relationship with both the NDVI and NDWI indices. The histogram (diagonal elements) indicated that all indices exhibited a distribution pattern close to normal. The upper triangle showed the Pearson correlation values between the variables, with the highest correlation between NDWI and NDVI (-0.968) and the lowest between NDWI and NDBI (-0.018).

Modelling with covariates

The LHBSTCARM G=3 was selected as the optimal model for predicting DHF based on environmental aspects indicated by the satellite imagery, whereby the elements of vegetation (NDVI), groundwater (NDWI) and building density (NDBI) were considered. This reasoning is expected to facilitate the prediction of DHF incidence, ensuring a faster, low-cost, rapid and efficient approach. To determine the optimal model, the model G=3 was evaluated using four different hyperprior settings for the inverse gamma (IG) distribution. Model 3, with hyperprior IG (0.75, 0.0005), yielded the lowest WAIC value (2457.207) indicating the best fit among the tested models (Table 4). In contrast, models 1, 2 and 4 showed higher WAIC values (2460.562, 2461.996 and 2466.959, respectively). Table 4 shows that in model 3, the coefficients for NDVI and NDWI were statistically significant and less than -2, while the NDBI coefficient was not significant. This suggests that vegetation and groundwater presence negatively influence DHF incidence, whereas building density (NDBI) does not show a statistical effect (Table 4 and Table 5).

The model assumes DHF cases follow a Poisson distribution. The odds ratio (OR) for NDBI was approximately 0.9186 (count: $OR_{NDBI} = e^{-0.0848} \approx 0.9186$), suggesting an 8.1% (count: 1 – 0.9186) decrease in DHF incidence for each unit increase in building density. However, as the 95% Confidence Interval (CI) for NDBI includes zero, the effect is not statistically significant and may serve more as a precautionary indicator than a definitive predictor. The covariate NDVI, representing an environmental vegetation element, has an odd ratio of 0.1555 (count: $OR_{NDBI} = e^{-1.8625}$), indicating that with a one-unit increase in NDVI, the average incidence rate decreases by approximately 84.5% (count: 1 – 0.1555). This represents a significant decrease (95% CI: -4.0963,-0.3619), suggesting that areas with higher vegetation (higher NDVI) tend to have lower incidence. The covariate NDWI, which is an element of water or environmental wetness, had an OR of 0.1857 (count: $OR_{NDBI} = e^{-1.6840}$), which indicates that with a one-unit increase in NDWI, the average incidence rate decreased by approximately 81.4% (count: 1 - 0.1857). This is a statistically significant result, suggesting that areas with higher humidity (as indicated by high NDWI values) or in close proximity to water bodies tend to have lower incidence rates.

Discussion

DHF is exclusively a consequence of dengue mosquito bites and is not transmitted from human to human (Islam *et al.,* 2021), with the exception of instances involving blood transfusions. Consequently, modelling DHF through environmental factors as the primary medium for mosquito breeding has been extensively researched. This study made use of environmental proxy variables derived from satellite imagery, specifically NDBI, NDVI, and NDWI. Consequently, the application of the resulting model must take the fundamental assumptions pertaining to the index domain of Makassar City as a location for the prototype model. The LHBSTCARM at G=3 was found to be a most suitable model for dengue cases in Makassar City through the covariates of NDBI, NDVI and NDWI variables.

Transmission dynamics of infectious diseases are profoundly shaped by environmental determinants as shown by many authors (Louis *et al.,* 2014; Yu *et al.,* 2016; Ziemann *et al.,* 2018; Islam *et al.,* 2021), a relationship corroborated by this study's empirical findings, where the critical role of ecological variables in modulating dengue fever epidemiology within urban tropical settings is underscored. This study further reinforced existing evidence that environmental conditions, particularly vegetation density and aquatic habitats, are pivotal drivers of DHF transmission. Theoretically, *Aedes* mosquito proliferation is contingent on the availability of stagnant water reservoirs, such as ephemeral puddles or artificial containers, which serve as primary breeding sites. Such microhabitats not only support oviposition but also foster vegetation growth, creating shaded, nutrient-rich environments that enhance larval survival rates (Islam *et al.,* 2021). Notably, the NDWI exhibited an optimal threshold (0.00 – 0.20) for larval viability, beyond which hydrological fluctuations may disrupt larval development, as posited in prior research (Islam *et al.,* 2021). These findings align with ecological frameworks linking land-use patterns to arboviral disease emergence, emphasizing the need for targeted vector control strategies in hydrologically dynamic regions.

Makassar was selected as the pilot study area due to its district's environmental and urban characteristics, which reflect both the commonalities and heterogeneity observed across Indonesian cities. While the city exhibits ecological patterns representative of broader Indonesian urban ecosystems (e.g., tropical climate, anthropogenic land-use shifts), its localized hydrological and vegetational heterogeneity necessitated model calibration to region-specific parameters. The model G=3 framework was thus operationalized to delineate spatially relevant zones for analysis using Makassar's geospatial indices, specifically by the thresholds for the NDBI, NDVI, and NDWI. This methodological simplification ensured alignment with Makassar's unique environmental gradients while maintaining analytical rigor.

The NDVI of Makassar displays a range from -0.053 to 0.578, indicating a predominantly dry and unhealthy vegetation cover. The distribution is right-skewed (Figure 7) and it can be concluded that the vegetation of Makassar spread from non-vegetation to moderately vegetated areas (Hatulesila *et al.,* 2019). The NDWI, on the other hand, demonstrates a near-normal distribution, with a median value of -0.21538, which indicates that the city of Makassar is characterized by dry conditions, with the presence of moist land and minimal stagnant water suggesting that a reduction in water coverage may lead to an increase in the number of dengue patients.

The NDBI has no statistically significant impact indicating that, in general, Makassar is situated in open land areas (building enclosures below medium), although there are some built-up or residential areas (NDBI > 0.20). The presence of a negative skewness indicates that there are more enclosed buildings than there are unenclosed ones. Consequently, the insignificant negative influence previously identified still has the potential to impact DHF presence in Makassar City. Low-density residential areas have been observed to contribute to the increase of DHF outbreaks due to the availability of unattended spaces, such as plastic cups, bar-

rels, and other containers, which can serve as breeding grounds for larvae.

Conclusions

The water (NDWI) and vegetation (NDVI) indices have a considerable influence on the occurrence of dengue outbreaks in Makassar City. However, it would also be prudent to consider the building enclosure index (NDBI) as a precautionary measure. The characteristics indicate that the a region has minimal water coverage and scarcity of vegetation but there are also areas with sparse vegetation cover that meet the criteria for healthy vegetation and few areas with healthy vegetation. The investigation found the LHBSTCARM with G=3 most appropriate for the analysis of dengue cases in Makassar City predicting that a one-unit increase in NDVI would reduce the incidence of DHF by 85.5%, and the same increase in NDWI would reduce the incidence of DHF by 81.4%. The research findings can become instrumental in providing an early warning of potential dengue incidence in the forthcoming months.

References

Alghamdi T, Elgazzar K, Sharaf T. 2021. Spatiotemporal traffic prediction using hierarchical bayesian modeling. Futur Internet 9:1–6.

Anggraeni F, Mahmudah N. 2021. Bayesian Spatial Survival Lognormal 3 Parameter models for event processes dengue fever in Tuban. IAENG Int J Appl Math 51:1–8.

Ardiansyah M, Wijayanto H, Kurnia A, Djuraidah A. 2023. Geo-additive mixed model with variable selection using the adaptive elastic net to handle nonresponse in official rice productivity survey. Spat Stat 56:100761.

Aswi A, Cramb S, Duncan E, Hu W, White G, Mengersen K. 2020a. Climate variability and dengue fever in Makassar, Indonesia: Bayesian spatio-temporal modelling. Spat Spatiotemporal Epidemiol. 33;100335.

Aswi A, Cramb S, Duncan E, Hu W, White G, Mengersen K. 2020b. Bayesian spatial survival models for hospitalisation of Dengue: A case study of Wahidin hospital in Makassar, Indonesia. Int J Environ Res Public Health 17:0878.

Aswi A, Sukarna S, Cramb S, Mengersen K. 2021. Effects of climatic factors on dengue incidence: a comparison of bayesian spatio-temporal models. J Phys Conf Ser 1863:012050.

Baharom M, Ahmad N, Hod R, Manaf MRA. 2022. Dengue early warning system as outbreak prediction tool: a systematic review. Risk Manag Healthc Policy 15:871–86.

Bergquist R, Manda S. 2019. The world in your hands: Geohealth then and now. Geospat Health 14:779.

BPS-Makassar. 2024. Kota Makassar dalam Angka 2024. Makassar.

Chen ZY, Deng XY, Zou Y, He Y, Chen SJ, Wang QT, Xing DG, Zhang Y. 2023. A Spatio-temporal Bayesian model to estimate risk and influencing factors related to tuberculosis in Chongqing, China, 2014–2020. Arch Public Heal 81:1–12.

Dian Maya Sari, Sarumpaet SM, Hiswani. 2018. Determinan Kejadian Demam Berdarah Dengue (DBD) di Kecamatan Medan Tembung. J Kesehat Pena Med 8:9–25.

Faridah L, Suroso DSA, Fitriyanto MS, Andari CD, Fauzi I, Kurniawan Y, Watanabe K. 2022. Optimal validated multi-factorial climate change risk assessment for adaptation planning and evaluation of infectious disease: a case study of dengue hemorrhagic fever in Indonesia. Trop Med Infect Dis 7:0172.

Fitri DJ, Djuraidah A, Wijayanto H. 2024. bayesian conditional negative binomial autoregressive model: a case study of stunting on Java Island 2021. Commun Math Biol Neurosci 2024:8281.

García JA, Acero FJ, Portero J. 2023. A Bayesian hierarchical spatio-temporal model for extreme temperatures in Extremadura (Spain) simulated by a regional climate model. Clim Dyn 61:1489–503.

Hatulesila JW, Mardiatmoko G, Irwanto I. 2019. Analisis Nilai Indeks Kehijauan (NDVI) pada Pola Ruang Kota Ambon, Provinsi Maluku. J Hutan Pulau-Pulau Kecil 3:55–67.

Hii YL, Zhu H, Ng N, Ng LC, Rocklo J. 2012. Forecast of dengue incidence using temperature and rainfall. PLoS Negl Trop Dis 6:1908.

Islam MT, Quispe C, Herrera-Bravo J, Sarkar C, Sharma R, Garg N, Fredes LI, Martorell M, Alshehri MM, Sharifi-Rad J, et al. 2021. Production, transmission, pathogenesis, and control of dengue virus: a literature-based undivided perspective. Biomed Res Int 2021:1–23.

Khaidir K, Zara N, Ikhsan R. 2022. Gambaran Penyakit Demam Berdarah Dengue di Poliklinik Umum Puskesmas Muara Batu Aceh Utara. Galen J Kedokt dan Kesehat Mhs Malikussaleh 1:44

Kurniawati Y, Wijayanto H, Kurnia A, Domiri DD, Susetyo B. 2023. Selection of multinomial logit models based on accuracy reclassification of the area sampling frame labels. Sci Technol Asia 28:18–30.

Kusumaningrum D, Wijayanto H, Kurnia A, Notodiputro KA, Ardiansyah M, Parvez IM. 2024. Four-parameter beta mixed models with survey and sentinel 2A satellite data for predicting paddy productivity. Smart Agric Technol 9:100525.

Lee D, Sarran C. 2015. Controlling for unmeasured confounding and spatial misalignment in long-term air pollution and health studies. Environmetrics 26:477-87.

Lee D, Lawson A. 2016. Quantifying the spatial inequality and temporal trends in maternal smoking rates in Glasgow. Ann Appl Stat 10:1427–46.

Lian Q, Sun W, Dong W. 2023. Hierarchical spatial-temporal neural network with attention mechanism for traffic flow forecasting. Appl Sci. 1:9729.

Louis VR, Phalkey R, Horstick O, Ratanawong P, Wilder-Smith A, Tozan Y, Dambach P. 2014. Modeling tools for dengue risk mapping - a systematic review. Int J Health Geogr 13:1–15.

Martheswaran TK, Hamdi H, Barty A Al, Zaid AA, Das B. 2022. Prediction of dengue fever outbreaks using climate variability and Markov chain Monte Carlo techniques in a stochastic susceptible - infected - removed model. Sci Rep 12:1–17.

Moreno-Madriñán MJ, Crosson WL, Eisen L, Estes SM, Estes MG, Hayden M, Hemmings SN, Irwin DE, Lozano-Fuentes S, Monaghan AJ, et al. 2014. Correlating remote sensing data with the abundance of pupae of the dengue virus mosquito vector, Aedes Aegypti, in Central Mexico. ISPRS Int J Geo-Information 3:732–49.

Muhsoni FF. 2015. Penginderaan Jaun (Remote Sensing). Bangkalan-Madura: UTMPRESS. Available from: https://msp.trunojoyo.ac.id/wp-content/uploads/2018/07/2015_Penginderaan-Jauh_Firman-Farid-Muhsoni.pdf.

Mukhopadhyay S, Ogutu JO, Bartzke G, Dublin HT, Piepho HP.

2019. Modelling spatio-temporal variation in sparse rainfall data using a hierarchical Bayesian regression model. J Agric Biol Environ Stat 24:369–393.

Rasjid A, Yudhastuti R, Notobroto HB, Hartono R. 2019. Climate change: an overview of the prevalence of dengue hemorrhagic fever in the South Sulawesi Province of Indonesia. Indian J Public Heal Res Dev 10:1982–6.

Silitonga P, Bustamam A, Muradi H, Mangunwardoyo W, Dewi BE. 2021. Comparison of dengue predictive models developed using artificial neural network and discriminant analysis with small dataset. Appl Sci 11:1–16.

Solís-Navarro M, Vargas-De-León C, Gúzman-Martínez M, Corzo-Gómez J. 2022. A Bayesian prediction spatial model for confirmed dengue cases in the State of Chiapas, Mexico. Hindawi J Trop Med 2022:1–13.

Song J, Liu J, Zhang X, Chen X, Shang Y, Gao F. 2024. Spatio-Temporal fluctuation analysis of ecosystem service values in northeast china over long time series: based on Bayesian hierarchical modeling. Land 13:1–20.

Suhet. 2015. SENTINEL-2: user handbook. Hoersch B, editor. European Space Agency.

Sukarna S, Notodiputro KA, Sartono B. 2023. Comparison between binomial GLMM and binomial GMET for temporary unemployment in West Java, Indonesia. Adv Comput Sci Res 109:198-209.

Sukarna, Wahyuni MS, Syam R. 2021. Comparison of Bayesian Spatio-temporal Models of Tuberculosis in Makassar, Indonesia. J Phys Conf Ser 2123012048.

Supangat U, Badriah DL, Mamlukah M, Suparman R. 2023. Faktor-Faktor yang Berhubungan dengan Kematian Kasus Demam Berdarah di Kota Tasikmalaya 2022. J Heal Res Sci 3:63–71.

Susanti, Suharyo. 2017. Hubungan Lingkungan Fisik dengan Keberadaan Jentik Aedes pada Area Bervegetasi Pohon Pisang. Unnes J Public Heal 6:271–7.

WHO. 2004. Disease Outbreak News: 2004 - Indonesia. WHO Homepage., forthcoming. [accessed 2024 Nov. 1]. https://www.who.int/emergencies/disease-outbreak-news/item/2004_05_11a-en.

WHO. 2024. Disease Outbreak News: Dengue - Global situation. WHO Homepage., forthcoming. [accessed 2024 Nov. 1]. https://www.who.int/emergencies/disease-outbreak-news/item/2024-DON518.

Yu H-L, Lee C-H, Chien L-C. 2016. A spatiotemporal dengue fever early warning model accounting for nonlinear associations with hydrological factors : a Bayesian maximum entropy approach. Stoch Environ Res Risk Assess 30:2127–41.

Ziemann A, Fairchild G, Conrad J, Manore C, Parikh N, Del Valle S, Generous N. 2018. Predicting dengue incidence in Brazil using broad-scale spectral remote sensing imagery. Int Geosci